Commun. Math. Phys. 107, 295–318 (1986)

Communications in
Mathematical
Physics

Symbolic Dynamics for the Renormalization Map of a Quasiperiodic Schrödinger Equation

Martin Casdagli

Mathematics Institute, University of Warwick, Coventry CV4 7AL, United Kingdom
La Jolla Institute, Center for Studies of Nonlinear Dynamics, 3252 Holiday Court, Suite 208, La Jolla, CA 92037, USA

Abstract. A rigorous analysis is given of the dynamics of the renormalization map associated to a discrete Schrödinger operator H on $l^2(\mathbb{Z})$, defined by $H\psi(n)=\psi(n+1)+\psi(n-1)+Vf(n\sigma)\psi(n)$, where V is a real parameter, f is a certain discontinuous period-1 function, and $\sigma=(-1+\sqrt{5})/2$ is the golden mean. The renormalization map for H is a diffeomorphism, T, of $\mathbb{R}^3$, preserving a cubic surface S_V. For $V \geqq 8$ we prove that the non-wandering set of the restriction of T to S_V is a hyperbolic set, on which T is conjugate to a subshift on six symbols. It follows from results in dynamical systems theory that the optimally approximating periodic operators to H have spectra which obey a global scaling law. We also define a set which we call the pseudospectrum" of the operator H. We prove it to be a Cantor set of measure zero, and obtain bounds on its Hausdorff dimension. It is an open question whether the pseudospectrum coincides with the spectrum of H.

Introduction

There has been much interest in Schrödinger operators with a quasiperiodic potential (see [18, 19, 26–28, 33] and references therein). These operators have numerous physical applications. For example, they describe the electron spectrum of periodic crystals in a magnetic field [15], and the electron and phonon spectrum of the recently discovered quasi-crystals [3]. They also arise in the linear stability of motions in classical mechanics [1]. Operators with quasi periodic potential also pose very interesting questions for the functional analyst [33]. They are in some sense intermediate between operators with periodic potential and operators with random potential. Periodic potentials are well known to lead to absolutely continuous "band spectra" and extended eigenstates [31], whereas random potentials lead to pure point spectra and localized eigenstates, in one dimension [20]. In the quasiperiodic case the general belief is that the spectra are Cantor sets. At present, the only theorems in this direction are for a generic set of potentials, which are very well approximated by periodic potentials [2]. In this case, the underlying irrational number generating the quasiperiodicity is a Liouville number. Such numbers form a set of measure zero, thus the result is not as general as it might seem.

The theory of operators with a quasiperiodic potential has a fundamental connection with the small-divisor problems of classical mechanics. Indeed it can be shown using ideas of KAM theory that for sufficiently weak analytic quasiperiodic potentials, most of the spectrum is absolutely continuous [12]. On the other hand, it can also be shown that in certain situations, if the quasiperiodic potential is strong enough, the spectrum has no absolutely continuous component [4, 14]. Thus in a one-parameter family of quasiperiodic potentials, one expects a so-called metal-insulator transition at a certain critical strength of the potential. At the critical value, numerical observations reveal that the spectra of the periodic operators which optimally approximate the quasiperiodic operator have beautiful scaling properties [15]. In the case where the quasiperiodicity is generated by the golden mean, this behavior has to some extent been explained by considering a fixed point of a non-linear renormalization map on a function space, though the theory is not yet rigorous [28].

In this paper, following [10, 17–19, 21, 26, 27], we study a discrete Schrödinger operator with specially chosen discontinuous quasiperiodic potential, dependent on a real parameter V. The number generating the quasiperiodicity is taken to be the golden mean, which has typical diophantine properties. Physically, the operator describes the propagation of phonons in a one-dimensional quasi-crystal [21]. The behavior of the operator is somewhat pathological. Numerical results reveal that its states are neither extended nor localized in the conventional sense [18, 19, 27], and in fact it is known rigorously not to have localized states [10]. It is thus a simple example of a one-parameter family of operators which always lies at criticality. The advantage of studying this operator is that its renormalization map reduces to a non-linear map on a two dimensional space. This fact makes it relatively easy to numerically establish connections between scaling properties of the spectrum and eigenvalues of the linearization of the map at its fixed points [18, 19, 27]. It is the purpose of this paper to make these ideas rigorous, and indicate how they can be extended, by giving a global analysis of the dynamics of the renormalization map.

In the first half of the paper we use geometric methods developed in [11, 25, 34] to show that the renormalization map has a hyperbolic non-wandering set, on which it is conjugate to a subshift on precisely six symbols. The result is restricted to the range $V \geqq 8$. However, we explain why we believe the result to be true for all $V > 0$. We also explain the occurrence of the six symbols by displaying them in the dynamics of the "exactly solvable" case, when $V = 0$. In the second half of the paper we use our results on the dynamics of the renormalization map to deduce properties of the spectrum of the operator. Finite symbol sequences are used to label the band spectra of the optimally approximating periodic operators with period given by the Fibonacci numbers. The scaling properties of these band spectra are naturally described in terms of the symbol sequences. In order to use our results on the renormalization map to deduce properties of the quasiperiodic operator itself, we define a set which we call the "pseudospectrum" of the operator. Hopefully the pseudospectrum coincides with the spectrum, but we have not been able to prove this. However, we show that the pseudospectrum is a Cantor set of measure zero. We then apply results from the ergodic theory of axiom-A diffeomorphisms to deduce the existence of new exponents governing global scaling properties and "ergodic" scaling properties of the periodic operators. We also obtain bounds on the Hausdorff dimension of the pseudospectrum of the quasiperiodic operator in terms of these exponents. Finally we obtain a relationship between symbol sequences and rotation numbers, which have also been used to characterize the spectra of Schrödinger operators [9, 14, 16, 27, 33].

From the point of view of functional analysis, our results are somewhat limited. The approach we use does not enable us to investigate the spectrum of the quasiperiodic operator directly. However, we believe it gives a useful insight into how Cantor set spectra can arise from the complicated dynamics of an underlying renormalization map. From the point of view of dynamical systems, the map we study is a simple example of a renormalization map with a non-trivial dynamical behavior. A renormalization map can usually be guessed to have non-trivial dynamics by an observation of the data it is designed to explain. This has lead to other, more ambitious, attempts at global renormalization schemes [13, 23]. However, we remark that from an observation of numerically obtained band spectra, it would be difficult to infer that our renormalization map requires precisely six symbols to describe it. Thus the renormalization map we have studied serves as a completely solved example, exhibiting non-trivial combinatorics, that may be relevant to the other global renormalization schemes.

In Sect. 1 we define the quasiperiodic operator to be studied, and review the renormalization technique used to analyze it. In Sect. 2 we collect our results on the symbolic dynamics of the renormalization map. In Sect. 3 we use these results to provide a symbol sequence labeling for the spectra of the optimally approximating periodic operators, and for the pseudospectrum of the quasiperiodic operator, which we deduce is a Cantor set of measure zero. In Sect. 4 we obtain a global scaling law for the spectra of the optimally approximating periodic operators. We also introduce the concept of an ergodic scaling law, and obtain bounds on the Hausdorff dimension of the pseudospectrum of the quasiperiodic operator. In Sect. 5 we relate our rigorous results to numerical work of others [27], by obtaining a relationship between symbol sequences and rotation numbers.

1. The Renormalization Technique

The discrete Schrödinger operator acting on $l^2(\mathbb{Z})$ is defined by (1.1),

$$H\psi(n)=\psi(n+1)+\psi(n-1)+v(n)\psi(n), \tag{1.1}$$

where $v(n)\in\mathbb{R}$ denotes the potential at site $n\in\mathbb{Z}$, and $\psi(n)\in\mathbb{C}$ denotes the wave function at site $n\in\mathbb{Z}$. Let $\mathbb{S}^1$ denote the unit circle. The operator H is said to be *quasiperiodic* if $v(n)$ is of the form $v(n)=f(g^n(\theta_0))$, where $\theta_0\in\mathbb{S}^1$, g is a homeomorphism of $\mathbb{S}^1$ with irrational rotation number, and f is a real valued function on $\mathbb{S}^1$.

Restrict attention to the special case where $\theta_0=0$, $g=R_\alpha$ (the rigid rotation through angle α), and f is discontinuous of the form (1.2)

$$f(\phi)=\begin{cases} V & -\sigma<\phi\leqq-\sigma^3 \\ -V & -\sigma^3<\phi\leqq 1-\sigma, \end{cases} \tag{1.2}$$

where $\sigma=(-1+\sqrt{5})/2$ is the golden mean, and $V \geqq 0$. The quasiperiodic operator, Q, to be analyzed is defined by taking $\alpha=\sigma$. The periodic operators, P_n, defined by taking $\alpha=F_{n-1}/F_n$, the optimal approximants to σ, will play a key role in what follows (F_n are the Fibonacci numbers: $F_{n+1}=F_n+F_{n-1}$, $F_1=F_0=1$).

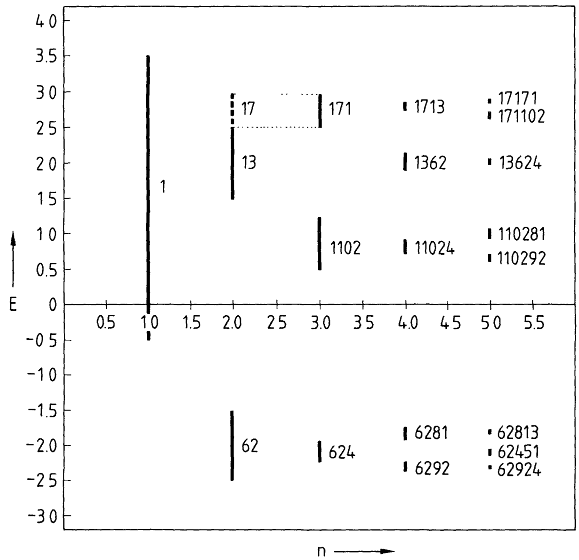

Fig. 1. The bold lines represent the band spectra of P_n for $n=1, \ldots, 5$. The case $V=1.5$ is shown. The bands have been labeled using the symbol sequence scheme of Sect. 3.1. The dotted line 17 illustrates the bifurcation of Sect. 3.1

As is well known [31], the spectrum, B_n, of the operator P_n is given by (1.3)

$$B_n=\{E \in \mathbb{R} \mid |\operatorname{trace} M(n)| \leqq 2\}, \tag{1.3}$$

where $M(n)=S(F_n-1) \ldots S(1)S(0)$ is the product of so-called transfer matrices

$$S(i)=\begin{bmatrix} E-v(i) & -1 \\ 1 & 0 \end{bmatrix}.$$

The identity (1.3) provides a simple criterion for computing B_n numerically. In this way one obtains the sequence of band spectra illustrated in Fig. 1. It was the remarkable self-similarity of this picture which provided the impetus for much of our research. Unfortunately there is no known identity analogous to (1.3) in the quasiperiodic case. However, much numerical work has been done using criteria similar to (1.3) [18, 19, 27]. This motivates us to define the "pseudospectrum" of the operator Q as follows.

