Symbolic Dynamics for the Renormalization Map of a Quasiperiodic Schrödinger Equation and Periodic Orbits for Dissipative Twist Maps.

Martin Casdagli

Mathematics Institute, University of Warwick,
Coventry, United Kingdom.

Submitted for the degree of Ph.D.

CONTENTS

Acknowledgements.

First and foremost, I would like to extend my gratitude to my supervisor, David Rand. He was always an inexhaustible source of ideas and enthusiasm. He encouraged me to read papers written both by physicists and by mathematicians, and to perform computer studies of dynamical systems. He suggested both of the problems which form this thesis, and gave me good ideas on how to start to solve them. I elaborate on this in the summary overleaf.

I also consider myself fortunate to have shared an office with Ben Mestel. As well as having many mathematical conversations with him, he helped me on the way to becoming a proficient computer programmer. During the period Sept '83 to May '84, I was lucky enough to be invited by David go to with him to Cornell University. Amongst many others there, I benefited from talks with P.Holmes, J.Hubbard, E.Siggia and S.Shenker.

Of those at Warwick University, I am especially grateful to R.MacKay, A.Manning, J.Stark, P.Walters, and E.C.Zeeman for answering various questions of mine. I would like to thank P.le Calvez for pointing out some errors in the original draft of my paper "Periodic Orbits For Dissipative Twist Maps", and for communicating his results on this topic to me prior to publication.

Finally, I would like to thank my parents for their support and encouragement over the past three years.

This work was supported by a U.K. Science and Engineering Research Council award.

Summary.

I study two quite different problems in this thesis. Both have been written up as papers, with their own detailed introductions. The thesis essentially consists of these two papers, although the first paper has been extended for this purpose to include extra explanatory material. In this summary I give a less formal introduction to the two papers. The aim is to give the reader some idea as to how these papers came into being.

1. Symbolic Dynamics for the Renormalization of a Quasiperiodic Schrödinger Equation.

The subject of dynamical systems first captured my imagination on a reading of Feigenbaum's renormalization theory of universality in period doubling. David suggested that I work on a renormalization theory for a one-dimensional Schrödinger equation with quasiperiodic potential, there being very little known about the problem. These equations are very interesting for applications, e.g. stability analysis of Saturn's rings, theory of an electron in a two-dimensional crystal with superimposed magnetic field, and electronic structure of quasicrystals, to mention but a few.

David pointed out a paper by Kadanoff concerning a particularly simple example of a quasi-periodic Schrödinger equation, for which the renormalization map is two-dimensional. He suggested that if the map could be shown to possess a Smale horseshoe, then the spectrum of the Schrödinger equation must contain a Cantor set. It was relatively easy to deduce that this picture was correct by performing some computer experiments on the renormalization map. In fact I found more to be true: the dynamics of the map could be completely described using six symbols. Strings of these six symbols can then be used to label the spectrum. The proof of these facts is geometrical in spirit. It is a lengthy exercise in applying standard techniques developed for proving the existence of invariant Cantor sets in non-linear maps. The fact that precisely six symbols are required for the description of the renormalization map seems to be the consequence of the occurrence of this symbolic dynamics in an "exactly solvable" map, related to the problem, which

I describe.

An interesting question then arises: what scaling properties of the spectra of the optimally approximating periodic equations can be deduced from the global dynamics of the renormalization map? It is, of course, well known that the existence of a fixed point in a renormalization map leads to a scaling law, with exponent governed by the expanding eigenvalue of the linearized map at the fixed point. A remark by Newhouse, that the theory topological pressure would be relevant to the problem, was very helpful at this point. I deduced the existence of a "global" scaling exponent, describing the total measure of the bands of the optimally approximating periodic systems. The exponent is obtained by taking a certain average of eigenvalues at all the periodic points of the renormalization map.

Using the multiplicative ergodic theorem, I deduced the existence of an "ergodic" scaling exponent, which measures how the length of a "typical" band in the spectrum of a periodic system decreases as the period increases. Finally, I applied a theorem of McCluskey and Manning to deduce bounds on the Hausdorff dimension of the spectrum of the quasiperiodic equation in terms of the two exponents mentioned above.

2. Periodic Orbits for Dissipative Twist Maps.

Periodic orbits can be proved to exist in area preserving maps of a cylinder by an elegant variational approach. The orbits are obtained as minima of a real valued function of many variables (the number of variables being equal to the period), subjected to periodic boundary conditions. Given that periodic orbits exist, from results of Hall and Katok one can deduce the existence of quasiperiodic orbits (with irrational rotation number), using only the twist hypothesis. David suggested that I look for periodic orbits in dissipative twist maps, via a variational approach devised by him. However the function that it was suggested I minimize was highly inhomogeneous in its variables, and it was inappropriate to apply periodic boundary conditions. Nevertheless, I explored the minimization problem on a computer, and found that results could be obtained using rigid boundary conditions, fixed by a parameter which is later varied. This lead naturally to a more powerful topological approach for deducing the existence of periodic orbits, which exploits the geometry of the twist and the topology of the cylinder in a particularly simple way. I thus obtained a theorem on the existence of periodic (and hence quasi-periodic) orbits in one parameter families of dissipative twist maps.

Using this topological approach, I also obtained results on the allowed periodic motions which can occur on an attractor of an individual dissipative twist map. The result relies heavily on a pioneering paper of Birkhoff's. A key step is to introduce the concept of an "attractor with the intersection property", which is a generalization of a strange attractor. After I had written up this result, I found that P.le Calvez had obtained a closely related result a few months previously, using a similar topological criterion for the existence of periodic orbits. The paper presented here was rewritten to take this into account.

Symbolic Dynamics for the Renormalization Map of a Quasiperiodic Schrödinger Equation

Martin Casdagli

Mathematics Institute
University of Warwick, Coventry CV4 7AL, U.K.

and

La Jolla Institute*
Center for Studies of Nonlinear Dynamics
3252 Holiday Court, Suite 208
La Jolla, California 92037

ABSTRACT

A rigorous analysis is given of the dynamics of the renormalization map associated to a discrete Schrödinger operator H on $l^2(\mathbf{Z})$, 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 $\sigma = (\sqrt{5}-1)/2$ is the golden mean. The renormalization map for H is a diffeomorphism, T, of $\mathbf{R}^3$, preserving a cubic surface S_V. For $V \geq 8$ 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.

* Current Address

Introduction.

There has been much interest in Schrödinger operators with a quasiperiodic potential (see [18,19,22,23,24,33] 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 quasiperiodic 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 from 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 [10,17,18,19,21,26,28], 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 [18,19,23], 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 [18,19,23]. 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 $V \geq 8$. 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 [9,14,16,23,33].

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 Section 1 we define the quasiperiodic operator to be studied, and review the renormalization technique used to analyze it. In Section 2 we collect our results on the symbolic dynamics of the renormalization map. In Section 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 Section 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 Section 5 we relate our rigorous results to numerical work of others [27], by obtaining a relationship between symbol sequences and rotation numbers.

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.

1. The Renormalization Technique.

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

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

where $v(n)\in \mathbf{R}$ denotes the potential at site $n\in \mathbf{Z}$, and $\psi(n)\in \mathbf{C}$ denotes the wave function at site $n\in \mathbf{Z}$. Let $\mathbf{S}^1$ 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 $\theta_0 \in \mathbf{S}^1$, g is a homeomorphism of $\mathbf{S}^1$ with irrational rotation number, and f is a real valued function on $\mathbf{S}^1$.

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)

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

where $\sigma = (\sqrt{5}-1)/2$ is the golden mean, and $V\ge 0$. 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: $F_{n+1} = F_n + F_{n-1}$, $F_1 = F_0 = 1$).

The object of fundamental mathematical interest associated to the linear operator H of (1.1) is its *spectrum*, *spec*(H), defined by $spec(H) = \{E\in \mathbf{C} \mid H - E\,Id \text{ is non-invertible}\}$, where Id is the identity operator. The operator H is characterized by its spectrum, just as a linear operator on a finite dimensional vector space is characterized by its eigenvalues. In fact any eigenvalue of H lies in its spectrum. The associated eigenvector is known as a *localized* state. However there may be elements of the spectrum which are not eigenvalues. In general, all that can be said of the operator H of (1.1) is that its spectrum is a closed subset of the real line.

In the case where H is a periodic operator, for which there exists a p such that $v(n+p) = v(n)$ for all $n\in Z$, there is a complete characterization of the spectrum in the following sense. Let $S_n(E)$ denote the 2x2 unimodular matrix $S_n(E) = \begin{bmatrix} E - v(n) & -1 \\ 1 & 0 \end{bmatrix}$. Then the spectrum of H is given by (1.3),

$$spec(H) = \{E \in \mathbf{R} \mid |\, \mathrm{trace} M_p(E)| \leq 2\} \tag{1.3}$$

where $M_p(E) = S_p(E) \cdots S_2(E) S_1(E)$.

We now briefly outline how the identity (1.3) can be obtained from a simple (non-rigorous) physical argument. Physically, the spectrum of the operator H is the set of "energies" E for which the second order difference equation $H\psi = E\psi$ has "physical" solutions ψ, in the following sense. The equation $H\psi = E\psi$ describes a quantum mechanical particle moving in a one dimensional lattice. The physical interpretation of a solution ψ to this equation is that the probability of finding a particle of energy E at the site n is proportional to $|\psi(n)|^2$. Thus for ψ to be a physical solution, it is required that $|\psi(n)|^2$ is bounded as $n \to \pm\infty$. Note that physical solutions need not lie in $l^2(\mathbf{Z})$; for example $\psi(n) = \exp(ikn)$ represents an *extended* state, which is equally likely to be found at any site.

Consider now the case of a periodic Schrödinger operator H. The equation $H\psi = E\psi$ can be written in the matrix form $\Psi(n+1) = S_n(E)\Psi(n)$, where $S_n(E)$ has been defined above, and $\Psi(n) = \begin{bmatrix} \psi_n \\ \psi_{n-1} \end{bmatrix}$. Since $v(n)$ has period p it follows that $\Psi(np) = (M_p(E))^n \Psi(0)$. Let λ_1 and λ_2 be the eigenvalues of $M_p(E)$. Since $M_p(E)$ is unimodular, we have $\lambda_1\lambda_2 = 1$. Suppose that λ_1 and λ_2 do not lie on the unit circle $\{z \in \mathbf{C} \mid |z| = 1\}$. Then it follows that $\psi(np) = \alpha_1 exp(n\lambda_1) + \alpha_2 \exp(n\lambda_2)$ is (exponentially) unbounded as either $n \to +\infty$ or $n \to -\infty$. Thus for E to lie in the "physical" spectrum, we require the eigenvalues of the matrix $M_p(E)$ to lie on the unit circle. This is precisely the condition (1.3) above. Note that when $\lambda_1 = \lambda_2 = 1$, the solutions $\psi(np)$ will in general only be polynomially bounded. From the physical point of view, this is not important, since it is a borderline case. However, it is essential that the associated value of E be admitted to the spectrum, since the spectrum is a closed set.

Restrict attention now to the periodic operators P_n, defined above, and let $B_n = spec(P_n)$. 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. For example, the lower half of Fig. 1, labeled by symbols of the form 6 2.., looks like a scaled down inverted copy of the upper half of Fig. 1, labeled by symbols of the form 1... Also, it looks as though the sets B_n are tending towards a Cantor set as $n \to \infty$. In fact it is a consequence of Sections 2 and 3 that, when $V > 8$, the sets B_n converge to a Cantor set of measure zero, in the Hausdorff metric. Unfortunately, there are not the necessary general theorems available to allow us to conclude that the spectrum of the quasiperiodic operator Q is a Cantor set. One of the difficulties is that 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)

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

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, $M_{F_n}(E)$ 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 [26,27,28,29], allows us to determine the detailed structure of the sets B_n and B_∞ using methods of dynamical systems theory. This theory shows that the matrix $M_{F_n}(E)$ is given by a matrix product of the form $BAABA...$, generated by n iterations of the "renormalization map" $R(A,B) = (BA,A)$, with the initial conditions

$A = \begin{bmatrix} E+V & -1 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} E-V & -1 \\ 1 & 0 \end{bmatrix}$. Following [29], we now show how the map R arises from a renormalization $\bar{\bar{T}}$ of the circle map R_σ. Consider the quasiperiodic operator Q. The potential v_n associated to it is obtained by considering the orbit of the point 0 under R_σ. Let b denote the subinterval $(-\sigma, -\sigma^3]$ of S^1, and a denote the subinterval $(-\sigma^3, 1-\sigma]$ of S^1. Then the orbit of 0 may be described by a "symbol sequence" *..baaba..* according to which subinterval its iterates lie in. We will show that this symbol sequence is invariant under the transformation $R': (a,b) \to (ba,a)$. We represent a circle map f by a pair (ξ,η), so that $f(x) = \xi(x)$ for $x \in b$, and $f(x) = \eta(x)$ for $x \in a$. Thus R_σ is represented by the pair (ξ,η), where $\xi(x) = x + \sigma$ and $\eta(x) = x + \sigma - 1$ (see Fig. 0). Then the renormalized map $\bar{\bar{T}}f$ is defined by $\bar{\bar{T}}f = \beta o g o \beta^{-1}$, where the map g is represented by the pair $(\eta, \eta o \xi)$, and β is the expanding map defined by $\beta = -\sigma^{-1}(x + \sigma^3) - \sigma^3$ (see Fig. 0). It is easily verified that $\bar{\bar{T}}f = f$, so that the orbit of 0 under f is the same as the orbit under $\bar{\bar{T}}f$. However, if we divide the interval a into the two subintervals $a_1 = (-\sigma^3, 0]$ and $a_2 = (0, 1-\sigma]$, then we have $\xi(b) = a_2$ and $\eta(a_2) \subset b \cup a_1$, so that $\eta o \xi(b) \subset b \cup a_1$. Thus given an orbit of f we can label the corresponding orbit of the map g by replacing the symbol ba by b and by leaving the other symbols (all a's) fixed. The construction of the map $\bar{\bar{T}}f$ involves rescaling the map g by a negative number, so that to obtain the corresponding symbol sequence for the orbit under $\bar{\bar{T}}f$ we replace the symbol ba by a, and the remaining symbols a by b. This establishes our claim. For an alternative derivation of the transformation $R(A,B) = (BA,A)$, see [21]. For the general theory of the renormalization of circle maps see [29].

