DOE-ET-53088-189 IFSR#189

TRANSPORT NEAR THE ONSET OF STOCHASTICITY

James D. Meiss
Institute for Fusion Studies
The University of Texas at Austin
Austin, Texas 78712

May 1985

TRANSPORT NEAR THE ONSET OF STOCHASTICITY

JAMES D. MEISS
Inst. Fusion Studies, University of Texas, Austin

Abstract For two-degree-of-freedom Hamiltonians, (e.g., a particle in a 2-D potential or the flow of magnetic field lines) an invariant torus in phase space acts as an absolute barrier for trajectories. When an invariant torus is destroyed by a perturbation, a remnant remains with gaps. This "cantorus" forms a formidable barrier even well into the stochastic regime. We show that correlation functions decay algebraically invalidating the common assumptions of chaos. The decay rate is given by a universal exponent, obtained from self-similar scaling.

INTRODUCTION

In colliding proton machines, the particle dynamics are to a good approximation given by Hamilton's equations. Similarly, in fusion devices such as the tokamak and stellerator, the single particle dynamics are believed to determine, to a great extent, the confinement properties of the device, though of course Coulomb collisions play an important role. I will describe in this talk techniques for computing particle lifetimes in such systems. For simplicity we consider the single particle dynamics in a given set of fields, such as the magnet lattice of the colliding beam storage device, or the toroidal and poloidal fields in a fusion confinement device.

Our method is inspired by one used by Wigner[1] to estimate chemical reaction rates. Wigner envisioned an N degree-of-freedom Hamiltonian system describing the dynamics of several reacting components: $H(\mathbf{z})$ such that $\mathbf{z}=(z^1,...z^{2N})$ are the coordinates $(\mathbf{p},\mathbf{q})$. As an example, the reaction might be of the form $A_1+A_2+A_3 \rightarrow A_1A_2+A_3$. He assumed that classical dynamics is a valid description. One can approximately separate the phase space into two regions corresponding to unreacted (all components far apart) and reacted (where A_1 and A_2 are close together). A similar division of phase space can be made in other applications as well, such as in confinement devices, namely the inside (the good region) and the outside (where particles are not supposed to go!).

The procedure is to construct a closed 2N-2 dimensional surface, $G(\mathbf{z})=0$, in the 2N-1 dimensional energy surface, which separates the unreacted region from the reacted region. The reaction rate can be estimated by computing the one way flux of trajectories across this surface

$$\mathfrak{J} \equiv \int dz^{2N}\, \dot{\mathbf{z}}\cdot\nabla G/|\nabla G|\ \delta(H(\mathbf{z})-E)\ \delta(G)\ \Theta(\mathbf{v}\cdot\nabla G)$$

where $\dot{\mathbf{z}}$ is the velocity in phase space, computed from Hamilton's equations, the first δ-function restricts the integration to the energy surface, the second to the $G=0$ surface, and the Θ-function requires that the flux be positive. Note that Liouville's theorem implies that the flux crossing G in the opposite direction has the same magnitude as $\mathfrak{J}$.

The reaction rate is certainly longer than the time it takes trajectories to cross the surface $G=0$, which if the system is found with equal à priori probability in the region $G<0$, at $t=0$, is bounded by

$$\tau_{Reaction} > \tau_G \equiv V/\mathfrak{F}$$

where V is the volume of the G<0 region. Wigner's idea is to vary the surface G, seeking to maximize τ_G, thereby obtaining the best possible estimate of the reaction time. If we hold V fixed, this is equivalent minimizing $\mathfrak{F}$.

What is the surface of minimum flux? The answer is unknown in the general case. Indeed, I believe it has turned out to be very difficult to implement this prescription even numerically. One of the difficulties, about which we will say more below, is that the extremum of flux turns out not to be an isolated minimum, rather there are many surfaces with the same flux.

The point of our work[2] is to give the proper answer to Wigner's implicit question for a simple case, that of the 2-degree-of-freedom Hamiltonian. The reason it is easier to solve this problem is that trajectories on the 3-D energy surface of an integrable 2 d.o.f. Hamiltonian, lie on 2-D tori--these tori divide 3-space into the requisite inside and outside. When integrability is destroyed by perturbing the system (so as to break the symmetry which leads to the second invariant), some of the 2-D tori are preserved: the KAM tori. When a KAM surface exists, the flux across it is zero, and the escape time is infinite. As the perturbation is increased KAM tori are destroyed and motion between regions of phase space previously separated by KAM surfaces becomes possible. Indeed in the next section we show that the "surfaces" of minimum flux in this case are the remnants of KAM surfaces; they are called cantori.[3]

The flux model, combined with knowledge of the universal way in which KAM surfaces are destroyed and replaced by cantori, gives a universal exponent, η, for the divergence of the transport time with perturbation parameter:[2] $\tau \sim (k-k_{cr})^{-\eta}$. This exponent agrees well with numerical experiments.