Definition. We define the *pseudospectrum*, B_∞, of the operator Q by (1.4)

$$B_\infty=\{E \in \mathbb{R} \mid |\operatorname{trace} M(n)| \text{ is bounded as } n \rightarrow \infty\}. \tag{1.4}$$

In this paper we give a comprehensive description of the structure of the pseudospectrum of the operator Q. We are optimistic that the pseudospectrum of

Q coincides with the spectrum of Q, however this is not clear. Firstly, $M(n)$ gives some information on the wavefunction on a subset of sites only. Secondly, it is not clear that for all E in the spectrum, the wavefunctions must be bounded. The usual result is that for almost all E, with respect to the spectral measure class, one of the wavefunctions is polynomially bounded.

A renormalization theory, developed in [26–29], allows us to determine the structure of the sets B_n and B_∞ using methods of dynamical systems theory. This theory shows that the matrix $M(n)$ is given by a matrix product of the form $BAABA\ldots$, generated by n iterations of the "renormalization map" $R(A,B) = (BA, A)$, with the initial conditions $A = \begin{pmatrix} E+V & -1 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} E-V & -1 \\ 1 & 0 \end{pmatrix}$. The map R, which acts on the six-dimensional space of pairs of unimodular matrices, has been studied numerically in [27]. However, to determine properties of the spectrum, it suffices to study a simpler map.

It was observed in [19] that the quantity $x_n = \frac{1}{2}\,\mathrm{trace}\, M(n)$, satisfies the "trace identity" (1.5),

$$x_{n+1} = 2x_n x_{n-1} - x_{n-2} \tag{1.5}$$

with initial conditions $x_1 = \dfrac{E+V}{2}$, $x_0 = \dfrac{E-V}{2}$, $x_{-1} = 1$. Thus the map $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by (1.6)

$$T(x, y, z) = (2xy - z, x, y) \tag{1.6}$$

determines B_n via (1.7),

$$B_n = \{E \in \mathbb{R} | \pi_1 T^{n-1}(L_V(E)) \in [-1, 1]\}, \tag{1.7}$$

where $L_V : \mathbb{R} \to \mathbb{R}^3$ is the linear map defined by (1.8),

$$L_V(E) = \left(\frac{E+V}{2}, \frac{E-V}{2}, 1 \right) \tag{1.8}$$

and π_1 is the projection in the x direction. It is the renormalization map T which will be studied in Sect. 2. The map T also determines the pseudospectrum B_∞, of the operator Q, by (1.9).

$$B_\infty = \{E \in \mathbb{R} | \pi_1 T^n(L_V(E)) \text{ is bounded as } n \to \infty\}. \tag{1.9}$$

2. Symbolic Dynamics of the Renormalization Map

In this section we introduce some concepts from symbolic dynamics, and give our results on the renormalization map T. The results will be used in subsequent sections to derive detailed information on the sets B_n and B_∞. We make use of the following simple properties of the map T [17].

(1) T is a volume preserving diffeomorphism of $\mathbb{R}^3$ and $T^{-1} = \varrho_{xz}^{-1} \circ T \circ \varrho_{xz}$, where ϱ_{xz} is the reflection in the $x = z$ plane.

(2) T preserves the family of cubic surfaces $\{S_V | V \in \mathbb{R}^+\}$ defined by (2.1).

$$S_V = \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 + z^2 - 2xyz = 1 + V^2\}. \tag{2.1}$$

The restriction of the map T to the surface S_V is denoted by T_V.

(3) A necessary condition for a bi-infinite sequence $\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots$ generated by (1.5) to remain bounded is that it has property $\mathbb{P}$: no two consecutive terms of the sequence have modulus greater than unity.

Our objective is to prove that for $V \geqq 8$, the non-wandering set, Ω_V, of T_V is a Cantor set with a hyperbolic structure. In fact, we believe this to be true for all $V>0$, as conjectured in [17], for reasons we give at the end of this section. The technique we use is well known, and has been reviewed in [25]. For an application to the Hénon map, see [11]; for convenience we summarize the main ideas here. Property (3) above is used to find a compact set R_V such that the orbit under T_V of any point lying outside R_V is unbounded. It follows that Ω_V is contained in the set $\Lambda = \bigcap_{n=-\infty}^{\infty} T_V^n(R_V)$. It turns out that the set R_V consists of a finite number of disjoint closed regions $R_1, \ldots, R_N$, whose images under the map T_V intersect the regions $R_1, \ldots, R_N$ in a manner similar to Smale's horseshoe construction [34]. Careful estimates on the size and shape of the regions $R_1, \ldots, R_N$ and their images enable us to deduce that the set Λ is a hyperbolic set each point of which may be uniquely coded by a bi-infinite sequence of symbols chosen from the set $\{1, \ldots, N\}$ according to which of the sets $R_1, \ldots, R_N$ contain its successive backward and forward iterates. It follows that the points of Λ may be put into correspondence with a Cantor set, and that the action of the map T_V on the set Λ is described by a "symbolic dynamics." It is then easy to construct a dense orbit for this symbolic dynamics, to deduce that $\Omega_V = \Lambda$, so that Ω_V is a Cantor set. We now make these ideas more precise.

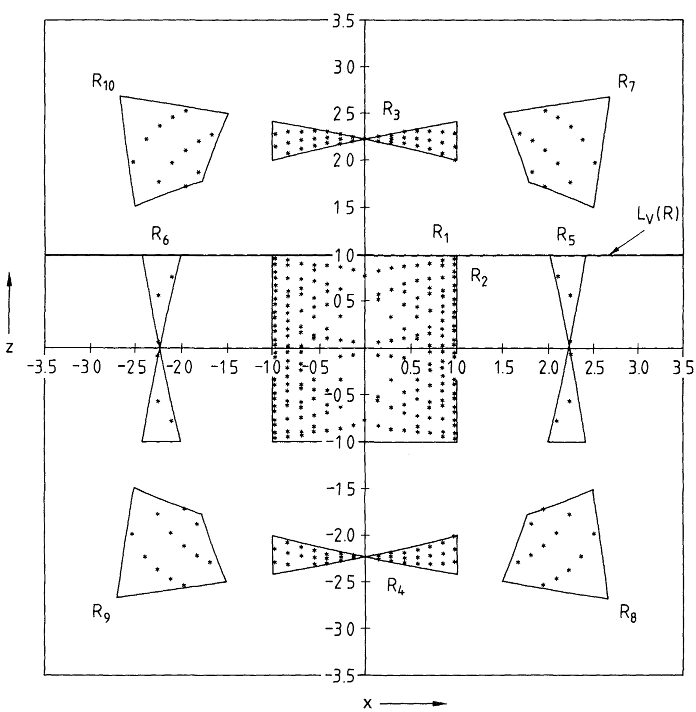

Fig. 2. The xz projections of $R_1, \ldots, R_{10}$ for $V=2$. Note that R_2 lies vertically below R_1 on the surface S_V. Also illustrated is the line $L_V(\mathbb{R})$, relevant to the operators P_n and Q

The set $\bar{R}_V$ is defined by (2.2),

$$R_V = \{(x, y, z) \in S_V | \mathbf{w}(x, y, z) \text{ has property } \mathbb{P}\}\,, \tag{2.2}$$

where $\mathbf{w}(x, y, z) = 2yz - x, x, y, z, 2xy - z$. We remark on this special choice of R_V at the end of this section. Note that $\mathbf{w}(x, y, z)$ is a subsequence of the bi-infinite sequence $\ldots x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots$ generated by the recurrence relation (1.4) using initial condition $x_{-1}, x_0, x_1 = x, y, z$. Thus it follows immediately from property (3) above that if $(x, y, z) \notin R_V$, then the orbit of (x, y, z) is unbounded, so that $\Omega_V \in R_V$. The set R_V is illustrated in Fig. 2; it consists of 10 disjoint regions $R_1, \ldots, R_{10}$ which are defined as follows. Let the symbols $L^-, s, L^+, *$ denote the intervals $(-\infty, 1], [-1, 1], [1, \infty)$, $(-\infty, \infty)$ respectively. The sets $R_1, \ldots, R_{10}$ are defined according to which of the intervals $L^-, s, L^+, *$ the coordinates of $\mathbf{w}(x, y, z)$ lie in, by Table 2.1.

Table 2.1

$R_1 = {*}sL^+s{*}$,	$R_2 = {*}sL^-s{*}$		
$R_3 = L^-ssL^+s$,	$R_4 = L^+ssL^-s$,	$R_5 = sL^+ssL^-$,	$R_6 = sL^-ssL^+$
$R_7 = sL^+sL^+s$,	$R_8 = sL^+sL^-s$,	$R_9 = sL^-sL^-s$,	$R_{10} = sL^-sL^+s$

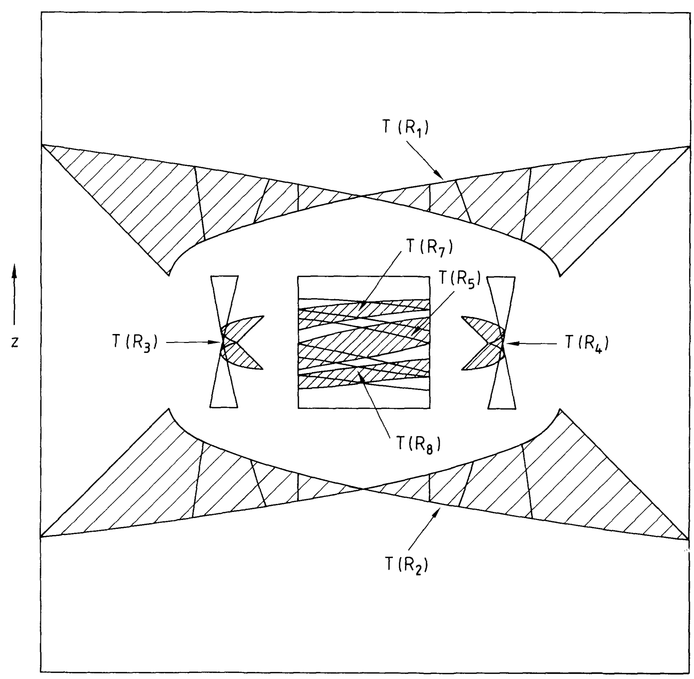

Fig. 3. The xz projections of the regions $T(R_1), \ldots, T(R_5)$, $T(R_7)$, $T(R_8)$ for $V=2$. The regions $T(R_6)$, $T(R_9)$, $T(R_{10})$ lie in the region R_2, and have not been shaded

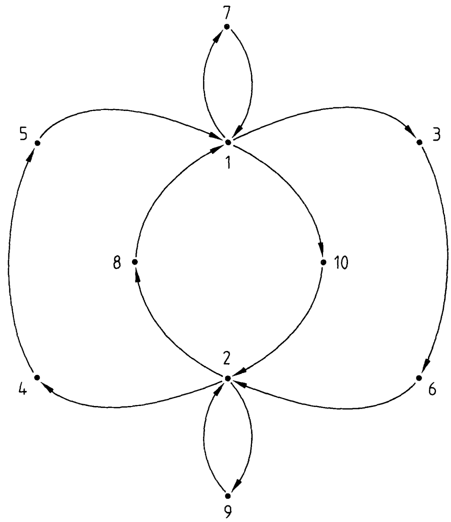

Fig. 4. The directed graph G, defining the subshift σ_A on 10 symbols

The images of the regions $R_1, \ldots, R_{10}$ under the map T_V are illustrated in Fig. 3. It may be verified by an inspection of Fig. 3 that the regions $R_1, \ldots, R_{10}$ satisfy $T(R_i) \cap R_j \neq \emptyset$ whenever there is a connection $i \to j$ in the directed graph G of Fig. 4. This motivates us to define a 10×10 matrix A by $A_{ij} = 1$ if there is a connection $i \to j$ in the graph G, and $A_{ij} = 0$ otherwise. The next lemma establishes some of the above observations.