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 $x_n = \frac{1}{2}\,\mathrm{trace}\, M_{F_n}(E)$, satisfies the "trace identity" (1.5),

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

with initial conditions $x_1 = \frac{E+V}{2}$, $x_0 = \frac{E-V}{2}$, $x_{-1} = 1$. This fact is easily verified. Let $X_n = M_{F_n}(E)$. We add the relationships (which follow immediately from the renormalization map above) $X_{n+1} = X_n X_{n-1}$ and $X_{n-2} = X_n X_{n-1}^{-1}$, to obtain $X_{n+1} + X_{n-2} = X_n(X_{n-1} + X_{n-1}^{-1})$. But X_{n-1} is a 2×2 unimodular matrix, and when added to its inverse, yields the matrix $2x_{n-1}Id$, where Id is the identity matrix. Taking the trace of the resulting matrix identity yields equation (1.5) as required. We have thus established that the map $T: R^3 \to R^3$ defined by (1.6)

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

determines B_n via (1.7),

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

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

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

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

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

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 $\mathbf{R}^3$ and $T^{-1} = \rho_{xz}^{-1} \circ T \circ \rho_{xz}$, where ρ_{xz} is the reflection in the $x = z$ plane.

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

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

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 **P**: no two consecutive terms of the sequence have modulus greater than unity.

Our objective is to prove that for $V \geq 8$, 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.

The set R_V is defined by (2.2),

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

where $\mathrm{w}(x,y,z,) = 2yz - x, x, y, z, 2xy - z$. We remark on this special choice of R_V at the end of this section. Note that $w(x,y,z,)$ is a subsequence of the bi-infinite sequence $..x_{-2},x_{-1},x_0,x_1,x_2,..$ 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 $R_1,..,R_{10}$ 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 $R_1,..,R_{10}$ are defined according to which of the intervals $L^-,s,L^+,*$ the coordinates of $w(x,y,z)$ lie in, by Table 2.1.

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

Table 2.1

The images of the regions $R_1,..,R_{10}$ under the map T_V are illustrated in Fig. 3. It may be verified by an inspection of Fig. 3 that the regions $R_1,..,R_{10}$ satisfy $T(R_i) \cap R_j \neq \varnothing$ whenever there is a connection $i \rightarrow j$ in the directed graph G of Fig. 4. This motivates us to define a 10x10 matrix A by $A_{ij} = 1$ if there is a connection $i \rightarrow j$ 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 $R_1,..,R_{10}$ are closed and disjoint, their union forms the whole of R_V, and $A_{ij} = 0$ implies $T(R_i) \cap R_j = \varnothing$.*

Proof. We first show that the union of the regions $R_1,..,R_{10}$ 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 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 (1.7).

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 $R_1,...,R_{10}$ 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 $R_1,...,R_{10}$ to overlap when $V > 2$. It suffices to show that if the point (x,y,z) is contained in R_V and $w(x,y,z)$ has a coordinate $w_i \in L^{\pm}$, then $|w_i| \geq |V-1|$, 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 $|w_i| \geq |V-1|$, 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 $|y| \geq |V-1|$, 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 $i \rightarrow j$ 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. □

Remark. In proving Theorem 2.1 below, we show that the regions $R_1,...,R_{10}$ 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" $\{1,..,m\}$ of m symbols, define the set Σ_m of two-sided symbol sequence by $\Sigma_m = \prod_{n=-\infty}^{\infty} \{1,..,m\}$. When $\{1,..,m\}$ 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* $\sigma : \Sigma_m \rightarrow \Sigma_m$ by $\sigma(s)_n = s_{n+1}$, where s_n is the nth symbol is s. Let A be the 10x10 matrix defined above. Then define $\Sigma(A)$ by (2.3),

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

and define the *subshift* σ_A to be $\sigma \mid_{\Sigma(A)}$. 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 $s \in \Sigma(A)$, define the sets V_{s^+} and H_{s^-} by $V_{s^+} = \bigcap_{n \in N} V_{s_0 s_1 \ldots s_n}$ and $H_{s^-} = \bigcap_{n \in N} H_{s_0 s_{-1} \cdots s_{-n}}$, where $V_{s_0 s_1 \ldots s_n} = R_{s_0} \cap T^{-1} R_{s_1} \cap \cdots \cap T^{-n} R_{s_n}$ and $H_{s_0 s_{-1} \ldots s_{-n}} = R_{s_0} \cap T R_{s_{-1}} \cap .. \cap T^n R_{s_{-n}}$. Then the map $x : \Sigma(A) \to R_V$ is defined by $x(s) = V_{s^+} \cap H_{s^-}$. It is an immediate consequence of this construction that if $x(s) \neq \emptyset$ then $T^n x(s) \in R_{s_n}$ for all $n \in \mathbf{Z}$, so that $x(s)$ has its past and future history coded by the symbol sequence 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 $x(s)$ 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 $V_{s_0 \cdots s_n}$ and $H_{s_0 \cdots s_{-n}}$. To state our results precisely, we introduce some geometrical concepts [25].

Let $I^2 = [a,b] \times [a,b]$ be a square in $\mathbf{R}^2$. 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 $|u(x_1) - u(x_2)| \leq \mu |x_1 - x_2|$ 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 $H = \{ (x,y) \in I^2 \mid x \in I, u_1(x) \leq y \leq u_2(x) \}$ 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 $V \geq 8$ and $s_0 \in \{1,2\}$, there exists $\mu \in (0,1)$ and $\nu \in (0,1)$ such that $V_{s_0 \ldots s_n}$ (resp. $H_{s_0 \ldots s_{-n}}$) is a μ-vertical (resp. μ-horizontal) strip of diameter $\leq \nu^n$, whenever $s_1 \ldots s_n$ satisfies*

$A_{s_i s_{i+1}} = 1$ *for all* $0 \le i \le n-1$. *Moreover if* $A_{s_i s_{i+1}} = 0$ *for some* $0 \le i \le n-1$, *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* $V \ge 8$, *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 $(1+\sigma, \sigma, -\omega, -\bar{\omega}, -1, 1$ 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 $A_{ij}^k > 0$ 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 $\sigma_A{}'$ 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 $V \ge 8$. In fact the restriction of Corollary 2.2 to the range $V \ge 8$ 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.

It was shown in [17] that when $V = 0$, the map T_V is conjugate to the map $A : S^2 \to S^2$, where $S^2 = \{(\theta,\phi) \in \mathbb{R}^2 \mid \theta \sim \theta+1, \phi \sim \phi+1, (\theta,\phi) \sim -(\theta,\phi)\}$ is a sphere, and A is the Anosov-like map $A(\theta,\phi) = (\theta+\phi, \theta)$. Fig. 5 illustrates a 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 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 $A^{\pm}B^{\pm}C^{\pm}$ [3]. 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 S^2, but that the members of the partition no longer overlap. Moreover numerical experiments reveal that the partion 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 $V \geq 8$ in subsequent sections.

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. 5). 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 $V_{s_0 \cdots s_n}$, 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 $s_0 \cdots s_n$ can be split up into a sequence of symbols drawn from S, so that the set $V_{s_0 \cdots s_n}$ is given by an intersection of the form $V_{t_0} \cap \phi^{-1} V_{t_1} \cap .. \cap \phi^{-m} V_{t_m}$, where $t_i \in S$.

Define B to be the matrix (2.5) with respect to the basis S,

$$B = \begin{bmatrix} 1&0&0&0&1&1 \\ 1&0&0&0&1&1 \\ 1&0&0&0&1&1 \\ 0&1&1&1&0&0 \\ 0&1&1&1&0&0 \\ 0&1&1&1&0&0 \end{bmatrix} \qquad (2.5)$$

and let $s_0, \ldots, s_n \in S$. Theorem 2.1 is thus equivalent to showing that there exists $\mu \in (0,1)$ and $\nu \in (0,1)$ such that $V_{s_0 \cdots s_n}$ (resp. $H_{s_0 \cdots s_n}$) is a μ-vertical (resp. μ-horizontal) strip of diameter $\leq \nu^n$ whenever $B_{s_i s_{i+1}} = 1$ for all $0 \leq i \leq n-1$. For $V \geq 8$, 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 $V \geq 8$, $\mu = 1/3$, and referring to the x, z coordinate system, the conditions are as follows.

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

(2) For all $s \in S$, ϕ maps vertical (resp. horizontal) boundaries of V_s to vertical (resp. 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 $S^+ = \{(\xi,\eta) \mid |\eta| \leq \mu |\xi| \}$ defined over the region $X = \{\bigcup_{s \in S} V_s\} \cap \{\bigcup_{s \in S} H_s\}$ 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 $|\xi_1| \geq \mu^{-1} |\xi_0|$. Also the cone field $S^- = \{(\xi,\eta) \mid |\eta| \geq \mu^{-1} |\xi| \}$ 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 $|\eta_1| \geq \mu^{-1} |\eta_0|$.

Using the symmetry $T^{-1} = \rho_{xz}^{-1} \circ T \circ \rho_{xz}$, 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 $L_1 = \{(1, V+t, t) \mid t \in [-1,1]\}$. 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 $C = \{(x(t), y(t), z(t)) \mid t \in [-1,1]\}$, 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 $\mu-$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 $\mu-$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 $\mu-$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 $H_s \cap V_t = \varnothing$, 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 $H_s \cap V_t \neq \varnothing$.

(4) The set X is covered by $\{V_s \mid s \in S\}$, 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 $\{(\xi, \zeta, \eta) \mid (\xi, \eta) \in \mathbf{R}^2\}$, 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).

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

We must show that whenever $(x,y,z)\in V_{17}$ and ξ_0,η_0 is such that $|\eta_0| \le |\xi_0|/3$, then $|\eta_1| \le |\xi_1|/3$ and $|\xi_1| \ge 3|\xi_0|$, 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 $\hat{v}$ given by (2.11),

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

satisfies $\hat{v}_3/\hat{v}_1 \subset [-1/3,1/3]$ and $\hat{v}_1 \cap (-3,3) = \varnothing$, where $\hat{y}$, $\hat{\zeta}$ are the intervals obtained by substituting $\hat{x}$, $\hat{z}$ for x,z in the equations (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 $V \ge 5$), and (2.12).

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

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

Remark. Some of the other computations require $V \ge 8$. 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 Section 2 to obtain qualitative results on the spectrum of the periodic operators P_n, and the pseudospectrum of the quasiperiodic operator Q.

3.1. The Periodic Operators.

We now describe how the symbolic dynamics of Section 2 gives rise to a labeling of the spectra B_n of the periodic operators P_n. The identity (1.7) of Section 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 $L_V(\mathbf{R})$ 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 $V_{s_0 \cdots s_{n-1}}$, where $s_0 \in \{1,6\}$. Also $T^{n-1}V_{s_0 \cdots s_{n-1}} \subset R_{s_{n-1}}$, 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 $\{L_V(E)\} \in V_{s_0 \cdots s_{n-1}}$, 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),

$$\Sigma_n'(A) = \{ s_0 .. s_{n-1} \in \prod_{i=0}^{n-1} \{1,..,10\} \mid s_0 \in \{1,6\}, s_{n-1} \in \{1,2,3,4\}, A_{s_i s_{i+1}} = 1 \text{ for } 0 \le i \le n-2 \} \quad (3.1)$$

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

$$b(s_0 .. s_{n-1}) = \{ E \in \mathbf{R} \mid \{L_V(E)\} \in V_{s_0 .. s_{n-1}} \} \qquad (3.2)$$

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. We first show by induction that $card(\Sigma_n'(A)) = F_n$. Define the space $\Sigma_n(A)$ of symbol sequences of length n by

$$\Sigma_n(A) = \{ s_0 .. s_{n-1} \in \prod_{i=0}^{n-1} \{1,..,10\} \mid s_0 \in \{1,6\}, A_{s_i s_{i+1}} = 1 \text{ for } 0 \le i \le n-2 \}$$

and let a_n, b_n, c_n denote , respectively, the number of symbol sequences in $\Sigma_n(A)$ which end with a symbol in $\{1,2\}$, $\{3,4\}$, $\{5,..,10\}$, respectively. It suffices to show that $a_n = F_{n-1}$, $b_n = F_{n-2}$ and $c_n = F_n$. This is easy to verify for $n = 2$. Suppose it is true for $n = N$. Then when $n = N+1$, we can deduce from the graph G that $a_{N+1} = c_N = F_N$, $b_{N+1} = a_N = F_{N-1}$ and $c_{N+1} = 2a_N + b_N = 2F_N + F_{N-1} = F_{N+1}$. Thus the hypothesis is also true for $n = N+1$, and hence by induction is true for all N, as required.