The knowledge of the flux through phase space does give a lower bound on reaction times, but more detailed knowledge of the evolution of trajectories requires a transport theory. Such a theory takes into account the local fluxes through many cantori to give a global, long time result. Below we discuss a simple transport theory obtained by retaining a discrete set of cantori, and assuming that the motion in regions between cantori is completely random.[4] This Markov model leads to the prediction of an algebraic decay of correlations in regions of phase space bounded by a KAM surface. The reason for this slow decay is that as an orbit approaches a KAM surface, the flux through the cantori (of which there are infinitely many) approaches zero. Orbits can become trapped for arbitrarily long times near a KAM surface.

AREA PRESERVING MAPS

In a two degree of freedom system motion takes place on a 3-D energy surface. The flow can be reduced two a 2-D map by watching the orbits pierce a surface-of-section defined as the intersection of the energy surface with some surface, say, $p_1=0$. The map which takes an initial condition on the surface and gives the next intersection of the orbit with the surface is area-preserving. Henceforth we talk only of area-preserving maps.

It is convenient to think of maps on a cylindrical phase space, $(x,y \mid x+1 \to x)$, with vertical coordinate y, representing a momentum, and angle x. Denote them by T_k: $(x,y) \to (x',y')$. It is assumed that T depends on a parameter, k, and that the $k=0$ case is an integrable map. We also assume T_k is a twist map: it causes rotation at different rates at different vertical positions on the cylinder. More formally the twist condition is $\partial x'/\partial y|_x > 0$. A useful example, which apparently has the properties of a typical twist map, is the standard map[5]

$$\begin{aligned} y' &= y - \frac{k}{2\pi}\sin(2\pi x) \\ x' &= x + y' \end{aligned} \tag{1}$$

The standard map will be used for illustration, but everything I say holds for any one parameter family of twist maps. We denote an orbit of $\mathcal{T}$ by the sequence $(x_t, y_t) = \mathcal{T}(x_{t-1}, y_{t-1})$ where t takes integer values.

As was already known by Poincaré,[6] the phase space of $\mathcal{T}$ is filled with a complicated mixture of regular and irregular (or chaotic) orbits. There are many classes of regular orbits on the cylinder: those which encircle the cylinder we call class zero. These orbits are the survivors of the invariants of the integrable twist map at k=0. Their rotation numbers are given by $\nu = \lim_{t\to\infty} x_t/t$ where x_t is not taken modulus unity for this calculation. If ν is irrational, the orbit densely fills a curve (see however below), otherwise it is periodic. By the Poincaré-Birkoff theorem, periodic orbits come in pairs--elliptic (stable, eigenvalues of modulus unity) and hyperbolic (unstable, real eigenvalues); more precisely these orbit have positive and negative index, or values of the residue, as defined by Greene.[7]

The regular orbits of nonzero class are found generally in the neighborhood of elliptic periodic orbits. They surround points on the periodic orbit, forming "island chains." Class one regular orbits encircle points of a class zero periodic orbit. Class c periodic orbits encircle points of class c-1 periodic orbits, forming islands about islands, ad infinitum.

When k is small enough, $\mathcal{T}$ is nearly integrable and the Kolmogorov-Arnold-Moser (KAM) theorem[6] is applicable. It implies that class zero regular orbits occupy a finite measure (approaching the entire space as $k\to 0$) of the cylinder. Those orbits preserved for k>0 have "sufficiently" irrational ν; they are the KAM surfaces or invariant circles. A second version of this theorem shows that the measure of the regular orbits near an elliptic orbit is finite. By contrast, 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 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.

As each circle of irrational ν is destroyed there appears in its place a new orbit with the same rotation number, though no longer dense on a smooth curve.[8] This orbit lies on a Cantor set: it is a "curve" with a deleted infinity of gaps, or open intervals (more precisely its projection on the x-axis is a Cantor set). Percival named these orbits cantori.[3] Crudely speaking, the gaps are caused by overlapping of neighboring islands. Since the frequency is irrational, once one gap appears in an invariant circle, there must be an infinity of gaps whose endpoints are the iterates of the first. Numerically (for simple examples) it appears that all the gaps in the cantorus are iterates of one another.[2]

FLUX

We can define the flux through the various types of orbits of $\mathcal{T}$.[2] The simplest case is a period-q orbit; consider the hyperbolic case. We wish to construct the surface which is the analogue of Wigner's G=0. One way to do this is to connect two neighboring points on the orbit with some arbitrary curve $\mathcal{C}$, see Fig. 1. It is convenient to draw $\mathcal{C}$ through one point on the elliptic orbit, as well. Iterate $\mathcal{C}$ once with the map, obtaining $\mathcal{TC}$ which again connects two points on the hyperbolic orbit and goes through a point on the elliptic orbit. If we iterate $\mathcal{C}$ q−1 times, the result is a curve

$$\mathcal{G} = \sum_{i=0}^{q-1} \mathcal{T}^i \mathcal{C}$$

which connects all the points on both orbits.