Lemma 2.1. *For $V > 2$ the regions $R_1, \ldots, R_{10}$ are closed and disjoint, their union forms the whole of R_V, and $A_{ij} = 0$ implies $T(R_i) \cap R_j = \emptyset$.*

Proof. We first show that the union of the regions $R_1, \ldots, R_{10}$ forms the whole of R_V. This amounts to establishing that Table 2.1 exhausts all the combinations of the symbols $L^\pm$ and s allowed as labels of R_V. By property $\mathbb{P}$, the symbols $L^\pm$ must be both preceded and followed by an s if they are to label a point of R_V. Also, it can be verified that the combinations L^+ssL^+ and L^-ssL^- are disallowed by taking $L^\pm, s, s$ as initial conditions in the recurrence relation (1.5). Finally, the combination sss is disallowed when $V > 2$, because a point $(x, y, z) \in S_V$ with $(x, y, z) \in (s, s, s)$ cannot satisfy $x^2 + y^2 + z^2 - 2xyz = 1 + V^2$. Thus Table 2.1 exhausts all the possible combinations.

To show that the regions $R_1, \ldots, R_{10}$ are disjoint, we first observe that they are labeled by distinct sequences of the symbols $L^\pm, s$. However, the intervals L^+ and L^- just overlap with the interval s. We must show that this does not cause the regions $R_1, \ldots, R_{10}$ to overlap when $V > 2$. It suffices to show that if the point (x, y, z) is contained in R_V and $\mathbf{w}(x, y, z)$ has a coordinate $w_i \in L^\pm$, then $|w_i| \geqq |V-1|$, since it then follows that $L^\pm, s$ could have been chosen to be the disjoint closed intervals $(-\infty, -V+1]$, $[-1, 1]$, $[V-1, \infty)$ without altering the definition of R_V. To show that $w_i \in L^\pm$ implies $|w_i| \geqq |V-1|$, we observe from Table 2.1 that w_i is necessarily a coordinate of a 3-vector (x, y, z), whose other two coordinates are represented by the symbol s. Without loss of generality, suppose $w_i = y$. Then since $(x, y, z) \in S_V$, we have $y = xz \pm (V^2 + (1 - x^2)(1 - z^2))^{1/2}$. Hence $|y| \geqq |V-1|$, as required.

Finally, we show that $A_{ij}=0$ implies $R_i \cap T(R_j)=\emptyset$. From Table 2.1 it can be verified that a necessary condition for $R_i \cap T(R_j)$ to be non-empty is that there is a connection $i \to j$ in the graph G, so that $A_{ij}=1$. Thus if $A_{ij}=0$, we must have $R_i \cap T(R_j)=\emptyset$ as required. $\square$

Remark. In proving Theorem 2.1 below, we show that the regions $R_1, \ldots, R_{10}$ are non-empty and that $A_{ij}=1$ implies $R_i \cap T(R_j) \neq \emptyset$.

We now introduce some definitions from symbolic dynamics [25]. Given an "alphabet" $\{1, \ldots, m\}$ of m symbols, define the set Σ_m of two-sided symbol sequence by $\Sigma_m = \prod_{n=-\infty}^{\infty} \{1, \ldots, m\}$. When $\{1, \ldots, m\}$ is endowed with the discrete topology, and Σ_m with the product topology, Σ_m is called a *shift space*, and it is homeomorphic to a Cantor set. Define the *shift* $\sigma: \Sigma_m \to \Sigma_m$ by $\sigma(\mathbf{s})_n = s_{n+1}$, where s_n is the n^{th} symbol is $\mathbf{s}$. Let A be the 10×10 matrix defined above. Then define $\Sigma(A)$ by (2.3),

$$\Sigma(\mathrm{A}) = \{\mathbf{s} \in \Sigma_{10} | A_{s_i s_{i+1}} = 1 \text{ for all } i \in \mathbb{Z}\}, \tag{2.3}$$

and define the *subshift* σ_A to be $\sigma|_{\Sigma(A)}$. Now consider the map T_V acting on R_V. Our objective is to conjugate the map σ_A to T_V; that is we must construct a map $x: \Sigma(A) \to R_V$ such that $x \circ \sigma_A = T_V \circ x$.

We define the map $x: \Sigma(A) \to R_V$ by labeling points of R_V using symbol sequences, as follows. For each $\mathbf{s} \in \Sigma(A)$, define the sets $V_{\mathbf{s}^+}$ and $H_{\mathbf{s}^-}$ by

$$V_{\mathbf{s}^+} = \bigcap_{n \in N} V_{s_0 s_1 \ldots s_n} \quad \text{and} \quad H_{\mathbf{s}^-} = \bigcap_{n \in N} H_{s_0 s_{-1} \ldots s_{-n}},$$

where

$$V_{s_0 s_1 \ldots s_n} = R_{s_0} \cap T^{-1} R_{s_1} \cap \ldots \cap T^{-n} R_{s_n}$$

and

$$H_{s_0 s_{-1} \ldots s_{-n}} = R_{s_0} \cap T R_{s_{-1}} \cap \ldots \cap T^n R_{s_{-n}}.$$

Then the map $x: \Sigma(A) \to R_V$ is defined by $x(\mathbf{s}) = V_{\mathbf{s}^+} \cap H_{\mathbf{s}^-}$. It is an immediate consequence of this construction that if $x(\mathbf{s}) \neq \emptyset$ then $T^n x(\mathbf{s}) \in R_{s_n}$ for all $n \in \mathbb{Z}$, so that $x(\mathbf{s})$ has its past and future history coded by the symbol sequence $\mathbf{s}$. Thus, by construction of the map x, we are guaranteed that $T_V \circ x = x \circ \sigma_A$. The main problem is now to show that the map x is well defined [i.e. that $x(\mathbf{s})$ consists of a single point in R_V], and is continuous. In order to do this, we obtain bounds on the sizes of the sets $V_{s_0 \ldots s_n}$ and $H_{s_0 \ldots s_{-n}}$. To state our results precisely, we introduce some geometrical concepts [25].

Let $I^2 = [a, b] \times [a, b]$ be a square in $\mathbb{R}^2$. Given $\mu \in (0, 1)$, we call a curve in I^2 a μ-horizontal curve if it is the graph, $gr(u)$, of a continuous function $u: I \to I$ satisfying $|u(x_1) - u(x_2)| \leqq \mu |x_1 - x_2|$ for all x_1, x_2 in I. If $gr(u_1)$ and $gr(u_2)$ are two such curves with $u_1(x) < u_2(x)$ for all x in I, then we call the set H defined by

$$H = \{(x, y) \in I^2 | x \in I,\ u_1(x) \leqq y \leqq u_2(x)\}$$

a μ-horizontal strip, with diameter $d(H) = \max_{x \in I} |u_2(x) - u_1(x)|$. μ-vertical strips are defined similarly. In the following we will refer to these concepts in the x, z coordinate system. We are now in a position to state our main result.

Theorem 2.1. *For $V \geqq 8$ and $s_0 \in \{1, 2\}$, there exists $\mu \in (0,1)$ and $\nu \in (0,1)$ such that $V_{s_0 \ldots s_n}$ (respectively $H_{s_0 \ldots s_n}$) is a μ-vertical (respectively μ-horizontal) strip of diameter $\leqq \nu^n$, whenever $s_1 \ldots s_n$ satisfies $A_{s_i s_{i+1}} = 1$ for all $0 \leqq i \leqq n-1$. Moreover if $A_{s_i s_{i+1}} = 0$ for some $0 \leqq i \leqq n-1$, then $V_{s_0 \ldots s_n}$ and $H_{s_0 \ldots s_n}$ are empty.*

The importance of Theorem 2.1 is that it allows us to apply the ideas of [25], outlined earlier in this section, to deduce the following corollary.

Corollary 2.2. *For $V \geqq 8$, the map $x : \Sigma(A) \to R_V$ is well defined, satisfies $T_V \circ x = x \circ \sigma_A$, and is a homeomorphism onto Ω_V. Moreover, Ω_V is a hyperbolic set, homeomorphic to a Cantor set.*

Remark. It may be shown that the matrix A has 6 non-zero eigenvalues ($1+\sigma$, σ, $-\omega$, $-\bar{\omega}$, -1, 1 where $\omega = (-1 + i\sqrt{3})/2$), and that the symbolic dynamics for T must therefore use at least 6 symbols [6]. Also A is irreducible (by inspection of G), and A is mixing (there is a k such that $A^k_{ij} > 0$ for all i, j), since it has a unique eigenvalue of largest modulus. These facts will be used later. An inspection of the graph G reveals that the subshift σ_A is conjugate to a subshift σ'_A obtained from a graph G', defined by identifying the symbols 7 and 8 with 5, and 9 and 10 with 6 in the graph G. Thus the map T_V is conjugate to a subshift on precisely six symbols.

Before embarking on the proof of Theorem 2.1, we make some remarks on the choice of the set R_V of (2.2) and the restriction to $V \geqq 8$. In fact the restriction of Corollary 2.2 to the range $V \geqq 8$ is related to the artificial choice of the region R_V. We chose this region so that we could apply the techniques of [25]. The particular choice in (2.2) leads to the simplest application of these techniques. However, there is a natural choice for R_V, which applies for all $V > 0$, and which gives a good insight into how the graph G arises. In fact, the dynamics of the graph G is subtly embedded in the dynamics of the map T_V for $V = 0$, as we now describe.

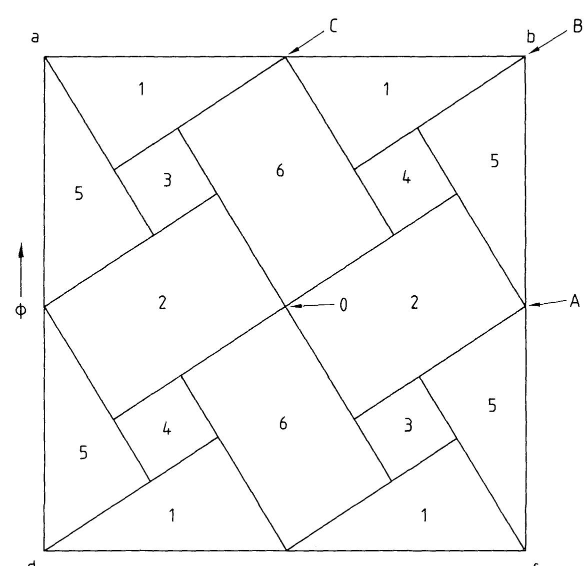

Fig. 5. The partition of S_V for T when $V=0$. Note that $ab \sim dc$ and $ad \sim bc$

It was shown in [17] that when $V=0$, the map T_V is conjugate to the map $A:\mathbb{S}^2\to\mathbb{S}^2$, where

$$\mathbb{S}^2=\{(\theta,\phi)\in\mathbb{R}^2|\theta\sim\theta+1,\ \phi\sim\phi+1,\ (\theta,\phi)\sim-(\theta,\phi)\}$$

is a sphere, and A is the Anosov-like map $A(\theta,\phi)=(\theta+\phi,\theta)$. Figure 5 illustrates a Markov partition for A which glues together under $\sim$ in a well-defined fashion, and has the symbolic dynamics of G', where G' is the graph with 6 symbols, equivalent to G. When $V=0$, paths in G' do not uniquely label points in $\mathbb{S}^2$, and indeed Ω_0 is not a Cantor set. However, as V is increased infinitesimally from 0, there is a bifurcation. The fixed point O bifurcates to a period 2 cycle $O^{\pm}$, and the period 3 cycle ABC bifurcates to a period 6 cycle $A^{\pm}B^{\pm}C^{\pm}$ [17]. It may then be verified that a partition of regions with boundaries made up of the local stable and unstable manifolds of $O^{\pm}, A^{\pm}, B^{\pm}, C^{\pm}$ has the same symbolic dynamics as the partition of $\mathbb{S}^2$, but that the members of the partition no longer overlap. Moreover numerical experiments reveal that the partition now has a hyperbolic structure and that all orbits falling outside the partition become unbounded, so that Ω_V becomes a Cantor set for all $V>0$. However we have been unable to make these ideas rigorous, as we do not have good enough bounds on the stable and unstable manifolds of $O^{\pm}, A^{\pm}, B^{\pm}, C^{\pm}$. We therefore implicitly assume $V\geqq 8$ in subsequent sections.