By Theorem 2.1 the sets $V_{s_0 \cdots s_{n-1}}$, such that $s_0 \cdots s_{n-1} \in \Sigma_n'(A)$, are disjoint non-empty vertical strips intersecting the line $L_V(\mathbf{R})$. Thus the images of distinct symbol sequences under the map b are disjoint non-empty closed intervals, and b is an injection. Since $card(\Sigma_n'(A)) = F_n$, 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 $h_n : \mathbf{R} \to \mathbf{R}$ 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. $\square$

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 $L_V(\mathbf{R})$ 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" $\{B_n \mid n \in \mathbf{Z}^+\}$. 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 $b(s_0 \cdots s_{n-1}s_n) \in B_n$ does not imply $b(s_0 \cdots s_{n-1}) \in B_{n-1}$, 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 6 2.

The above labeling is important, because in Section 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 $s = s_0 \cdots s_{n-1}$ and $t = t_0 \cdots t_{n-1}$ 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 $s = 1 s_1 .. s_{i-2} \gamma_1 s_{i+2} .. s_{n-1}$ and $t = 1 s_1 .. s_{i-2} \gamma_2 t_{i+2} .. t_{n-1}$ where $\gamma_1 = 171$ and $\gamma_2 = 1102$. It suffices to show that $\pi_1 V_{1 s_1 .. s_{i-2} \gamma_1} > \pi_1 V_{1 s_1 .. s_{i-2} \gamma_2}$. Observe that $V_{1 s_1 .. s_{i-2} \gamma} = T^{-(i-2)}(H \cap V_\gamma)$, where $H = H_{1 s_{i-2} .. s_1 1}$ 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 $j \in \{2, ..., i\}$ 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. 5). 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 $s_1 .. s_{i-2}$. □

3.2. The Quasiperiodic Operator.

The labeling of the band structure $\{B_n \mid n \in \mathbf{Z}^+\}$ 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)

$$\Sigma'(A) = \{s_0 .. s_n .. \in \prod_{n=0}^{\infty} \{1, .., 10\} \mid s_0 \in \{1,6\}, A_{s_i s_{i+1}} = 1 \text{ for all } i \geq 0\} \tag{3.3}$$

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

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

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 $V_{s_0 \cdots s_n \cdots}$ such that $s_0 .. s_n .. \in \Sigma'(A)$, are distinct non-empty vertical curves intersecting the horizontal line $L_V(\mathrm{R})$ in distinct points. Let $L_V(q(s_0 .. s_n ..)) = V_{s_0 \cdots s_n ..} \cap L_V(\mathrm{R})$ be such an intersection point. It follows from (1.8) that $q(s_0 \cdots s_n ..)$ consists of a unique point. By definition of $V_{s_0 .. s_n ..}$ we have that $\pi_1 T^n L_V(q(s_0 .. s_n ..))$ is bounded as $n \to \infty$. Then from the dynamical equation (1.9) for B_∞, we have $q(s_0 .. s_n ..) \in B_\infty$. Thus the map q is an injection into B_∞. Let V be the union of the sets $V_{s_0 \cdots s_n ..}$, such that $s_0 \cdots s_n .. \in \Sigma'(A)$. To show that the map q is a surjection, we observe that all points of $L_V(\mathrm{R})$ not intersecting 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 $\mathrm{s,t} \in \Sigma'(A)$, where $\mathrm{s} = s_0 s_1 ..$ and $\mathrm{t} = t_0 t_1 ...$ The topology on $\Sigma'(A)$ is metrizable, and we choose a metric d defined by $d(\mathrm{s,t}) = 2^{-i}$, 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 $d(\mathrm{s,t}) < \delta$ implies s and t agree in their first $n+1$ places, where $n > \dfrac{\log \varepsilon}{\log \nu}$. Then $d(\mathrm{s,t}) < \delta$ implies $|\, \pi_1 L_V(q(\mathrm{s})) - \pi_1 L_V(q(\mathrm{t}))\,| < d(V_{s_0 .. s_n}) < \nu^n < \varepsilon$. It follows from (1.8) that $|\, q(\mathrm{s}) - q(\mathrm{t})| < 2\varepsilon$, 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 $\mathrm{V} \subset W^s(\Omega_V)$, it follows that V has Lebesgue measure zero. The vertical curves of V intersect $L_V(\mathrm{R})$ transversally at $L_V(B_\infty)$, and it follows that $L_V(B_\infty)$ has Lebesgue measure zero in $L_V(\mathrm{R})$, 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 Section 2 that for $V \geq 8$, 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 $n \in \mathbf{N}$ by (4.1).

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

The results of Section 2 also allow us to conclude the following. The line $L_V(\mathbf{R})$, 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 $L_V(q(\mathbf{t}))$, 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 $\Sigma_1^{\pm}$ 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]).

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 Section 2 that for $V \geq 8$, 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 $n \in \mathbf{N}$ by (4.1).

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

The results of Section 2 also allow us to conclude the following. The line $L_V(\mathbf{R})$, 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 $L_V(q(\mathbf{t}))$, 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 $\Sigma_1^{\pm}$ 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),

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

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),*

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

where 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 $\phi^u(x) = -\log |dT_x|_e$. In fact there is numerical evidence that the scaling of $m(B_n)$ is geometrical [19], which is a stronger property. The topological pressure $P(f)$, is defined as follows. Let (X,d) be a compact metric space, $T: X \to X$ a continuous map, and f a member of $C(X,\mathrm{R})$, the continuous real valued functions on X. Then the pressure of f with respect to T, $P(f)$, is defined by, $P(f) = \lim_{\varepsilon \to 0} \mathit{limsup}_{n \to \infty} \frac{1}{n} \log P(f,n,\varepsilon)$, where

$$P(f,n,\varepsilon) = \inf\{ \sum_{x \in F} \exp(S_n f)(x) \mid F \text{ is a } (n,\varepsilon) \text{ spanning set for } X \}$$

with $(S_n f)(x) = \sum_{i=0}^{n-1} f(T^i x)$, and a subset F of X is said to (n,ε) span X if for all $x \in X$, there exists $y \in F$ with $\max_{0 \le i \le n-1} d(T^i(x), T^i(y)) \le \varepsilon$. It may easily be shown that the topological pressure is well defined, though it may be infinite [35]. If the map T is conjugate to a topologically mixing subshift of finite type, then using the expanding properties of the map T, it can be shown that $P(f) = \lim_{n \to \infty} \frac{1}{n} \log \sum_{x \in Fix\, T^n} \exp(S_n f)(x)$ [32]. Taking f to be the function ϕ^u, we have $(S_n \phi^u)(x) = -\log |dT_x^n|_e$, so that from (4.4) we have $\lambda_g = P(\phi^u)$, which therefore exists.

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 in the underlying space whose orbits remain within ε of Ω from time 0 to n decays like $\exp nP(\phi^u)$ [5]. To see the connection with our result, consider a point $x' \in B_n$, with n large. Then x' lies very close to a point of B_∞, so that $x = L_V(x')$ is close to $W^s(\Omega_V)$, the stable manifold of the non-wandering set. Thus the images of x are soon within ε of Ω_V. Also the successive images of x remain ε- close to Ω_V for a long time, since we must ultimately have $\pi_1 T^n(x) = [-1,1]$. Thus there is a close relationship between the Lebesgue measure of those points in S_V whose orbits remain within ε of Ω_V from time 0 to n, and $m(B_n)$. However, for completeness, we will prove our result from first principles.

Finally, we remark that the theory of topological pressure was originally introduced for quite different reasons. The topological pressure satisfies a variational principle, which can be used to supply many interesting examples of invariant measures for T, called equilibrium states. The study of equilibrium states in turn leads to an understanding of the asymptotic behavior of the orbits of a set of points of full measure, at least for Axiom-A systems. They are also used in mathematical formulations of statistical mechanics. See [5,32,35] for details and further references.

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, $|v_{m+1} - dT_{x_m} v_m| \le K \, |v_m|^2$, where $K = \sup_{z \in \Omega} |d^2 T_z|$. 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 $|e_m| \le K |v_m|^2$. 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 $m > M_\varepsilon$,

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

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 $n \in \mathrm{N}$, $|v_{m+n}| / |v_m| = e^{n\xi} |dT^n_{x_m}|_c$ for some $\xi \in (-\varepsilon/2, \varepsilon/2)$. Thus $|v_{m+n}| = K_m e^{n\xi} |dT_x^{n+m}|_c |v_0|$, where $K_m = |v_m| / |dT_x^m|_c |v_0|$. 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 $n \in \mathrm{N}$. Then for all $\varepsilon > 0$, there exists N_ε, such that $n > N_\varepsilon$ implies $|a_n - b_n| = e^{n\xi} |dT_x^{-n}|_c |x - y|$ 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 $\lambda-$ lemma [30], there exists M_ε so that $T^{-r}\Sigma$ is $\varepsilon/2 - C^1$ close to $W^s(x_r)$ and $T^m L$ is $\varepsilon/2 - C^1$ close to $W^u(x_r)$ whenever $r, m \ge M_\varepsilon$. Fix $m = M_\varepsilon$, and take $n \ge 2M_\varepsilon$. 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 $b(s_0..s_n)$ and $q(s)$ with their images under the the parametrization map $E \to L_V(E)$. We apply Lemma 4.3 to the renormalization map T, taking $x = x(\sigma^n s)$, $L = L_V(E)$, $a_n = q(s_0..s_n..)$, and $\Sigma = \Sigma_1^{\pm}$. Let $s_0..s_{n-1} \in \Sigma_n'(A)$. Then $[\pi_1 b_n^-, \pi_1 b_n^+] = b(s_0..s_{n-1})$, where $b_n^{\pm}$ are the points in $T^{-n}\Sigma_1^{\pm} \cap L_V(E)$ nearest to $q(s_0..s_{n-1}..)$. We conclude that for all $\varepsilon > 0$, there exists N_ε such that $n > N_\varepsilon$ implies (4.7),

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

for some $\xi \in (-\varepsilon/2, \varepsilon/2)$, where $x_n = T^{-n}x = x(s)$. Since Ω_V is compact, N_ε may be chosen independently of $s_0..s_{n-1}$, and we can sum (4.7) over all $s_0..s_{n-1} \in \Sigma_n'(A)$, 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 $V_{s_0..s_{n-1}}$ such that $s_0..s_{n-1} \in \Sigma_n'(A)$. It remains to show that for sufficiently large n, $m(B_n) = e^{n\xi}\beta_n$ for some $\xi \in (-\varepsilon, \varepsilon)$, where $\beta_n = \sum_{x \in FixT^n} (\,|dT^n_x|_e)^{-1}$.

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 $\{\,|dT^m_x|_e \mid x \in \Omega\}$ is bounded, there exists $N_\varepsilon(m) \in \mathrm{N}$ such that $n > N_\varepsilon(m)$ implies $|dT^n_x|_e = e^{n\xi'(x)}$, where $\xi'(x) \in (-\varepsilon/2, \varepsilon/2)$ for all $x \in \Omega$. Thus from (4.8) we deduce (4.9).

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

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

From the mixing property of A, there exists $M_\varepsilon \in \mathrm{N}$ such that if $m = M_\varepsilon$, and n is sufficiently large, then $Card\,(FixT^{n+m} \cap V_s) = e^{n\xi(s)}$, where $\xi(s) \in (-\varepsilon, \varepsilon)$ for all

$s = s_0..s_{n-1} \in \Sigma'_n(A)$. 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 $x(s) \in FixT^{n+m}$ with $s_0 \in \{1,6\}$ can be dealt with in this fashion. Thus $\beta_{n+m} \le Ke^{n\varepsilon}e^{-n\xi}m(B_n)$, 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. □

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 $q(s_0..s_{n-1}..)$, with exponent λ_e, if the following limit exists

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

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

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

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

Proof. It follows from equation (4.7) that

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

where $t \in \Sigma(A)$ has the same tail as $s_0..s_{n-1}..$. By the multiplicative ergodic theorem [35], the limit (4.12) exists for μ – almost all $t \in \Sigma(A)$, 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 $\mu = \lim_{n\to\infty} (N_n)^{-1} \sum_{x \in FixT^n} \delta_x$, where $N_n = Card(FixT^n)$. 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).

$$\lambda_e = (N_n)^{-1} \sum_{x \in FixT^n} \log(|dT_x^n|_e)^{-1} \tag{4.13}$$

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 Section 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),*

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

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 $\delta = HD(W^u(x) \cap \Lambda)$ is independent of $x \in \Lambda$, and satisfies (4.15),

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

where h_{top} is the topological entropy of f, $\phi^u: W^u(\Lambda) \to R$ 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 equation (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_*, 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 Section 4.2. Finally, it can be shown that there is a Lipshitz map between $\Lambda \cap L_V(E)$ and $\Lambda \cap W^u_{loc}(x)$, and since HD is preserved under Lipshitz maps, we have $\delta = HD(B_\infty)$. □

5. Rotation Numbers.

In theoretical studies of Schrödinger operators with quasiperiodic potentials, one of the basic tools is the rotation number $\rho(E)$ [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 labeling of the gaps of the spectrum, due to the following result [10]: $2\rho(E)$ 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 $k(E) = 2\rho(E)$. 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 Section 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 [10]. Let H be the operator given by (1.1), and let $\psi(0), \psi(1), \cdots, \psi(n), ..$ be a solution of the equation $H\psi = E\psi$ for $n \geq 0$ with initial condition $\psi(0) = \cos\theta, \psi(1) = \sin\theta$.

