Commun. Math. Phys. 135, 379–399 (1991)

Communications in
Mathematical
Physics

Spectral Properties of a Tight Binding Hamiltonian with Period Doubling Potential

Jean Bellissard[1,2], **Anton Bovier**[3], **and Jean-Michel Ghez**[2,4]

[1] Centre de Physique Théorique, CNRS, Luminy Case 907, F-13288 Marseille Cedex 9, France
[2] Université de Provence, Marseille, France
[3] Physikalisches Institut, Universität Bonn, W-5300 Bonn, Federal Republic of Germany
[4] PHYMAT, Département de Mathématiques, Université de Toulon et du Var, F-83957 La Garde Cedex, France

Received May 7, 1990; in revised form July 16, 1990

Abstract. We study a one dimensional tight binding hamiltonian with a potential given by the period doubling sequence. We prove that its spectrum is purely singular continuous and supported on a Cantor set of zero Lebesgue measure, for all nonzero values of the potential strength. Moreover, we obtain the exact labelling of all spectral gaps and compute their widths asymptotically for small potential strength.

I. Introduction

The discovery of the quasi-crystalline phase in AlMn by Schechtman et al. [1] (see' e.g. [2] for a review) has provoked an increasing interest in physical systems that are neither periodic nor random, i.e. neither crystalline nor amorphous. Besides the host of experimental studies, there have been considerable efforts to study theoretically the properties of such systems. Most results are so far confined to the one-dimensional case, most notably to the one dimensional tight-binding hamiltonian

$$H_V = -\Delta + \hat{V}, \tag{1.1}$$

where $\hat{V}$ is a diagonal matrix whose diagonal elements are given by some aperiodic sequence V_n. The problem of studying the spectral properties of such operators turns out to be an extremely interesting mathematical problem in itself, and there is by now a considerable amount of literature, both heuristic and rigorous, dealing with it. There is a considerable amount of general results concerning the case where V_n is quasiperiodic [3, 1, 4]. Among these, the Fibonacci sequence has attracted special attention [5], and it has been proven by Sütö [6] and by Bellissard et al. [7] that, for any nonzero value of the potential strength, the spectrum of H_V in this case is purely singular continuous. The proofs make use of the existence of what is called the trace map [8], a recursion formula for the trace of the transfer matrix (to be discussed below) that in some sense expresses the existence of an exact renormalization group structure in this model. This property is shared by a large class of sequences, in particular sequences obtained from substitutions [9] (or "automatic sequences"). The example of the Thue-Morse sequence has been mostly studied [10, 11], but in a recent paper Luck [12] has considered a number of examples and has presented some interesting conjectures concerning the spectral properties of the associated hamiltonians, including rather detailed information on the labelling and opening of gaps.

From these studies one is led to expect the following picture for all models of this class: First, they should have singular continuous spectrum as soon as the potential is turned on; the particular structure of the Cantor set supporting the spectrum depends on the sequence chosen, with spectral gaps labelled by some set of algebraic numbers. This set can be determined using K-theory, as shown by Bellissard [13].

In this note we present a rather complete and detailed analysis for the case of the "period doubling sequence" [14]. Apart from proving that the general picture above is correct in this particular case, we give a precise description of the spectral gaps. This will be seen to confirm the predictions of Luck [12] for this example.

As in previous rigorous works, our basic tool in investigating the spectrum of H_V will be the "trace map," a recursion formula for the transfer matrix associated with H. We will show that the investigation of this dynamical system will suffice to obtain all desired information on the spectrum of H_V.

The period doubling sequence belongs to a class of "automatic sequences" obtained from a certain substitution rule on the letters of a given alphabet (see, e.g. [9]). In the case of the period doubling sequence, the alphabet consists of two letters, say A and B and the substitution rule is given as

$$A \to AB, \qquad B \to AA. \tag{1.2}$$

Thus, starting from the letter A or B, respectively, after n applications of the substitution rule we arrive at sequences of length 2^n, denoted by $A^{(n)}$ and $B^{(n)}$. The period doubling sequence is the limiting infinite sequence $A^{(\infty)}$ obtained from this process. Finally, a numerical sequence ε_i above is obtained from this one by replacing A with $+1$ and B with -1 in $A^{(\infty)}$. The diagonal elements V_i of the potential in (1.1) are then defined as $V_i = V_{-1-i} = \varepsilon_i V$.

By construction the partial sequences satisfy the recursion relation

$$\begin{aligned} A^{(n)} &= A^{(n-1)} B^{(n-1)}, \\ B^{(n)} &= A^{(n-1)} A^{(n-1)}, \end{aligned} \tag{1.3}$$

which will be seen to furnish the basic tool to analyze our Hamiltonian. For, writing the eigenvalue equation for H_V in the conventional vector form,

$$\begin{pmatrix} \psi_E(n+1) \\ \psi_E(n) \end{pmatrix} = P(V_n, E) \begin{pmatrix} \psi_E(n) \\ \psi_E(n-1) \end{pmatrix}, \tag{1.4}$$

where

$$P(V_n, E) \equiv \begin{pmatrix} E - V_n & -1 \\ 1 & 0 \end{pmatrix}, \tag{1.5}$$

we see that the behaviour of solutions of (1.1) is governed by the transfer matrix

$$T_N(E) \equiv \prod_{k=1}^{N} P(V_k, E). \tag{1.6}$$

The recursion relations (1.3) imply a recursion relation for the transfer matrix where $N = 2^n$. Defining $T_n^{(A)}$ and $T_n^{(B)}$ as the transfer matrices corresponding to the sequences $A^{(n)}$ and $B^{(n)}$, respectively, one verifies that they satisfy

$$\begin{aligned} T_{n+1}^{(A)} &= T_n^{(B)} T_n^{(A)}, \\ T_{n+1}^{(B)} &= (T_n^{(A)})^2. \end{aligned} \tag{1.7}$$

Remarkably, this recursion for the transfer matrices gives rise to a similar relation for their traces [8]. Making use of the fact that any unimodular 2×2-matrix M satisfies

$$M^2 - \operatorname{tr}(M) M + id = 0, \tag{1.8}$$

a straightforward computation shows that if $x_n \equiv \operatorname{tr}(T_n^{(A)})$ and $y_n \equiv \operatorname{tr}(T_n^{(B)})$,

$$\begin{aligned} x_{n+1} &= x_n y_n - 2, \\ y_{n+1} &= x_n^2 - 2. \end{aligned} \tag{1.9}$$

Equation (1.9) is called the trace map. Notice that the initial values $x = x_0$ and $y = y_0$ are related to E and V by

$$x_0 = E - V, \qquad y_0 = E + V. \tag{1.10}$$

The trace map is the fundamental tool for investigating the spectrum of H_V; its analog has been used in the study of the Thue-Morse and Fibonacci sequences [11, 6, 7]. For general results on the existence and nature of trace maps, see [8].

The remainder of this paper is organized as follows: In Sect. 2 we will show that the spectrum of H_V is the complement of the set of unstable points for the dynamical system (1.9); by this we understand the set of initial conditions for which the modulus of x_n will "eventually" be larger than two. This will be used to prove that the spectrum of H_V is purely singular continuous, for any $V \neq 0$. In fact, this will follow from a general theorem stating that for a class of aperiodic sequences, comprising in particular sequences obtained from substitutions with "primitive" [9] substitution rules, the spectrum of the corresponding hamiltonians is always of measure zero, provided the Lyapunov exponent vanishes on the spectrum. This theorem is based on a lemma of Kotani [15] and a generalization of a result of Avron and Simon [3] for almost periodic sequences, using a lemma of Herman [16]. This part of our analysis is similar to those of Sütö [6] and Bellissard et al. [7].

In Sect. 3 we analyze the spectral gaps of our hamiltonian in detail. We show that there are two families of spectral gaps, one labelled by dyadic numbers and opening linearly for small V, the other by dyadic numbers divided by 3 and opening exponentially [i.e. like $\exp(-1/|V|)$]. Moreover, we show that the spectral gaps are bounded by at least differentiable curves, and we establish precise asymptotic formulas for them. Notice that we will prove directly that these are all gaps, without making reference to K-theory, as was done in a similar analysis of the gaps of the Thue-Morse hamiltonian in [11].

II. The Trace Map and the Spectrum of H_V

In this section we establish a precise relation between the trace map and the spectrum of our hamiltonian. Some soft analysis will then show that the spectrum is purely singular continuous.

We begin with the definition of the unstable set for the dynamical system given by the trace map.

Definition. A point $(x, y) \in \mathbf{R}^2$ is called unstable, if there is n_0 s.t. if $x_0 = x$, $y_0 = y$, for all $n \geqq n_0$, $|x_n| > 2$. We denote by $\mathscr{U}$ the set of all unstable points.