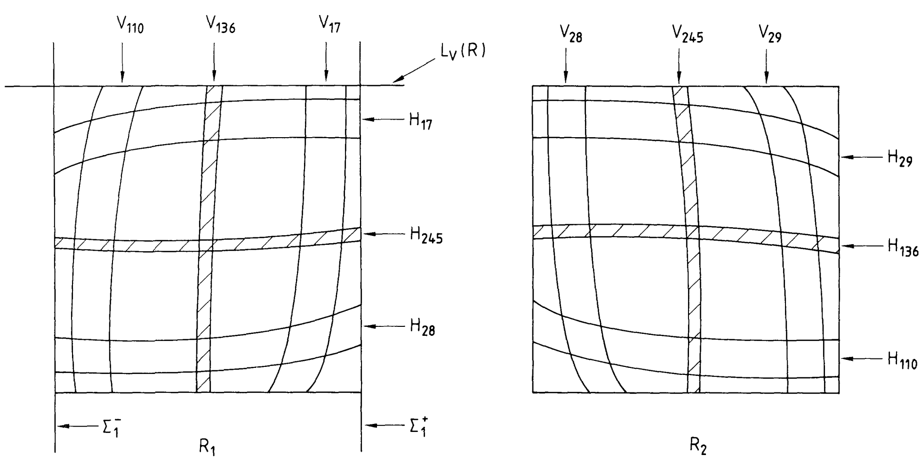

Fig. 6. An illustration of the vertical and horizontal strips on which the map ϕ acts. Also illustrated is the line $L_V(\mathbb{R})$ and the bifurcation lines $\Sigma_1^{\pm}$ of Sect. 4.1

Proof of Theorem 2.1. The second part of the theorem is an immediate consequence of Lemma 2.1. The bulk of the proof amounts to a careful manipulation of inequalities on the map T_V. We consider a map ϕ which embodies the dynamics of the map T_V in more manageable form (see Fig. 6). The map ϕ is defined on $\bigcup_{s\in S} V_s$, where $S=\{17,110,28,29,136,245\}$ as follows:

$$\phi(x)=\begin{cases}T_V^2(x) & x\in V_{17}\cup V_{110}\cup V_{28}\cup V_{29}\\ T_V^3(x) & x\in V_{136}\cup V_{245}\end{cases}. \tag{2.4}$$

The advantage of studying the map ϕ is that all of its dynamics is concentrated in the regions R_1 and R_2, where the notions of strips being vertical and horizontal in the $x-z$ coordinate system works well. Note that all of the dynamics of the map T_V is embedded in the dynamics of the map ϕ. For example, if a vertical strip $V_{s_0 \dots s_n}$, with $s_0, s_n \in \{1,2\}$, is to be non-empty, then by Lemma 2.1 we must have $A_{s_i s_{i+1}} = 1$ for all i. By an inspection of the graph G, this implies that the symbol sequence $s_0 \dots s_n$ can be split up into a sequence of symbols drawn from S, so that the set $V_{s_0 \dots s_n}$ is given by an intersection of the form $V_{t_0} \cap \phi^{-1} V_{t_1} \cap \dots \cap \phi^{-m} V_{t_m}$, where $t_i \in S$. Define B to be the matrix (2.5) with respect to the basis S,

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \end{pmatrix} \tag{2.5}$$

and let $s_0, \dots, s_n \in S$. Theorem 2.1 is thus equivalent to showing that there exists $\mu \in (0,1)$ and $\nu \in (0,1)$ such that $V_{s_0 \dots s_n}$ (respectively $H_{s_0 \dots s_n}$) is a μ-vertical (respectively μ-horizontal) strip of diameter $\leqq \nu^n$ whenever $B_{s_i s_{i+1}} = 1$ for all $0 \leqq i \leqq n-1$. For $V \geqq 8$, we claim that this is true for $\mu = 1/3$ and $\nu = \mu(1-\mu)^{-1} = 1/2$. Simple generalizations of theorems in [25] allow us to reduce the proof to checking 4 conditions for the map ϕ. Fixing $V \geqq 8$, $\mu = 1/3$, and referring to the x, z coordinate system, the conditions are as follows:

(1) For all $s \in S$, V_s (respectively H_s) are non-empty disjoint μ-vertical (respectively horizontal) strips satisfying $\phi(V_s) = H_s$.

(2) For all $s \in S$, ϕ maps vertical (respectively horizontal) boundaries of V_s to vertical (respectively horizontal) boundaries of H_s.

(3) For $s, t \in S$, $H_s \cap V_t \neq \emptyset$ if and only if $B_{st} = 1$.

(4) The cone field $S^+ = \{(\xi, \eta) | |\eta| \leqq \mu |\xi|\}$ defined over the region $X = \left\{ \bigcup_{s \in S} V_s \right\} \cap \left\{ \bigcup_{s \in S} H_s \right\}$ is mapped into itself by $d\phi_x$ for all $x \in X$, in such a way that if $(\xi_0, \eta_0) \in S^+$ and $(\xi_1, \eta_1) = d\phi_x(\xi_0, \eta_0)$, then $|\xi_1| \geqq \mu^{-1} |\xi_0|$. Also the cone field $S^- = \{(\xi, \eta) | |\eta| \geqq \mu^{-1} |\xi|\}$ defined over X is mapped into itself by $d\phi_x^{-1}$ for all $x \in X$, in such a way that if $(\xi_0, \eta_0) \in S^-$ and $(\xi_1, \eta_1) = d\phi_x^{-1}(\xi_0, \eta_0)$, then $|\eta_1| \geqq \mu^{-1} |\eta_0|$.

Using the symmetry $T^{-1} = \varrho_{xz}^{-1} \circ T \circ \varrho_{xz}$, it suffices to check the above conditions on the vertical strips and on S^+.

(1) We check this for V_{17}, the other calculations being similar. From Table 2.1 it can be seen that $V_{17} = R_1 \cap T^{-2} R_1$. The region R_1 is represented by the symbols $*sL^+s*$, and a simple calculation reveals that the right-hand vertical boundary of R_1 is given by the line $L_1 = \{(1, V+t, t) | t \in [-1, 1]\}$. The right-hand vertical boundary of V_{17} is given by $R_1 \cap C$, where $C = T^{-2} L_1$ is the curve defined by $C = \{(x(t), y(t), z(t)) | t \in [-1, 1]\}$, where

$$(x(t), y(t), z(t)) = (t, 2t(t+V) - 1, 4t^3 + 4Vt^2 - 3t - V). \tag{2.6}$$

The curve C intersects R_1 in a non-empty curve C', since $z(1/2 + 1/(2V)) > 1$ and $y(t) > V-1$ if $t \in [1/2, 1/2 + 1/(2V)]$. Also C' is a μ-vertical curve, since

$dz(t)/dx(t) > 4V$ if $t \in [1/2, 1/2 + 1/(2V)]$. Similarly, the left-hand vertical boundary of V_{17} is a μ-vertical curve. It may be verified that there are no other intersections of the boundary of $T^{-2}R_1$ with R_1, and thus that V_{17} is a non-empty μ-vertical strip. It follows that the horizontal boundaries of V_{17} are given by pieces of the horizontal boundaries of R_1. The V_s are disjoint because the R_i are disjoint.

(2) The vertical boundaries of the V_s are given by pre-images of the vertical boundaries $\partial_v R_i$ of the R_i, $i \in \{1,2\}$, and hence will be mapped by ϕ to $\partial_v R_i$. A calculation similar to (1) above shows that the $\partial_v R_i$ define the vertical boundaries of the H_s, and the result follows.

(3) If $B_{st} = 0$ then $H_s \cap V_t = \emptyset$, since the R_i are disjoint, by Lemma 2.1. If $B_{st} = 1$, then calculations for the boundaries of H_s and V_t as in (1), will reveal that $H_s \cap V_t \neq \emptyset$.

(4) The set X is covered by $\{V_s | s \in S\}$, and we perform the necessary computations on the cone field S^+ over V_{17}, the other cases being similar. The tangent plane to R_1 at (x,y,z) is given by $\{(\xi,\zeta,\eta) | (\xi,\eta) \in \mathbb{R}^2\}$, where $\zeta = \zeta(\xi,\eta)$ satisfies (2.7),

$$(x - yz)\xi + (y - xz)\zeta + (z - xy)\eta = 0\,, \tag{2.7}$$

and $y = y(x,z)$ is given by (2.8),

$$y(x,z) = xz + (V^2 + (1-x^2)(1-z^2))^{1/2}\,. \tag{2.8}$$

The Jacobian matrix $M(x,y,z)$ of dT at (x,y,z) is given by (2.9),

$$M(x,y,z) = \begin{pmatrix} 2y & 2x & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{2.9}$$

We must show that whenever $(x,y,z) \in V_{17}$ and ξ_0, η_0 is such that $|\eta_0| \leqq |\xi_0|/3$, then $|\eta_1| \leqq |\xi_1|/3$ and $|\xi_1| \geqq 3|\xi_0|$, where ξ_1 and η_1 are given by (2.10),

$$(\xi_1, \zeta_1, \eta_1) = M(T(x,y,z))M(x,y,z)(\xi_0, \zeta_0, \eta_0)\,. \tag{2.10}$$

By linearity we may assume that $\xi_0 = 1$ and $\eta_0 \in [-1/3, 1/3]$. In performing the calculation for (1), it was established that if $(x,y,z) \in V_{17}$, then $x \in \hat{x} = [1/2 - 1/(2V), 1/2 + 1/(2V)]$ and $z \in \hat{z} = [-1, 1]$. Thus, using the formalism of interval arithmetic [24], it suffices to show that the vector of intervals $\hat{\mathbf{v}}$ given by (2.11),

$$\hat{\mathbf{v}} = \hat{M}^2(1, \hat{\zeta}, [-1/3, 1/3]) \tag{2.11}$$

satisfies $\hat{v}_3/\hat{v}_1 \subset [-1/3, 1/3]$ and $\hat{v}_1 \cap (-3,3) = \emptyset$, where $\hat{y}, \hat{\zeta}$ are the intervals obtained by substituting $\hat{x}, \hat{z}$ for x, z in Eqs. (2.7), (2.8), and $\hat{M}^2 = M(T(\hat{x},\hat{y},\hat{z}))M(\hat{x},\hat{y},\hat{z})$ is a matrix of intervals. After some computation we arrive at $\hat{y} \subset [V - 1/2 - 1/(2V),\ V + 1/2 + 1/(2V)]$, $\hat{\zeta} \subset [-4, 4]$ (using $V \geqq 5$), and (2.12).

$$\hat{M}^2 \subset \begin{pmatrix} [4V-8, 4V+8+8/V] & [-2/V,\ 7/(3V)] & [1-1/V,\ 1+1/V] \\ [2V-1-1/V,\ 2V+1+2/V] & [1-1/V,\ 1+1/V] & -1 \\ 1 & 0 & 0 \end{pmatrix}. \tag{2.12}$$

Finally, substituting into Eq. (2.11), and using $V \geqq 5$, it may be verified that $\hat{\mathbf{v}}$ has the required properties. □

Remark. Some of the other computations require $V \geqq 8$. We could relax this requirement by covering the set X more efficiently, with a large number of small rectangles. A non-constant cone field could also be used, and the resulting interval arithmetic could be performed rigorously on a computer.