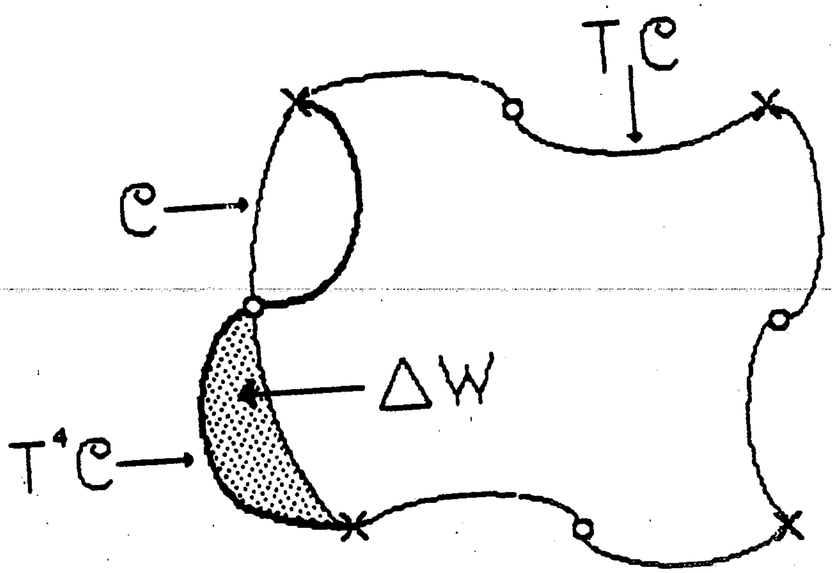

Fig. 1 Turnstile in a frequency 1/4 orbit. X indicates hyperbolic orbit and O elliptic.

Now imagine iterating $\mathcal{G}$ once more. All parts of $\mathcal{G}$ iterate onto $\mathcal{G}$ except the last segment, $\mathcal{T}^{q-1}\mathcal{C}$, which becomes $\mathcal{T}^q\mathcal{C}$ and connects the original two x-points. This last segment will not lie on top of $\mathcal{C}$ (except in the exceptional integrable case) for this would imply that $\mathcal{G}$ is an entire curve of periodic orbits. In fact the curves $\mathcal{T}^q\mathcal{C}$ and $\mathcal{C}$ typically intersect only at the three points on the periodic orbits, since otherwise more than two periodic orbits would exist. This yields the figure-of-eight structure show in Fig. 1, which we call the turnstile.[2]

The area in one lobe of the turnstile is the flux: the area escaping through $\mathcal{G}$ on each iterate of the map. By area preservation, the two lobes have the same area. Note that we have localized the entire flux through the orbit to one region. However, it is not difficult to see that the value of the flux is actually independent of way $\mathcal{G}$ is constructed. To show this, we compute the flux as

$$\mathcal{F} = \int_{\text{turnstile}} dx\,dy = \int_{x,T^q\mathcal{C}}^{0} y\,dx - \int_{x,\mathcal{C}}^{0} y\,dx$$

This integral can be related to the action of the orbits, by use of the Lagrangian variational principle. Here the action of an orbit is defined as $W = \int L\,dt$ where the integral is along the orbit. It has an obvious discrete time analogue, given by simply integrating along the Hamiltonian orbit in the energy surface between two intersections of the surface-of-section. This calculation shows that the flux is given by the difference of action between the elliptic and the hyperbolic orbits:[2]

$$\mathcal{F} = \Delta W \equiv W(\text{elliptic}) - W(\text{hyperbolic}) \tag{2}$$

Equation (2) demonstrates our assertion: the flux through a periodic orbit is a uniquely defined quantity, the curve $\mathcal{G}$ need never have been drawn. It also points out the problem with Wigner's variational principle: any curve $\mathcal{G}$ through the periodic orbits has the same flux, and so the minumum flux does not define a unique $\mathcal{G}$.

We can now attempt to find which orbits of the map have the minimum flux. Figure 2 gives the flux through a host of periodic orbits for the standard map. This figure is constructed using the Farey tree procedure to generate the rationals, corresponding to the frequencies of the periodic orbits.[9] We first pick two rationals, say

1/3 and 1/2, and call them level zero. A level one rational between these is obtained by adding numerators and denominators, yielding 2/5. Now take this number and add to it the numerators and denominators of the two level 0 rationals to get two new rationals, 1/4 and 2/5. In general, each level j rational yields two daughter rationals at level j+1 by adding it (in this peculiar way) to each of the two nearest level j-1 rationals. This procedure eventually generates all the rationals between the original pair. In Fig. 2 we have connected the two daughter rationals to their parent, giving a binary tree.

The remarkable observation is that ΔW is a monotonically decreasing function of level along a path of connected branches.[9] Where do these paths go? The denomonators of the rationals increase with level, and following a path corresponds to approaching some definite irrational. Along most paths to irrationality, ΔW approaches a non-zero limit. This approach to a limit begins geometrically, but as the limit is approached the rate accelerates. The evidence of Fig. 2 is that every irrational has smaller flux than nearby rationals. Of course the orbit that corresponds to such an irrational is just the cantorus. In fact one of the proofs of existence of the cantori involves the statement that the limit as $p/q \to \nu$ of ΔW exists and is positive just when the KAM surface is destroyed.[10] When there is a circle the ΔW's converge to zero.