Remark. In [11], $\mathscr{U}$ was defined as the interior of the set of unstable points. We will see, however, that $\mathscr{U}$ is an open set.

The following lemma provides a very convenient characterization of $\mathscr{U}$ for our system. Let us introduce the sets

$$D_{\pm}^{(0)} = \{(x, y) | y > 2, \pm x > 2\}. \tag{2.1}$$

Lemma 1. *$(x, y) \in \mathscr{U}$ if and only it there exists n such that either $(x_n, y_n) \in D_{+}^{(0)}$ or $(x_n, y_n) \in D_{-}^{(0)}$.*

Proof. We show first that $(x_n, y_n) \in D_{\pm}^{(0)}$ implies instability. By definition $|x_n| > 2$ if $(x_n, y_n) \in D_{\pm}^{(0)}$, so we only need to show that $(x_{n+k}, y_{n+k}) \in D_{\pm}^{(0)}$ for all k. Now, if $(x_n, y_n) \in D_{\pm}^{(0)}$, then $|x_{n+1}| \geqq |x_n y_n| - 2 > 4 - 2$, and $y_{n+1} = x_n^2 - 2 > 4 - 2$, so $(x_{n+1}, y_{n+1}) \in D_{\pm}^{(0)}$. Thus (x, y) is unstable.

To prove that the images of any unstable point eventually enter $D_{+}^{(0)}$ or $D_{-}^{(0)}$, let (x, y) be unstable. Then there is n such that $|x_{n+k}| > 2$ for all $k \geqq 0$, and in particular $|x_{n+1}| > 2$. But if $|x_n| > 2$, then $y_{n+1} > 2$, and since $|x_{n+1}| > 2$, (x_{n+1}, y_{n+1}) is either in $D_{+}^{(0)}$ or $D_{-}^{(0)}$, which proves the lemma. □

As a first application of Lemma 1 we have:

Lemma 2. *The set of unstable points, $\mathscr{U}$, is open.*

Proof. By Lemma 1, $(x, y) \in \mathscr{U}$ implies that for some finite n, $(x_n, y_n) \in D_{+}^{(0)} \cup D_{-}^{(0)}$. Since for n fixed, x_n and y_n are analytic functions of x, y, and $D_{+}^{(0)} \cup D_{-}^{(0)}$ is open, there is a neighborhood of (x, y) those n^{th} image is in $D_{+}^{(0)} \cup D_{-}^{(0)}$, and thus unstable. □

The set of unstable points determines completely the spectrum of H_V.

Theorem 1. *E is in the spectrum of H_V, if and only if $x = E + V$ and $y = E - V$ are such that $(x, y) \in \mathscr{U}^c$. Moreover, H_V has no eigenvalues, no generalized eigenfunction tends to zero at infinity.*

Proof. In [11] it was shown that if $(x, y) \in \mathscr{U}$, then E is in the complement of the spectrum of H_V. We will now show that the converse is also true, i.e. stable points are in the spectrum. The main ideas of the proof are those used by Sütö [6]; on a technical level much more work is required. Theorem 1 will follow from the following lemma:

Lemma 3. *Assume $(x, y) \in \mathscr{U}^c$. Then $T_n^{(A)}$ does not converge to a projection.*

The proof of Lemma 3 is, unfortunately, very cumbersome and will be postponed to Appendix A. We will now show that this lemma implies the theorem.

Assume that $E \notin \sigma(H_V)$. Then the inhomogeneous equation

$$(H_V - E)\psi_E = \delta_0$$

has a solution in $l^2(\mathbf{Z})$. Then, since $\psi_E(-1)$, $\psi_E(0)$ and $\psi_E(1)$ cannot be all zero, either $\begin{pmatrix}\psi_E(1)\\ \psi_E(0)\end{pmatrix}$ or $\begin{pmatrix}\psi_E(0)\\ \psi_E(-1)\end{pmatris}$ is nonzero. In the case $\begin{pmatrix}\psi_E(1)\\ \psi_E(0)\end{pmatrix} \neq 0$, notice that

$$\begin{pmatrix}\psi_E(2^{n_l}+1)\\ \psi_E(2^{n_l})\end{pmatrix} = T_{n_l}^{(A)} \begin{pmatrix}\psi_E(1)\\ \psi_E(0)\end{pmatrix}. \tag{2.2}$$

But then

$$\|\psi_E\|_2^2 \geqq \frac{1}{2}\sum_{l=0}^{\infty} \left|\begin{pmatrix}\psi_E(2^{n_l}+1)\\ \psi_E(2^{n_l})\end{pmatrix}\right|^2 = \frac{1}{2}\sum_{l=0}^{\infty} \left|T_{n_l}^{(A)}\begin{pmatrix}\psi_E(1)\\ \psi_E(0)\end{pmatrix}\right|^2. \tag{2.3}$$

But by Lemma 3, if (x, y) were not in the unstable set, there would exist no vector Ψ_0 such that the sequence $|T_n^{(A)}\Psi_0|$ converges to zero, and thus the series $\sum \left|T_{n_l}^{(A)}\begin{pmatrix}\psi_E(1)\\ \psi_E(0)\end{pmatrix}\right|^2$ would diverge, and ψ_E would not be in $l^2(\mathbf{Z})$, contrary to our assumption.

The case $\begin{pmatrix}\psi_E(0)\\ \psi_E(-1)\end{pmatrix} \neq 0$ can be excluded by the same argument, using the reflection symmetry of the potential, i.e. $V_{-n-1} = V_n$. Thus we conclude that $(x, y) \in \mathscr{U}$.

At the same time, if $E \in \sigma(H_V)$, E cannot be an eigenvalue. For otherwise an l^2 solution of the homogeneous equation $(H_V - E)\psi_E = 0$ would exist, which is excluded by the arguments given above; it follows even that $\psi_E(n)$ does not decay at infinity. □

For some particular points we can even make more precise statements concerning the generalized eigenfunctions:

Lemma 4. *Let (x, y) be such that for some n, $x_n = y_n = 2$, and let n denote the smallest such number. Then*

(i) $T_{n+k}^{(A)} = [T_n^{(A)}]^{2^k+1}$, *and*
(ii) $E \in \sigma(H_V)$, *E is not an eigenvalue and there exists a periodic solution ψ_E of the Schrödinger equation with period 2^n. Unless $V = 0$, there is only one such solution.*

Proof. The cases $n = 0$ and $n = 1$ are trivial, as they correspond to $E = \pm 2$ and $V = 0$ and thus the free laplacian. For $n \geqq 2$ notice that $x_n = y_n = 2$ implies that $x_{n-2} = 0$. Thus, using (1.8) $(T_{n-2}^{(A)})^2 = -id$ and hence $T_{n-1}^{(B)} = -id$, $T_n^{(A)} = -T_{n-1}^{(A)}$, and $T_n^{(B)} = (T_n^{(A)})^2$. It is thus obvious that $T_{n+k}^{(A)}$ will be a power of $T_n^{(A)}$, and a simple calculation shows that the power is the one given in (i).

Now, since $\mathrm{tr}(T_n^{(A)}) = 2$, there is a basis such that

$$T_n^{(A)} = \begin{pmatrix}1 & a_n\\ 0 & 1\end{pmatrix},$$

and thus

$$T_{n+k}^{(A)} = \begin{pmatrix}1 & (2^k+1)a_n\\ 0 & 1\end{pmatrix}.$$

With this explicit form for $T_n^{(A)}$, one repeats the proof of Theorem 1 to show that E is in the spectrum and not an eigenvalue; moreover, since $T_{n+k}^{(A)}$ has an eigenvalue one, (the same for all k!), a solution of the Schrödinger equation with period 2^n can be constructed. There is only one eigenvector and thus only one periodic solution, unless $a_n = 0$. But notice that $T_{n-1}^{(A)} = -id$ implies $T_{n-2}^{(A)} T_{n-2}^{(B)} = T_{n-2}^{(A)} T_{n-2}^{(A)}$, and thus $T_{n-2}^{(A)} = T_{n-2}^{(B)}$; a fortiori, $y_{n-2} = x_{n-2} = 0$. This, however, can only arise if $x_0 = y_0$, and thus $V = 0$. □

Remark 1. We will see in Sect. 3 that the set verifying the assumptions of Lemma 4 forms part of the boundaries of the spectral gaps.

Remark 2. It is interesting to note that in a rather different, hierarchical model, the generalized eigenstates belonging to some gap boundaries have a very similar characterization [17].