3. Qualitative Properties of the Spectrum

In this section we apply the results of Sect. 2 to obtain qualitative results on the spectrum of the periodic operators P_n, and the pseudospectrum of the quasi-periodic operator Q.

3.1. The Periodic Operators

We now describe how the symbolic dynamics of Sect. 2 gives rise to a labeling of the spectra B_n of the periodic operators P_n. The identity (1.7) of Sect. 1 states that E is in the spectrum B_n if the vector $T^{n-1}L_V(E)$ has x-coordinate of modulus less than one. Observe that $L_V(\mathbb{R})$ is a line in S_V intersecting vertical boundaries of the sets R_1 and R_6 (see Fig. 2). It therefore intersects the sets $V_{s_0 \ldots s_{n-1}}$, where $s_0 \in \{1, 6\}$. Also $T^{n-1} V_{s_0 \ldots s_{n-1}} \subset R_{s_{n-1}}$, and the only regions R_i with x-coordinates of modulus less than one are the regions R_1, R_2, R_3, and R_4. Thus E is in the spectrum B_n if $\{L_V(E)\} \in V_{s_0 \ldots s_{n-1}}$, where $s_0 \in \{1, 6\}$ and $s_{n-1} \in \{1, 2, 3, 4\}$. This motivates us to define a space $\Sigma'_n(A)$ of symbol sequences of length n by (3.1),

$$\Sigma'_n(\mathrm{A}) = \left\{ s_0 \ldots s_{n-1} \in \prod_{i=0}^{n-1} \{1, \ldots, 10\} | s_0 \in \{1, 6\},\ s_{n-1} \in \{1, 2, 3, 4\}, \right.$$
$$\left. A_{s_i s_{i+1}} = 1 \text{ for } 0 \leqq i \leqq n-2 \right\} \tag{3.1}$$

and a map $b : \Sigma'_n(A) \to B_n$ by (3.2),

$$b(s_0 \ldots s_{n-1}) = \{E \in \mathbb{R} | \{L_V(E)\} \in V_{s_0 \ldots s_{n-1}}\}. \tag{3.2}$$

The next lemma states that the spectrum B_n is completely described by the map b.

Lemma 3.1. *The map b is an injection, and the images of distinct symbol sequences under b are disjoint non-empty closed intervals in B_n.*

Proof. By Theorem 2.1 the sets $V_{s_0, \ldots, s_{n-1}}$, such that $s_0 \ldots s_{n-1} \in \Sigma'_n(A)$, are disjoint non-empty vertical strips intersecting the line $L_V(\mathbb{R})$. Thus the images of distinct symbol sequences under the map b are disjoint non-empty closed intervals, and b is an injection. Also, it may be deduced from the combinatorics of the graph G that $\mathrm{card}(\Sigma'_n(A)) = F_n$, so that we have constructed F_n bands in B_n by the above method. To show that the map b is a surjection, we must show that no other bands arise. Define the polynomial $h_n : \mathbb{R} \to \mathbb{R}$ of degree F_n by $h_n(E) = \pi_1 T^{n-1} L_V(E)$ so that $B_n = h_n^{-1}[-1, 1]$. The pre-image of an interval under a polynomial of degree d consists of at most d disjoint intervals. Thus the F_n bands constructed above exhaust the spectrum, and no other bands can arise. □

Remark. The spectrum B_n is labeled by ten symbols rather than the optimal six required to label Ω_V. This is because in the optimal labeling, the symbols 9 and 10 are identified with 6, but the line $L_V(\mathbb{R})$ only intersects R_6, and not R_9 or R_{10}. Thus in the application of the map T to the description of the spectrum, the choice (2.2) for the set R_V is in some sense a natural one. Indeed the regions R_i just overlap when $V=1.5$, and this bifurcation is reflected in the "band structure" $\{B_n|n\in\mathbb{Z}^+\}$. For example the band $b(171)$ just overlaps the band $b(13)$ when $V=1.5$ (see Fig. 1). The above labeling of B_n by symbol sequences also explains the lack of nesting of the band structure. This lack of nesting results because $b(s_0\dots s_{n-1}s_n)\in B_n$ does not imply $b(s_0\dots s_{n-1})\in B_{n-1}$, since $s_n\in\{1,2,3,4\}$ does not imply $s_{n-1}\in\{1,2,3,4\}$. For example $s_{n-1}s_n$ could be the pair 62.

The above labeling is important, because in Sect. 4 the scaling properties of B_n to be deduced from the dynamics of T are stated directly with respect to this labeling. We have labeled Fig. 1 using an ordering property of the map b. We define an ordering on $\Sigma_n'(A)$ so that if $s,t\in\Sigma_n'(A)$ satisfy $s>t$ then $\pi_1 V_s>\pi_1 V_t$. Then from (3.2) and the definition (1.8) of L_V, it follows that $b(s)>b(t)$, so that the map b is order preserving. The ordering on $\Sigma_n'(A)$ is defined as follows.

Definition. Let $s=s_0\dots s_{n-1}$ and $t=t_0\dots t_{n-1}$ be distinct symbol sequences in $\Sigma_n'(A)$, and let i be the first place in which s differs from t. Then define $s>t$ if either (a) $i=0$ and $(s_0,t_0)=(1,6)$, or (b) $s_{i-1}=t_{i-1}=1$ and $(s_i,t_i)\in\{(7,3),(7,10),(3,10)\}$, or (c) $s_{i-1}=t_{i-1}=2$ and $(s_i,t_i)\in\{(8,4),(8,9),(4,9)\}$

Lemma 3.2. *The map b is order preserving, and $s>t$ implies $\pi_1 V_s>\pi_1 V_t$ for all $s,t\in\Sigma_n'(A)$.*

Proof. By the above remarks, it suffices to prove the second half of the lemma. The proof splits into a number of cases, one of which is performed here. We take $s=1s_1\dots s_{i-2}\gamma_1 s_{i+2}\dots s_{n-1}$ and $t=1s_1\dots s_{i-2}\gamma_2 t_{i+2}\dots t_{n-1}$, where $\gamma_1=171$ and $\gamma_2=1102$. It suffices to show that $\pi_1 V_{1s_1\dots s_{i-2}\gamma_1}>\pi_1 V_{1s_1\dots s_{i-2}\gamma_2}$. Observe that $V_{1s_1\dots s_{i-2}\gamma}=T^{-(i-2)}(H\cap V_\gamma)$, where $H=H_{1s_{i-2}\dots s_1 1}$ and $\gamma\in\{\gamma_1,\gamma_2\}$. The successive pre-images of H under either T^{-2} or T^{-3} will fall in R_1 or R_2. A straightforward calculation reveals that for all $j\in\{2,\dots,i\}$ the order of $\pi_1 T^{-(j-2)}(H\cap V_{\gamma_2})$ and $\pi_1 T^{-(j-2)}(H\cap V_{\gamma_1})$ is preserved whenever $T^{-(j-2)}(H)\subset V_{171}\cup V_{292}$ and reversed otherwise (see Fig. 6). An inspection of the graph G then reveals that for any path joining 1 to 1, there are an even number of occurrences of the symbols 3, 5, 8, and 10. Hence the V_γ will undergo an even number of order reversals, and thus $\pi_1 V_s>\pi_1 V_t$ follows from $\pi_1 V_{\gamma_1}>\pi_1 V_{\gamma_2}$, independently of $s_1\dots s_{i-2}$. □

3.2. The Quasiperiodic Operator

The labeling of the band structure $\{B_n|n\in\mathbb{Z}^+\}$ extends to the pseudospectrum, B_∞, of the quasiperiodic operator Q as follows. Let $\Sigma'(A)$ be the space of one-sided symbol sequences defined by (3.3),

$$\Sigma'(A)=\left\{s_0\dots s_n\dots\in\prod_{n=0}^{\infty}\{1,\dots,10\}|s_0\in\{1,6\},\ A_{s_is_{i+1}}=1 \text{ for all } i\geqq 0\right\} \tag{3.3}$$

endowed with the product topology, and define a map $q:\Sigma'(A)\to B_\infty$ by (3.4),

$$q(s_0\dots s_n\dots)=\{E\in\mathbb{R}|\{L_V(E)\}\cap V_{s_0\dots s_n\dots}\neq\emptyset\}. \tag{3.4}$$

The next theorem uses properties of the map q to obtain information on the pseudospectrum of the operator Q.

Theorem 3.3. *The map q is a homeomorphism, and B_∞ is a Cantor set of measure zero.*

Proof. By Theorem 2.1 the sets $V_{s_0 \ldots s_n \ldots}$, such that $s_0 \ldots s_n \ldots \in \Sigma'(A)$, are distinct non-empty vertical curves intersecting the horizontal line $L_V(\mathbb{R})$ in distinct points. Let $L_V(q(s_0 \ldots s_n \ldots)) = V_{s_0 \ldots s_n \ldots} \cap L_V(\mathbb{R})$ be such an intersection point. It follows from (1.8) that $q(s_0 \ldots s_n \ldots)$ consists of a unique point. By definition of $V_{s_0 \ldots s_n \ldots}$ we have that $\pi_1 T^n L_V(q(s_0 \ldots s_n \ldots))$ is bounded as $n \to \infty$. Then from the dynamical equation (1.9) for B_∞, we have $q(s_0 \ldots s_n \ldots) \in B_\infty$. Thus the map q is an injection into B_∞. Let $\mathbb{V}$ be the union of the sets $V_{s_0 \ldots s_n \ldots}$, such that $s_0 \ldots s_n \ldots \in \Sigma'(A)$. To show that the map q is a surjection, we observe that all points of $L_V(\mathbb{R})$ not intersecting $\mathbb{V}$ are eventually mapped by T outside of the region R_V. Thus their positive semi-orbits become unbounded, and they do not correspond to points in the pseudospectrum B_∞. Hence the map q is a bijection, and since $\Sigma'(A)$ is compact, to show that q is a homeomorphism it suffices to show that it is continuous. Let $\mathbf{s}, \mathbf{t} \in \Sigma'(A)$, where $\mathbf{s} = s_0 s_1 \ldots$ and $\mathbf{t} = t_0 t_1 \ldots$. The topology on $\Sigma'(A)$ is metrizable, and we choose a metric d defined by $d(\mathbf{s}, \mathbf{t}) = 2^{-i}$, where i is the smallest integer such that $s_i \neq t_i$. Take $\varepsilon > 0$, and let ν be the nesting constant of Theorem 2.1. We choose $\delta(\varepsilon)$ so that $d(\mathbf{s}, \mathbf{t}) < \delta$ implies $\mathbf{s}$ and $\mathbf{t}$ agree in their first $n+1$ places, where $n > \dfrac{\log \varepsilon}{\log \nu}$. Then $d(\mathbf{s}, \mathbf{t}) < \delta$ implies $|\pi_1 L_V(q(\mathbf{s})) - \pi_1 L_V(q(\mathbf{t}))| < d(V_{s_0 \ldots s_n}) < \nu^n < \varepsilon$. It follows from (1.8) that $|q(\mathbf{s}) - q(\mathbf{t})| < 2\varepsilon$, so that the map q is continuous. Thus B_∞ is homeomorphic to $\Sigma'(A)$, which is itself homeomorphic to a Cantor set.