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

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

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

It is shown in [10] 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 $\psi(0)/\psi(1) = \cot\mathrm{an}(\theta)$.

Definition. The *integrated density of states* of H is the map $k:\mathbf{R}\to\mathbf{R}$ 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 [10] that $k(E) = 2\rho(E)$. 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 $(M_{F_n}(E))_{11} = 0$, where $M_{F_n}(E)$ is the the product of E-dependent transfer matrices described in Section 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 shows how to translate the labeling of B_∞ by symbol sequences into a labeling by rotation number. The symbol sequence is first translated into a sequence of 0's and 1's according to Table 5.1 (for example the symbol sequence $1\,7\,1\,7\,1\,3\,6\cdots$ is translated into 0101001...). The resulting sequence is then translated to yield the rotation number according to the "irrational expansion" (5.4). Precisely how Table 5.1 is arrived at is explained in the proof of Corollary 5.1.

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

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

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

s_i	1 7	1 3 6	1 10	2 8	2 4 5	2 9
d_{i+1}	0 1	0 0 1	0 0	0 1	0 0 1	0 0

Table 5.1

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 $E = q(s_0..s_n..) \in B_\infty$. Given any $N \in \mathbf{N}$, it is possible to find an $n > N$ such that $s_{n-1} \in \{1,2,3,4\}$, so that $q(s_0..s_n..) \in b(s_0..s_{n-1}) \in B_n$. 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 $b(t_0..t_{n-1})$, with $t_0 = 6$, lying below $b(s_0..s_{n-1})$. Similarly, it can be shown that there are F_{n-i-3} bands in B_n having labeling starting with $s_0..s_{i-1}1\ 10$, and F_{n-i-4} bands in B_n having labeling starting with $s_0..s_{i-1}1\,3\,6$. Hence if $s_i = 1$ and the block 1 7 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 $2\rho(E) = \lim_{n\to\infty} (m-1)/F_n$, and the result follows. □

The next corollary expresses some of the scaling results found in Section 4 in the language of rotation number.

Corollary 5.2 *Let $E = q(s_0..s_n..)$ 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 $\rho(E)$ 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 $E \in \rho^{-1}(X)$, 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 $B_\infty \cap (E,E+\delta) = \varnothing$ or $B_\infty \cap (E-\delta,E) = \varnothing$. Consider the former case. It follows that if $E = q(s_0 \cdots s_n ..)$ is at a gap edge, then there exists N such that $q(s_0 \cdots s_N s_{N+1} ..) > q(s_0 \cdots s_N t_{N+1} ..)$ whenever $t_{N+1} .. \neq s_{N+1} ...$ Using the ordering property of Lemma 3.2, it follows that $s_{N+1} ..$ has tail 1 7 1 7... Thus by the remarks of Section 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 $q(s_0 \cdots s_n ..)$ where $s_0 \cdots s_n ..$ has a period-2 tail 2 9 2 9...

(2) If $\rho(E)$ has an irrational expansion (5.4) with a periodic tail, then $E = q(s_0 \cdots s_n ..)$ where $s_0 \cdots s_n ..$ has a periodic tail. Thus, by the remarks of Section 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 Section 4.3. Then it can be shown, using the relationship $k = 2\rho$, and (5.3), that for all $\lambda \in [0,1/2]$, $\mu'(\rho^{-1}[0,\lambda]) = \lambda$, so that Lebesgue measure is induced on $\rho(B_\infty)$ by μ'. □

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 No.470. Springer-Verlag, New York 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. Am. Math. Soc. Providence 1975

[7] Bowen, R., Ruelle, D.: The ergodic theory of axiom A flows. Inventiones 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 quasi-periodic 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 theoreme 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 differérences 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., Manning, A.: Hausdorff dimension for horseshoes. Ergod. Th. and Dynam. Sys., 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. 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. Vol T9, 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 8D, 303-342 (1983)

[30] Palis, J., de Melo, W.: Geometric Theory of Dynamical Systems. Springer-Verlag, New York 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. Addison-Wesley, Reading 1978

[33] Simon, B.: Almost Periodic Schrödinger operators; A Review. Adv. App. 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. Springer-Verlag, New York 1982

Figure Captions.

Fig. 0. The geometric construction of the renormalized map $\bar{\bar{T}}f$ from the cirle map f.

Fig. 1. The bold lines represent the band spectra of P_n for $n = 1,...,5$. The case $V = 1.5$ is shown. The bands have been labeled using the symbol sequence scheme of Section 3.1. The dotted line 1 7 illustrates the bifurcation of Section 3.1.

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

Fig. 3. The xz projections of the regions $T(R_1),...,T(R_5),T(R_7),T(R_8)$ 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.

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

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

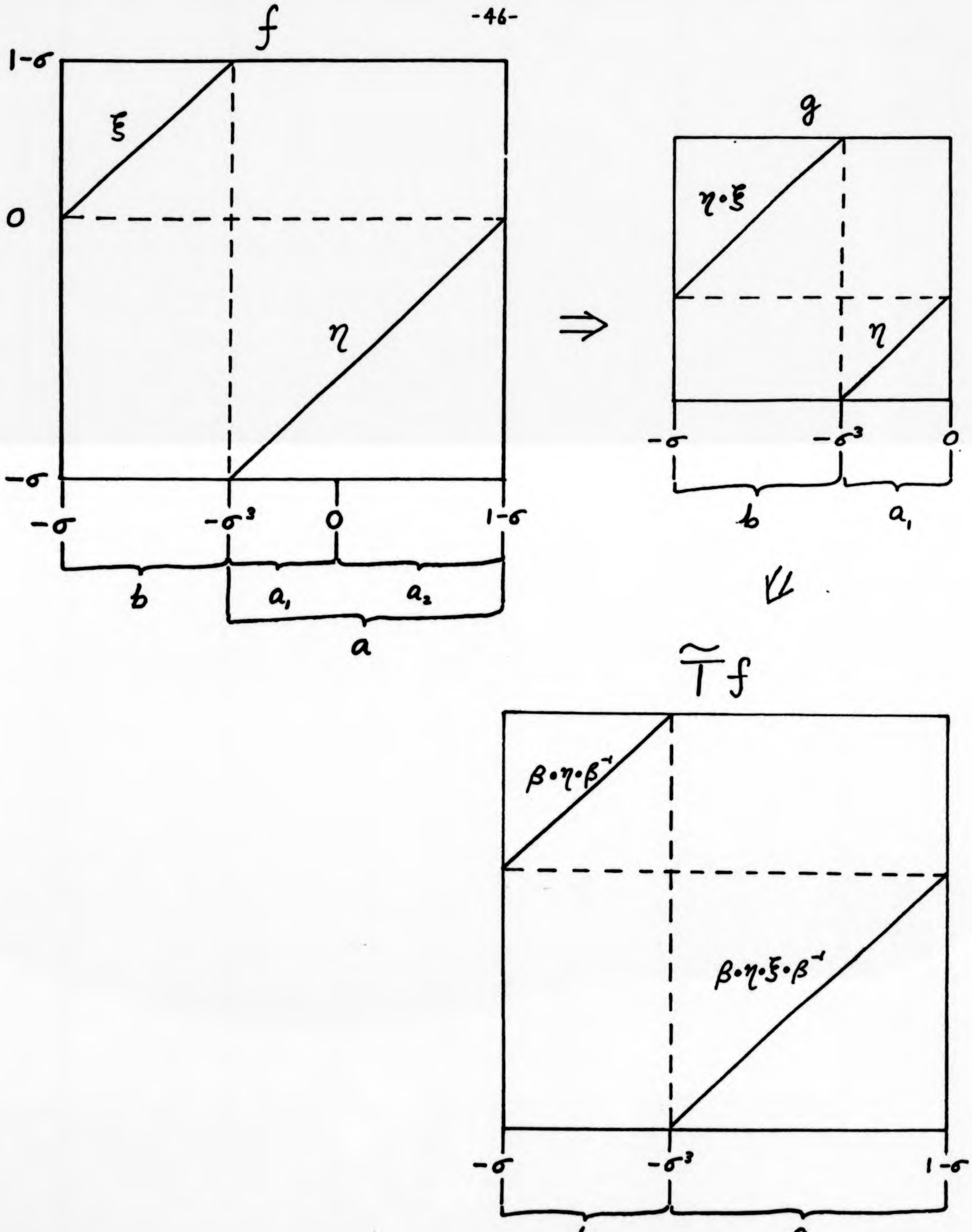
f
$1-\sigma$
ξ
0
η
$-\sigma$
$-\sigma$
$-\sigma^3$
0
$1-\sigma$
b
a_1
a_2
a
g
$\eta\circ\xi$
η
$-\sigma$
$-\sigma^3$
0
b
a_1
$\widetilde{T}f$
$\beta\circ\eta\circ\beta^{-1}$
$\beta\circ\eta\circ\xi\circ\beta^{-1}$
$-\sigma$
$-\sigma^3$
$1-\sigma$
b
a

Figure O.

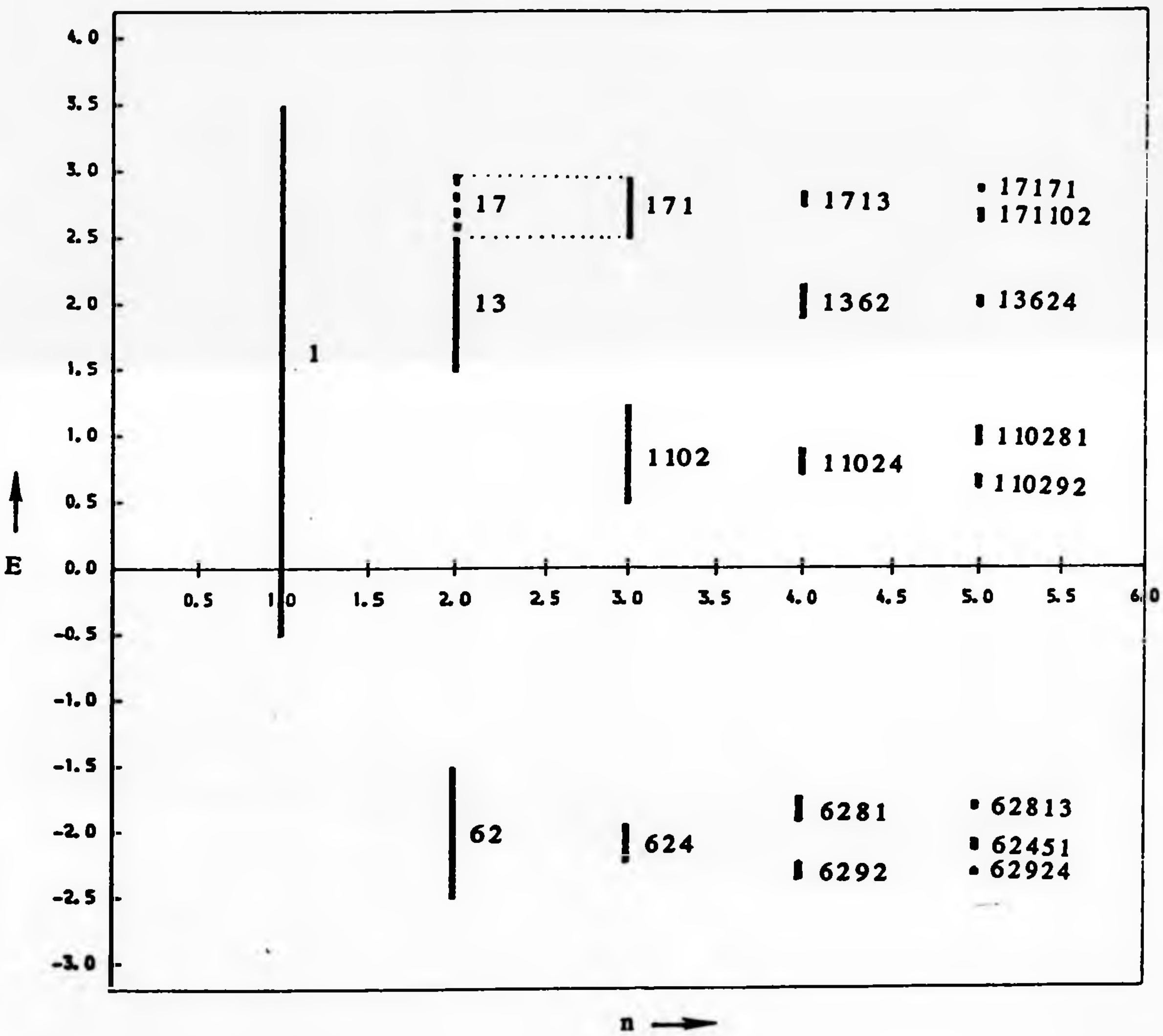
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
E
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
n
1
17
13
171
1102
1713
1362
11024
17171
171102
13624
110281
110292
62
624
6281
6292
62813
62451
62924

