

Markov Tree Model of Transport in Area Preserving Maps

James D. Meiss* and Edward Ott#

Inst. Theor. Physics

University of California

Santa Barbara, Ca. 93106

Abstract: Transport in an area preserving map with a mixture of regular and chaotic regions is described in terms of the flux through invariant cantor sets called cantori. A model retaining a discrete set of cantori approaching a boundary circle gives the Markov chain description of Hanson, Cary and Meiss. The inclusion of cantori surrounding island chains, and islands about islands, etc. gives a Markov tree model with a slower decay rate. The survival probability distribution is shown to decay asymptotically as a power law. The decay exponent agrees reasonably well with the computations of Karney and of Chirikov and Shepelyanski.

*Permanent address: Institute for Fusion Studies, University of Texas, Austin, TX 78712

#Permanent address: Dept. of Electrical Engineering and Dept. of Physics, University of Maryland, College Park, MD 20742

§1. Introduction

Transport as a result of chaotic motion in Hamiltonian systems has been the subject of much recent interest¹⁻⁵. This interest is stimulated both by fundamental implications (such as those for statistical mechanics⁶), as well as practical applications (such as particle heating and confinement in fusion plasma devices⁷). In this paper we discuss the transport properties of area-preserving maps, which can be thought of as resulting from a two-degree-of-freedom Hamiltonian system by surface-of-section.

Typically the phase space of such a map has both regular and chaotic trajectories. Chaotic motion is characterized by exponential instability that precludes any description based on individual orbits; one can at best hope for a statistical description of families of orbits. Yet because of the mixture of regular and chaotic motion, the simplest statistical approximations are bound to fail. Indeed, recent numerical experiments by Karney⁸, and Chirikov and Shepelyanski⁹ show that motion in the chaotic region exhibits long correlations whenever the phase space has a regular component.

Chirikov and Shepelyanski study a map for which the parameters are chosen so that the chaotic region is bounded by a KAM surface which has golden mean rotation number. They compute what we will call the survival probability, $F(t)$, which is the probability that a particle initially "near" the KAM surface, will still be "near" that surface at time t (up to normalization this is the Poincaré recurrence distribution of Ref. [9]). They find

$$F(t) \approx t^{-z} \quad (1)$$

where $z \approx 1.34$, for $1 \ll t \leq 10^5$. Karney studies the quadratic map, initializing particles in the chaotic region outside the stable island for this system. Here long correlations are induced by the outermost KAM surface surrounding the island, which does not have the golden rotation number. A roughly power-law decay was again obtained (albeit with additional substructure) with a similar exponent $z \approx 1.45$ for $1 \ll t \leq 10^8$. The explanation of this slow (i.e. algebraic rather than exponential) decay is the subject of this paper.

One model¹⁰ for explaining this behavior is that of a continuous diffusion coefficient that approaches zero as $(y-y_0)^\delta$ where y_0 represents the position of the boundary KAM surface. Algebraic decay results from the small diffusion in the neighborhood of the KAM surface. In contrast to this smoothly varying

diffusion coefficient, numerical experiments show that transport tends to occur in steps as the particle suddenly crosses certain leaky barriers¹ (called cantori) across which the transport is particularly slow. The quasi-discrete nature of transport is the basis for the model of Hanson et al¹¹, which was also discussed by Greene¹². In this model, knowledge of the flux across individual cantori^{1,14} is combined with an assumption of loss of memory between crossing times to give a one-dimensional Markov chain. For the case of a critical golden circle the scaling results of the renormalization theory¹³ imply that the transition probabilities of the Markov chain are self-similar.

The solution of the self-similar chain gives a power law decay, in agreement with numerical experiments; however, the exponent obtained ($z \approx 3.05$) is too large by about a factor of two when compared with the numerically obtained values.

In this paper we present an improved model, which was suggested in Ref. 1, and examine its predictions. Specifically, note that between barrier cantori there are island chains, and that particles which come near the KAM surfaces surrounding one of these islands can get stuck there for a long time as well. Furthermore, for every island, there are islands surrounding it. Clearly this situation cannot be described by a purely one-dimensional chain model. To take the essential features into account, we discuss in §3 and §4, a model based on a self-similar random walk on a tree.¹ Roughly speaking, branching in the tree model corresponds to a particle's "decision" whether to cross the next cantorus barrier (as in the one-dimensional case) or to become trapped around an island situated between the two cantori. The solution of the tree model is given in §4, while further discussion and conclusions are given in §5. The obtained exponent agrees reasonably well with the value found in the numerical experiments.

§ 2. Properties of Area Preserving Maps

In this section we briefly discuss some properties of area-preserving maps, without taking care to be especially rigorous. It is convenient to think of maps on a cylindrical phase space, the vertical coordinate y represents momentum and the angle is x with period 1. We denote such maps $T:(x,y) \rightarrow (x',y')$. T is assumed to be a twist map: it causes rotation at different rates at different vertical positions on the cylinder. More formally the twist condition is $\partial x'/\partial y|_x > 0$. An orbit of the map is a sequence $\{x_t, y_t\}$; t represents time and takes integer values.

A useful example, which apparently has the properties of a typical twist map, is the standard map,

$$\begin{aligned} y_{t+1} &= y_t - k \sin(2\pi x_t) / 2\pi, \\ x_{t+1} &= x_t + y_{t+1} \pmod{1}, \end{aligned} \quad (2)$$

with the single parameter k . The standard map will be used for illustration, but everything in the model holds for any one parameter family of twist maps.

§2.1 Rotation Numbers, Diffusion and Correlation

As was already known by Poincaré, the phase space of a typical area preserving map is filled with a complicated mixture of regular and irregular (or chaotic) orbits. Regular orbits are invariant curves on which the motion has a definite rotation frequency. There are two types of regular orbits on the cylinder, those which encircle the cylinder, and those which do not. The encircling curves are the survivors of the invariant circles of the integrable twist map at $k=0$. We will call these "orbits of class zero." The rotation frequency of a class zero orbit is defined by

$$\nu \equiv \lim_{t \rightarrow \infty} (x_t^* - x_0) / t,$$

where x^* is a "lift" of the x -orbit (i.e., do not take the mod 1 in Eq. (1)). When k is small enough, the KAM theorem implies that the class zero regular orbits occupy a finite measure on the cylinder. Other, "nonzero class," regular orbits are found in the neighborhood of elliptic periodic orbits. They surround the periodic orbit, forming "island chains." A second version of the KAM theorem shows that the measure of the regular orbits near an elliptic orbit, is finite. The irregular motion is "generated" by transversal intersections of the stable and unstable manifolds of hyperbolic orbits. Such intersections guarantee that there are no global, smooth, invariant manifolds near the hyperbolic orbits. As k increases the apparent area occupied by the irregular orbits increases, though it is not known whether this area is actually nonzero.