Another interesting feature of Fig. 2 is that paths which oscillate in direction, first moving to the next level towards higher frequency, and then to the level beyond towards lower frequency, are those that tend to lead to the smallest values of ΔW. Oscillating paths lead to a special type of irrational, named noble by Percival. Nobles are defined in terms of their continued fraction representations

$$\nu = a_0 + 1/(a_1 + 1/(a_2 + \ldots \equiv [a_0, a_1, a_2, \ldots] \qquad (3)$$

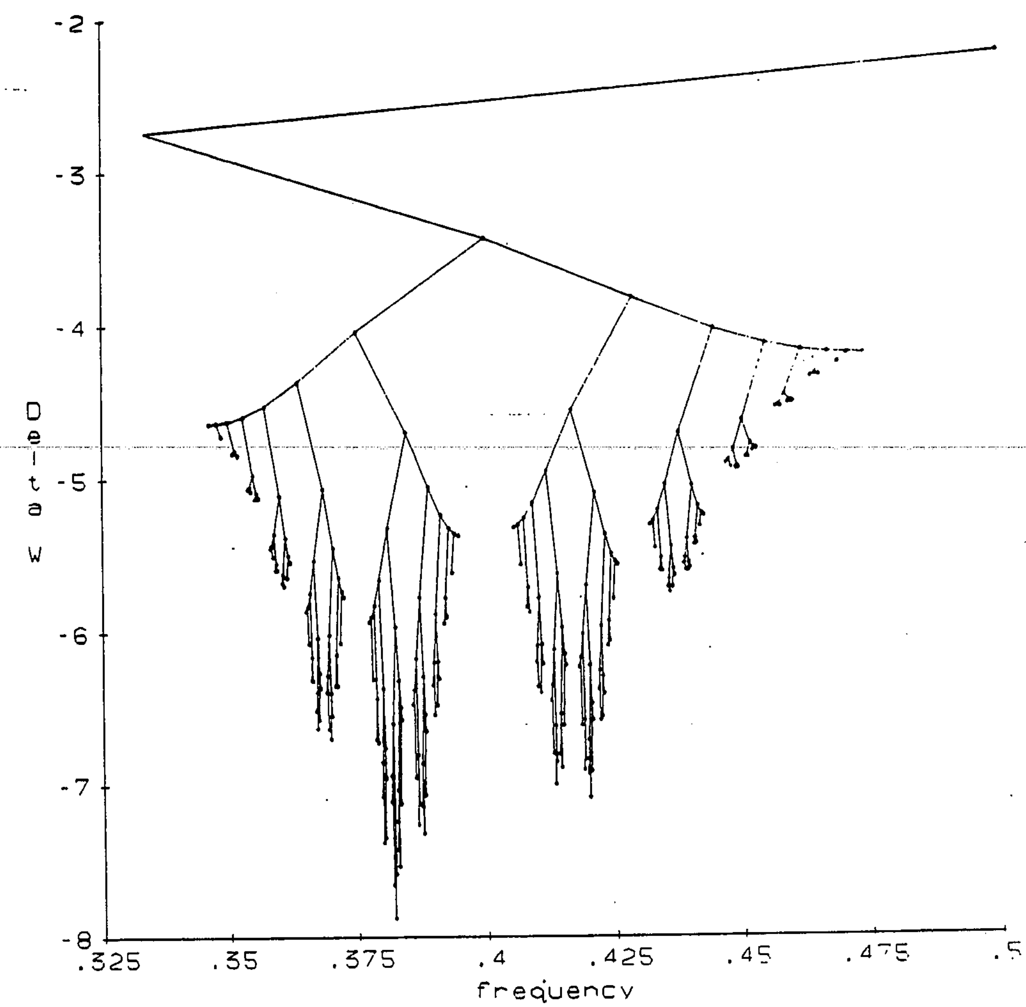

Fig. 2 Farey Tree showing approach of ΔW_ν to irrationals for the standard map at k=0.971635406, when there is exactly one class zero circle. Its frequency is $(3-\sqrt{5})/2$ = [0,2,1,1,1,...] ≈0.38197. $\mathrm{Log}_{10}(\Delta W)$ is shown for levels zero through 8.

where the a_i are integers. A noble has a continued fraction of the form: $\nu = [a_0, a_1, \ldots a_m, 1, 1, 1, \ldots]$, with an infinite tail of one's. That nobles tend to have the smallest values of ΔW in any interval of frequency indicates that the noble cantori are those orbits which minimize the flux. This is the answer to Wigner's query.

MARKOV MODEL FOR TRANSPORT

So far we have directed all our attention to the flux, which is defined by iterating a curve in phase space once with the map. The more difficult problem is the long time properties of the motion in stochastic regions. This is the problem of transport. I first will discuss the usual formulation, and indicate its successes and failures

One of the quantities of interest is the autocorrelation function

$$C(t) \equiv \langle (y_{t+1}-y_t)(y_1-y_0) \rangle = (k^2/4\pi^2)\, \langle \sin(2\pi x_t)\sin(2\pi x_0) \rangle \qquad (4)$$

Here the brackets imply an average over some, as yet unspecified statistical ensemble. From C(t), the diffusion coefficient for y can be obtained

$$D \equiv \lim_{t\to\infty} \frac{\langle (y_t-y_0)^2 \rangle}{2t} = C(0)/2 + \sum_{j=1}^{\infty} C(j) \qquad (5)$$

The last expression is valid only if the correlations fall off faster than 1/t. Note that this formulation gives a global diffusion coefficient, averaged over infinite time. If the motion is bounded between KAM surfaces on the cylinder, then D is identically zero, even though there is still transport within the connected stochastic region.