Figure 1

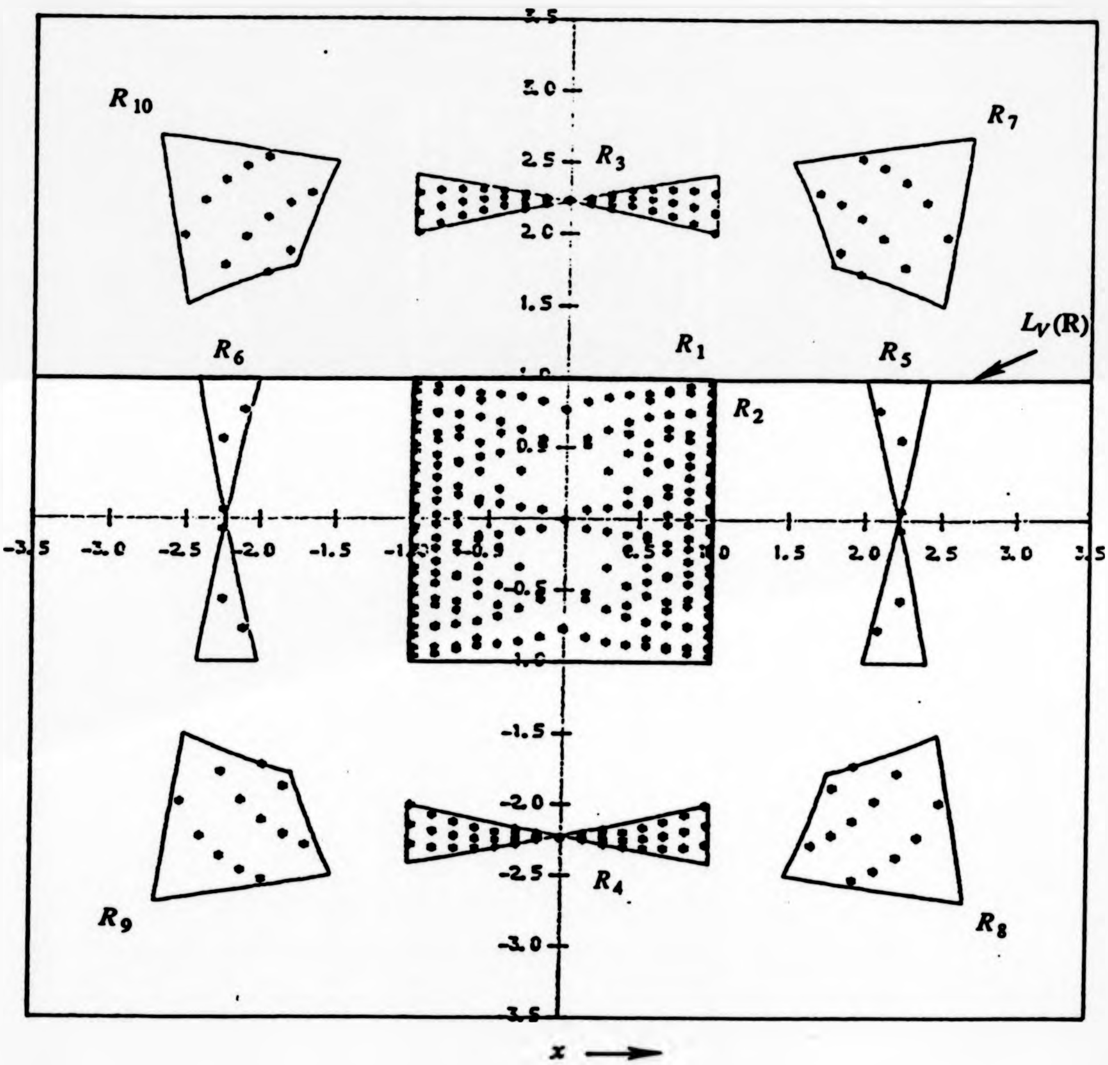

Figure 2

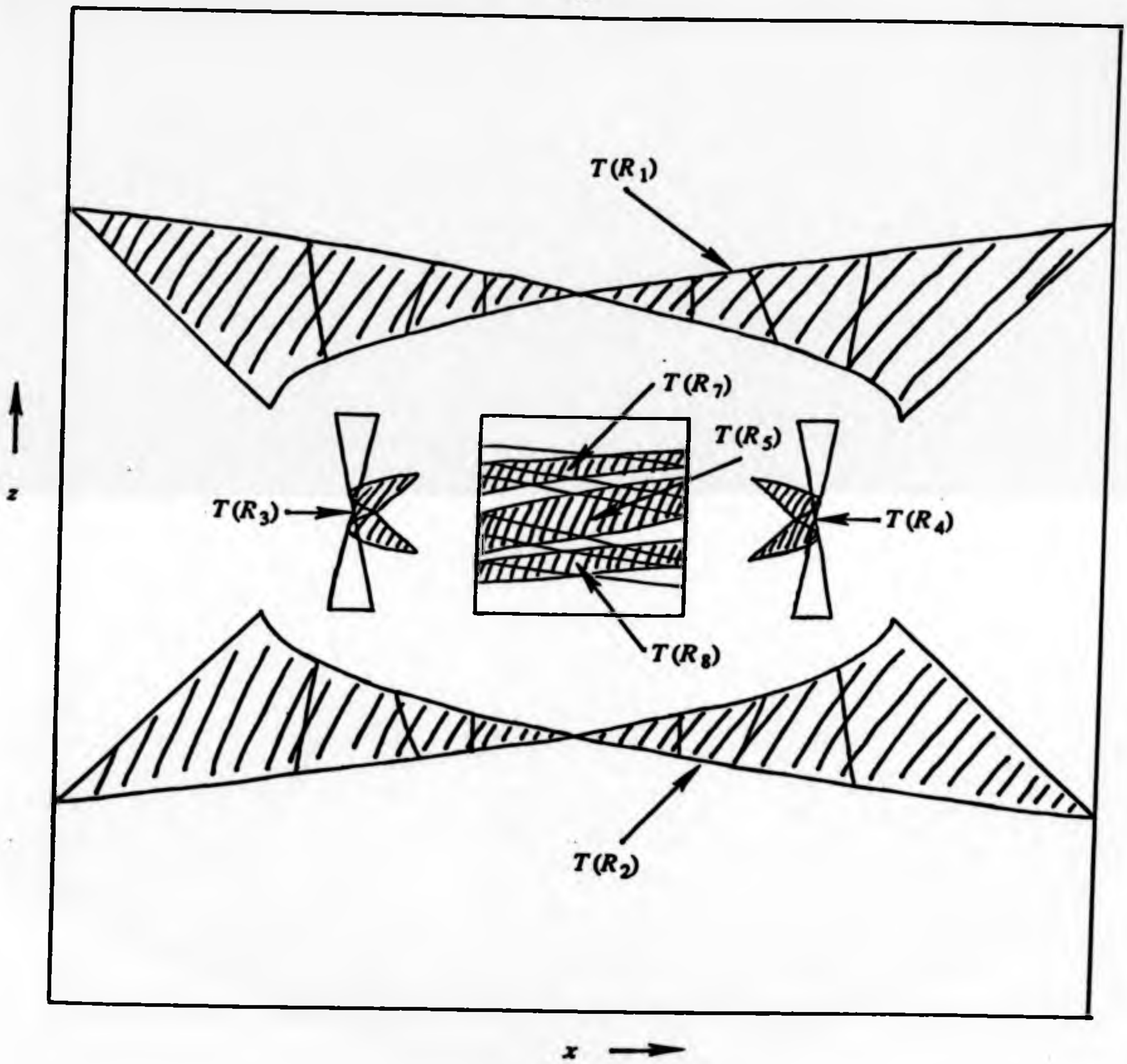

Figure 3

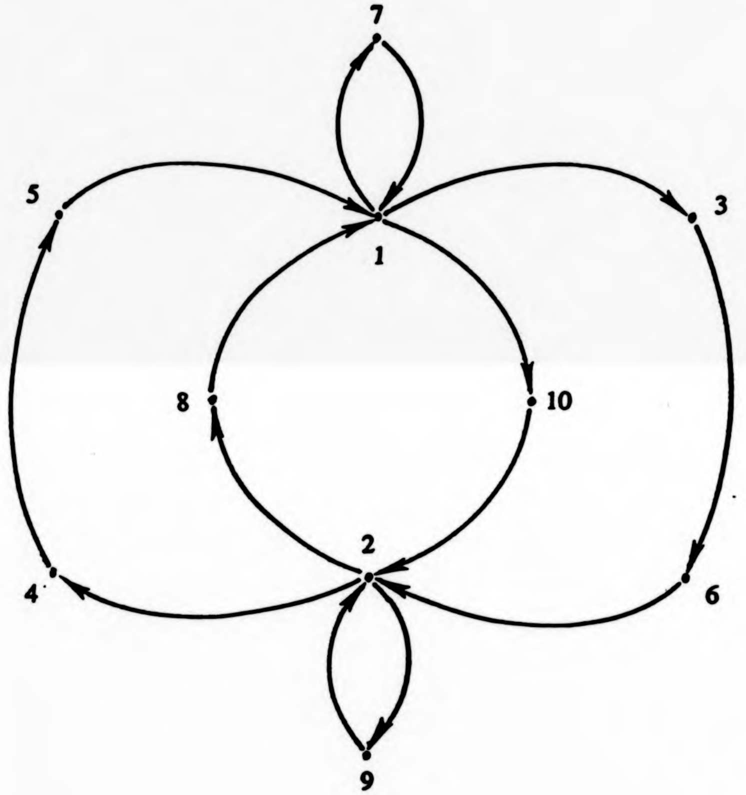

Figure 4

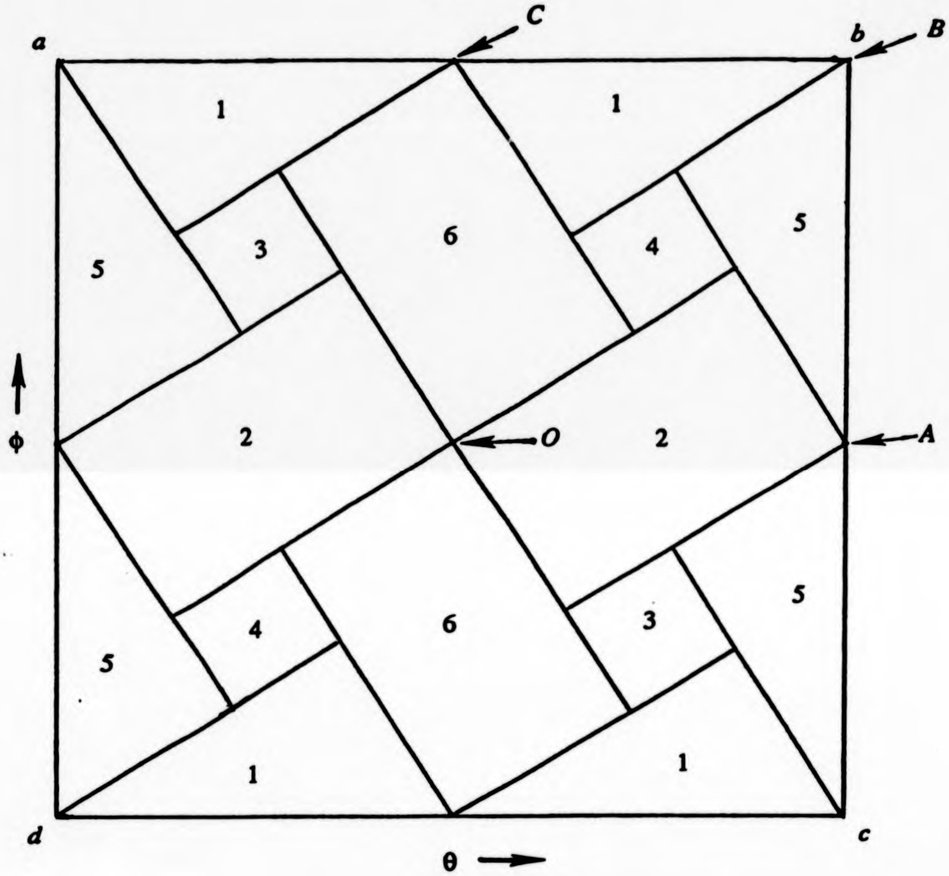

Figure 5

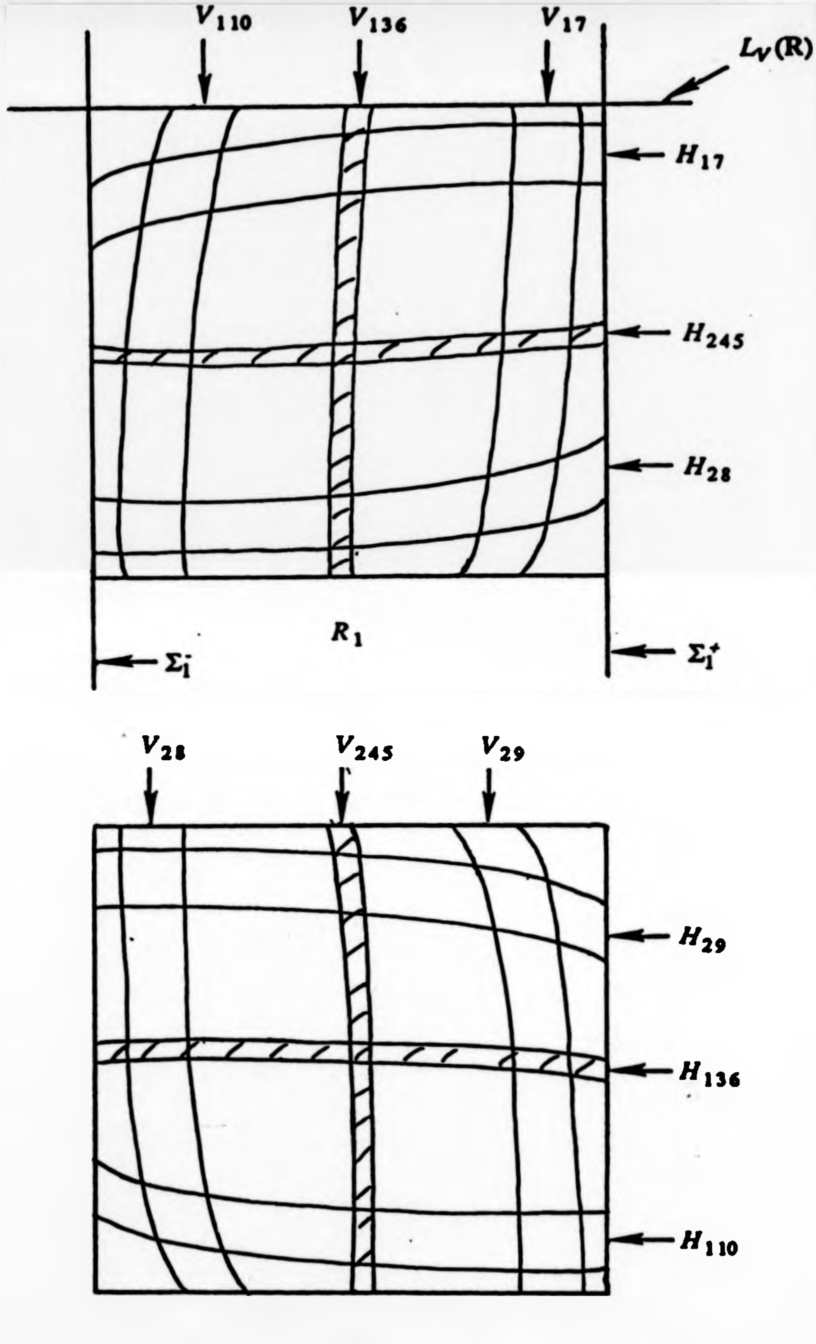
V_{110}
V_{136}
V_{17}
$L_V(R)$
H_{17}
H_{245}
H_{28}
R_1
Σ_1^-
Σ_1^+
V_{28}
V_{245}
V_{29}
H_{29}
H_{136}
H_{110}
R_2

Figure 6

Periodic Orbits for Dissipative Twist Maps

Martin Casdagli[*]

Mathematics Institute,
University of Warwick, Coventry CV4 7AL, U.K.

ABSTRACT

We develop simple topological criteria for the existence of periodic orbits in maps of the annulus. These are applied to one-parameter families of dissipative twist maps of the annulus and their attractors. It follows that many of the motions found by variational methods in area preserving twist maps also occur in the dissipative case.

[*] Current Address: La Jolla Institute, 3252 Holiday Ct., Suite 208, La Jolla, CA 92037.

Introduction.

The study of twist maps of the annulus was initiated by Poincaré, in connection with the 3-body problem. In this case the map is area preserving, and Birkhoff [4] was able to use a variational formulation to show that it has many periodic orbits. The variational approach was refined by Aubry [3] and Mather [19], who deduced the existence of well-ordered periodic orbits, and of quasiperiodic orbits which are either invariant curves or Cantor sets.

In this paper we are concerned with the existence of such motions in dissipative twist maps, which have been used as models for various physical processes [2,7,14,21]. As a consequence of the work of Katok [17,18], and Hall [12], it suffices to consider periodic orbits. In the first part of this paper we use a very simple new technique to deduce the existence of periodic orbits in one-parameter families of dissipative twist maps. The technique is topological in character, and reduces the two dimensional fixed point problem of finding periodic orbits to two essentially one dimensional problems. Our results are reminiscent of those of Chenciner [10,11] on the degenerate Hopf bifurcation, though our methods are different.

In the second part of this paper we focus attention on the orbit structure of the attractors of dissipative twist maps. It is known that if the map has transverse homoclinic points, then its attracting set contains orbits whose rotation numbers form a non-trivial interval [2,14]. In [5] Birkhoff showed how to define internal and external rotation numbers for attracting sets of a dissipative twist map, and gave an example where these two numbers enclose a non-trivial interval. By analogy with the area preserving case, it is natural to ask if there are orbits on the invariant set of all rotation numbers lying in this interval. We show that this is indeed the case if the invariant set satisfies an intersection property which we define. Again, topological techniques are used to prove this theorem. Using the topology of the annulus, and the geometry of the twist we develop a criterion for the existence of periodic orbits. This criterion represents an improvement on the "radially translated curve theorem" of Poincaré-Birkhoff [2].

In contrast to proving the existence of strange attractors (which necessarily satisfy the intersection property), it is easy to prove the existence of invariant sets with the intersection property. We show that every dissipative twist map contains such a set which is weakly attracting.

In Section 1 we collect definitions and notation used throughout the paper, and study periodic orbits of parameterized families. In Section 2 we prove our result on rotation intervals for invariant sets with an intersection property. In Section 3 we prove that every dissipative map of the annulus has an invariant set with the intersection property, which is also weakly attracting.

Acknowledgements.

I would like to thank my supervisor, David Rand, for suggesting this problem, and bringing Birkhoff's paper [5] to my attention in Feb '84. I would also like to thank R.S.MacKay and E.C.Zeeman for useful discussions. During the course of this work it was found that P. le Calvez had previously obtained some closely related results [8]. I would like to thank him for pointing out some errors in my original draft. This work was supported by a U.K. Science and Engineering Research Council award, and by the La Jolla Institute Contract, U.S. Department of Energy, DE-AC03-84ER40182.