There have been many numerical studies of irregular motion, and attempts to describe it as "random" in some sense. The most widely used construct is a

diffusion coefficient.² For the standard map a formal series for the y -diffusion coefficient can be obtained using an ensemble average over phase space.^{3,4} This coefficient can be written as a sum of correlation functions:

$$D = \lim_{t \rightarrow \infty} \langle (y_t - y_0)^2 \rangle / 2t = C_0/2 + \sum_{j=1}^{\infty} C_j,$$

The last equality is valid only if the force correlation functions,

$$C_j = k^2 \langle \sin(2\pi x_{t+j}) \sin(2\pi x_t) \rangle / 4\pi^2$$

and

fall off fast enough for convergence. The series for C_j and D agree well with numerical studies for $k \gg 1$ providing there are no apparent islands in the phase space. By contrast, islands cause an enhanced correlation, leading to a long time tail, and bring into doubt even the convergence of the sum for the the diffusion coefficient!

§ 2.2 Flux, Cantori, and Boundary Circles

To develop a theory for situations with both regular and chaotic motions, the notion of uniform diffusion is abandoned in favor of a more local description. Here we describe irregular motion in terms of the flux of orbits through curves in phase space.^{1,14} In particular the enhanced correlation due to a regular region may be attributed to the existence of remnants of invariant circles which have a small flux through them. These remnants are called cantori:¹⁵ they are cantor sets with a given irrational rotation number. Mather, Aubry, and Katok¹⁶ prove the existence of an invariant set, which is either a circle or a cantorus, for every irrational frequency ν (providing the map satisfies the twist condition about the periodic orbit under consideration and the frequency is in the range of twist). Cantori can be visualized as circles with an infinity of gaps caused by the nearby island chains attempting to overlap.

Motion in an irregular region can be separated into motion along the cantori and leakage through their gaps. The leakage can be computed by the flux of orbits, defined as the area which leaks one way through the cantorus on one iteration of the map¹. One way to determine it, is to draw some curve

connecting all the gaps of the cantor. Iterating this curve once yields another curve which also fills in the gaps of the cantor. The upward flux, is defined as the area above the first and below the second curve. By area preservation there is an equal downward flux: the net flux is zero. Another construction of two curves involves using the stable and unstable manifolds of the cantor.¹ In this case they differ only in one gap of the cantor; and form a figure-of-eight structure called the turnstile;¹ however, it is not hard to see that the value of the flux is independent of the choice of curve, providing it fills the gaps and does not intersect itself. In fact, using the action variational principle for the map, one can show¹ that the flux is equal to $\Delta W = W_h - W_c$, where W_c is the action of the cantor and W_h is the action of the orbit homoclinic to the cantor (e.g. the orbit which asymptotes to the cantor in both directions of time). This shows that ΔW depends only on the cantor and its homoclinic orbit. The only caveat in this result is that ΔW is the algebraic area of the turnstile. This equals the geometric area crossing the cantor if the turnstile has the figure-of-eight topology¹. In every example we have tried this is true, though it is possible that some exceptions could be constructed. In conclusion, Mather proves the existence of the cantor, and its homoclinic orbit, as well as the finiteness of ΔW . The result of Ref. 1 is that ΔW has the interpretation of flux.

While ΔW exactly gives the flux crossing a cantor on each iteration of the map, the subsequent motion of this area is very complicated: after several iterations the curve defined above stretches and wiggles along unstable manifolds, and can recross the cantor. To develop a model for long time transport, we need to understand the subsequent motion of these orbits. One aspect of this problem is that there are many cantori in the irregular region: in fact one for every irrational frequency.

It is clear that the most important cantori will be those with small flux. Mather has shown¹⁷ that ΔW is identically zero for an invariant circle (which is clear from our interpretation of ΔW as a flux) and is a continuous function on the irrationals; therefore, cantori very close to invariant circles have fluxes which can be arbitrarily small. This implies that orbits in the neighborhood of such invariant circles will be stuck in this region for long periods of time. The invariant circles which form the boundary of an irregular region are called "boundary circles."

Boundary circles have some special properties.¹⁸ It is believed that they are critical invariant curves: they are at the threshold of destruction, in the

sense that there is a neighborhood of the parameter k for which an invariant circle with the given winding number does not exist. The argument for this is as follows. By the KAM theorem we know that the interval of the parameter for which an invariant circle of sufficiently irrational frequency is smooth, is open. Furthermore Birkoff's theorem¹⁹ implies that the interval for which it doesn't exist is also open. It is conjectured that the boundary of these two intervals is the same point: an invariant curve loses its smoothness just at the threshold of its destruction. It is known that smooth circles have other invariant circles in every neighborhood (this is obtained by the renormalization theory combined with the KAM theorem²⁰). Circles that have a neighborhood containing no other circles, such as boundary circles which have none on one side, must therefore be critical.

In general the boundary of the irregular region consists of an infinite number of boundary circles. These can be separated into classes. The zeroth class are those that encircle the cylinder. Consider for definiteness a region bounded from above by one such circle, of frequency ν_0 . Typically there will be elliptic, class zero, periodic orbits in the stochastic region below ν_0 , e.g. with rotation numbers $p_0/q_0 < \nu_0$. Encircling these orbits are the island chains, these we call class one orbits. The rotation frequency of a class one orbit is defined as the average number of rotations about the elliptic fixed point, per q_0 iterations of the map:

$$\nu_1 \equiv \lim_{t \rightarrow \infty} \{\theta(T^{q_0 t}(x,y)) - \theta(x,y)\} / 2\pi t,$$

where θ is the angle of rotation about one of the points on the p_0/q_0 orbit. Invariant circles around p_0/q_0 actually are q_0 separate circles which are iterates of each other. The outermost class one invariant circle is the boundary circle of the island; call its irrational rotation number ν_1 . Outside of this boundary circle there are elliptic class one periodic orbits with frequencies $p_1/q_1 < \nu_1$. The outermost curves of the island chains surrounding these (the islands about islands) form boundary circles of class two, and so forth.

There are other orbits which are not included in the class hierarchy, which however, can be stable and form island chains. An example is periodic orbits born by tangent bifurcation. Formally, these islands can be included in the model developed below if their properties are known. In practice we neglect them.

§ 3. The Tree Model

§ 3.1 Tree States

Figure 1 shows a schematic illustration of the situation just described. The uppermost circle represents a class zero boundary circle. Above the class zero circle is a region, shown cross hatched, not accessible from below. The stickiness of the ν_0 boundary circle can be attributed, in part, to the cantori with $\nu < \nu_0$. Below the boundary circle, all the irrational orbits are cantori, since the boundary is by definition the lowest frequency non-destroyed circle. The most important cantori in any frequency interval will be those with the minimum value of ΔW in that interval; we call these the "minimizing" cantori. To the extent that ΔW , as shown in Fig. 2, has sharp local minima, it is a reasonable approximation to neglect all cantori but the locally minimizing ones. There are an infinity of such minimizing cantori as ν_0 is approached: they correspond to rational approximations of the continued fraction for ν_0 (see §4). We refer to this sequence of cantori as class zero "levels," with level $\rightarrow \infty$ indicating approach to the boundary circle.