A local formulation of transport coefficients will be possible only when we use the flux ideas.

Transport has been traditionally calculated by assuming that the motion is random in some sense. The simplest such assumption is that the angle is completely decorrelated from itself after one iteration, e.g. for the standard map for large k. This complete randomness assumption yields the so-called quasilinear value for the diffusion coefficient D_{QL} (= $k^2/4$ for Eq. 1).[5] This calculation can be made more rigorous, by letting the brackets in Eq. (4) denote an average over initial conditions (in the simplest formulation equally distributed over phase space).[11] Correlations can be calculated as a series of products of Bessel functions, and the diffusion coefficient is obtained of the form[11,12] $D = D_{QL}(1 - 2J_2(k) + ...)$. For $k >> 1$ this series apparently agrees well with numerical calculations. Importantly, however, both the calculations and the theory retain only a finite number of time steps in the correlations. As longer calculations are done, one observes anomalies in the diffusion coefficient:[13,14] for many values of k it never seems to converge. These anomalies are most easily seen when there are moderate size islands present in the system. As Chirikov[5,14] has explained, there are elliptic periodic orbits in the standard map for arbitrarily large k! It is unknown whether there is <u>any</u> interval of k for which no elliptic orbits are present.

It is therfore essential to treat transport when there is a mixture of regular and irregular regions. Furthermore, it would be nice to have a formulation which gives local transport coefficients. I believe that the flux description provides such a formulation, though it has yet to be made rigorous.

The problem is to combine the effect of many turnstiles.[2] The simplest model is to assume that only a discrete, well separated set of cantori are important barriers to the motion. This is motivated by

Fig. 2, which shows that to some extent ΔW_{ν} has sharp local minima. The cantorus corresponding to the minumum ΔW in a region of phase space is the limiting barrier for transport through that region. If all the other barriers in the region have much larger values of ΔW, then orbits move through these relatively rapidly. Observations of the motion show that orbits rapidly tend to "fill out" finite regions of phase space, becoming trapped for long periods behind the locally minimizing cantori. Loss of memory in the regions between the minimizing cantori occurs at a rate given by the Lyapunov exponent h, which for the standard map is of order $\ln(k/2)$, $k>1$. Assuming $h\times$(transition time)$<<1$, the transition probability for moving across the minimizing cantorus should be nearly independent of the details of the past history of the motion; it depends only on the orbit being in the region near the cantorus. This is the Markov approximation.

To construct the Markov chain, we divide phase space into states bounded by locally minimizing cantori, and labeled by an index j. The probability of a transition from state i to state j per iteration is

$$p_{ij} = \Delta W_{ij}/A_i \tag{6}$$

where A_i is the area of of the stochastic region in state i. The evolution of the density of particles in each state is given by $N_j(t+1)-N_j(t) = \Sigma\, N_i(t)p_{ij}$. If the transistion probabilities are small compared to 1, then the difference can be replaced by a derivative giving the Markov process:

$$dN_j/dt = \Sigma\, N_i p_{ij} \tag{7}$$

UNIVERSAL EXPONENT FOR CANTORUS CROSSING

Our first application of the Markov model is to compute the transition time across a recently destroyed KAM surface.[2] When a circle is destroyed the flux increases smoothly from zero.[9] The break-up is similar to a second order phase transition, and indeed a version of the renormalization group can be applied to determine properties of the map in the neighborhood of the critical value of the parameter corresponding to destruction.[9]

The renormalization is best understood in terms of convergents of the continued fraction for the frequency. Convergents are obtained by truncating the continued fraction:

$$\nu_m = p_m/q_m = [a_0,a_1,\ldots a_m] \tag{8}$$

Convergents are alternately larger than and smaller than ν. Associated with each convergent is a periodic orbit with its chain of islands. Simply put, the renormalization transformation[9] contracts areas by a factor ξ^{-1}, and maps one island of the m^{th} convergent onto one island of the $(m+1)^{st}$:

$$\mathcal{R}: \begin{cases} \nu_m \to \nu_{m+1} \\ (x,y) \to \mathbf{M}(x,y) : \det(\mathbf{M}) = \xi^{-1} \end{cases}$$

The coordinates (x,y) are measured from a point on the circle, and the matrix M may depend on m. Application of $\mathcal{R}$ many times is equivalent to studying the neighborhood of the circle. For k less than some critical value, k_{cr}, $\mathcal{R}^n\mathcal{T}$ approaches the integrable map, while for $k = k_{cr}$, the value at which the KAM surface is just about to be destroyed, the limit is a non-trivial map with exactly one class zero circle. When k is near k_{cr} the structure of $\mathcal{T}$ in the neighborhood of the circle is determined by the linearization of $\mathcal{R}$. The largest eigenvalue, δ, determines how the effective value of k changes upon renormalization:

$$(k-k_{cr}) \rightarrow \delta(k-k_{cr})$$

The parameters ξ,δ depend on the class of frequencies to which ν belongs. The simplest case is that of the noble numbers, since each successive renormalization after a finite number merely adds a 1 to the continued fraction. In this case the parameters approach constants

$$\xi = 4.339144 \qquad \delta = 1.627950 \tag{9}$$

The noble frequencies yield a self-similar structure at criticality. Recall from Fig. 2, in any region of phase space the last KAM surface to be destroyed is noble. Thus the transport through a region of phase space containing a recently destroyed circle, should scale with noble parameters.