Theorem 1 allows us to compute (in principle) the spectrum of H_V by studying the dynamical system (1.9). We have already seen that as a by-product of our proof we obtained that the spectrum is continuous. We will see that with rather little effort we can show that it is purely singular continuous. To do this, we first establish that the Lyapunov exponent vanishes on the spectrum. Remember that the Lyapunov exponent $\gamma(E, V)$ is defined as

$$\gamma(E, V) = \lim_{n \to \infty} \frac{1}{n} \ln \| T_n(E, V) \| , \tag{2.4}$$

provided this limit exists.

Proposition 1. *Let $\mathcal{O}_V = \{E | \gamma(E, V) = 0\}$. Then $\sigma(H_V) = \mathcal{O}_V$.*

Proof. The positivity of the Lyapunov exponent outside the spectrum of H_V is a general result based on the exponential decay of the Green's function for E in the resolvent set (see e.g. [18]). What is left to show is that the Lyapunov exponent vanishes whenever E is in the spectrum. This will be based on the following simple bound on x_n.

Lemma 5. *Let $(x, y) \in \mathcal{U}^c$. Then for any $\alpha > \sqrt{2}$, there exist constants c, d and an integer n_0, such that for all $n \geqq n_0$,*

$$|x_n| \leqq c e^{d\alpha^n} . \tag{2.5}$$

Proof. By Lemma 1, $(x, y) \in \mathcal{U}^c$ implies that (x_n, y_n) will never enter $D_{\pm}^{(0)}$. If $|x_n| \leqq 2$ for all n, there is nothing to prove. If for some n, $|x_n| > 2$, then $y_{n+1} > 2$, and we must thus have $|x_{n+1}| = |x_n y_n - 2| \leqq 2$. But then

$$|x_{n+2}| = |(x_n y_n - 2)(x_n^2 - 2) - 2| \leqq 2x_n^2 + 6 . \tag{2.6}$$

From this the bound (2.5) follows by induction. □

The bound obtained for x_n will now be used to get a bound on the norm of $\| T_n^{(A)} \|$.

Lemma 6. *Let $(x, y) \in \mathcal{U}^c$. Then for any $\beta > \frac{\sqrt{5}+1}{2}$, there exist constants c, d, and an integer n_0, such that for all $n \geqq n_0$,*

$$\| T_n^{(A)} \| \leqq c e^{d\beta^n} . \tag{2.7}$$

Proof. Notice that (1.7) together with (1.8) implies that

$$\begin{aligned} \| T_{n+1}^{(A)} \| &= \| (T_{n-1}^{(A)})^2 T_n^{(A)} \| = \| (T_{n-1}^{(A)} x_{n-1} - T_n^{(A)}) T_n^{(A)} \| \\ &\leqq |x_{n-1}| \, \| T_n^{(A)} \| \, \| T_{n-1}^{(A)} \| + \| T_n^{(A)} \| . \end{aligned} \tag{2.8}$$

From (2.8) the bound (2.7) follows by induction, using the bound (2.5) for x_{n-1} with some $\sqrt{2}<\alpha<\beta$. □

To conclude the proof of the proposition, notice that it follows from the definition of the period doubling sequence that if $n=\sum_{i=0}^{s} 2^{l_i}$, then

$$T_n(E,V)=\prod_{i=0}^{s} T_{l_{s-i}}^{(A)}. \tag{2.9}$$

Using the Schwarz inequality and the bound (2.7) we get that

$$\frac{1}{n}\ln\|T_n(E,V)\|\leqq\frac{1}{n}\sum_{i=0}^{s}\ln\|T_{l_i}^{(A)}\|\leqq\frac{1}{n}\sum_{i=0}^{s}[d\beta^{l_i}+\ln c], \tag{2.10}$$

which, since β can be chosen less than 2, implies

$$\lim_{n\to\infty}\frac{1}{n}\ln\|T_n(E,V)\|=0. \ \square$$

Remark. Proposition 1 together with the fact that $\mathscr{U}$ is in the complement of the spectrum in fact allows to show that $\mathscr{U}^c=\sigma(H_V)$ without using Lemma 3. Notice simply that on the complement of the spectrum the Lyapunov exponent is positive, while we show that it vanishes on the complement of $\mathscr{U}$. Thus $\mathscr{U}^c$ is contained in the spectrum and the spectrum is contained in it. Unfortunately, this argument does not yield the continuity of the spectrum, as subexponentially decaying eigenfunctions may exist.

Proposition 1 has an important consequence for the nature of the spectrum of H_V, due to a general theorem, that we will now derive. Let v_n be any aperiodic sequence taking a finite number of values and denote by Ω its "hull," that is the closure in the product topology of the set of translates of this sequence. Let ω denote an element of Ω and let $H_V(\omega)$ be the hamiltonian with potential given by the sequence ω. Let μ be any probability measure on Ω such that the translations, T, are an ergodic, measure preserving group. Thus (Ω, T, μ) is an ergodic dynamical system. We call the system uniquely ergodic, if there is no other invariant probability measure.

Theorem 2. *Assume the sequence v_n is aperiodic and (Ω, T, μ) is uniquely ergodic. Then, if $\sigma(H_V)=\{E|\gamma(E,V)=0\}$, $\sigma(H_V)$ is supported on a set of zero Lebesgue measure.*

Proof. Our proof follows the ideas of Sütö [6] who proved this result for the Fibonacci sequence. The main new ingredient needed for the generalization is a lemma of Herman [16].

We will need two basic lemmas. Denote by $\gamma_\omega(E,V)$ the Lyapunov exponent corresponding to $H_V(\omega)$. The ergodicity of translations imply that $\gamma_\omega(E,V)$ exists μ-almost surely and equals a constant, $\gamma_\mu(E,V)\equiv\int\mu(\omega)\gamma_\omega(E,V)$. Let N_μ denote the set

$$N_\mu\{E|\gamma_\mu(E,V)=0\}.$$

Kotani [15] showed the following:

Lemma 7 (Kotani). *Let v_n be a sequence taking a finite number of values, and let μ be an invariant measure on its hull. Then the set N_μ has Lebesgue measure zero, unless the support of μ is a finite set.*

Proof. See [15].

Obviously, suppμ can be finite only for periodic sequences, so Lemma 7 implies that in our case, N_μ has measure zero, unless $V=0$. The set N_μ can in principle be different from the set $N_\omega \equiv \{E|\gamma_\omega(E, V)=0\}$ for a specific sequence $\omega \in \Omega$. However, the difference between the two sets must have zero Lebesgue measure:

Lemma 8. *If (Ω, T, μ) is uniquely ergodic, then for any $\omega \in \Omega$, the set $S \equiv N_\mu \Delta N_\omega$* [1] *has Lebesgue measure zero.*

Lemma 8 for almost periodic sequences follows from a result of Avron and Simon [3] and was used in [6]. The generalization above is proven along the same lines, using a lemma due to Herman [16]. We give the proof in Appendix B.

Obviously, the two lemmas imply Theorem 2. □

Now, in [9] (see Chap. V, in particular Theorems V.2 and V.13) it is shown that the assumption of Lemma 8 is verified in the case of automatic sequences under some weak hypothesis on the substitution rule. In particular, the dynamical system obtained from the period doubling sequence is uniquely ergodic. Combining thus Theorem 2 with Proposition 1 and Theorem 1, we obtain as an immediate corollary for the period doubling sequence:

Theorem 3. *For any $V \neq 0$, $\sigma(H_V)$ is a Cantor set of zero Lebesgue measure and the spectrum is purely singular continuous.*

Remark. Theorem 3 presumably holds for a large class of automatic sequences. Notice that the only hard part is to show the continuity of the spectrum, which requires Lemma 3, while the fact that the spectrum has measure zero follows by Theorem 2 from the rather simple estimates leading to Proposition 1. In the case of the Thue-Morse sequence we have verified that all main results of this section can also be obtained, as was conjectured in [11]. In fact, in this case the symmetry of the substitution rule (the words A^{4^n} are mirror symmetric) allows even to prove Lemma 3 in a much simpler manner.

III. Detailed Structure of the Spectral Gaps

In the previous section we established the link between the spectrum of H_V and the set of unstable points of the dynamical system given by the trace map. A rather simple analysis of this dynamical system allowed us to conclude that the spectrum of H_V is singular continuous for all $V \neq 0$. In the present section we give a detailed analysis of the dynamical system that will allow us to give a complete and detailed description of the spectral gaps for H_V. For comparison and orientation, Fig. 1 represents the complement of the unstable set, that is the spectrum, as obtained from numerical calculations.