Between any two adjacent minimizing class zero cantori there are, in general, many class zero periodic orbits, some of which are elliptic in character. The cross-hatched regions in Fig. 1 represent areas (or "islands") that are completely enclosed by boundary circles encircling elliptic periodic orbits. For most of our subsequent considerations we shall assume that between any two adjacent flux minimizing class zero cantori only one such island chain is "significant" and neglect all others. The most significant chain will be that with the largest area between the adjacent minimizing cantori. We emphasize, however, that this assumption is merely for convenience and simplicity of presentation; our analysis is easily extended to an arbitrary number of significant chains (see §5.3).

A class one boundary circle, ν_1 , forms the outer boundary of the island chain. Since the rotation frequency about an island goes to zero at the island separatrix (which, however, is broken and contains a turnstile¹), there are an infinity of class one cantori outside this boundary circle, corresponding to all irrational numbers in the interval $0 < \nu < \nu_1$. In accord with our modeling we retain only a discrete set of these cantori: those which are local minima of the flux. These cantori form a sequence of levels at class one.

Between each pair of successive class one cantori there is again one most significant, class one periodic orbit whose situation (although on a smaller scale) can be regarded as analagous to those of lower class: it has a sequence of class two cantori approaching its boundary circle, and so on ad infinitum.

Since the leakage through the minimizing cantori is slow, it is useful to distinguish the areas they bound. We call such areas "states," and we now describe a convenient method of labeling these states. First choose some reference class zero state and call it the null state, or " \emptyset ." In Fig. 1 the lowest visible state is labeled \emptyset . Other states are identified by a symbol sequence specifying their location relative to \emptyset (States "outside" the lower cantorus bounding \emptyset do not enter our consideration; we suppose the particle is lost forever when it traverses this cantorus). In particular, a state, S, is specified by a sequence of two symbols, say "1"s and "2"s:

$$S = \{\sigma_1, \sigma_2, \dots, \sigma_N\}. \quad (3)$$

Here $\sigma_1=1$ if the first step in the direct path from \emptyset to S increases the level by one; e.g. it is a move toward ν_0 by crossing the upper class zero cantorus bounding \emptyset . If instead the step increases the class by one; e.g. it is a move toward the class one boundary circle ν_1 by crossing its outermost, class one cantorus (see Fig. 1), then $\sigma_1=2$. To determine σ_2 , we now view the current state, $\{\sigma_1\}$, just as we viewed state \emptyset before: if $\sigma_1=1$ then another increase in level gives $\sigma_2=1$ as well, while crossing the level zero, class two cantorus around the class one island chain in state $\{1\}$ would give $\sigma_2=2$. Thus the states can be viewed as being located on a binary tree, or "Bethe lattice," as shown in Fig. 3.

Some additional notation will be useful below. Let DS be the symbol sequence obtained by deleting the last entry of S. Hence, if S has N components, as in Eq. (3), then

$$DS = \{\sigma_1, \sigma_2, \dots, \sigma_{N-1}\}. \quad (4)$$

Thus DS deotes the state just "outside" S in Fig. 1, or just "below" S on the tree. Each state has a unique predecessor, DS, but two daughters, which are the two states just above S on the tree (Fig. 3); we denote the daughters of S by

$$S1 = \{\sigma_1, \sigma_2, \dots, \sigma_{N,1}\}; \quad S2 = \{\sigma_1, \sigma_2, \dots, \sigma_{N,2}\}. \quad (5)$$

In terms of Fig. 1, the numerical experiments described in §1 are performed by initializing a large number of points in $\{1\}$ and then asking what fraction $F_1(t)$ of these has never entered \emptyset after t iterates of the map. If a particle is regarded as being "killed" when it enters \emptyset , $F_1(t)$ represents the fraction that are "alive" at time t . Clearly to determine information pertaining to F_1 we need information on how often particles make transitions between adjacent states on the tree. This is the subject of the next subsection.

§3.2 Transition Probabilities

Because the motion in a given state is chaotic between transitions, and because the transitions are assumed rare (i.e., the minimum flux cantori are not too leaky), we can characterize particle motions by a transition probability $p_{S \rightarrow S'}$, which is the "probability" that a particle in S will be in S' after one iterate of the map. We assume here that exponential divergence of orbits (at the rate given by a local Lyapunov exponent) causes memory loss on a time scale short compared to the transition time (this has not yet been verified in detail). Complete loss of memory implies that the motion can be described by Markov transition probabilities: they depend only on the current state, and not on the complete motion history. Furthermore we assume that the transition probabilities per step are small: in particular the total probability for leaving state S

$$p_S \equiv \sum_{S' \neq S} p_{S \rightarrow S'} \ll 1. \quad (6)$$

This implies that the Markov system can be treated using a continuous time description.

Consider two adjacent states on the tree, and denote them S and DS . Let A_S and A_{DS} be the phase space areas of the accessible chaotic regions enclosed by these states (we assume this is not zero!). Now iterate these areas once under the map. As mentioned in §2 some of A_S will be mapped into DS , and vice versa;

these two areas will be of equal size. We denote this transferred area by $\Delta W_{DS,S}$. The transition probabilities are

$$\begin{aligned} P_{S \rightarrow DS} &= \Delta W_{DS,S} / A_S , \\ P_{DS \rightarrow S} &= \Delta W_{DS,S} / A_{DS} . \end{aligned} \quad (7)$$

As we will see in §4, these probabilities obey approximate scaling relations on the tree due to the self-similarity of the dynamical system in the neighborhood of critical invariant curves. We denote the scaling factors by

$$a_1 \equiv A_{S1} / A_S \quad a_2 \equiv A_{S2} / A_S , \quad (8a)$$

$$w_1 \equiv \Delta W_{S,S1} / \Delta W_{DS,S} \quad w_2 \equiv \Delta W_{S,S2} / \Delta W_{DS,S} , \quad (8b)$$

$$\varepsilon_1 \equiv w_1 / a_1 \quad \varepsilon_2 \equiv w_2 / a_2 , \quad (8c)$$

where the a 's, w 's and ε 's are scaling constants to be obtained in §4. The coefficient a_1 represents the ratio of areas of successive level states with the same class, while a_2 represents the ratio of neighboring states of successive classes. To determine the flux and area at S , let

$$\begin{aligned} \lambda[S] &= (\text{the number of 1's in the symbol sequence of } S), \\ \rho[S] &= (\text{the number of 2's in the symbol sequence of } S). \end{aligned}$$

Thus $\rho[S]$ is the "class" of state S . Then from (8) we obtain

$$\begin{aligned} \Delta W_{DS,S} &= W_0 w_1^{\lambda[S]} w_2^{\rho[S]} , \\ A_S &= A_0 a_1^{\lambda[S]} a_2^{\rho[S]} , \end{aligned}$$

where W_0 and A_0 are constants that depend on the choice of the reference state, \emptyset . Using (7) the transition probabilities are

$$\begin{aligned} P_{S \rightarrow DS} &= P_0 \varepsilon_1^{\lambda[S]} \varepsilon_2^{\rho[S]} , \\ P_{S \rightarrow S1} &= w_1 P_{S \rightarrow DS} , \\ P_{S \rightarrow S2} &= w_2 P_{S \rightarrow DS} , \end{aligned} \quad (9)$$

where $p_0 = W_0/A_0$. Equations (9) and Fig. 3 define a self-similar random walk on a tree and constitute the basic model of this paper.