To compute the scaling of the transition probability, we note that ΔW is an area and should scale with ξ. Thus

$$\Delta W_\nu(k_{cr}+\Delta k) = \xi^{-1} \Delta W_\nu(k_{cr}+ \delta\Delta k)$$

This implies that ΔW has the form

$$\Delta W_\nu = C(\Delta k)^\eta \tag{10}$$

times a function periodic in log(Δk) with period log(δ). Here the critical exponent is

$$\eta = \log(\xi)/\log(\delta) = 3.011722 \tag{11}$$

This exponent is independent of the particular map under study. It depends only on the statement that the last circle to be destroyed in the region has noble rotation number.

We estimate the time to cross the noble cantorus from (6) as

$$\tau = 1/p_{ij} = C'(\Delta k)^{-\eta} \tag{12}$$

since the area in a stochastic region does not change dramatically with Δk. This compares favorably with numerical experiments on the golden mean circle of the standard map. Chirikov[5] estimated that the crossing time goes to infinity with an exponent 2.55 computed from data over a range $1<k<4$. We expect that the universal exponent obtains in a somewhat smaller range of k. Similarly, re-analysis[15] of the data for D given by Rechester et al[12] shows that $D \propto (\Delta k)^{\eta}$ is an extremely good fit between $.98<k<2$. The numerical coefficient, C', is given by the area in the stochastic region in which the orbit begins divided by C (= 0.0176). If we estimate the stochastic area as 0.5 of the area of phase space (1.0), then $C' \approx 28$, which is about a factor of 3 below the measured value. Similarly, we find $D_{theory} \approx 3.5 \times D_{exp}$. It is interesting that we have found the same numerical ratio of the simplest theoretical estimate to experiment holds for other maps and other cantori as well.[16]

It is clear that the simple estimate for the crossing time must be below the actual crossing time, since it is necessary to go through the turnstile in the golden mean cantorus. This agrees with the finding that $\tau_{exp} = 3.5\ \tau_{theory}$. However, the large discrepancy in value seems to be due to the neglect of the infinity of other cantori in the neighborhood of the golden mean cantorus. Mather's theorem implies that nearby cantori have nearly the same value of ΔW. Unfortunately, as more of these cantori are included, the simple Markov picture of transport seems no longer valid. In particular, two very close cantori have turnstiles which actually overlap, thus transitions between them are certainly not independent. It appears that it is about 3.5 times harder to get through the structure consisting of a minimizing cantorus and all its neighbors than would be estimated by considering the minimizing cantorus alone!

CORRELATION DECAY, MARKOV MODEL

A typical stochastic region is bounded by KAM surfaces which limit the extent of the motion. These boundary circles have a profound effect on correlations of stochastic particles. In this section we develop a model which includes exactly one boundary circle.[4] We show that even though nearby orbits diverge exponentially, correlation functions do not decay exponentially.

The invariant circle which bounds a stochastic region must be a critical circle because by definition there are no nearby invariant circles on one side.[17] This implies that the critical scaling results of the renormalization group apply in its neighborhood.

On one side of the critical circle there are chains of islands corresponding to the continued fraction convergents, Eq. (8). Supposing the stochastic region corresponds to frequencies less than ν, we define the level of approximation as j=m/2, for m even since only the even convergents are in the region of interest. Between successive levels there are an infinity of cantori corresponding to irrationals in the range $\nu_j<\omega<\nu_{j+1}$. One of these cantori, call it ω_j, will have the minimum flux for that region. In the spirit of the Markov model, suppose that it is the only important cantorus in that region. The j^{th} state in the Markov chain corresponds to the stochastic region contained between ω_{j-1} and ω_j; it has area A_j. For now (see however below) we neglect the effect of the island chain in the j^{th} region. There are an infinity of states with successively smaller fluxes corresponding to $j \to \infty$, see Fig. 3.

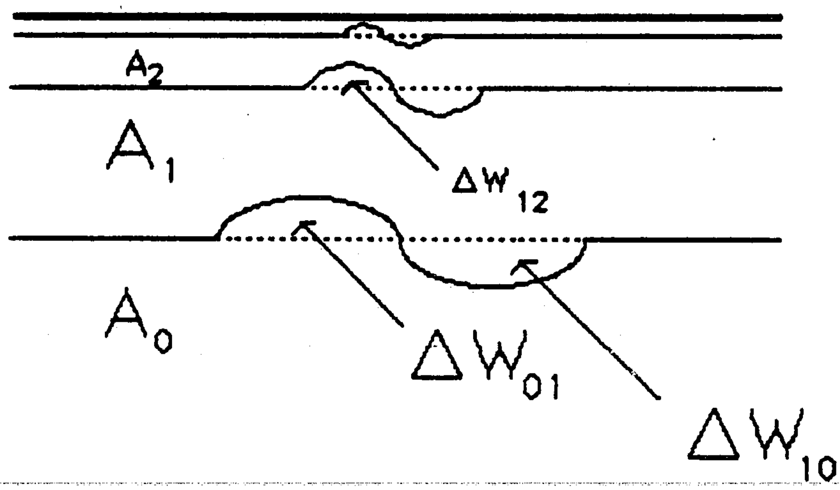

Fig. 3 Markov Chain near a critical circle.