---

[1] Here Δ denotes the symmetric difference

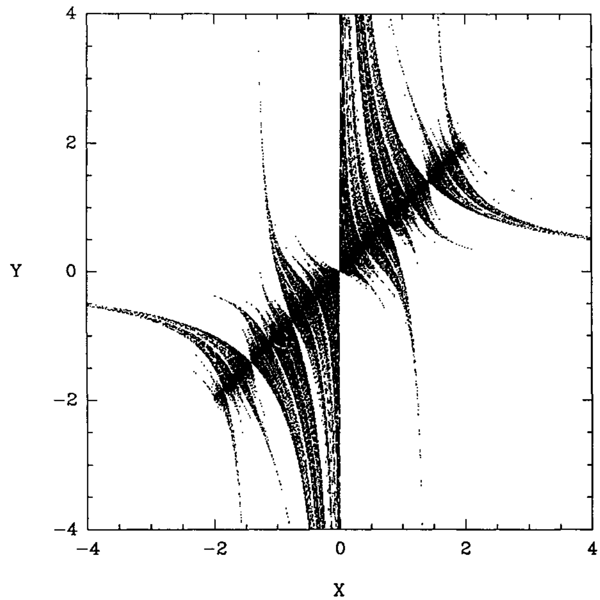

Fig. 1. The complement of the unstable region for the trace map

A similar analysis has been given in [11] for the case of the Thue-Morse potential. The gap structure for the period doubling sequence was studied previously by Luck [12] using perturbation theory and numerical computations. We will see that our results confirm and refine his findings.

The starting point of our investigation is Lemma 1 from Sect. II. It implies that the set of unstable points, $\mathcal{U}$, consists of the pre-images of the two sets $D_+^{(0)}$ and $D_-^{(0)}$. We are thus led to consider the inverse of the trace map.

Lemma 9. *The trace map* (1.9) *has the following properties:*

(i) *Any point* $(x, y) \in \mathbf{R}^2 \backslash \{(-2, -2)\}$ *has two inverse images, given by*

$$\tau_+(x, y) = \left(\sqrt{y+2}, \frac{x+2}{\sqrt{y+2}}\right), \tag{3.1}$$

and

$$\tau_-(x, y) = \left(-\sqrt{y+2}, -\frac{x+2}{\sqrt{y+2}}\right). \tag{3.2}$$

(ii) *The inverse image of the point* $(-2, -2)$ *is the line*

$$L_0 = \{(x, y) | x = 0\}.$$

The proof of this lemma is by simple inspection. The next lemma collects some useful properties of the maps $\tau_\pm$.

Lemma 10.

(i) *The map* τ_+ *has a unique fixpoint,* (2, 2), *and the tangent map,* T_+, *at this point is given by*

$$T_+ = \begin{pmatrix} 0 & 1/4 \\ 1/2 & -1/4 \end{pmatrix}, \tag{3.3}$$

whose eigenvalues and eigenvectors are

$$\lambda^+_{1,2}=\frac{-1\pm 3}{8}, \quad e^+_{1,2}=\begin{pmatrix}1\\ \frac{-1\pm 3}{2}\end{pmatrix}.$$

(ii) *The map τ_- has a unique fixpoint, $(-1,-1)$, and the tangent map, T_-, at this point is given by*

$$T_-=\begin{pmatrix}0 & -1/2\\ -1 & 1/2\end{pmatrix}, \tag{3.4}$$

whose eigenvalues and eigenvectors are

$$\lambda_{1,2}=\frac{1\pm 3}{4}, \quad e^-_{1,2}=\begin{pmatrix}1\\ \frac{-1\mp 3}{2}\end{pmatrix}.$$

This lemma may again be verified by a simple computation.

Remark. The fixpoints and the tangent maps essentially determine the opening of the gaps for small V. We will see later that the exponential opening of a class of gaps, as predicted by Luck, is tied to the fact that T_- has an eigenvalue one.

We will see that τ_+ and τ_- give rise to two invariant sets, $D^{(\infty)}_+$ and $D^{(\infty)}_-$, whose interiors are contained in $\mathcal{U}$. In fact, these sets are the limits of the sets $D^{(n)}_\pm \equiv \tau^n_\pm(D^{(0)}_\pm)$, the fixpoints of $\tau_\pm$ are in the boundary of $D^{(\infty)}_\pm$ and correspond to the intersection of these sets with the spectrum of the free laplacian.

To construct the set $D^{(\infty)}_+$ explicitly, we introduce the following family of functions, which will be seen to describe the boundary of $D^{(n)}_+$. Let $g_0:(\sqrt{2},\infty)\to(0,\infty)$ be given by

$$g_0(x)=\begin{cases}\left(\frac{4}{x^2-2}+2\right)/x, & \text{for } \sqrt{2}<x<2\\ 2, & \text{for } x=2\\ 4/x, & \text{for } x>2,\end{cases} \tag{3.5a}$$

and define, for $n>0$, $g_n:(\sqrt{2},\infty)\to(0,\infty)$ recursively by

$$g_n(x)=\frac{g^{-1}_{n-1}(x^2-2)+2}{x}. \tag{3.5b}$$

Proposition 2. *The g_n defined above have the following properties:*

(i) *The functions g_n are monotonous, strictly decreasing convex functions on their domain; they are infinitely differentiable except at $x=2$.*
(ii) *For all $\sqrt{2}<x<\infty$, and all $n>0$,*

$$g_n(x)\leqq g_{n-1}(x).$$

(iii) *$g_n(x)$ form the lower boundary of the regions $D^{(n+2)}_+$, i.e.*

$$D^{(n+2)}_+=\{(x,y)|x>\sqrt{2}, y>g_n(x)\}. \tag{3.6}$$

(iv) *The sequence of functions converges to a function* $g_\infty(x)$ *that is uniformly decreasing, convex and continuously differentiable. It satisfies the functional equation*

$$g_\infty(x)=\frac{g_\infty^{-1}(x^2-2)+2}{x}, \tag{3.7}$$

and is the unique solution of this equation with the above properties.

$$D_+^{(\infty)}=\{(x,y)|x>\sqrt{2}, y>g_\infty(x)\}. \tag{3.8}$$

(v) *Near* $x=2$, g_∞ *permits an asymptotic expansion given by*

$$g_\infty(2+\varepsilon)=2-2\varepsilon-\frac{3}{a\ln 2}\varepsilon^2\ln|\varepsilon|+\varepsilon^2 p(\varepsilon)+o(\varepsilon^2), \tag{3.9}$$

where $p(\varepsilon)$ *is a bounded function satisfying* $p(-2\varepsilon)=p(\varepsilon)$.

Proof. Point (iii) is clearly the rationale for defining the sequence g_n. To establish point (i), notice that g_0 has all the properties stated; monotonicity and differentiability properties for g_n are easily seen to follow from those of g_{n-1}. To establish convexity of g_n, we compute the second derivative of g_n:

$$g_n''(x)=4xg_{n-1}^{-1''}(x^2-2)-\frac{2g_{n-1}^{-1'}(x^2-2)}{x}+\frac{2(g_{n-1}^{-1}(x^2-2)+2)}{x^3}. \tag{3.10}$$

If g_{n-1} is decreasing and convex, then so is g_{n-1}^{-1}, and thus by (3.10), $g_n''(x)$ is positive wherever it exists, which is everywhere except at $x=2$. Thus the two branches of g_n for $x<2$ and $x>2$ are convex, and we only have to show that the left and right tangents of the two branches enclose an angle smaller than or equal to π. But this is true for g_0 (check!), and as $g_n(2)=2$ for all n, the behaviour of the tangents is governed by the map T_+ given in Lemma 10, and a simple calculation shows that this remains true for all n. Moreover, one finds that the slope of both left and right tangents converge to -2 as n tends to infinity (i.e. they follow the direction of the eigenvector with largest eigenvalue of T_+). Thus we have proven (i).

To prove that $g_n(x)$ for any x is a decreasing sequence, notice that $g_n(x) \leqq g_{n-1}(x)$ is equivalent to

$$g_{n-1}(xg_{n-1}(x)-2)\leqq(x^2-2).$$

Thus assuming this to hold, we see that

$$g_n(xg_n(x)-2)\leqq g_{n-1}(xg_n(x)-2)=x^2-2,$$

and thus $g_{n+1}(x)\leqq g_n(x)$. It remains to show that $g_1(x)\leqq g_0(x)$, which one may check by hand.

Using the monotonicity of the g_n established above, it is easy to show that they form indeed the boundaries of $D_+^{(n+2)}$. For $n=0$, this is true by construction (check!) and for general n it is proven by induction: If (x,y) are such that $y>g_{n-1}(x)$ and $x>\sqrt{2}$, then for $(x',y')=\tau_+(x,y)$, $x'=\sqrt{y+2}$ and $y'=\frac{x+2}{x'}$. Since g_{n-1} is decreasing, $y>g_{n-1}(x)$ implies $x>g_{n-1}^{-1}(y)$, and thus

$$y'>\frac{g_{n-1}^{-1}(x'^2-2)+2}{x'}\equiv g_n(x'). \tag{3.11}$$

Moreover, $y=g_{n-1}(x)$ is mapped to $y'=g_n(x')$, and since g_n is below g_{n-1}, the continuity of τ_+ implies the result. Note also that $D_+^{(n-1)}\subset D_+^{(n)}$.