In previous models of transport, such as that of Hanson et al.,¹¹ islands were not taken into account. In our notation this means that states with any "2" in their symbol sequence were neglected. Thus in this case all states are of the form $S=\{1,1,1,\dots,1\}$ and the tree degenerates into a chain. The chain transition probabilities are then given by Eq. (9) with $w_2=p=0$.

If, on the other hand, M class c islands between two adjacent class c minimizing cantori are taken into account (see §5.3), then states are still defined by symbol sequences as in Eq.(3), but the σ_i would range through the $M+1$ values, $1,2,\dots,M+1$.

§ 4. Self-Similarity and Boundary Circles

In this section we describe how renormalization theory can be used to obtain the scaling coefficients in Eq. (9). We will just require a few simple properties of this rescaling transformation, and refer to Refs. 13, 18, 21, and 22 for more convincing and detailed discussions.

§ 4.1 Level Scaling

To understand the structure of the phase space of a map in the neighborhood of an invariant circle, consider the continued fraction expansion for its frequency:

$$\nu = n_0 + 1/(n_1 + 1/(n_2 + \dots \equiv [n_0, n_1, n_2, \dots]). \quad (10)$$

where the partial quotients, n_i , are positive integers. Recall that for an irrational number, the continued fraction has an infinite set of partial quotients. Truncating the expansion at some finite order, m , gives a rational approximation to ν

$$p_m/q_m = [n_0, n_1, \dots, n_m]. \quad (11)$$

These "convergents" are alternately larger than and smaller than ν . Since we are interested in approaching a boundary circle from the "chaotic" side (say the low frequency side), we consider only every other convergent, and refer to these as the levels of approximation ($l=m/2$).

The most remarkable recent advance in dynamics is the discovery of renormalization properties of dynamical systems. A renormalization transformation is a rescaling of phase space and time. For the case of area preserving maps in the neighborhood of critical circles the l^{th} application of this transformation is defined so that points on the $(l+1)^{\text{st}}$ (elliptic and hyperbolic) convergent orbits are mapped onto the l^{th} , and one iteration of the rescaled map corresponds to n_{2l} iterations of the original.

For special frequencies, in particular those whose partial quotients are periodic for large i , the rescaling also approaches a periodic limit: the map in the neighborhood of such a critical circle is self-similar. The most important case is that of the so-called "noble" frequencies; these are defined as numbers with continued fraction representations ending in an infinite sequence of ones,

$$\nu_{\text{noble}} = [n_0, n_1, \dots, n_j, 1, 1, 1, 1, \dots].$$

In this case the scaling theory implies the ratio of the area of one island of the q_{2l} component chain at level $l+1$, to one at level l (for $2l \gg j$) is given by¹³

$$w_1 \approx 0.053112. \quad (12)$$

Nobles are important because they appear to form the most robust invariant circles in phase space.

The scaling of islands is also reflected in the scaling of the cantori between them. In particular, between the level l and $l+1$ island chains, there are an infinity of cantori of the same class. Of those, we identify the one with the smallest value of ΔW as the minimizing cantorus which separates the two levels. Near a critical noble circle this cantorus has the frequency

$$(p_{2l} + \delta p_{2l+2}) / (q_{2l} + \delta q_{2l+2})$$

where γ is the golden mean, $\gamma = (1 + \sqrt{5})/2$. Thus, a "state" is the connected chaotic region surrounding the island at one level between the minimizing cantori on each side. Since under the rescaling transformation the minimizing cantorus separating levels l and $l+1$, is mapped onto the one separating $l-1$ and l , the flux through the minimizing cantori (which is an area) scales with the same coefficient, w_1 .

Another important scaling is the rate at which denominators of the successive convergent frequencies approach infinity. For a noble frequency this is

$$\lim_{l \rightarrow \infty} (q_{2l+2}/q_{2l}) \equiv \varepsilon_1^{-1} = \gamma^2 \approx 2.6180 \quad (13)$$

Thus the number of islands in the island chain corresponding to level l increases as γ^{2l} . For our purposes, the important implication of this is that the total area in the level l state scales as the number of islands in the chain, times the area of one island or:

$$a_1 = w_1/\varepsilon_1 \approx 0.13905 \quad (14)$$

The self-similarity is not exact in general because boundary circles are not typically noble. In fact, boundary circle rotation numbers typically take the form¹⁸

$$[\dots, 1, n_{2l-1}, 1, n_{2l+1}, 1, \dots] \quad (15)$$

in which the odd partial quotients are not always one but vary between one and five in an apparently random way (90% of the time it is a 1, 2, or 3 and 10% of the time there is a 4 or 5; furthermore, 10% of the even quotients are 2's). The convergents on the chaotic side of the boundary circle are those which (most of the time) end in a one, and hence "look noble" in some sense. On the other side of the boundary circle there may be other circles arbitrarily close.

Boundary circle frequencies can be determined using the "residue" criterion of Greene¹³; it is a refinement of Chirikov's overlap criterion. To apply this one computes the residue of various periodic orbits encircling the island. According to the criterion, if the residue of a periodic orbit is significantly larger than $1/4$ it is typically outside the island boundary circle, while if it is significantly

smaller than $1/4$ it is inside. (The residue for convergent orbits to the boundary circle approaches the critical value 0.250089 in the noble case). Using the Farey tree procedure to construct rationals one can obtain the continued fraction representation for the boundary frequency.¹⁸

Greene, Stark and MacKay¹⁸ obtain a generalized notion of scaling for boundary circles

$$\Delta W_{2l} \approx C(q_{2l})^{-3.05} .$$

This can be rewritten to obtain a relation between ε_1 , defined as a ratio of q 's as in (13), and w_1 , defined as a ratio of ΔW 's as in (8b),

$$w_1 = \varepsilon_1^{3.05} . \tag{16}$$

It is surprising that even though the partial quotients vary between 1 and 5, the scaling exponent in (16) holds to within a few percent. Note that the generalized scaling reduces to the noble case when $\varepsilon_1 = \gamma^{-2}$. In general, however, ε_1 and consequently w_1 and a_1 will fluctuate with level as the n_i vary. Roughly speaking, ε_1 fluctuates about its mean by about 30%. To actually treat the general boundary circle case, the $p_{S \rightarrow S'}$ should be given with a probability distribution, corresponding to that of the partial quotients in the continued fraction expansion. We make the approximation in this paper that ε_1 is constant and leave the treatment of these fluctuations to a future paper.