To show that B_∞ is of Lebesgue measure zero, we use the following result [5, 7]: For a C^2 axiom A diffeomorphism of a surface, the set $W^s(\Omega)$ of points that approach the non-wandering set Ω has Lebesgue measure zero. Since in our case $\mathbb{V} \subset W^s(\Omega_V)$, it follows that $\mathbb{V}$ has Lebesgue measure zero. The vertical curves of $\mathbb{V}$ intersect $L_V(\mathbb{R})$ transversally at $L_V(B_\infty)$, and it follows that $L_V(B_\infty)$ has Lebesgue measure zero in $L_V(\mathbb{R})$, and hence that the Lebesgue measure of B_∞ is zero, as required. Note that if V is sufficiently large, the nesting constant ν of Theorem 2.1 can be shown to satisfy $\nu \in (0, \sigma)$, and the result can be proved from first principles, using a nesting argument. □

4. Quantitative Properties of the Spectrum

The map T was originally introduced as a renormalization map, and we now pursue the analysis from this standpoint. This will lead to quantitative results of a global nature for the spectra of the periodic operators P_n. First we summarize the simpler deductions that can be made from a local analysis of the map T.

4.1. Local Scaling Laws

We have deduced in Sect. 2 that for $V \geqq 8$, T_V has a period-p point $x(\mathbf{s}) \in R_V$, corresponding to each period-p symbol sequence $\mathbf{s}$ in $\Sigma(A)$, and that Ω_V is a hyperbolic set. Define the lines $\Sigma_n^{\pm}$ for $n \in \mathbb{N}$ by (4.1),

$$\Sigma_n^{\pm} = \{(x, y, z) \in S_V | \pi_1 T^{n-1}(x, y, z) = \pm 1\}. \tag{4.1}$$

The results of Sect. 2 also allow us to conclude the following. The line $L_V(\mathbb{R})$, parametrized by E, cuts the stable manifold, $W^s(x(\mathbf{s}))$, of $x(\mathbf{s})$ transversally with non-zero velocity, at all the points $L_V(\mathbf{t})$), such that $\mathbf{t} \in \Sigma'(A)$ has the same period-p tail as $\mathbf{s}$. Also T acts on the lines $\Sigma_n^{\pm}$ by $T(\Sigma_n^{\pm}) = \Sigma_{n-1}^{\pm}$, and $\Sigma_1^{\mp}$ cuts the unstable manifold, $W^u(x(\mathbf{s}))$, of $x(\mathbf{s})$ transversally. Let $|b(\mathbf{t}_n)|$ denote the length of the nearest interval in B_n to the point $q(\mathbf{t})$ of B_∞. It is determined by an intersection of $\Sigma_n^{\pm}$ with $L_V(E)$. The above geometry allows us to conclude, as in [8], that $|b(\mathbf{t}_n)|$ obeys the scaling relation (4.2),

$$\lim_{n \to \infty} \frac{|b(\mathbf{t}_n)|}{|b(\mathbf{t}_{n+p})|} = |dT_x^p|_e, \tag{4.2}$$

where $|dT_x^p|_e$ is the expanding eigenvalue of dT^p at $x = x(\mathbf{s})$. We say that there is a *local scaling law* at points in B_∞ with period-p tail, governed by a period-p point of T (compare [18]).

4.2. A Global Scaling Law

There is an obvious exponent, λ_g, that can be thought of as measuring the scaling of the entire band structure, defined by (4.3),

$$\lambda_g = \lim_{n \to \infty} \frac{1}{n} \log m(B_n) \tag{4.3}$$

where $m(B_n)$ denotes the (Lebesgue) measure of B_n. The next theorem shows how to obtain the exponent λ_g from a knowledge of the quantities $|dT_x^n|_e$ at the periodic points of T.

Theorem 4.1. *The exponent* λ_g *is given by* (4.4),

$$\lambda_g = \lim_{n \to \infty} \frac{1}{n} \log \sum_{x \in \operatorname{Fix} T^n} (|dT_x^n|_e)^{-1}, \tag{4.4}$$

where $\operatorname{Fix} T^n$ *denotes the set of fixed points of* T^n.

Remark. The limit in Theorem 4.1 exists, and is equal to the topological pressure, $P(\phi^u)$, of T with respect to the function $\phi^u(x) = -\log|dT_x|_e$ [32]. In fact there is numerical evidence that the scaling of $m(B_n)$ is geometrical [19], which is a stronger property. Theorem 4.1 should be compared to the following "escape rate" result of Bowen and Ruelle. For a C^2-diffeomorphism, the Lebesgue measure of those points whose orbits remain within ε of Ω from time 0 to n decays like $\exp nP(\phi^u)$ [5].

In order to prove Theorem 4.1, we will need the following two lemmas.

Lemma 4.2. *Let T be a C^2 diffeomorphism of a surface with compact hyperbolic non-wandering set Ω. Let $y \in W^s(x)$, where $x \in \Omega$. Then for all $\varepsilon > 0$, there exists N_ε such that $n > N_\varepsilon$ implies $|T^n x - T^n y| = e^{n\xi}|dT_x^n|_c|x - y|$ for some $\xi \in (-\varepsilon, \varepsilon)$, where $|dT_x^n|_c$ denotes the eigenvalue of dT_x^n in the contracting direction.*

Proof. Define $x_n = T^n x$, $y_n = T^n y$, and $v_n = y_n - x_n$. The strategy of the proof is to first map y into the "linear region," and then dominate subsequent contractions by the linear part of T. By Taylor's theorem, $|v_{m+1}-dT_{x_m}v_m|\leqq K|v_m|^2$, where $K=\sup_{z\in\Omega}|d^2T_z|$. Therefore

$$v_{m+1}=dT_{x_m}t_m+dT_{x_m}(v_m-t_m)+e_m\,, \tag{4.5}$$

where t_m is a tangent vector to $W^s(x_m)$ at x_m of length $|v_m|$, and $|e_m|\leqq K|v_m|^2$. By hypothesis $y\in W^s(x)$, hence $|v_m|\to 0$ and $|v_m-t_m|/|v_m|\to 0$ as $m\to\infty$. Thus there exists an M_ε so that $|v_m|<C\varepsilon/4K$ and $|v_m-t_m|/|v_m|<C\varepsilon/4E$ whenever $m>M_\varepsilon$, where $C=\inf_{z\in\Omega}|dT_z|_c$ and $E=\sup_{z\in\Omega}|dT_z|_e$. Then from (4.5) it follows that for $m>M_\varepsilon$,

$$|v_{m+1}|\leqq|dT_{x_m}|_c|v_m|+|dT_{x_m}|_e|v_m-t_m|+K|v_m|^2\leqq(1+\varepsilon/2)|dT_{x_m}|_c|v_m|\,. \tag{4.6}$$

There is a similar inequality in the other direction, thus $m>M_\varepsilon$ implies $|v_{m+1}|/|v_m| =e^{\xi}|dT_{x_m}|_c$ for some $\xi\in(-\varepsilon/2,\varepsilon/2)$. Multiplying such equations together for the values $m=m, m+1, \ldots, m+n$, and using the chain rule, it follows that for $m>M_\varepsilon$ and any $n\in\mathbb{N}$, $|v_{m+n}|/|v_m|=e^{n\xi}|dT^n_{x_m}|_c$ for some $\xi\in(-\varepsilon/2,\varepsilon/2)$. Thus $|v_{m+n}| =K_me^{n\xi}|dT^{n+m}_x|_c|v_0|$, where $K_m=|v_m|/|dT^m_x|_c|v_0|$. We now choose n so large that $K_m=e^{n\xi'}$ for some $\xi'\in(-\varepsilon/2,\varepsilon/2)$, and the result follows. □

Lemma 4.3. *Let T satisfy the hypotheses of Lemma 4.2, and in addition suppose there is a point $x\in\Omega$ with one-dimensional stable and unstable manifolds $W^s(x)$ and $W^u(x)$. Let Σ be a curve that intersects $W^u(x)$ transversally at a point y, and let L be a curve having non-empty transverse intersection with $W^s(T^{-n}(x))$ for all $n\in\mathbb{N}$. Then for all $\varepsilon>0$, there exists N_ε, such that $n>N_\varepsilon$ implies $|a_n-b_n|=e^{n\xi}|dT^{-n}_x|_c|x-y|$ for some $\xi\in(\varepsilon,-\varepsilon)$, where $a_n\in L\cap W^s(T^{-n}(x))$, and b_n is the closest point in $L\cap T^{-n}\Sigma$ to a_n.*

Proof. By Lemma 4.2, it is sufficient to show that $|a_n-b_n|=e^{n\xi}|x_n-y_n|$ for sufficiently large n, where $x_n=T^{-n}x$ and $y_n=T^{-n}y$. Let $r=n-m$. By the λ-lemma [30], there exists M_ε so that $T^{-r}\Sigma$ is $\varepsilon/2-C^1$ close to $W^s(x_r)$ and T^mL is $\varepsilon/2-C^1$ close to $W^u(x_r)$ whenever $r, m\geqq M_\varepsilon$. Fix $m=M_\varepsilon$, and take $n\geqq 2M_\varepsilon$. Then the points x_r, y_r, a_r, b_r lie at the corners of a curvilinear region which is $\varepsilon/2-C^1$ close to a small parallelogram touching $W^u(x_r)$ and $W^s(x_r)$ at the point x_r. Thus $|a_r-b_r|/|x_r-y_r|$ is $\varepsilon/2$-close to 1. Also $|a_n-b_n|/|x_n-y_n|=K_m|a_r-b_r|/|x_r-y_r|$ for some constant K_m independent of n. By choosing n sufficiently large, we can ensure that $K_m=e^{n\xi}$ for some $\xi\in(-\varepsilon/2,\varepsilon/2)$, and the result follows. □

Proof of Theorem 4.1. For ease of notation, we identify $b(s_0\ldots s_n)$ and $q(\mathbf{s})$ with their images under the parametrization map $E\to L_V(E)$. We apply Lemma 4.3 to the renormalization map T, taking $x=x(\sigma^n\mathbf{s})$, $L=L_V(E)$, $a_n=q(s_0\ldots s_n\ldots)$, and $\Sigma=\Sigma_1^\pm$. Let $s_0\ldots s_{n-1}\in\Sigma_n'(A)$. Then $[\pi_1b_n^-,\pi_1b_n^+]=b(s_0\ldots s_{n-1})$, where $b_n^\pm$ are the points in $T^{-n}\Sigma_1^\pm\cap L_V(E)$ nearest to $q(s_0\ldots s_{n-1}\ldots)$. We conclude that for all $\varepsilon>0$, there exists N_ε such that $n>N_\varepsilon$ implies (4.7),

$$|b(s_0\ldots s_{n-1})|=e^{n\xi}(|dT^n_{x_n}|_e)^{-1} \tag{4.7}$$

for some $\xi\in(-\varepsilon/2,\varepsilon/2)$, where $x_n=T^{-n}x=x(\mathbf{s})$. Since Ω_V is compact, N_ε may be chosen independently of $s_0\ldots s_{n-1}$, and we can sum (4.7) over all $s_0\ldots s_{n-1}\in\Sigma_n'(A)$, to deduce (4.8),

$$m(B_n)=\sum_{x\in S_n}e^{n\xi(x)}(|dT^n_x|_e)^{-1}\,, \tag{4.8}$$

where $\xi(x) \in (-\varepsilon/2, \varepsilon/2)$ for all $x \in S_n$, and S_n is any subset of Ω_V containing precisely one point in each vertical strip $V_{s_0 \dots s_{n-1}}$ such that $s_0 \dots s_{n-1} \in \Sigma_n'(A)$. It remains to show that for sufficiently large n, $m(B_n) = e^{n\xi}\beta_n$ for some $\xi \in (-\varepsilon, \varepsilon)$, where $\beta_n = \sum_{x \in \operatorname{Fix} T^n} (|dT_x^n|_e)^{-1}$.