By (ii) and the fact that $g_n(x)\geqq 0$, the $g_n(x)$ converge pointwise and being monotone even uniformly on compact subsets. Moreover, the limiting function is convex and as such has left and right derivatives at every point. We will show that they must in fact coincide everywhere. Let $g_\infty^{L,R}(x)$ denote the left and right derivative at x. Put $d(x)=g_\infty^R(x)-g_\infty^L(x)$. Since g_∞ must satisfy (3.7), one verifies that

$$g_\infty^{L,R}(x)=\frac{2}{g_\infty^{R,L}(g_\infty^{-1}(x^2-2))}-\frac{g_\infty(x)}{x}. \tag{3.12}$$

Notice that $g_\infty^{-1}(x^2-2)=\tau_+^{-1}(x)$. From (3.12) we may then obtain, using that $g_\infty^{L,R}(x)\leqq 0$, the inequality

$$g_\infty^{L,R}(x)g_\infty^{R,L}(\tau_+^{-1}(x))\geqq 2. \tag{3.13}$$

On the other hand, from (3.12) we derive

$$\begin{aligned} d(x)&=\frac{2}{g_\infty^R(\tau_+^{-1}(x))g_\infty^L(\tau_+^{-1}(x))}\,d(\tau_+^{-1}(x))\\ &=\frac{2}{g_\infty^R(\tau_+^{-1}(x))g_\infty^L(\tau_+^{-1}(x))g_\infty^R(\tau_+^{-2}(x))g_\infty^L(\tau_+^{-2}(x))}\,d(\tau_+^{-2}(x))\\ &\leqq d(\tau_+^{-2}(x)). \end{aligned} \tag{3.14}$$

Now since by convexity $d(x)\geqq 0$, if for some x other than the fixpoint of τ_+, $d(x)=\delta>0$, then for all k $d(\tau_+^{-2k}(x))>\delta$, which will force the slope of g_∞ to be positive for x large enough, contradicting the fact that g_∞ is decreasing. This proves that $d(x)=0$ except possibly at $x=2$. However, as remarked above, the properties of the tangent map T_+ ensure that the left and right derivatives at $x=2$ both equal -2. Thus g_∞ is in fact a C^1 function.

We want to prove that g_∞ is the unique positive, decreasing solution of the functional Eq. (3.7). Note that any solution of (3.7) crossing the line $x=y$ must do so at $x=2$. Since the branch $x<0$ is determined by the branch $x>2$, the idea is to show that only one solution of (3.7) can be bounded between zero and two for all $x>2$. Assume thus that there are two solutions, f and g, of (3.7) such that $f(2)=g(2)=2$. Assume now that for a given $x_0>2$, $f(x_0)-g(x_0)=\delta>0$. Put

$$x_n=x_{n-1}f(x_{n-1})-2.$$

Notice that (3.7) implies

$$f(xf(x)-2)=g(xg(x)-2),$$

and thus

$$f(xf(x)-2)-g(xf(x)-2)=g(x(g(x)-2)-g(xf(x)-2)=x(g(x)-f(x))g'(z),$$

for some z in $[xg(x)-2, xf(x)-2]$. Using this relation with $x=x_1$, iterating twice and using convexity to bound the derivative of g, we obtain

$$f(x_2)-g(x_2)\geqq\delta x_1x_0g'(x_1)g'(x_2)\geqq 4\sqrt{2}\delta,$$

where the last bound made use of (3.13) and the obvious bounds $x_0>2$, $x_1>\sqrt{2}$. Iterating this bound shows that eventually $f(x_n)-g(x_n)>2$, which is impossible if g

is positive and f monotone decreasing. Thus δ must equal zero, which we wanted to prove.

What is left is to compute the asymptotic expansion for g_∞ near $x=2$. To do so, we put

$$g(2+\varepsilon)=2-2\varepsilon+h(\varepsilon), \tag{3.15}$$

and use that by (3.7)

$$g_\infty(xg_\infty(x)-2)=x^2-2. \tag{3.16}$$

In terms of h this gives

$$0=3\varepsilon^2-(4+2\varepsilon)h(\varepsilon)-h(-2\varepsilon-2\varepsilon^2+(2+\varepsilon)h(\varepsilon)). \tag{3.17}$$

Keeping only the leading order terms, this simplifies to

$$0=3\varepsilon^2-4h(\varepsilon)+h(-2\varepsilon). \tag{3.18}$$

One verifies that

$$h(\varepsilon)=\varepsilon^2\ln\left(|\varepsilon|^{-\frac{3}{4\ln 2}}p(\varepsilon)\right), \tag{3.19}$$

is a solution for any $p(\varepsilon)$ satisfying $p(\varepsilon)=p(-2\varepsilon)$. Such a function is necessarily bounded at zero. It will be determined by higher order terms, which, in principle may of course be computed. We will content ourselves with the present expression. This completes the proof of Proposition 2. □

We see thus that $D_+^{(\infty)}$ is a convex region asymptotically bounded by the x-axis and the line $x=\sqrt{2}$, whose boundary intersect the straight line $x=y$ (i.e. $V=0$) at $x=y=2$ ($E=2$) with slope -2. Apparently, this region determines the upper boundary of the spectrum of H_V. Before we come to the gaps generated by $D_+^{(\infty)}$, we construct the corresponding domain $D_-^{(\infty)}$. As we will see, the boundaries of the regions $D_-^{(n)}$ are formed by two functions u_n and d_n, where

$$u_n:(-\sqrt{2},\xi_n]\to[\zeta_n,\infty)$$
$$d_n:(-\infty,\xi_n]\to[\zeta_n,0)$$

are recursively defined as follows:

$$d_0=\frac{2-\sqrt{2}}{x},$$

and

$$u_0=\frac{2+\dfrac{2-\sqrt{2}}{x^2-2}}{x}, \tag{3.20}$$

$$\xi_0=-\sqrt{1+\sqrt{2}},\qquad \zeta_0=\frac{\sqrt{2}-2}{\sqrt{1+\sqrt{2}}}.$$

Then

$$d_n(x)=\frac{u_{n-1}^{-1}(x^2-2)+2}{x},$$

$$u_n(x)=\frac{d_{n-1}^{-1}(x^2-2)+2}{x}, \tag{3.21}$$

and

$$(\xi_n, \zeta_n) = \tau_-(\xi_{n-1}, \zeta_{n-1}). \tag{3.22}$$

Proposition 3. *The functions d_n and u_n defined above have the following properties:*

(i) *d_n and u_n are monotonous, strictly decreasing and infinitely differentiable functions on their respective domains.*

(ii) *The following relations hold wherever left- and right-hand side are defined:*

$$\begin{aligned} d_n(x) &\leqq u_n(x), \\ d_n(x) &\leqq d_{n-1}(x), \\ u_n(x) &\geqq u_{n-1}(x). \end{aligned} \tag{3.23}$$

(iii) *$d_n(x)$ and $u_n(x)$ form the boundaries of the region $D_-^{(n+4)}$, i.e.*

$$D_-^{(n+4)} = \{(x, y) | x < \xi_n, d_n(x) < y < u_n(x)\}. \tag{3.24}$$

(iv) *The functions d_n and u_n converge to monotone decreasing functions with bounded derivatives,*

$$d_\infty : (-\infty, -1] \to [-1, 0),$$
$$u_\infty : (-\sqrt{2}, -1] \to [-1, \infty),$$

satisfying the functional equations

$$\begin{aligned} d_\infty(x) &= \frac{u_\infty^{-1}(x^2-2)+2}{x}, \\ u_\infty(x) &= \frac{d_\infty^{-1}(x^2-2)+2}{x}. \end{aligned} \tag{3.25}$$

Moreover, they are the unique solution of these equations with the stated properties.

(v) *d_∞ and u_∞ permit an asymptotic expansion at -1 to the left, i.e. for $0 \leqq \delta \ll 1$,*

$$\begin{aligned} d_\infty(-1-\delta) &= -1 + 2\delta - \delta^2 + O(\delta^3), \\ u_\infty(-1-\delta) &= -1 + 2\delta - \delta^2 + O(\delta^3). \end{aligned} \tag{3.26}$$

Moreover,

$$u_\infty(-1-\delta) - d_\infty(-1-\delta) = \exp\left(-\frac{\ln 2}{\delta}\right) \delta^{\ln 2} f(\delta), \tag{3.27}$$

where f is a function that satisfies for any $\varepsilon > 0$, $\delta^\varepsilon < f(\delta) < \delta^{-\varepsilon}$.