§ 4.2 Class Scaling

The largest island chain between two levels of class c minimizing cantori is that chain with frequency p_{2l}/q_{2l} , corresponding to the convergent of the boundary circle. The boundary circle of this island chain will also have, in general, a frequency with a continued fraction of the form given by Eq. (15). Noting that islands are typically destroyed by period-doubling bifurcation when their central rotation frequency reaches $1/2$, the boundary circle frequency will be considerably less than $1/2$, e.g., the second partial quotient will larger than one. Each orbit corresponding to a convergent of the continued fraction forms a higher level island chain encircling the class c island. The outermost convergent

has the frequency $[0,n,1] \equiv 1/q$ (unless the third partial quotient is 2; then the outermost convergent has the form $2/q$, with q odd).

The obvious question arises as to the relation between the structure of the outermost class $c+1$ island to the class c island it encircles. One can define a "class renormalization" analagous to that for levels discussed in the previous section. This rescaling transformation is a generalization of the period doubling bifurcation case.²¹ For period doubling one considers the bifurcation of period 2 orbits from a fixed point, or in our notation $q=2$. More generally one can consider period p/q -tupling bifurcations corresponding to the birth of a frequency p/q orbit. A renormalized map is defined as

$$N(T) = B^{-1} T^q B$$

where B is a coordinate change which moves the origin to a point on the higher class orbit of frequency p/q and rescales phase space areas. In the period doubling case it is well known that the renormalization transformation has a critical fixed point corresponding to a self-similar class-sequence of period two orbits. For this critical fixed point the Jacobian of B^{-1} , which determines the area scaling of the transformation, is 0.015209. At the fixed point the period two orbits are all unstable, and so do not correspond to island chains, and give no contribution to the transport model.

However, higher order bifurcations can lead to fixed points where all the members of the bifurcation sequence are stable. For example consider the class zero, frequency $1/3$ island of the standard map. The critical fixed point of septupling occurs at $k \approx 1.0420961$. Here the $1/7 = [0,6,1]$ class one island chain is the outermost convergent of the boundary circle of the class zero chain; it has residue 0.2755. Similarly the $1/7$ class two orbit encircling this class one chain is the outermost convergent to the class one boundary circle, with residue 0.2677. The residues of the higher class $1/7$ orbits approach the value 0.26580, which is slightly above the boundary circle value. One finds the boundary circle has the continued fraction expansion $[0,6,1,4,1,5,1,\dots]$ for the universal map at the septupling fixed point.

The area scaling of an island can be computed by considering ΔW of the periodic orbit, defined as the difference in action between the elliptic and hyperbolic orbits. This is equivalent to the Jacobian of the coordinate change B . The ratio of island areas scales as the ratio of the ΔW 's. We find that at the $1/7$ fixed point the area of a higher class island is smaller than the area of the

island it encircles by the asymptotically constant factor

$$w_2 \approx 0.014158 \quad (18)$$

This scaling factor can also be identified as the flux ratio, Eq.(3.6d): the ratio of the ΔW for crossing the "outermost" class $c+1$ minimizing cantorus, to that for crossing the lower bounding class c cantorus of the state S . This is true by self-similarity at the fixed point: the phase space surrounding the class c island is identical to that surrounding the class $c+1$ island under rescaling.

To determine the ratio of areas of the class $c+1$ state to the class c state, this ratio should be multiplied by the ratio of number of islands at class $c+1$ to class c , which for this fixed point is exactly 7:

$$\varepsilon_2 = 1/7 \quad (19)$$

which gives the remaining scaling coefficient a_2 :

$$a_2 = w_2/\varepsilon_2 = 0.099105 \quad (20)$$

In general, any particular island is not at the parameter value which is a fixed point of class renormalization, e.g. q is not constant for higher renormalizations. Remarkably a version of self-similarity still survives with good accuracy. We find that the integer q which defines the first convergent to the boundary circle typically varies between 5 and 8, but the area of the class c island always decreases according to the relation:

$$\Delta W \approx C(q)^{-2.19}$$

where q is the length of the level zero, class c orbit, and ΔW is the area of one of the class c islands. The exponent is, to within a few percent, the same for all islands and parameter values providing $1/q$ is the frequency of the outermost convergent to the boundary circle. This equation gives a relationship between ε_2 and w_2 :

$$w_2 = \varepsilon_2^{2.19} \quad (21)$$

As in the level renormalization, the variation of q implies that the

self-similarity we assume is not exact, rather the coefficients w_2 and ε_2 vary with class.

§ 5. Distribution of First Return Times

As mentioned in §3.1 the numerical experiments effectively measure the quantity $F_{\sigma_1}(t)$, the fraction of particles starting in state $S=\{\sigma_1\}$ which are still "alive" (i.e., have never entered \emptyset) at time t . Here it will be convenient to study a slightly different quantity, $R_{S \rightarrow S'}(t)$, defined as the probability that a particle in state S at time $t=0$ first reaches S' at time t . It is easy to see that R is related to F through

$$R_{\sigma_1 \rightarrow \emptyset}(t) = dF_{\sigma_1}/dt. \quad (22)$$

where we make the continuous time approximation implied by (6). Thus, if F_{σ_1} decays algebraically with exponent z , Eq.(1), then

$$R_{\sigma_1 \rightarrow \emptyset} \sim t^{-(z+1)}. \quad (23)$$

§ 5.1 Fundamental Integral Equation for $R_{\sigma_1 \rightarrow \emptyset}$

It is also convenient to introduce another quantity $R_{S \rightarrow S'}^d(t)$, which is the probability that a particle in state S at time zero reaches state S' at time t without having been in any other state between. Following Hanson et al.,¹¹ we call $R_{S \rightarrow S'}$ the "first passage time distribution," and $R_{S \rightarrow S'}^d$ the direct first passage time distribution. Clearly $R_{S \rightarrow S'}^d$ is zero unless S' is either DS , $S1$ or $S2$. In particular,

$$R_{S \rightarrow S'}^d = p_{S \rightarrow S'} \exp(-p_S t), \quad (24)$$

where p_S , Eq.(6), is the probability that a particle leaves S on one iterate

$$p_S = p_{S \rightarrow DS} + p_{S \rightarrow S1} + p_{S \rightarrow S2}. \quad (25)$$