Let us assume for simplicity that the boundary circle has noble frequency (Stark and MacKay[18] discuss the general boundary circle case, and show that the scaling coefficients are not so different from Eq.(9)). The scaling of the transition probabilities follows from the definition (6), and Eq. (8). The function ΔW scales with the area factor ξ

$$\Delta W_{j,j+1} = \Delta W_{0,1}(\xi)^{-2j} \tag{13}$$

where the two follows from m=2j. Since $\mathcal{R}^2$ maps one island at level j onto one at level j+1, the total area in the stochastic region scales as the area coefficient times the number of islands

$$A_j = A_0(g/\xi)^{-2j} \tag{14}$$

Here the g is the golden mean obtained from the result that

$$q_{m+1}/q_m \to g \equiv (1+\sqrt{5})/2 \qquad (15)$$

This yields the probability scaling

$$p_{j,j+1}/p_{j-1,j} = \epsilon_L = g^{-2} = 0.381966 \qquad (16)$$

The probability of becoming more deeply trapped is much less than that of escaping:

$$p_{j,j+1}/p_{j,j-1} \equiv w_L = \zeta^{-2} = 0.053112 \qquad (17)$$

The self-similar Markov chain implied by Eqs. (7),(16), and (17) can be solved exactly. We obtain expressions for the first return distribution,

$$R_{1,0}(\tau) \equiv \mathrm{Prob}(j=0 \text{ at } t=\tau \mid j=1 \text{ at } t=0 \text{ and } j \neq 0 \text{ for } t<\tau)$$

which describes how long a particle initially trapped near the critical circle, remains near the circle. The self similarity implied by Eqs. (16) and (17) shows that higher level return distributions are related to $R_{1,0}$ by

$$R_{j+1,j}(t) = \epsilon_L R_{j,j-1}(\epsilon_L t) \qquad (18)$$

Using this relation one can show[4] that the Laplace transform of $R_{1,0}$ obeys a particularly simple, quadratic equation. The solution has a branch cut at s=0 implying algebraic decay in time:

$$R_{10}(\tau) \to \tau^{-\beta-1} p(\log\tau/\log\epsilon) \qquad (19)$$

where p(z)=p(z+1) is a known periodic function. The decay exponent is given by the equation

$$w_L \epsilon_L^{-\beta} = 1 \Rightarrow \beta = \log(w_L)/\log(\epsilon_L) = 3.0500 \qquad (20)$$

Other return distributions, R_{j0}, decay asymptotically with the same exponent. The survival probability, defined by

$$P_{10}(t) = \int_t^\infty R_{10}(\tau)\, d\tau \rightarrow t^{-\beta} \qquad (21)$$

It is the probability that at time t the orbit is still trapped. This quantity is called the recurrence distribution in Ref. 20. Finally a correlation function is[19]

$$C(t) = \alpha \int_t^\infty P_{10}(\tau)\, d\tau \rightarrow t^{-\beta+1} \qquad (22)$$

where α is fixed to set C(0)=1. This function gives the fraction of particles trapped near the KAM surface at time t, and thus represents those particles which are still correlated with each other. A diffusion coefficient can be defined from C, by Eq. (5). Karney[19] shows that this D represents the contribution to the diffusion from an island in the stochastic sea. Note that D exists only when $\beta>2$, as is true for this case.

MARKOV TREE MODEL

In a typical map any stochastic region is bounded by an infinity of invariant circles. These correspond to the various classes of circles surrounding stable periodic orbits. For example, two class zero circles, with frequencies ν_b and ν_t, will restrict the motion to some segment of the cylinder. Between these frequencies are stable class zero periodic orbits, $\nu_b < p/q < \nu_t$, which will each have a chain of q outermost invariant class one circles. The rotation number with respect to a periodic orbit can be defined as the average angle an orbit rotates about the elliptic point upon iteration with T^q. The island rotation number goes to zero as one moves away from the fixed point towards the separatrix (which is broken, and has a turnstile). This implies that outside a boundary circle of frequency ν_1, there are class one cantori with the island rotation numbers $\nu < \nu_1$. Similarly there are class one periodic orbits with island rotation numbers $p/q < \nu_1$. The structure repeats.