Proof. The proof of Proposition 3 is largely anologous to that of Proposition 2, and we will spare the reader most of the details. The proofs of points (i) through (iii) proceed like those of the analogous statements in Proposition 2; notice however that we do not show that the d_n and u_n are convex (or concave) functions, since this is in general not true. One sees from the asymptotic expansion of u_∞ and the fact that it must diverge at $x = -\sqrt{2}$ that this function must possess a point of inflection; musing the functional equations (3.25), it is possible to show that indeed both d_∞ and u_∞ possess an infinity of inflection points. The lack of convexity requires some modifications in the proof of point (iv). In particular, we have not been able to show that the derivatives of u_∞ and d_∞ are continuous. The boundedness of the derivatives follows however from the monotonicity and Eqs. (3.25), which imply that

$$\begin{aligned} d'_\infty(x) &= \frac{2}{u'_\infty(\tau_-^{-1}(x))} - \frac{d_\infty(x)}{x}, \\ u'_\infty(x) &= \frac{2}{d'_\infty(\tau_-^{-1}(x))} - \frac{u_\infty(x)}{x}. \end{aligned} \tag{3.28}$$

The proof of the uniqueness of the solutions of (3.21) is also somewhat different. Note first that the solutions obtained as the limits of d_n and u_n are necessarily the largest and smallest ones, respectively, since they bound the unstable set $D_-^{(\infty)}$. Any solution d of (3.21) must then satisfy the bounds

$$\frac{1}{x} \leqq d(x) \leqq \frac{2-\sqrt{2}}{x}, \tag{3.29}$$

and by computing the successive images of a point (x, y) under the twice iterated trace map, one shows that in order that (3.29) remains satisfied, y is determined uniquely as a function of x.

Point (v) is of particular interest as it contains the information pertinent to the opening of gaps. To derive the asymptotic expansion, we put, for $\delta \geqq 0$,

$$\begin{aligned} u_\infty(-1-\delta) &= -1+2\delta+\Phi(\delta), \\ d_\infty(-1-\delta) &= -1+2\delta+\Gamma(\delta). \end{aligned} \tag{3.30}$$

Writing (3.25) in the form

$$u_\infty(xd_\infty(x)-2) = d_\infty(xu_\infty(x)-2) = x^2-2, \tag{3.31}$$

we obtain for Γ and Φ the equations

$$\begin{aligned} 3\delta^2 + 2\Gamma(\delta)(1+\delta) + \Gamma(\delta+2\delta^2+\Phi(\delta)(1+\delta)) &= 0, \\ 3\delta^2 + 2\Phi(\delta)(1+\delta) + \Phi(\delta+2\delta^2+\Gamma(\delta)(1+\delta)) &= 0. \end{aligned} \tag{3.32}$$

Neglecting terms of order δ^3 this simplifies to

$$\begin{aligned} 3\delta^2 + 2\Phi(\delta) + \Gamma(\delta) &= 0, \\ 3\delta^2 + 2\Gamma(\delta) + \Phi(\delta) &= 0, \end{aligned} \tag{3.33}$$

which implies (3.26). We are particularly interested in the difference between u_∞ and d_∞. Substracting the two Eqs. (3.32), and putting $\Delta(\delta) = \Phi(\delta) - \Gamma(\delta)$, we obtain

$$\begin{aligned} &2\Delta(\delta)(1+\delta) - \Delta(\delta+2\delta^2+\Phi(\delta)(1+\delta)) \\ &\quad + \Phi(\delta+2\delta^2+\Phi(\delta)(1+\delta)) - \Phi(\delta+2\delta^2+\Gamma(\delta)(1+\delta)) = 0, \end{aligned} \tag{3.34}$$

or, by the mean value theorem

$$\begin{aligned} &2\Delta(\delta)(1+\delta) - \Delta(\delta+2\delta^2+\Phi(\delta)(1+\delta)) \\ &\quad + \Phi'(\delta+2\delta^2+\Phi(\delta)(1+\delta)+s\Delta(\delta)(1+\delta))(1+\delta)\Delta(\delta) = 0. \end{aligned} \tag{3.35}$$

Keeping only the leading orders in δ, this simplifies to

$$2\Delta(\delta)(1-2\delta^2) = \Delta(\delta+\delta^2). \tag{3.36}$$

Taking logarithms of this equation, and setting

$$\ln \Delta(\delta) \equiv -\frac{\ln 2}{\delta} + \varrho(\delta), \tag{3.37}$$

we find that to leading order in δ, $\varrho(\delta) = \ln 2 \ln \delta$, which gives (3.27), and concludes the proof of the proposition. □

Remark. The region $D_-^{(\infty)}$ constructed above is the first spectral gap, opening at $x=y=-1$, corresponding to $E=-1$, $V=0$. Equation (3.27) tells us how the gap opens asymptotically for V small. δ is related to V by $\delta = \frac{2V}{3}$, so that the width of the gap is given by

$$w(V) \sim \exp\left(-\frac{3 \ln 2}{2V}\right) V^{\ln 2}. \tag{3.38}$$

Note that the result of Luck [12] which predicts such a behaviour with the exponent of V near 0.5 on the basis of perturbation theory and numerical computations is compatible with our exact formula.

We will now show how the totality of spectral gaps is constructed. First, all images of $D_-^{(\infty)}$ under an arbitrary sequence of the maps τ_+ and τ_- are spectral gaps, intersecting the line $V=0$ at the image point of $(-1, -1)$. The second family of gaps in associated to $D_+^{(\infty)}$. Note first that a simple computation shows that

$$\tau_-(D_+^{(\infty)}) = -D_+^{(\infty)}, \tag{3.39}$$

which corresponds obviously to the lower boundary of the spectrum. Now no point with $y < -2$ has a real image under τ_+ or τ_-, and the image of $(-2, -2)$ is the y-axis. Putting these together, we get the following lemma, describing the first gap associated to $D_+^{(\infty)}$:

Lemma 11. *Let*

$$\tilde{D} \equiv \{(x, y) | x < -2, -g_\infty(-x) > y > -2\},$$

then the region $G_+ \equiv \tau_+(\tilde{D}) \cup \tau_-(\tilde{D})$ is a spectral gap, opening at $x=y=0$. It is bounded by the y-axis and the images of the curve $y = -g_\infty(-x)$, with $x < -2$. The gap opens linearly with an opening angle $\alpha = \arctan 2$.

As before, this gap gives rise to an infinite family of spectral gaps that open at the image points of $(0, 0)$. This exhausts all gaps.

Theorem 4. *Let ω denote a sequence $(\sigma_0, \sigma_1, \ldots, \sigma_n)$, with $\sigma_i = \pm 1$, and put $\tau_\omega = \tau_{\sigma_n} \ldots \tau_{\sigma_0}$. Then*

(i) *Let (x, y) be such that $E \notin \sigma(H_V)$. Then there exists a sequence ω such that (x, y) is contained in either $\tau_\omega(D_+^{(\infty)})$ or $\tau_\omega(D_-^{(\infty)})$.*

(ii) *The corresponding regions intersect the interior spectrum of the free Laplacian at the points $\tau_\omega(0,0)$ and $\tau_\omega(-1, -1)$, respectively. That is, spectral gaps open at exactly those energies, and no others.*

(iii) *The gaps at the points $\tau_\omega(0,0)$ open linearly, the opening angle being typically of the order $2^{-|\omega|}$. The other gaps open exponentially according to the relation given in* (3.38).

(iv) *The value of the density of states at the energies where gaps open is either a dyadic number or one third of a dyadic number.*

Proof. Points (i) and (ii) summarize our findings above. For point (iii), we have seen in Lemma 11 that the gap opening at (0, 0) opens with an angle arctan 2. The opening angles of the subsequent gaps can be computed by considering the images of straight lines with given slope near $x=y$. Denoting the slope after n applications of τ_+ or τ_- by a_n, one gets the recursion[2]

$$a_{n+1}=-1+\frac{2}{a_n}. \tag{3.40}$$

Starting with two lines of slope a_0, b_0, the difference becomes

$$|a_n-b_n|=2^n\prod_{k=0}^{n-1}\frac{1}{|a_kb_k|}|a_0-b_0|. \tag{3.41}$$

One checks easily that the recursion (3.40) has two fixpoints, 1 and -2, of which -2 is stable and 1 unstable. Thus a_n and b_n will in general converge to -2, and asymptotically we get from (3.41) that $|a_n-b_n|\sim 2^{-n}|a_0-b_0|$, from which follows (iii).