The first passage time distribution $R_{1 \rightarrow \emptyset}(t)$ obeys

$$R_{1 \rightarrow \emptyset}(t) = R_{1 \rightarrow \emptyset}^d(t) + \int_0^t dt' \left\{ R_{1 \rightarrow 12}^d(t') R_{12 \rightarrow \emptyset}(t-t') + R_{1 \rightarrow 11}^d(t') R_{11 \rightarrow \emptyset}(t-t') \right\}. \quad (26)$$

This equation follows from the fact that particles arriving in \emptyset can do so either directly (the term $R_{1 \rightarrow \emptyset}^d(t)$) or else by first making an "up transition" from $\{1\}$ to $\{12\}$ or to $\{11\}$. In the this latter case, the contribution to $R_{1 \rightarrow \emptyset}(t)$ is given by an integral over the time t' at which the first up transition out of $\{1\}$ is made. Similarly, we have

$$R_{12 \rightarrow \emptyset}(t) = \int_0^t dt'' R_{12 \rightarrow 1}(t'') R_{1 \rightarrow \emptyset}(t-t''),$$

$$R_{11 \rightarrow \emptyset}(t) = \int_0^t dt'' R_{11 \rightarrow 1}(t'') R_{1 \rightarrow \emptyset}(t-t''). \quad (27)$$

Equations (26) and (27) follow from the definition of Markovian dynamics on the Bethe lattice.

The scaling relations, Eqs. (9), can be used to obtain relations between the different first passage distributions. In particular, we see that the transition probabilities governing $R_{11 \rightarrow 1}$ are the same as those governing $R_{1 \rightarrow \emptyset}$ if the latter are multiplied by ε_1 . That is, time is stretched by the factor ε_1^{-1} , or

$$R_{11 \rightarrow 1}(t) = \varepsilon_1 R_{1 \rightarrow \emptyset}(\varepsilon_1 t). \quad (28)$$

Similarly,

$$R_{12 \rightarrow 1}(t) = \varepsilon_2 R_{1 \rightarrow \emptyset}(\varepsilon_2 t) . \quad (29)$$

Finally we can express $R_{2 \rightarrow \emptyset}$ in terms of $R_{1 \rightarrow \emptyset}$. Note from Eq.(9) that

$$\frac{P\{1,S\} \rightarrow \{1,S'\}}{P\{2,S\} \rightarrow \{2,S'\}} = \frac{\varepsilon_1}{\varepsilon_2} .$$

Thus

$$R_{2 \rightarrow \emptyset}(t) = \varepsilon_2 R_{1 \rightarrow \emptyset}(\varepsilon_2 t / \varepsilon_1) / \varepsilon_1 , \quad (30)$$

and hence from (29)

$$R_{12 \rightarrow 1}(t) = \varepsilon_2 R_{1 \rightarrow \emptyset}(\varepsilon_2 t) . \quad (31)$$

Now consider the direct first passage time distributions appearing in Eq.(26). From Eqs. (24) and (25) and again utilizing (9) we obtain

$$R_{1 \rightarrow \emptyset}^d(t) = p_{1 \rightarrow \emptyset} \exp(-p_1 t) = p_0 \varepsilon_1 \exp(-\alpha p_0 \varepsilon_1 t) ,$$

$$R_{1 \rightarrow 12}^d(t) = p_{1 \rightarrow 12} \exp(-p_1 t) = p_0 \varepsilon_1 w_2 \exp(-\alpha p_0 \varepsilon_1 t) ,$$

$$R_{1 \rightarrow 11}^d(t) = p_{1 \rightarrow 11} \exp(-p_1 t) = p_0 \varepsilon_1 w_1 \exp(-\alpha p_0 \varepsilon_1 t) ,$$

$$\alpha \equiv 1 + w_1 + w_2 . \quad (32)$$

From Eq.(30), we see that it suffices to find either $R_{1 \rightarrow \emptyset}$ or $R_{2 \rightarrow \emptyset}$ since one may be obtained from the other. For the sake of making the symmetry $1 \leftrightarrow 2$ manifest, we introduce a function $h(t)$ given by

$$R_{1 \rightarrow \emptyset}(t) \equiv p_0 \varepsilon_1 h(\alpha p_0 \varepsilon_1 t) , \quad (33)$$

which from Eq.(30) implies

$$R_{2 \rightarrow 0}(t) \equiv p_0 \varepsilon_2 h(\alpha p_0 \varepsilon_2 t) . \quad (34)$$

Letting $\tau = \alpha p_0 \varepsilon_1 t$, be a scaled time variable, and inserting Eqs. (27)-(34) into Eq. (26), we obtain the fundamental integral equation for the tree model,

$$h(\tau) = e^{-\tau} + \int_0^{\tau} \int_0^{\tau-\tau'} e^{-\tau''} k(\tau'') h(\tau-\tau'-\tau'') d\tau' d\tau'' , \quad (35)$$

$$k(\tau) \equiv [w_1 \varepsilon_1 h(\varepsilon_1 \tau) + w_2 \varepsilon_2 h(\varepsilon_2 \tau)] \alpha^{-2} . \quad (36)$$

Note that the kernel, k , is symmetric under the interchange $1 \leftrightarrow 2$. Alternatively we may write (35) as

$$h = e^{-\tau} + e^{-\tau} \circledast k \circledast h , \quad (37)$$

where \circledast signifies the convolution operation. Equation (35) is a quadratically nonlinear integral equation for h . In the case of the one-dimensional chain model,¹¹ only one of the two terms in (36) is present. Then k reduces to

$$k(\tau) = w_1 \varepsilon_1 \alpha^{-2} h(\varepsilon_1 \tau) \quad (38)$$

Surprisingly, Hanson et al.¹¹ were able to obtain a general solution to (35) in this case. We shall not be able to do the same for the tree. However, our interest is in extracting the characteristic long-time asymptotic behavior of $h(\tau)$, and, as shown in the next section, this can be done, modulo some reasonable assumptions on the character of the solution.

§ 5.2 Dispersion Relation

Since the integral equation (37) has a convolution structure, it is clearly simplified by Laplace transform:

$$H(s) \equiv \mathcal{L}[h(\tau)] \equiv \int_0^{\infty} \exp(-s\tau) h(\tau) d\tau .$$

Noting the identity $\mathcal{L}[h(\varepsilon\tau)] = \varepsilon^{-1}H(s/\varepsilon)$, Eq. (35) becomes upon application of \mathcal{L}

$$H(s) \left\{ s+1 - \alpha^{-2} [w_1 H(s/\varepsilon_1) + w_2 H(s/\varepsilon_2)] \right\} = 1. \quad (39)$$

The long time behavior of $h(\tau)$ is determined by the singularity of $H(s)$ with the smallest real part. In particular, a power law dependence, as in Eq.(23) is obtained from a branch point of $H(s)$ at $s=0$. For example, if $H(s)$ has an expansion near $s=0$ of the form

$$H(s) = f(s) + s^z g(s), \quad (40)$$

where f and g are analytic at $s=0$ ($g(0) \neq 0$ by definition), then the asymptotic evaluation of $\mathcal{L}^{-1}[H]$ yields Eq.(23). We now assume (30) and determine the exponent z . Evidence for the validity of (30) is obtained from the one-dimensional case, where the exact solution expanded around $s=0$ does indeed have this form. Furthermore one can use (30) to generate power series for f and g , in the present case, as was done in the one-dimensional case. Unfortunately, these series are complicated, and so we are content to find only the coefficients of s^0 and s^z .