1. Families of Dissipative Twist Maps.

Notation. We will use the following notation throughout this paper. S^1 denotes the unit circle, and $\bar{A}$ denotes the unit cylinder $S^1 \times R$. We define $\bar{H}$ to be the set of compact connected sets $\bar{\Lambda}$ separating $S^1 \times R$, i.e. $S^1 \times R - \bar{\Lambda}$ consists of two unbounded components $\bar{\Lambda}_{int}$ and $\bar{\Lambda}_{ext}$, where $\bar{\Lambda}_{int}$ (resp. $\bar{\Lambda}_{ext}$) contains the lower (resp. upper) end of the cylinder. There is a partial order on $\bar{H}$ defined by $\bar{A} < \bar{B}$ if $\bar{A} \subset \bar{B}_{int}$. An annular region is a set $\bar{B} \subset \bar{A}$ with frontier consisting of two disjoint sets $\partial^+\bar{B}, \partial^-\bar{B} \in \bar{H}$ such that $\partial^+\bar{B} > \partial^-\bar{B}$. A trapping region for a map $\bar{f}$ of $\bar{A}$ is a closed annular region $\bar{B}$ such that $\bar{f}(\bar{B}) \subset interior(\bar{B})$. The lift to R^2 of a set $\bar{S} \subset \bar{A}$ is denoted by S, and all of the above notation lifts in the obvious way.

Definition. A diffeomorphism $f: R^2 \to R^2$ is called a *twist map* if:

(1) f commutes with T where $T: R^2 \to R^2$ is the unit translation $T(\theta,r)=(\theta+1,r)$. It follows that f is the lift of a unique map $\bar{f}: \bar{A} \to \bar{A}$ of the cylinder.

(2) f is orientation preserving, and $\bar{f}$ is end preserving.

(3) There exists $\delta>0$ such that for all $(\theta,r) \in R^2$, $\dfrac{\partial(\pi_1 f(\theta,r))}{\partial r} > \delta$, where $\pi_1: R^2 \to R$ is the projection $\pi_1(\theta,r)=\theta$.

We call f a *dissipative* twist map if in addition:

(4) There exists $\lambda \in (0,1)$ such that for all $x \in R^2$, $0 < Det\ Df(x) \le \lambda$.

(5) There exists $M \in R^+$ such that for all $N \ge M$, $S^1 \times [-N,N]$ is a trapping region for $\bar{f}$. Note that conditions (1) to (4) do not necessarily imply condition (5).

In this paper we will be working with a fixed lift f of $\bar{f}$, so that rotation numbers are uniquely defined.

As a motivating example, consider the two parameter family $f_{a,k}$ of dissipative twist maps given by (1.1), where $\lambda \in (0,1)$ is fixed.

$$f(\theta,r) = (\theta + \omega + \lambda r - \frac{k}{2\pi}\sin(2\pi\theta)\,,\, \lambda r - \frac{k}{2\pi}\sin(2\pi\theta)) \tag{1.1}$$

If there is a $(\theta,r) \in \mathbf{R}^2$ such that $f^q_{\omega,k}(\theta,r) = (\theta+p, r)$ where $p \in \mathbf{Z}, q \in \mathbf{Z}^+$, we call (θ,r) a p/q periodic point. Then define the "Arnold Tongues" $I_{p/q}$ by (1.2)

$$I_{p/q} = \{(\omega,k) \in \mathbf{R}^2 \mid f_{\omega,k} \textit{ has a } p/q \textit{ periodic orbit}\} \tag{1.2}$$

We will also be interested in the sets $I_\sigma, \sigma \in \mathbf{R}-\mathbf{Q}$ defined by (1.3)

$$I_\sigma = \{(\omega,k) \in \mathbf{R}^2 \mid f_{\omega,k} \textit{ has an Aubry–Mather set of rotation number } \sigma\} \tag{1.3}$$

(An Aubry-Mather set is a closed f,T invariant set M which is minimal, and on which π_1 is injective, and f is order preserving: for all $x,x' \in M$, $\pi_1(x) < \pi_1(x')$ implies $\pi_1(f(x)) < \pi_1(f(x'))$. It follows that M has a well defined rotation number, and is either an invariant curve or a Cantor set [17,18]).

When $k = 0$, the circle $r = 0$ is an attracting invariant set for $f_{\omega,k}$. It is also normally hyperbolic, and therefore persists for $|k|$ sufficiently small. It may then be deduced that in this range of k, $I_{p/q} \cap \mathbf{R}\times\{k\} \neq \varnothing$ for all $p/q \in \mathbf{Q}$, and that the I_σ are curves of the form $\omega = u(k)$, where u is a continuous function. When $|k|$ is large, it is known that smooth invariant curves do not exist [6]. However it is a consequence of the next theorem that all the sets $I_\omega, \omega \in \mathbf{R}$ persist for arbitrarily large k, though they may overlap. This is illustrated in Fig 1.