Finally point (iv) follows from the known formula for the density of states of the free Laplacian,

$$N(E)=\frac{1}{\pi}\arccos(-E/2) \tag{3.42}$$

(see, e.g. [11]). □

Remark. In Sect. 2 we have proven that the support of the spectrum is a Cantor set of zero measure. It would seem natural to prove this fact directly by computing the measure of the gaps, e.g. for small V. This turns out, however, to be rather difficult and we have not been able to do this in a rigorous way. In fact, what seems to happen is slightly different from naive intuition. Naively, we would expect that, for small V, the linear gaps are the most important ones, and that they have full measure. This is further supported by our finding that the opening angles of the n^{th} hierarchy of gaps behave like 2^{-n}, from which one would be tempted to say that each hierarchy contributes a measure of the order of a constant. However, this is wrong: the region of linear opening shrinks exponentially fast with n and is followed by a region where the gap width remains constant. This is due to the fact that the principal gap, for large V, is contained between the axis $y=0$ and $y=\pm\sqrt{2}$. A rough estimate indicates that indeed in the limit $V=0$, the total measure of the linear gaps is zero! Thus, the exponential gaps carry full measure, a fact that is less surprising if one considers the fact that the region $D_-^{(\infty)}$ is widening linearly for large V.

Appendix A

In this appendix we give the proof of Lemma 3. Since we have to show only that $T_n^{(A)}$ does not converge to a projection on any given subspace, we may choose a basis for which this amounts to show that

$$T_n^{(A)}\not\to\begin{pmatrix}0 & *\\ 0 & *\end{pmatrix}.$$

---

[2] Here we use a linear analysis, assuming that $y\neq 2$

To do this, since $T_n^{(\Delta)}$ is unimodular with trace x_n, we write

$$T_n^{(\Delta)} = \begin{pmatrix} \varepsilon_n & \gamma_n/\delta_n \\ \delta_n & x_n - \varepsilon_n \end{pmatrix}, \tag{A.1}$$

where $\gamma_n = \varepsilon_n(x_n - \varepsilon_n) - 1$. We will show that if we choose $\Delta > 0$ small enough, then there exists a sequence, n_l, of integers tending to infinity, such that either $|\varepsilon_{n_l}| > \Delta$, or $|\delta_{n_l}| > \Delta$. The construction of this sequence will be given in an algorithmic form.

Notice first that for $(x, y) \in \mathscr{U}^c$, it is impossible that for some n, both x_n and x_{n+1} have modulus bigger than 2. For $|x_n| > 2$ implies $y_{n+1} > 2$, and thus, if $|x_{n+1}|$ were bigger than 2, (x, y) would be unstable by Lemma 1.

Let us start by choosing a $\tilde{n}_0$ such that $|x_{\tilde{n}_0}| \leqq 2$. If at least one of $\varepsilon_{\tilde{n}_0}$, $\delta_{\tilde{n}_0}$, $\varepsilon_{\tilde{n}_0+1}$, $\delta_{\tilde{n}_0+1}$ has modulus larger than Δ, set $n_0 = \tilde{n}_0$ or $n_0 = \tilde{n}_0 + 1$, respectively, and put $\tilde{n}_1 = n_0 + 1$ or $\tilde{n}_1 = n_0 + 2$ (depending on whether $|x_{n_0+1}| \leqq 2$, or $|x_{n_0+2}| \leqq 2$). This way we produce a sequence with the desired property, until we reach a point, say $\tilde{n}_l$, for which all four epsilons and deltas have modulus less than Δ. At this point, we will make use of the fact that (1.7) allows to compute $\varepsilon_{\tilde{n}_l+2}$ and $\delta_{\tilde{n}_l+2}$. One gets (we let for notational simplicity $n \equiv \tilde{n}_l$):

$$\begin{aligned} \varepsilon_{n+2} &= \varepsilon_{n+1}\varepsilon_n^2 + \varepsilon_{n+1}\gamma_n + \gamma_n x_n \frac{\delta_{n+1}}{\delta_n}, \\ \delta_{n+2} &= x_n \delta_n \varepsilon_{n+1} + \delta_{n+1}\gamma_n + (x_n - \varepsilon_n)^2 \delta_n . \end{aligned} \tag{A.2}$$

Moreover, computing the trace of $T_{n+2}^{(\Delta)}$ in this way and comparing with the trace map yields an extra equation:

$$\begin{aligned} x_{n+2} &= \gamma_n x_n \frac{\delta_{n+1}}{\delta_n} + \gamma_{n+1} x_n \frac{\delta_n}{\delta_{n+1}} + \gamma_n x_{n+1} x_n^2 x_{n+1} \\ &\quad - x_n^2 \varepsilon_{n+1} - 2x_n \varepsilon_n (x_{n+1} - \varepsilon_{n+1}) + \varepsilon_n^2 x_{n+1} \\ &= x_n^2 x_{n+1} - 2x_{n+1} - 2 . \end{aligned} \tag{A.3}$$

In the following, we assume that $|\varepsilon_n| < \Delta$, $|\delta_n| < \Delta$, $|\varepsilon_{n+1}| < \Delta$, $|\delta_{n+1}| < \Delta$. Note that this implies that $\gamma_n = -1 + O(\Delta)$. Thus (A.3) can be simplified to

$$\begin{aligned} -x_{n+1} - 2 &= -x_n \frac{\delta_{n+1}}{d_n} + \gamma_{n+1} x_n \frac{\delta_n}{\delta_{n+1}} \\ &\quad - 2x_n \varepsilon_n (x_{n+1} - \varepsilon_{n+1}) + \varepsilon_n^2 x_{n+1} . \end{aligned} \tag{A.4}$$

Now, from (A.2) we see that

$$\varepsilon_{n+2} = x_n \frac{\delta_{n+1}}{\delta_n} + O(\Delta),$$

and thus $|\varepsilon_{n+2}| > \Delta$, unless

$$\left| x_n \frac{\delta_{n+1}}{\delta_n} \right| \leqq \Delta . \tag{A.5}$$

Thus, either we are done and can put $n_l = n + 2$, or (A.5) holds. In the latter case we can compute ε_{n+3} to leading orders, to get

$$\begin{aligned} \varepsilon_{n+3} &= \gamma_{n+1} x_{n+1} \frac{\delta_{n+2}}{\delta_{n+1}} + \gamma_{n+1}\varepsilon_{n+2} + o(\Delta) \\ &= \gamma_{n+1} x_{n+1} \left[-1 + \frac{x_n^2 \delta_n}{\delta_{n+1}} + O(\Delta) \right] + o(\Delta) . \end{aligned} \tag{A.6}$$

Notice that the quantity in brackets can be small (e.g. $<o(1)$), only if $\frac{x_n^2\delta_n}{\delta_{n+1}}=1+o(1)$. Combining this with (A.4), one finds that this can be true only if $|x_n^3|\leqq\Delta$. Keeping this in mind we use the trace condition Eq. (A.4) to get that, to leading order

$$\varepsilon_{n+3}=x_{n+1}[-\gamma_{n+1}-x_n(x_{n+1}+2)]. \tag{A.7}$$

We now distinguish three cases:

Case 1. $|x_{n+1}|\leqq\Delta^{1/2}$.

In this case, $x_{n+2}=x_{n+1}(x_n^2-2)-2\approx-2$. We may thus start anew with $n+1$ replacing n, and can be sure not to fall back into this case.

Case 2. $\Delta^{-1/2}\geqq|x_{n+1}|\geqq\Delta^{1/2}$.

In this case, $\gamma_{n+1}=-1+o(\Delta^{1/2})$, and

$$\varepsilon_{n+3}=x_{n+1}[1-x_n(x_{n+1}+2)],$$

will be bigger than Δ, unless the bracket is smaller than $\Delta^{1/2}$. But this requires, by the remark above, $|x_n|<\Delta^{1/3}$, and thus $|x_{n+1}|\approx\Delta^{-1/3}$, which implies that $|x_{n+2}|\gg 2$ as well, which is impossible on the stable set.

Case 3. $|x_{n+1}|>\Delta^{-1/2}$.

By the same reasoning as above, ε_{n+3} will have to be large, unless, possibly, $\gamma_{n+1}x_{n+1}$ is small. But this implies that $x_{n+1}\approx\varepsilon_{n+1}^{-1}$. But for the same reason as in Case 2, x_n cannot be very small on the stable set, and we find that again $|\varepsilon_{n+3}|\approx|x_nx_{n+1}^2|$ is much larger than Δ.