Substituting (30) into (29) and setting $s=0$, we obtain a quadratic equation for $f_0 = f(0)$, viz.,

$$f_0 [1 - \alpha^{-2}(w_1 + w_2)f_0] = 1.$$

Noting that $\alpha = 1 + w_1 + w_2$, we may express the two solutions for f_0 as

$$f_0 = 1 + w_1 + w_2 = \alpha, \quad (41)$$

$$f_0 = 1 + (w_1 + w_2)^{-1}. \quad (42)$$

Noting that $w_1 + w_2 < 1$ for the map, we will see that Eq.(42) yields a solution for H which does not decay with time ($z < 0$) and hence is not physically relevant (the physical solution for h , must be normalizable since it is a probability).

Using (30) in (29), the coefficient of s^z can be written

$$-f_0\alpha^{-2}[w_1\varepsilon_1^{-Z} + w_2\varepsilon_2^{-Z}] + [1 - (w_1+w_2)\alpha^{-2}f_0] = 0 .$$

Substituting for f_0 from Eq.(31), we obtain a "dispersion relation" for z ,

$$w_1\varepsilon_1^{-Z} + w_2\varepsilon_2^{-Z} = 1 . \quad (43)$$

This relation is the primary result of our analysis.

§ 5.3 Multiply Branched Trees

In Fig. 1 we have only shown one island family between adjacent minimizing cantori. Our subsequent analysis has been limited to this case. To include M relevant island families between adjacent cantori, one must use $M+1$ symbols for the state labels, e.g. $\sigma_i = 1, 2, \dots, M+1$. Correspondingly the number of branches at each node of the tree increases. All of the analysis of the previous sections goes through for general M . The result is that the dispersion relation (43) becomes

$$\sum_{j=1}^{M+1} w_j \varepsilon_j^{-Z} = 1 . \quad (44)$$

§ 5.4 Long-Time Behavior

Equation (43) and correspondingly (44) have a single purely real root for z and, in addition, an infinity of complex roots. The uniqueness of the real root follows from the fact that the left hand side of (44) increases monotonically with z from zero at $z=-\infty$ to infinity at $z=\infty$ (this only requires $\varepsilon_j < 1$).

For the one-dimensional chain model ($M=0$ in Eq. (44)), solution of the dispersion relation is

$$z = z_k = (\ln w_1 + 2\pi i k) / \ln \varepsilon_1 , \quad (45)$$

where k is an integer. Using the numerical values of w_1 and ε_1 , Eqs. (12)-(13), the real root for the chain model is

$$z = z_0 = \ln w_1 / \ln \varepsilon_1 \approx 3.05 . \quad (46)$$

Since (43) is related to the dispersion relation for the chain by the addition of the positive term, $w_2 \varepsilon_2^{-z}$, we see that z_0 is necessarily smaller when islands are included. Indeed using the numerical values, Eqs.(18)-(19), gives

$$z_0 \approx 1.96 , \quad (47)$$

which is considerably closer to the value (1.3-1.5) obtained numerically. Furthermore, inclusion of more island families would result in a smaller z_0 value. This is physically reasonable since the island cantori make an island effectively "sticky" (c.f. §2). That is, if a particle is very close to a critical curve, it takes a long time to get away from it. In other words, introduction of another island family into the model opens up more possible places for the particle to go before entering the state \emptyset .

Computation of the complex roots of Eq. (44) is more difficult when $M > 0$. To illustrate the situation for $M=1$, consider the case when the term $w_2 \varepsilon_2^{-z}$ is small compared to 1 (note that unfortunately this approximation is not valid for the actual numerical values of w_2 and ε_2). In this case a simple perturbative calculation yields

$$\begin{aligned} z_k &= z_k^0 - r e^{-ik\theta} , \\ r &= -w_2 \varepsilon_2^{-z_0^0} / \ln \varepsilon_1 , \\ \theta &= 2\pi \ln \varepsilon_2 / \ln \varepsilon_1 , \end{aligned} \quad (48)$$

where z_k^0 is the $M=0$ solution given in Eq.(45). These roots are schematically illustrated in Fig. 4. In general, the real root is also the root with the smallest real part, though there are arbitrarily many roots with real parts arbitrarily close to z_0 .

With an infinity of roots The Laplace transform of the first passage distribution takes the form

$$H(s) = f(s) + \sum_k s^{z_k} g_k(s) . \quad (49)$$

Evaluation of the long-time behavior from the inverse Laplace transform of $H(s)$ gives contributions to $h(t)$ from $z = z_k$ which scale like

$$t^{-(1+z_k)} = t^{-(1+x_k)} e^{-iy_k \ln t} . \quad (50)$$

where x_k and y_k are the real and imaginary parts of z . In the $M=0$ case, $y_k = 2\pi k / (\ln \varepsilon_1)$ and $x_k = z_0$. This gives a solution of the form

$$h(t) \sim t^{-(1+z_0)} F(\ln t) .$$

where F is a periodic function with period $\ln(1/\varepsilon_1)$. This is precisely the form of the exact solution obtained in Ref. 11. On a log-log plot, this solution appears as a straight line decay with slope $(1+z_0)$ on which is superimposed periodic oscillations in the distribution.

In the more general case the log-periodicity of the oscillations would no longer hold, however we still expect an overall decay with slope $(1+z_0)$ on which will be superimposed a component possessing a mixture of different periods. Indeed in the numerical experiments of Karney such an irregular oscillatory component is present.

§ 6. Conclusions

The primary result of this paper is that island chains reduce the decay rate of correlation functions. In our model the probability a particle is stuck in a certain region of phase space for a time t decays asymptotically as

$$t^{-z} \times \text{an oscillatory function of } \ln(t) .$$

The exponent z is determined by equation (44). We define a probabilistic state, S_i , as a region of phase space bounded by cantori enclosing an island chain which is a continued fraction convergent to a critical invariant circle. The coefficient w_i represents the relative area of an island in the state S_i which is a

convergent one "level" closer to the critical circle, and ε_1^{-1} represents the number of islands in S_1 relative to the number in S . The coefficient w_2 represents the relative area of the outermost higher "class" island encircling the island chain in S and ε_2^{-1} is the relative number of islands in this higher class chain. The state which is the region around this higher class island we call S_2 .