Our strategy is now to compare $m(B_n)$ to β_{n+m} for fixed m, and use the mixing property of the matrix A. Since $\{|dT_x^m|_e | x \in \Omega\}$ is bounded, there exists $N_\varepsilon(m) \in \mathbb{N}$ such that $n > N_\varepsilon(m)$ implies $|dT_x^n|_e = e^{n\xi'(x)}$, where $\xi'(x) \in (-\varepsilon/2, \varepsilon/2)$ for all $x \in \Omega$. Thus from (4.8) we deduce (4.9),

$$m(B_n) = e^{n\xi} \sum_{x \in S_n} (|dT_x^{n+m}|_e)^{-1}\,. \tag{4.9}$$

Since A is mixing, there exists $m \in \mathbb{N}$ such that it is possible to take $S_n \subset \operatorname{Fix} T^{n+m}$, and therefore $m(B_n) \leqq e^{n\xi}\beta_{n+m}$. If we can show that $\beta_{n+m} \leqq e^{3n\varepsilon} m(B_n)$, the proof will be complete.

From the mixing property of A, there exists $M_\varepsilon \in \mathbb{N}$ such that if $m = M_\varepsilon$, and n is sufficiently large, then $\operatorname{Card}(\operatorname{Fix} T^{n+m} \cap V_s) = e^{n\xi(s)}$, where $\xi(s) \in (-\varepsilon, \varepsilon)$ for all $s = s_0 \dots s_{n-1} \in \Sigma_n'(A)$. It follows that $e^{n\xi(s)}$ terms in the sum defining β_{n+m} can be grouped together, and bounded by the product of $e^{n\xi}$ with a term in the sum (4.9) for $m(B_n)$. All the $x(\mathbf{s}) \in \operatorname{Fix} T^{n+m}$ with $s_0 \in \{1, 6\}$ can be dealt with in this fashion. Thus $\beta_{n+m} \leqq K e^{n\varepsilon} e^{-n\xi} m(B_n)$, where K is a constant factor for bounding the contributions from the other terms of β_{n+m}. This inequality is evidently of the required form if n is sufficiently large. $\square$

4.3. Ergodic Scaling Laws

The concept of ergodic scaling was introduced in [29]. The idea is that if μ is an ergodic measure for T, then the Liapunov exponent of T with respect to μ will determine an "ergodic" scaling law at the points of intersection of $W^u(X)$ with the family of interest, for a set X of full μ-measure. In our situation there are many ergodic measures for T on Ω_V, for example all Markov measures on $\Sigma(A)$, characterized by a pair $\mu = (\mathbf{p}, P)$, where $\mathbf{p} \in \mathbb{R}^{10}$ is a probability vector, and P is a 10×10 irreducible stochastic matrix with $P_{ij} = 0$ whenever $A_{ij} = 0$ (see [35], note that a measure on $\Sigma(A)$ automatically defines a measure on Ω_V).

We say that there is an *ergodic scaling law* at $q(s_0 \dots s_{n-1} \dots)$, with exponent λ_e, if the following limit exists

$$\lambda_e = \lim_{n \to \infty} \frac{1}{n} \log |b(s_0 \dots s_{n-1})|\,, \tag{4.10}$$

where $b(s_0 \dots s_{n-1})$ is the band in B_n closest to $q(s_0 \dots s_{n-1} \dots)$. Given a Markov measure μ on $\Sigma(A)$, the next theorem states that there is an ergodic scaling law at $q(s_0 \dots s_n \dots)$ for μ'-almost all $s_0 \dots s_n \dots \in \Sigma'(A)$, where μ' is the measure induced by μ on $\Sigma'(A)$.

Theorem 4.4. *The limit* (4.10) *exists for μ'-almost all $s_0 \dots s_{n-1} \dots \in \Sigma'(A)$, and is given by* (4.11)

$$\lambda_e = \lim_{n \to \infty} \frac{1}{n} \int \log (|dT_x^n|_e)^{-1} d\mu(x)\,. \tag{4.11}$$

Proof. It follows from Eq. (4.7) that

$$\lim_{n\to\infty}\frac{1}{n}\log|b(s_0 \ldots s_{n-1})| = \lim_{n\to\infty}\frac{1}{n}\log(|dT^n_{x(\mathbf{t})}|_e)^{-1}, \tag{4.12}$$

where $\mathbf{t}\in\Sigma(A)$ has the same tail as $s_0 \ldots s_{n-1} \ldots$. By the multiplicative ergodic theorem [35], the limit (4.12) exists for μ-almost all $\mathbf{t}\in\Sigma(A)$, and is given by (4.11). The result follows from the definition of the induced measure μ' on $\Sigma'(A)$. □

Example. A natural example to take for the ergodic measure μ is the distribution on periodic points defined by $\mu = \lim\limits_{n\to\infty}(N_n)^{-1}\sum\limits_{x\in\mathrm{Fix}\,T^n}\delta_x$, where $N_n = \mathrm{Card}(\mathrm{Fix}\,T^n)$. It can be shown that μ is a Markov measure with certain transition probabilities, obtainable from the left and right eigenvectors of A [35]. The exponent λ_e is given by (4.13),

$$\lambda_e = \lim_{n\to\infty}(N_n)^{-1}\sum_{x\in\mathrm{Fix}\,T^n}\log(|dT^n_x|_e)^{-1}. \tag{4.13}$$

The probabilistic interpretation of λ_e is that if we step from a band in B_i to a band in B_{i+1} using these transition probabilities, we would expect to obtain a band of length of order $\exp n\lambda_e$ at stage n for sufficiently large n.

4.4. *Hausdorff Dimension of B_∞*

We know from Sect. 3.2 that B_∞ is a Cantor set of measure zero. The following theorem further characterizes B_∞.

Theorem 4.5. *The Hausdorff dimension* (HD) *of B_∞ satisfies* (4.14),

$$-\lambda_A/\lambda_e \leqq \mathrm{HD}(B_\infty) \leqq \lambda_A/(\lambda_A-\lambda_g), \tag{4.14}$$

where λ_g and λ_e are given by (4.4) *and* (4.13) *respectively, and $\lambda_A = \log(1+\sigma)$ is the logarithm of the largest eigenvalue of the matrix A.*

Proof. We use Corollary 3 of [22]: Let Λ be a basic set for a C^1 axiom-A diffeomorphism $f: M^2\to M^2$ with (1,1) splitting $T_\Lambda M = E^s \oplus E^u$. Then $\delta = \mathrm{HD}(W^u(x)\cap\Lambda)$ is independent of $x\in\Lambda$, and satisfies (4.15),

$$-h_{\mathrm{top}}/m(\phi^u) \leqq \delta \leqq h_{\mathrm{top}}/(h_{\mathrm{top}} - P(\phi^u)), \tag{4.15}$$

where h_{top} is the topological entropy of f, $\phi^u: W^u(\Lambda)\to\mathbb{R}$ is the function defined by $\phi^u(x) = -\log|dT_x|_e$, $m(\phi^u)$ is the integral of ϕ^u with respect to the measure of maximal entropy, and $P(\phi^u)$ is the topological pressure of T with respect to the function ϕ^u.

The renormalization map T_V satisfies the above hypotheses, with $\Lambda = \Omega_V$. Thus we can substitute the following into Eq. (4.14). h_{top} is given by λ_A the logarithm of the largest eigenvalue of the matrix A [35]. $m(\phi^u)$ is given by λ_e, since the measure of maximal entropy of a subshift is the measure on periodic points [35]. $P(\phi^u)$ is given by λ_g, as remarked in Sect. 4.2. Finally, it can be shown that there is a Lipshitz map between $\Lambda\cap L_V(E)$ and $\Lambda\cap W^u_{\mathrm{loc}}(x)$, and since HD is preserved under Lipshitz maps, we have $\delta = \mathrm{HD}(B_\infty)$. □

5. Rotation Numbers

In theoretical studies of Schrödinger operators with quasiperiodic potentials, one of the basic tools is the rotation number $\varrho(E)$ [9, 14, 16, 33]. Roughly speaking, it measures the average rate of rotation of the phase of an eigenstate over the lattice. The rotation number yields a labelling of the gaps of the spectrum, due to the following result [9]: $2\varrho(E)$ lies in the frequency module of the quasiperiodic potential whenever E lies outside the spectrum. The numerical scaling results of [27, 28] are also stated in the language of rotation number. In this section, we attempt to translate our labeling of the pseudospectrum B_∞ of the operator Q by symbol sequences, to a labeling by rotation number. In order to calculate the rotation number, we use the well known relationship between the integrated density of states, $k(E)$, and the rotation number, namely $k(E)=2\varrho(E)$. We calculate $k(E)$ by using the periodic operators P_n to approximate the operator Q. This procedure is convergent, and we believe that the resulting expression for $k(E)$ is correct, though this remains to be proven. Finally, restating the scaling results of Sect. 4 in terms of rotation numbers, we recover the numerical results of [27], together with some extensions of their results.

We now recall the definitions of the rotation number and integrated density of states for discrete Schrödinger operators [9]. Let H be the operator given by (1.1), and let $\psi(0), \psi(1), \ldots, \psi(n), \ldots$ be a solution of the equation $H\psi=E\psi$ for $n \geqq 0$ with initial condition $\psi(0)=\cos\theta$, $\psi(1)=\sin\theta$.

Definition. The *rotation number* of H is the map $\varrho: \mathbb{R} \to \mathbb{R}$ defined by (5.1),

$$\varrho(E)=\lim_{L\to\infty} \frac{1}{2L} N_L(E,\theta), \tag{5.1}$$

where $N_L(E,\theta)$ is the number of changes of sign in $\psi(n)$ for $1 \leqq n \leqq L$.

It is shown in [9] that the limit exists and is independent of θ. Now consider the restriction H_L of the operator H to the set $\{1, \ldots, L\}$ with boundary condition $\psi(0)/\psi(1)=\operatorname{cotan}(\theta)$.