We see thus that either ε_{n+2}, ε_{n+3} or (if Case 1 applies) ε_{n+4} has modulus larger than Δ, and we may choose the corresponding value as the new n_l. Obviously, in this way we construct the desired sequence and prove that $T_n^{(A)}$ does not converge to a projection on a given subspace. □

Appendix B

In this appendix we give the proof of Lemma 8. As we will see, it is a extension to the case of potentials generated by automatic sequences of a result of Avron and Simon [3] for almost periodic potentials. The main ingredient of the proof is the following result of Herman [16]:

Lemma B.1. *Let Y a compact space, G an homeomorphism on Y. Let $\psi\in C^0(Y,\mathbf{R})$ such that there exists λ such that, for any v G-invariant probability measure on Y, $\int\psi dv=\lambda$. Then, if $n\to\infty$, $\frac{1}{n}\sum_{i=0}^{n-1}\psi\circ G^i$ converges uniformly to λ.*

For sake of completeness, we reproduce the very nice proof of Herman.

Proof. Denote by $F=\{\eta-\eta\circ G|\eta\in C^0(Y,\mathbf{R})\}$. Obviously, a Radon measure μ on Y is G-invariant if and only if $\int fd\mu=0$ for any $f\in F$. Thus, $F^{\perp}$ is the set of G-invariant Radon measures on Y. Moreover, any G-invariant Radon measure μ has a unique decomposition $\mu=\mu_+-\mu_-$, where μ_+ and μ_- are G-invariant positive measures. Therefore the hypothesis of the lemma implies that $\int(\psi-\lambda)d\mu=0$ for any $\mu\in F^\perp$. Thus $\psi-\lambda\in F^{\perp\perp}$ which, by Hahn-Banach theorem, is nothing but the closure of F for the topology of uniform convergence. This means that there exists a sequence $\eta_i-\eta_i\circ G\in F$ converging uniformly to $(\psi-\lambda)$. But for $\eta_i-\eta_i\circ G\in F$,

$$\frac{1}{n}\sum_{k=0}^{n-1}(\eta_i-\eta_i\circ G)\circ G^k=\frac{\eta_i-\eta_i\circ G^n}{n}$$

converges to zero uniformly, and putting both observations together, the claim of the lemma is immediate. □

Now the proof of Lemma 8 follows the lines of [3] and [6].

Lemma B.2. *Assume Ω admits a unique translation invariant measure μ. Then, for $t>0$,*

$$\lim_{n\to\infty}\frac{1}{2n}\operatorname{Tr}(\chi_{(-n,n)}(x)\,e^{-tH_V(\omega)})=L_\omega(t)$$

converges uniformly and the limit is independent of ω.

Proof. Fix $t>0$. For $(\omega,x)\in\Omega\times\Omega$, define

$$g(\omega,x)=(e^{-tH_V(\omega)})\,(x,x).$$

Since $g(\omega,x)=g(T_x(\omega),0)$, where $T_x(\omega)=\omega+x$, we can denote $g(\omega,x)$ as $\tilde{g}(T_x\omega)$.

Now we can apply Herman's Lemma B.1 with $Y=\Omega$, $G=T_x$, $\psi=\tilde{g}$ which is continuous on Ω and μ the unique translation invariant measure on Ω.

We obtain that

$$\frac{1}{2n}\operatorname{tr}(\chi_{(-n,n)}(x)\,e^{-tH_V(\omega)})=\frac{1}{2n}\sum_{i=-n}^{n}\tilde{g}(T_x^i\omega)$$

converges uniformly to

$$\int_{\omega\in\Omega}\operatorname{tr}(e^{-tH_V(\omega)}(x,x))d\mu\,. \quad \square$$

The main consequence of Lemma B.2 is that the integrated density of states for $H_V(\omega)$ is independent of ω. Thus one can write it as $k(E)$. Let [3]

$$\hat{\gamma}(E,V)=\int\ln(E-E')dk(E'),$$

and define the approximants of the Lyapunov exponent,

$$\gamma_\omega^n(E,V)=\frac{1}{n}\ln\|T_{n,\omega}(E,V)\|,$$

where $T_{n,\omega}(E,V)$ is the transfer matrix for $H_V(\omega)$.

Avron and Simon showed in [3] that the uniform convergence of L_ω implies that $\gamma_\omega^n(E,V)$ converges to $\hat{\gamma}(E)$ in $L^2(dE)$, uniformly in ω. Obviously, this proves Lemma 8. □

---

[3] This is known as Thouless formula [19]

References

1. Shechtman, D., Blech, I., Gratias, D., Cahn, J.V.: Metallic phase with long-range orientational order and no translational symmetry. Phys. Rev. Lett. **53**, 1951–1953 (1984)
2. See e.g. Steinhardt, P.J., Ostlund, S.: The physics of quasicrystals. Singapore: World Scientific 1987
3. Avron, Y., Simon, B.: Almost periodic Schrödinger operators. II. The integrated density of states. Duke Math. J. **50**, 369–391 (1983)
4. Bellissard, J., Bessis, D., Moussa, P.: Chaotic states of almost periodic Schrödinger operators. Phys. Rev. Lett. **49**, 701–704 (1982)
5. Kohmoto, M., Kadanoff, L.P., Tang, C.: Localization problem in one dimension: mapping and escape. Phys. Rev. Lett. **50**, 1870–1872 (1983);
Ostlund, S., Pandit, R., Rand, D., Schnellnhuber, H.J., Siggia, E.D.: Schrödinger equation with an almost periodic potential. Phys. Rev. Lett. **50**, 1873–1876 (1983);
Kohmoto, M., Oono, Y.: Cantor spectrum for an almost periodic Schrödinger equation and a dynamical map. Phys. Lett. **102**A, 145–148 (1984);
Casdagli, M.: Symbolic dynamics for the renormalization map of a quasiperiodic Schrödinger equation. Commun. Math. Phys. **107**, 295–318 (1986)
6. Sütö, A.: The spectrum of a quasiperiodic Schrödinger operator. Commun. Math. Phys. **111**, 409–415 (1987);
Sütö, A.: Singular continuous spectrum on a Cantor set of zero Lebesgue measure for the Fibonacci hamiltonian. J. Stat. Phys. **56**, 525–531 (1989)
7. Bellissard, J., Iochum, B., Scoppola, E., Testard, D.: Spectral properties of one dimensional quasi-crystals. Commun. Math. Phys. **125**, 527–543 (1989)
8. Axel, F., Peyrière, J.: C. R. Acad. Sci. Paris **306** Serie II, 179–182 (1988)
9. Queffelec, M.: Substitution dynamical systems. Spectral analysis. Lecture Notes in Mathematics, vol. 1294. Berlin, Heidelberg, New York: Springer 1987
10. Axel, F., Allouche, J.P., Kléman, M., Mendès France, M., Peyrière, J.: Vibrational modes in a one-dimensional "quasi alloy." The Morse case. J. Phys. C **3**, 181–187 (1986);
Riklund, R., Severin, M., Youyan Liu: The Thue-Morse aperiodic crystal, a link between the Fibonacci quasicrystal and the periodic crystal. Int. J. Mod. Phys. B **1**, 121–132 (1987)
11. Bellissard, J.: Spectral properties of Schrödinger's operator with a Thue-Morse potential. In: Théorie des nombres et Physique. New York: plenum Press (to appear)
12. Luck, J.-M.: Cantor spectra and scaling of gap widths in deterministic aperiodic systems. Phys. Rev. B **39**, 5834–5849 (1989)
13. Bellissard, J.: *K*-theory of C^*-algebras in solid state physics. Lecture Notes in Physics, vol. 257. Berlin, Heidelberg, New York: Springer 1986
14. Collet, P., Eckmann, J.-P.: Iterated maps on the interval as dynamical systems. Boston: Birkhäuser 1980
15. Kotani, S.: Jacobi matrices with random potentials taking finitely many values. Rev. Math. Phys. **1**, 129–133 (1990)
16. Herman, M.: Une méthode pour minorer les exposants de Lyapounov et quelques exemples montrant le caractère local d'un théorème d'Arnold et de Moser sur le tore de dimension 2. Comment. Math. Helvetici **58**, 453–502 (1983)
17. Kunz, H., Livi, R., Sütö, A.: Cantor spectrum and singular continuity for a hierarchical hamiltonian. Commun. Math. Phys. **122**, 643–679 (1989)
18. Martinelli, F., Scoppola, E.: Introduction to the mathematical theory of Anderson localization. Rivista del Nuovo Cimento **10**, (1987)
19. Thouless, D.J.: A relation between the density of states and the range of localization in one dimensional random systems. J. Phys. C **5**, 77–81 (1972);
Herbert, D., Jones, R.: Localized states in disordered systems. J. Phys. C **4**, 1145–1161 (1971)

Communicated by T. Spencer