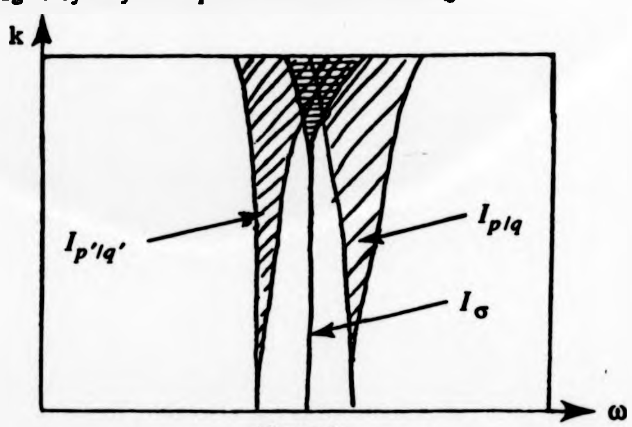

Figure 1

Theorem 1.1 *Let R_ω:$\mathbf{R}^2 \rightarrow \mathbf{R}^2$ be the rigid rotation $R_\omega(\theta,r)=(\theta+\omega,r)$, and let f be a dissipative twist map. Then for all $p/q \in \mathbf{Q}$ and $\theta_0 \in \mathbf{R}$, there exists $\omega \in \mathbf{R}$ such that $R_\omega o f$ has a p/q periodic point on the line $\{\theta_0\}\times\mathbf{R}$. Moreover, for all $\sigma \in \mathbf{R}-\mathbf{Q}$, there exists $\omega \in \mathbf{R}$ such that $R_\omega o f$ has an Aubry-Mather set of rotation number σ.*

Proof. Fix $\theta_0 \in \mathbf{R}$, and let $f_\omega = R_\omega o f$. For $\omega_1 < \omega_2$, define the closed set D_{pq} by

$$D_{pq} = \{(\omega,r) \in [\omega_1,\omega_2]\times\mathbf{R} \mid \pi_1 f^q_\omega(\theta_0,r)=\theta_0+p\} \tag{1.4}$$

and define the continuous function Δ:$\mathbf{R}^2 \rightarrow \mathbf{R}^2$ by

$$\Delta(\omega,r)=(\omega,\pi_2 f^q_\omega(\theta_0,r)-r) \tag{1.5}$$

We claim that (see Fig 2):

(i) D_{pq} contains a connected subset D^*_{pq} intersecting both $\{\omega_1\}\times\mathbf{R}$ and $\{\omega_2\}\times\mathbf{R}$

(ii) There exists $K>0$ such that if $\omega_1 < -K$ and $\omega_2 > K$, then $\Delta(D^*_{pq})$ intersects the set $\mathbf{R}\times\{0\}$ in at least one point $(\omega^*,0)$.

Then from the definition of D_{pq} and Δ, it follows that f_{ω^*} has a p/q periodic point on the line $\{\theta_0\}\times\mathbf{R}$.

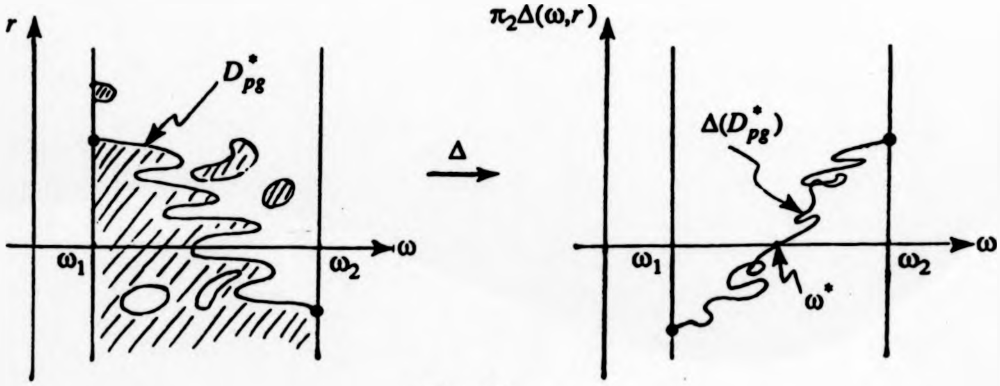

Figure 2

To verify (i), consider the open set O_1 defined by (1.6)

$$O_1 = \{(\omega,r) \in [\omega_1,\omega_2]\times\mathbf{R} \mid \pi_1 f^q_\omega(\theta_0,r) < \theta_0+p\} \tag{1.6}$$

The set O_1 is shown shaded in Fig 2. From the twist condition, $\bar{f}_\omega$ rotates points arbitrarily much in the positive direction near the upper end of $\tilde{A}$, and arbitrarily much in the negative direction near the lower end of $\tilde{A}$. Thus there exists $N \in \mathbf{R}$ such that $[\omega_1,\omega_2]\times[N,\infty)\subset O_1^c$ and $[\omega_1,\omega_2]\times(-\infty,-N]\subset O_1$. Let O_2 be the component of O_1 containing $[\omega_1,\omega_2]\times(-\infty,-N]$ and O_3 be the component of $([\omega_1,\omega_2]\times\mathbf{R})-Cl(O_2)$ containing $[\omega_1,\omega_2]\times[N,\infty)$. Then O_3 is simply connected, and therefore has a connected frontier F, so that $D_{pq}^* = D_{pq} \cap F$ has the required properties.

Let $M \in \mathbf{R}$ satisfy condition (5) for the map f. It follows that M satisfies condition (5) for all of the maps $f_\omega, \omega \in \mathbf{R}$. To verify (ii), we first show that if ω is sufficiently negative, and $(\omega,r)\in D_{pq}$, then $r \geq M$. Suppose by contradiction that $r < M$. Let $\pi_1 f_0(\theta_0,M)=L$. Then $\pi_1 f_0(\theta_0,r)<L$, and $\pi_1 f_\omega(\theta_0,r)<L+\omega$. Since $f_\omega(S^1\times\{M\})<S^1\times\{M\}$, we have by induction that $\pi_1 f_\omega^q(\theta_0,r)<q(L+\omega)$. Now choose ω so that $q(L+\omega)<p$, to contradict the hypothesis that $(\omega,r)\in D_{pq}$. Thus if ω_1 is sufficiently negative, it follows from the definition of M that $\pi_2 f_{\omega_1}^q(\theta_0,r)<r$, so that $\Delta(D_{pq}^*)$ intersects the lower half-plane. Similarly, if ω_2 is sufficiently positive, $\Delta(D_{pq}^*)$ intersects the upper half-plane. Then since Δ is continuous, and D_{pq}^* is connected, it follows that $\Delta(D_{pq}^*)$ is connected, and must intersect the line $\mathbf{R}\times\{0\}$.

It remains to deduce the existence of the Aubry-Mather sets. Following [18], define an orbit $\tilde{\Gamma}=\{(\theta_n,r_n)=\tilde{f}^n(\theta_0,r_0) \mid n \in \mathbf{Z}\}\subset\tilde{A}$ to be a *special* orbit if there exists a homeomorphism $g:S^1\to S^1$ such that $g^n(\theta_0)=\theta_n$ for all $n \in \mathbf{Z}$. $\tilde{\Gamma}$ has a well defined rotation number $\rho(\tilde{\Gamma})=\rho(g)$. A result of Hall [12] asserts that if $\tilde{f}$ is a twist map, and has a p/q periodic orbit, then $\tilde{f}$ has a special orbit of rotation number p/q. Hence from the first part of the theorem, given any $\sigma\in\mathbf{R}-\mathbf{Q}$, and a sequence of rationals $\{p_m/q_m \mid m \in \mathbf{Z}^+\}$ converging to σ, there exists a sequence of reals $\{\omega_m \mid m \in \mathbf{Z}^+\}$ such that $\tilde{f}_{\omega_m}$ has a special orbit $\tilde{\Gamma}_m=\{\tilde{f}_{\omega_m}^n(\theta_m,r_m) \mid n \in \mathbf{Z}\}$ of rotation number p_m/q_m. By condition (5) the sets $\tilde{\Gamma}_m, m \in \mathbf{Z}^+$ all lie in the compact subset $S^1\times[-M,M]$ of $S^1\times\mathbf{R}$, and the $\omega_m, m \in \mathbf{Z}^+$ are bounded. Thus, if necessary by going to a subsequence, the sequence $\{(\theta_m,r_m,\omega_m) \mid m \in \mathbf{Z}^+\}$ can be chosen to converge to a point (θ,r,ω) as $m\to+\infty$. If we can show that $\tilde{\Gamma}=\{\tilde{f}_\omega^n(\theta,r) \mid n \in \mathbf{Z}\}$ is a special orbit of rotation number σ, the proof of Theorem 1.1 will be complete, since the ω-limit set of the orbit $\bar{\Gamma}$ is an Aubry-Mather set of rotation number σ.

The proof is now almost identical to that in [18]. The function $\tilde{\psi}_m : \pi_1\tilde{\Gamma}_m \to [a,b]$, which assigns $\pi_2\tilde{f}^n_m(\theta_m,r_m)$ to $\pi_1\tilde{f}^n_m(\theta_m,r_m)$ is Lipshitz, with Lipshitz constant L independent of m. Extend $\tilde{\psi}_m$ by linear interpolation to a map $\tilde{\psi}'_m : S^1 \to [a,b]$, also with Lipshitz constant L. The space $\tilde{L}$ of Lipshitz maps from S^1 to $[a,b]$ with Lipshitz constant L is compact in the topology of uniform convergence. Thus, if necessary by going to a subsequence, we can find $\tilde{\psi}' \in \tilde{L}$ such that $\lim_{m\to\infty}\tilde{\psi}'_m = \tilde{\psi}'$. By the definition of convergence in $\tilde{L}$, we have $\tilde{\Gamma} \subset graph(\tilde{\psi}')$. The circle map $g = \pi_1 o \tilde{f}\, o(id \times \tilde{\psi}')$ is order preserving, so that $\tilde{\Gamma}$ is special. By construction, $\lim_{m\to\infty}\rho(\tilde{\Gamma}_m) = \sigma$, so that for every k, the order of the first k points for the rigid rotation by R_σ is the same as the rotation by $R_{\rho(\tilde{\Gamma}_m)}$, for sufficiently large m. Thus since $\rho(\bar{\Gamma})$ is well defined, it follows that $\rho(\bar{\Gamma}) = \sigma$. $\square$

2. Rotation Intervals.

In this section we explore the possibility of periodic orbits coexisting on invariant sets of a dissipative twist map f. Let $\Gamma \in \mathbf{H}$ be an invariant set for f. Then as in [5], we associate with Γ two real numbers, its *internal* and *external rotation numbers* $\rho_{int}(\Gamma)$ and $\rho_{ext}(\Gamma)$, defined as follows. Let $L = \{(\theta,r) \in \mathbf{R}^2 |\ \theta \in \mathbf{R}, r = r_0\}$ be a horizontal line lying above Γ. Denote by Γ' those points x in Γ for which the vertical line joining x to L contains no other points of Γ,and let π be the (bijective) vertical projection from L to Γ'. Then $\rho_{ext}(\Gamma)$ is defined to be $-\rho(g)$ where $\rho(g)$ is the rotation number of the map $g: L \to L$ defined by $g(\theta) = \pi^{-1} o f^{-1} o \pi(\theta)$. It can be shown that $f^{-1}(\Gamma') \subset \Gamma'$, so that g is well defined [5]. Moreover it can be seen that g is the lift of a circle map, and that g is monotonic, so that $\rho(g)$ is well defined [20]. Similarly, $\rho_{int}(\Gamma)$ is defined by considering pre-images of points on Γ vertically accessible from a horizontal line lying below Γ.

In [5] Birkhoff gave an example of a dissipative twist map with an invariant set $\Gamma \in \mathbf{H}$ for which $\rho_{int}(\Gamma) \neq \rho_{ext}(\Gamma)$. By analogy with area preserving twist maps and non-invertible circle maps

[16], it is natural to ask whether f has p/q periodic orbits for all $p/q \in [\rho_{int}(\Gamma), \rho_{ext}(\Gamma)]$. That this need not be the case is shown in [14]: there are examples of dissipative twist maps constructed as Poincaré return maps of differential equations which have the following orbit structure. Referring to Fig 3, in which ab is to be identified with cd, the map f has an invariant curve C with rotation number $\omega \neq 0$, and an unstable hyperbolic fixed point x, with stable and unstable manifolds $W^s(x), W^u(x)$. All other points are either attracted to another fixed point y (shaded region), or are attracted to the invariant curve C. Then the invariant set $\Gamma = C \cup W^u(x)$ has $[\rho_{int}(\Gamma), \rho_{ext}(\Gamma)] = [0,\omega]$, but there are no p/q periodic orbits for $p/q \in (0,\omega)$.

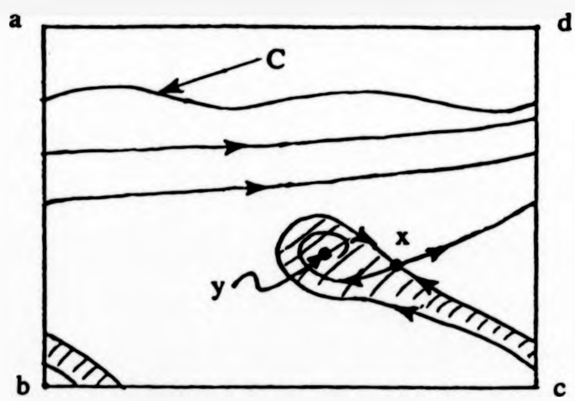

Figure 3

The above example motivates the following definition.

Definition. We say that an f-invariant set $\Gamma \in$ H has *the intersection property* if given any set $C \in$ H with $C \cap \Gamma \neq \varnothing$, then $C \cap f(C) \neq \varnothing$.

Remark. If $\Gamma \in$ H has a dense orbit, then it will necessarily satisfy the intersection property.

The next theorem states that if an invariant set has the intersection property then it possesses an "interval of rotation numbers".