Neglecting all islands, one obtains an exponent, $z=3.05$, which is larger than the numerical experiments by a factor of two. Including the largest island chain reduces this exponent to $z=1.96$ which is closer to the numerical value.

There are several approximations in the model which could be responsible for the remaining discrepancy between numerical experiments and theory. In order of what we feel is decreasing importance these are:

1) There is more than one chain of islands (of the same class) in state S , corresponding to the various rational frequencies between those of the bounding cantori. These can be included with coefficients w_j and ε_j which represent area and time scaling factors for these chains. As shown in Eq. (44) inclusion of more chains of islands will reduce the exponent z , and one could reasonably hope the importance of these island chains decreases sufficiently rapidly that the exponent converges to the numerically determined value of about 1.3-1.5.

2) Inclusion of the fluctuations of the scaling of $p_{S \rightarrow S'}$ might be necessary. These fluctuations are due to the fact that boundary frequencies are typically not noble, and islands are typically not at the q -tupling fixed point. We don't expect this to greatly change z , because of the scaling relations (16) and (21) which hold for arbitrary boundary circles and islands. In particular, for the one-dimensional Markov chain, Eq. (44) implies that z depends only on the exponent in Eq. (16) and not on w_1 and ε_1 independently. Unfortunately, for a several branched Markov tree, this is no longer true, and fluctuations in the ε 's may give a different average exponent.

3) Finally, our model takes into account only a discrete set of cantorus frequencies, out of the existing continuum. Inclusion of more and more cantori leads to breakdown of the Markov hypothesis since fluxes through neighboring cantori are not independent. In fact since the quantity ΔW is a continuous function on the irrationals, the "turnstiles" of close cantori overlap. This suggests constructing a continuum model for transport, although we have not yet been able to formulate this satisfactorily. The simplest continuum limit (in the one-dimensional chain model) gives a diffusion equation with a spatially

dependent diffusion coefficient. This model gives exactly the same result for z as the discrete model. One can further argue, that, even if the cluster of cantori around the minimizing one gives a significantly different effective transition probability, this cluster will have the same scaling properties as a single cantorus, and the result for z is unaffected.

Acknowledgements

The authors thank J. Stark and R. S. MacKay for stimulating discussions and for sharing the results of research with J. M. Greene on boundary circles prior to publication. This collaboration was undertaken at the Institute for Theoretical Physics, with the support of the National Science Foundation under grant No. PHY82-17853 supplemented with funds from NASA. Support by U.S. Department of Energy is also acknowledged.

References

- 1). J.D. Meiss, R.S. MacKay, and I.C. Percival, Phys. Rev. Lett. **52** 697 (1984); Physica **13D**, 55 (1984).
- 2). B. V. Chirikov, Phys. Reports **52**, 265 (1979).
- 3). A.B. Rechester, M.N. Rosenbluth, and R.B. White, Physical Review, **23A**, 2664 (1981).
- 4). J.R. Cary and J.D. Meiss, Physical Review **24A**, 2664 (1981).
- 5). J. D. Meiss, J.R. Cary, C. Grebogi, J.D. Crawford, A.N. Kaufman, H.D.I. Abarbanel, Physica **6D**, 360 (1983).
- 6). J. Ford, "The Statistical Mechanics of Classical Analytic Dynamics", in Fundamental Problems in Statistical Mechanics III, E. Cohen (ed.), pp.215-255 (North-Holland, Amsterdam 1975).
- 7). R.B. White, "Collisional Diffusion in a Torus with Imperfect Magnetic Surfaces," Princeton Plasma Physics Laboratory Report # PPL-1984, (1983); P.T. Bonoli and E. Ott, Phys. Fluids, **25**,359 (1982).
- 8). C.F.F. Karney, Physica **8D**, 360 (1983).

- 9). B.V. Chirikov and D.L. Shepelyanski, *Physica* **13D**, 395 (1984).
 - 10). B.V. Chirikov, in Lecture Notes in Physics no. 179, p. 29 (1983); B.V. Chirikov and D. L. Shepelyanskyy, in Proc. 9th International Conf. on Nonlinear Oscillations (Kiev 1981), (Kiev, Naukaova Dumka, 1983) v. II, p. 137 (English Translation is Princeton Plasma Physics Laboratory Report #PPPL-Trans-133).
 - 11). J.D. Hanson, J.R. Cary, J.D. Meiss, *J.Stat. Phys.* **39**, 327 (1985).
 - 12). J.M. Greene, "A Simple Transport Model," GA Technologies Report #GA-A17511(1984).
 - 13). L. Kadanoff, *Phys. Rev. Lett* **47**, 1641 (1981); R.S. MacKay, "Renormalization in Area Preserving Maps," Ph.D. Thesis Princeton University (University Microfilms, Ann Arbor, Michigan, 1982); *Physica* **7D**, 283 (1983).
 - 14). D. Bensimon and L.P. Kadanoff, *Physica* **13D**, 82 (1984).
 - 15). I.C. Percival, in Nonlinear Dynamics and the Beam-Beam Interaction, M. Month and J.C. Herrera (eds.), American Inst. of Physics Conf. Proceedings No. 57, p. 302 (1979).
-
- 16). J.N. Mather, *Topology* **21**, 457 (1982); S. Aubry and P.Y. Le Daeron, *Physica* **8D**, 381 (1983); A. Katok, *Ergodic theory and Dyn. Sys.* **2**, 185 (1982).
 - 17). J.N. Mather, "A Criterion for the Non-existence of Invariant Circles," preprint, Princeton University (1982).
 - 18). J.M. Greene, J. Stark and R.S. MacKay, "Boundary Circles in Area Preserving Maps" preprint, University of Warwick, 1985.
 - 19). M. R. Herman, *Astérisque*, **103-104**, (1983).
 - 20). D. Rand, personal communication, 1985.
 - 21) J.M. Greene, R.S. MacKay, F. Vivaldi, and M.J. Feigenbaum, *Physica* **3D**, 468 (1981).
 - 22). J.D. Meiss, "Class Renormalization," in preparation.

Fig. 1 Schematic of the connected stochastic region of an area preserving map near a boundary circle. Inaccessible areas are shaded. Boundaries of the labeled states are minimum flux cantori barriers, with their turnstiles shown. Each class one island shown represents one member of an island chain.

Fig 2. ΔW as a function of frequency for rationals near a critical circle of frequency $\nu = \alpha^{-2} \approx 0.382$. The rationals are on the Farey tree for $[1/3, 1/2]$. The lowest value of ΔW shown is for $\nu = 55/144$, which is the 10th convergent to the critical frequency. ΔW for irrational ν is approximated by the lower envelope.

Fig 3. Tree associated with Fig. 1 up to the two symbol level.

Fig 4. Roots of Eq. (48), showing the rates, z_i , which contribute to the decay of the first return distribution.