Definition. The *integrated density of states* of H is the map $k: \mathbb{R} \to \mathbb{R}$ defined by (5.2),

$$k(E)=\lim_{L\to\infty} \frac{1}{L} M_L(E,\theta), \tag{5.2}$$

where $M_L(E,\theta)$ is the number of eigenvalues of the operator H_L less than or equal to E.

It is shown in [9] that $k(E)=2\varrho(E)$. Taking H to be the operator Q, it may be verified that E is an eigenstate of the operator H_{F_n} for $\theta=0$ whenever E satisfies $(M(n))_{11}=0$, where $M(n)$ is the product of E-dependent transfer matrices described in Sect. 1. Thus we expect a close relationship between the spectrum of H_{F_n} and the spectrum of the periodic operators P_n. This leads us to the following conjecture.

Conjecture. Let $P_n(E)$ be the number of bands in the spectrum of the operator P_n bounded above by E. Then the integrated density of states is given by (5.3).

$$k(E)=\lim_{n\to\infty} \frac{1}{F_n} P_n(E). \tag{5.3}$$

Assuming the truth of the above conjecture, we have the following corollaries.

Corollary 5.1. *The rotation number $\varrho(q(s_0 \ldots s_n \ldots))$ of a point $q(s_0 \ldots s_n \ldots) \in B_\infty$, with $s_0 = 1$ satisfies* (5.4),

$$2\varrho(q(s_0 \ldots s_n \ldots)) = \sigma^2 + \sum_{i=1}^{\infty} d_i \sigma^i, \tag{5.4}$$

where the d_i are obtained by decomposing $s_0 \ldots s_n \ldots$ into blocks of length 2 *or* 3 *and according to Table 5.1.*

Table 5.1

s_i	17	136	110	28	245	29
d_{i+1}	01	001	00	01	001	00

Remark. A similar result holds for $s_0 = 6$.

Proof. We will use the ordering property of the map b stated in Lemma 3.1. Let $E = q(s_0 \ldots s_n \ldots) \in B_\infty$. Given any $N \in \mathbb{N}$, it is possible to find an $n > N$ such that $s_{n-1} \in \{1, 2, 3, 4\}$, so that $q(s_0 \ldots s_n \ldots) \in b(s_0 \ldots s_{n-1}) \in B_n$. We use Lemma 3.2 to calculate the quantity $P_n(E)$, as follows. It can be shown by induction that there are F_{n-2} bands $b(t_0 \ldots t_{n-1})$, with $t_0 = 6$, lying below $b(s_0 \ldots s_{n-1})$. Similarly, it can be shown that there are F_{n-i-3} bands in B_n having labeling starting with $s_0 \ldots s_{i-1}110$, and F_{n-i-4} bands in B_n having labeling starting with $s_0 \ldots s_{i-1}136$. Hence if $s_i = 1$ and the block 17 occurs, $F_{n-i-4} + F_{n-i-3} = F_{n-i-2}$ symbols will necessarily have been "climbed over" at stage i. Using similar calculations for the other possibilities, by the definition of the d_i, we have deduced that E lies in the m^{th} highest band of B_n, where $m = F_{n-2} + \sum_{i=1}^{n} d_i F_{n-i}$. Thus by definition, $P_n(E) = m - 1$, and using the above conjecture we have $2\varrho(E) = \lim_{n \to \infty} (m-1)/F_n$, and the result follows. □

Corollary 5.2. *Let $E = q(s_0 \ldots s_n \ldots)$ be a point in B_∞. Then*

(1) *If E is a gap edge of B_∞ there is a local scaling law at E, governed by a period-2 point of T.*

(2) *If $\varrho(E)$ has an "irrational expansion"* (5.4) *with a periodic tail, then there is a local scaling law at E governed by a periodic point of T.*

(3) *There is a set $X \subset [0, 1/2]$, of full Lebesgue measure, such that if $E \in \varrho^{-1}(X)$, there is an ergodic scaling law at E with exponent λ_e given by* (4.13).

Proof. (1) By definition, a gap edge of B_∞ is a point $E \in B_\infty$ for which there exists $\delta > 0$ such that either $B_\infty \cap (E, E + \delta) = \emptyset$ or $B_\infty \cap (E - \delta, E) = \emptyset$. Consider the former case. It follows that if $E = q(s_0 \ldots s_n \ldots)$ is at a gap edge, then there exists N such that $q(s_0 \ldots s_N s_{N+1} \ldots) > q(s_0 \ldots s_N t_{N+1} \ldots)$ whenever $t_{N+1} \ldots \neq s_{N+1} \ldots$. Using the ordering property of Lemma 3.2, it follows that $s_{N+1} \ldots$ has tail 1717 .... Thus by the remarks of Sect. 4.1, there is a local scaling law at E governed by a period-2 point of T. Similarly, in the other case, the gap edge is represented by $q(s_0 \ldots s_n \ldots)$ where $s_0 \ldots s_n \ldots$ has a period-2 tail 2929 ....

(2) If $\varrho(E)$ has an irrational expansion (5.4) with a periodic tail, then $E=q(s_0 \dots s_n \dots)$ where $s_0 \dots s_n \dots$ has a periodic tail. Thus, by the remarks of Sect. 4.1, there is a local scaling law at E governed by a periodic point of T.

(3) Let μ' be the measure on B_∞ induced by the measure on periodic points, as defined in Sect. 4.3. Then it can be shown, using the relationship $k=2\varrho$, and (5.3), that for all $\lambda \in [0, 1/2]$, $\mu'(\varrho^{-1}[0, \lambda])=\lambda$, so that Lebesgue measure is induced on $\varrho(B_\infty)$ by μ'. $\square$

Acknowledgements. I would like to thank my supervisor, David Rand, for suggesting this problem. I have also had useful discussions with A. Manning, B. Mestel, and P. Walters. This work was supported by a U.K. Science and Engineering Research Council award, and by the La Jolla Institute Grant, U.S. Department of Energy, DE-ACO3-84ER40182.

References

1. Avron, J., Simon, B.: Almost periodic Hill's equation and the rings of Saturn. Phys. Rev. Lett. **46**, 1166–1168 (1981)
2. Avron, J., Simon, B.: Almost periodic Schrödinger operators. Commun. Math. Phys. **82**, 101–120 (1982)
3. Bak, P.: Phenomenological theory of icosahedral incommensurate order in Mn-Al alloys. Phys. Rev. Lett. **54**, 1517–1519 (1985). See also p. 1520–1526
4. Bellissard, J., Lima, R., Testard, T.: A metal-insulator transition for the almost Mathieu model. Commun. Math. Phys. **88**, 207–234 (1983)
5. Bowen, B.: Equilibrium states and the Ergodic theory of Anosov diffeomorphisms. In: Lecture Notes in Mathematics, Vol. 470. Berlin, Heidelberg, New York: Springer 1975
6. Bowen, R., Lanford, O.E.: Zeta functions of restrictions of the shift transformation. In: Proceedings of a conference on global analysis, pp. 43–49. Vol. 14. Providence, RI: Am. Math. Soc. 1975
7. Bowen, R., Ruelle, D.: The ergodic theory of axiom A flows. Invent. Math. **29**, 181–202 (1975)
8. Collet, P., Eckmann, J.-P., Lanford, O.E.: Universal properties of maps of an interval. Commun. Math. Phys. **76**, 211–254 (1980)
9. Delyon, F., Souillard, B.: The rotation number for finite difference operators and its properties. Preprint, l'Ecole Polytechnique 1982
10. Delyon, F., Petritis, D.: Absence of localisation in a class of Schrödinger operators with quasiperiodic potential. Commun. Math. Phys. **103**, 441–444 (1986)
11. Devaney, R., Nitecki, Z.: Shift automorphisms in the Hénon mapping. Commun. Math. Phys. **67**, 137–146 (1979)
12. Dinaburg, E.I., Sinai, Ya.G.: The one dimensional Schrödinger equation with a quasiperiodic potential. Funct. Anal. Appl. **9**, 8–21 (1975)
13. Farmer, J.D., Satija, I.: Renormalization of the quasiperiodic transition to chaos for arbitrary winding numbers. Preprint, Los Alamos 1984
14. Herman, M.R.: Une methode pour minorer les exposants de Lyapounov et quelques exemples montrant le charactere local d'un théorème d'Arnold et de Moser sur le tore de dimension 2. Preprint, l'Ecole polytechnique 1982
15. Hofstader, D.R.: Energy levels and wave functions of Bloch electrons in rational and irrational magnetic fields. Phys. Rev. B **14**, 2239–2249 (1976)
16. Johnson, R., Moser, J.: The rotation number for almost periodic potentials. Commun. Math. Phys. **84**, 403–438 (1982)
17. Kadanoff, L.: Analysis of cycles for a volume preserving map. Preprint, University of Chicago 1983
18. Kohmoto, M.: Cantor spectrum for an almost periodic Schrödinger equation and a dynamical map. Preprint, University of Illinois 1984

19. Kohmoto, M., Kadanoff, L., Tang, C.: Localisation problem in one dimension: Mapping and escape. Phys. Rev. Lett. **50**, 1870–1872 (1983)
20. Kunz, H., Souillard, B.: Sur le spectre des opérateurs aux différences finies aléatoires. Commun. Math. Phys. **78**, 201–246 (1980)
21. Luck, J.M., Petritis, D.: Phonon spectra in one dimensional quasicrystals, Preprint CEN (Saclay) 1985)
22. McCluskey, H., Mannings, A.: Hausdorff, A.: Hausdorff dimension for horseshoes. Ergodic Theory Dyn. Syst. **3**, 251–260 (1983)
23. MacKay, R.S., Percival, I.C.: Self-similarity of the boundary of Seigel domains for arbitrary rotation number. In preparation
24. Moore, R.: Interval Analysis. In: Series in automatic computation. New York: Prentice-Hall 1966
25. Moser, J.: Stable and random motions in dynamical systems. In: Annals of mathematical studies 77. Princeton: Princeton University Press 1973
26. Ostlund, S., Kim, S.: Renormalization of quasiperiodic mappings. Physica Scripta T**9**, 193–198 (1985)
27. Ostlund, S., Pandit, R., Rand, R., Shellnhuber, H., Siggia, E.: One-dimensional Schrödinger equation with an almost periodic potential. Phys. Rev. Lett. **50**, 1873–1876 (1983)
28. Ostlund, S., Pandit, R.: Renormalization group analysis of the discrete quasiperiodic Schrödinger equation. Preprint, Cornell University 1984
29. Ostlund, S., Rand, D., Sethna, J., Siggia, E.: Universal properties of the transition from quasi-periodicity to chaos in dissipative systems. Physica **8**D, 303–342 (1983)
30. Palis, J., de Melo, W.: Geometric theory of dynamical systems. Berlin, Heidelberg, New York: Springer 1982
31. Reed, M., Simon, B.: Methods of modern mathematical physics. IV. Analysis of operators. London, New York: Academic Press 1978
32. Ruelle, D.: Thermodynamic formalism. p. 139. Reading, MA: Addison-Wesley 1978
33. Simon, B.: Almost periodic Schrödinger operators; A review. Adv. Appl. Math. **3**, 463–490 (1982)
34. Smale, S.: Diffeomorphisms with many periodic points. In: Differential and combinatorial topology. Princeton: Princeton University Press 1965
35. Walters, P.: An introduction to Ergodic theory. In: Graduate texts in mathematics, Vol. 79. Berlin, Heidelberg, New York: Springer 1982

Communicated by B. Simon

Received August 1, 1985; in revised form May 30, 1986