Theorem 2.1 *Let f be a twist map, and Γ be an invariant set with the intersection property. Then f has a p/q orbit for all $p/q \in [\rho_{int}(\Gamma), \rho_{ext}(\Gamma)]$, and an Aubry-Mather set of rotation number ω for all $\omega \in [\rho_{int}(\Gamma), \rho_{ext}(\Gamma)]$*

The essential idea in the proof of Theorem 2.1 is to consider the sets C_{pq}, $p \in \mathbf{Z}$, $q \in \mathbf{Z}^+$, defined by (2.1)

$$C_{pq} = \{(\theta,r) \mid \pi_1 f^q(\theta,r) = \theta + p\} \tag{2.1}$$

We first prove a lemma which gives a criterion for the existence of periodic points in terms of the sets C_{pq}. Define P_{pq} to be the set of p/q periodic points of f.

Lemma 2.1 *Let f be a twist map, then for all $p \in \mathbf{Z}$, $q \in \mathbf{Z}^+$ we have $P_{pq} = C_{pq} \cap f(C_{pq})$.*

Proof. Evidently $P_{pq} \subset C_{pq} \cap f(C_{pq})$. To show the converse, take a point (θ_0, r_0) in $C_{pq} \cap f(C_{pq})$ and consider its orbit $\{(\theta_n, r_n) \mid n \in \mathbf{Z}\}$. By hypothesis $\theta_q = \theta_0 + p$ and $\theta_{q-1} = \theta_{-1} + p$, and it remains to show that $r_q = r_0$. Since f is a twist map, (θ_0, r_0) is given by the unique intersection of the bi-infinite vertical line through θ_0 with the image of the bi-infinite vertical line through θ_{-1}. But f commutes with the translation T^p, hence the same process applied to the bi-infinite vertical lines through θ_q and θ_{q-1} must yield $r_q = r_0$. Thus $C_{pq} \cap f(C_{pq}) \subset P_{pq}$ as required. □

Remark. After proving Lemma 2.1, it was pointed out to us that similar ideas are used in the "radially translated curve theorem" of Poincaré-Birkhoff [1]. In fact they derive $P_{pq} = C_{pq} \cap f^q(C_{pq})$, but this is only true for sufficiently small nonlinearity (dependent on q). In this sense Lemma 2.1 is a more powerful criterion for the existence of periodic orbits.

The next lemma indicates that Lemma 2.1 will be useful, by establishing a topological property of the sets C_{pq}.

Lemma 2.2 *Let f be a twist map and take $p \in \mathbf{Z}$, $q \in \mathbf{Z}^+$. Then the set C_{pq} is non-empty, and has a component C_{pq}^* in* $\mathbf{H}$.

Proof. Consider the open set $O_1 \subset \mathbf{R}^2$ defined by (2.2)

$$O_1 = \{(\theta,r) \mid \pi_1 f^q(\theta,r) < \theta + p\} \tag{2.2}$$

From the twist condition, $\bar{f}$ rotates points arbitrarily much in the positive direction near the upper end of $\bar{A}$, and arbitrarily much in the negative direction near the lower end of $\bar{A}$. Thus there exists $N \in \mathbf{R}$ such that $\mathbf{R} \times [N, \infty) \subset O_1^c$ and $\mathbf{R} \times (-\infty, -N] \subset O_1$. Let O_2 be the component of O_1 containing $\mathbb{R}\times(-\infty,-N]$ and O_3 be the component of $\mathbb{R}^2-Cl(O_2)$ containing $\mathbb{R}\times[N,\infty)$. Then O_3 is simply connected, and therefore has a connected frontier $C^*_{pq}\subset C_{pq}$. Since O_1 is invariant under the translation T, so is C^*_{pq}, and it follows that $C^*_{pq}\in$ H. □

Proof of Theorem 2.1 Let $p/q\in[\rho_{int}(\Gamma),\rho_{ext}(\Gamma)]$. We first show that $C^*_{pq}\not\subset\Gamma_{ext}$. From the definition of $\rho_{ext}(\Gamma)$, there exists a point $y=(\theta,r)\in\Gamma'$ satisfying (2.3).

$$\pi_1 f^{-q}(y)\le\theta_0-q\rho_{ext}(\Gamma) \tag{2.3}$$

Let L be the vertical line through y given by $L=\{(\theta,r)\in\mathbb{R}^2\mid\ \theta=\theta_0, r\ge r_0\}$, and consider its image under f^{-q}. If C^*_{pq} were a subset of Γ_{ext}, then $f^{-q}(L)$ would necessarily intersect it at some point $z=(\theta,r)$. Since f is a twist map, one deduces that $\theta<\pi_1 f^{-q}(y)$ [13]. Then using inequality (2.3), and $p/q\le\rho_{ext}(\Gamma)$, we derive $\pi_1 f^q(z)>\pi_1(z)+p$, which contradicts $z\in C^*_{pq}$. Hence $C^*_{pq}\not\subset\Gamma_{ext}$.

Similarly $C^*_{pq}\not\subset\Gamma_{int}$, and we obtain $C^*_{pq}\cap\Gamma\ne\emptyset$. By Lemma 2.2, $C^*_{pq}\in$ H, and since Γ is assumed to satisfy the intersection property, we have $C^*_{pq}\cap f(C^*_{pq})\ne\emptyset$. Thus by Lemma 2.1, we conclude that f has a p/q periodic point. As in Theorem 1.1, the existence of the Aubry-Mather sets follows directly from the results of Hall [12], and Katok [17]. □

3. Sets with the Intersection Property.

In this section we construct invariant sets with the intersection property for a dissipative twist map. We will also be interested in the attracting properties of these sets; to be precise we make some definitions.

Definition. We say that a closed f - invariant set Γ is *weakly attracting* if for any neighborhood N of Γ, there is a neighborhood M of Γ contained in N such that $f(M)\subset M$. If in addition M may be chosen so that $\Gamma=\bigcap_{n=0}^{\infty} f^n(M)$, we say that Γ is *attracting*.

Let B be a trapping region for a dissipative twist map f. Then the invariant set $\Gamma(B)=\bigcap_{n=0}^{\infty} f^n(B)$, is attracting and is a member of H.

Definition. The set $\beta(f) = cl(\Gamma_{int}(B)) \cap cl(\Gamma_{ext}(B))$ is called the *Birkhoff attractor* of f.

It may be verified that $\beta(f)$ is well defined (i.e. independent of the choice of B), is a member of H, and satisfies the intersection property [8]. However $\beta(f)$ need not be even weakly attracting [9].

We propose to focus attention instead on another set which turns out to have some nicer properties. Define $H^{\pm}$ by $H^{+} = \{C \in H \mid f(C) \subset C_{ext}\}$ and $H^{-} = \{C \in H \mid f(C) \subset C_{int}\}$. Since f is dissipative, given $C^{\pm} \in H^{\pm}$, we have $C^{+} < C^{-}$, so that $C^{\pm}$ bound an annular region $R(C^{-}, C^{+})$. The set $\alpha(f)$ that we are interested in is defined to be the intersection over all annular trapping regions, i.e. $\alpha(f) = \bigcap R(C^{-}, C^{+})$ where the intersection runs over all $C^{\pm}$ in $H^{\pm}$.

Theorem 3.1 *The set* $\alpha(f)$ *is a member of* H, *satisfies the intersection property, and is weakly attracting.*

Proof. To exploit compactness properties, we work in $\tilde{A}$, but for ease of notation, drop the superscript "~". We verify the required three properties for $\alpha(f)$.

(i) $\alpha(f) \in H$. By construction $\alpha(f)$ is a closed and bounded set, thus is compact. The complement of $\alpha(f)$ is given by $R_{int} \cup R_{ext}$ where $R_{int} = \bigcup_{C \in H^{+}} C_{int}$ and $R_{ext} = \bigcup_{C \in H^{-}} C_{ext}$. Since f is dissipative, we deduce that $R_{int} \cap R_{ext} = \emptyset$, and that $\alpha(f)$ disconnects $\tilde{A}$ into these two components only. It remains to show that $\alpha(f)$ is connected. From its definition, $\alpha(f)$ is seen to be equivalent to a nested intersection of compact connected sets, and is therefore connected [15].

(ii) $\alpha(f)$ satisfies the intersection property. Let $C \in H$ be such that $\alpha(f) \cap C \neq \emptyset$. If we can show that C is not a member of $H^{+} \cup H^{-}$, then it follows that $C \cap f(C) \neq \emptyset$ as required. Suppose by contradiction that $C \in H^{-}$, and take a point $x \in \alpha(f) \cap C$. But then x would lie above $f(C)$ which itself is a member of H^{-}. This contradicts the hypothesis that $x \in \alpha(f)$. Similarly, we cannot have $C \in H^{+}$.

(iii) $\alpha(f)$ is weakly attracting. Take any neighborhood N of $\alpha(f)$. We must show that there exists a neighborhood M of $\alpha(f)$ such that $M \subset N$ and $f(M) \subset interior(M)$. Assume, by gong if necessary to a subset of N, that N is an annular neighborhood of $\alpha(f)$. We construct M as follows. Take a point x in $\partial^+ N$. Then x is not in $\alpha(f)$, so there exists $C(x) \in$ H such that $x \in C(x)_{ext}$. Let $\partial^+ N(x)$ be a non-empty open connected subset of $\partial^+ N$ containing x such that $C(x) \cap \partial^+ N(x) = \varnothing$. Since $\partial^+ N$ is compact, $\partial^+ N = \bigcup_{i=1}^{n} N(x_i)$ for some n and $x_1, \ldots, x_n$ in $\partial^+ N$. Consider the set $S = \bigcap_{i=1}^{n} C(x_i)_{int}$. By construction $S \cap \partial^+ N = \varnothing$, and $f(S) \subset S$. Then $\partial^+ M$ is taken to be the frontier of S. Defining $\partial^- M$ similarly, the set M is taken to be the open set bounded by $\partial^{\pm} M$. By construction $M \subset N$, and $f(M) \subset interior(M)$. □

References.

[1] V.S.Arnold & A.Avez. *Ergodic problems of Classical Mechanics.* pp.88-90. Pub. Benjamin: New York 1968.

[2] D.G.Aronson, M.A.Chory, G.R.Hall & R.P.McGehee. *Bifucations from an invariant circle for two-parameter families of maps of the plane: A computer assisted study.* Commun. Math. Phys., 83 (1982), 303-354.

[3] S.Aubry & P.L.Le Daeron. *The Discrete Frenkel-Kontorova Model and its extensions.* Physica 8D (1983), 381-422.

[4] G.D.Birkhoff. *On the periodic motions of dynamical systems.* Acta Mathematica, 50 , 359-379. Reprinted in *Collected Mathematical Papers,* Vol.II, American Math. Soc., New York 1950.

[5] G.D.Birkhoff. *Sur Quelques Courbes Fermées Remarkables.* Bull. Soc. Math. de France, 60 (1932), 1-26. Reprinted in *Collected Mathematical Papers,* Vol.II, American Math. Soc., New York 1950.

[6] T.Bohr. *A bound for the existence of invariant circles in a class of two-dimensional dissipative maps.* Preprint, Cornell University (1984).

[7] T.Bohr, P.Bak & M.H.Jensen, *Transition to chaos by interaction of resonances in dissipative systems.* Preprint, Cornell University (1984).

[8] P.Le Calvez. *Existence d'orbites quasi-periodiques dans les attracteurs de Birkhoff.* Preprint, Université Paris, Orsay (1985).

[9] P.Le Calvez. Private communication.

[10] A.Chenciner. *Orbites périodiques et ensembles de Cantor invariant d' Aubry-Mather au voisinage d'une bifurcation de Hopf dégénérée de difféomorphismes de* $\mathbf{R}^2$. C.R.A.S. t 297 , Série I (1983), 465-467.

[11] A.Chenciner. *Sur un énoncé dissipative du théorèm géometrique de Poincaré-Birkhoff.* C.R.A.S. t 294 , Série I (1982), 243-245.

[12] G.R.Hall. *A topological version of a theorem of Mather on twist maps.* Preprint, University of Wisconsin, Madison (1983).

[13] M.Herman. *Sur les courbes invariantes par les difféomorphismes de l'anneau.* Vol.I, Ch.1. Astérisque, 103-104 (1983).

[14] K.Hockett & P.Holmes. *Josephson's junction, annulus maps, Birkhoff attractors, horseshoes and rotation sets.* Preprint, Cornell University (1985).

[15] J.G.Hocking & G.S.Young, *Topology.* Addison-Wesley: Reading 1961.

[16] R.Ito. *Rotation sets are closed.* Math. Proc. Comb. Phil. Soc. 89 (1981), 107-111.

[17] A.Katok. *Some remarks on Birkhoff and Mather twist map theorems.* Ergod. Th. and Dynam. Sys., 2 (1982), 185-194.

[18] A.Katok. *Periodic and quasiperiodic orbits for twist maps.* Proceedings, Sitges 1982 L.Garrido (ed.), Spriger-Verlag, Berlin.

[19] J.N.Mather. *Existence of quasiperiodic orbits for twist homomorphisms.* Topology, 21 (1982), 457-467.

[20] Z.Nitecki. *Differentiable dynamics.* M.I.T. Press: (1971).

[21] S.Ostlund, D.Rand, J.Sethna & E.Siggia, *Universal properties of the transition from quasiperiodicity to chaos in dissipative systems.* Physica 8D (1983), 303-342.