The phase space has a structure of islands about islands, with each class of islands contributing a sequence of cantori approaching its boundary circle, as shown in Fig. 4. In the spirit of our Markov

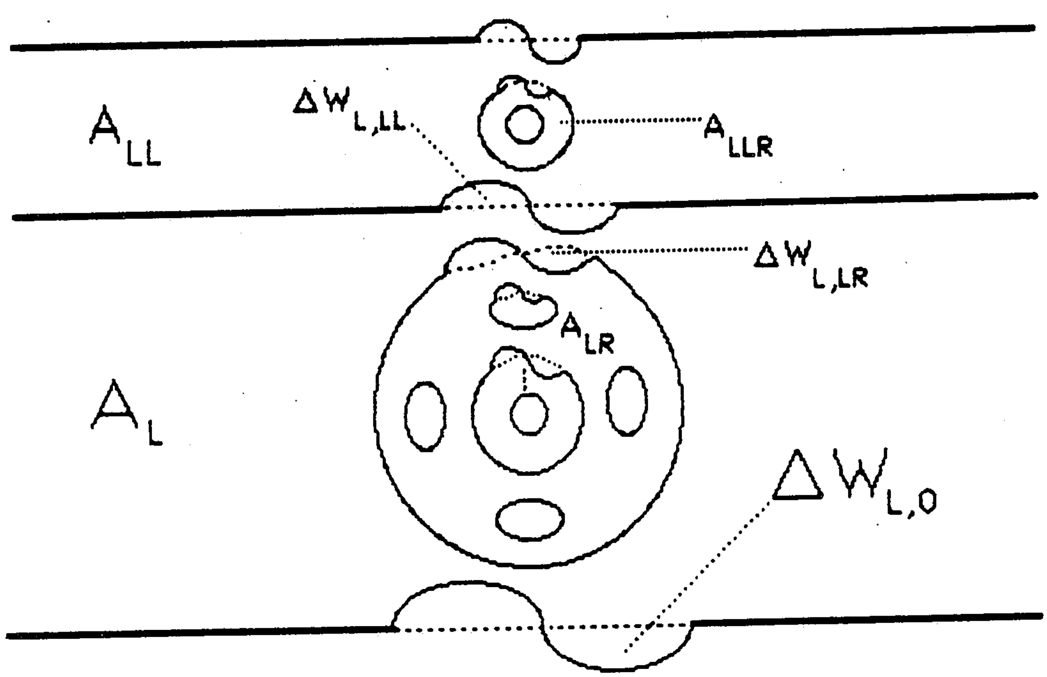

Fig. 4 Tree structure of cantori and islands.

model, we keep only a discrete set of these cantori. The most important islands near a class c boundary circle are the convergents to the boundary frequency. Between each pair of convergents there is a class c cantorus with minimum flux. Their turnstiles connect class c states in the Markov system. The class c+1 boundary circles surrounding islands have sequences of class c+1 convergents. Between each of these convergents there is a class c+1 minimizing cantorus. The outermost such minimizing cantorus is the partial barrier which connects a class c state in the Markov system with the class c+1 state.

The Markov system has the topology of a tree.[21] If there is only one important island between every pair of minimizing cantori, the tree is binary, Fig. 5. A state can be uniquely specified by a sequence of symbols, LLRLRRL..., which determine the path from an

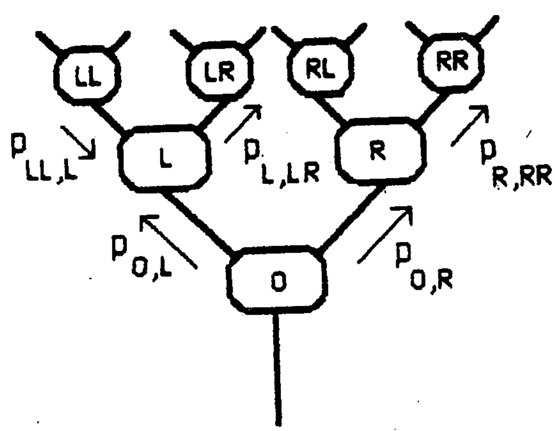

Fig. 5 Markov Tree, several transition probabilities labeled.

initial state, 0. In this picture moving to the left and upward corresponds to a transition from one level of a class c state to the next level of the same class state, or S→SL where S denotes any sequence of L's and R's. Moving up and to the right corresponds to a transition to the outermost level at class c+1, e.g. becoming trapped about an island, S→SR. Note that once the class increases to c+1, an orbit can never escape from the island without going back to S. This is a topological property of area-preserving maps, and would not be expected to hold in higher dimensions.

The parameters in Eqs.(16)-(17) describe the scaling for moving to the left which is why we have labeled them with subscript "L" ("class conserving" transitions). There are two new parameters analagous to these which describe the branching to the right. In the new notation

$$p_{L,LL}/p_{0,L} \equiv \epsilon_L \qquad p_{R,RR}/p_{0,R} \equiv \epsilon_R$$
$$p_{L,LL}/p_{L,0} \equiv w_L \qquad p_{L,LR}/p_{L,0} \equiv w_R \tag{22}$$

The parameter ϵ_R is the time scaling when moving from class c to class c+1, and w_R is the ratio of areas of the class c turnstile to the outermost class c+1 turnstile. These parameters have not yet been determined with the accuracy of the previous ones; however, I believe that there is a rescaling theory corresponding to class changes analagous to the previous one for level, but this has not been shown as of yet. A crude measurement gives

$$w_R \approx 0.03 \qquad \epsilon_R \approx 0.16 \tag{23}$$

The solution for the self-similar binary Markov tree, governed by the scaling Eq. (22) is similar to the previous case. Now there are two important first return distributions, $R_{L,0}(t)$, and $R_{R,0}(t)$, which correspond to particles initialized in two different states. We find that these are related by the scaling

$$f(t) \equiv \epsilon_R\, R_{L,0}(\epsilon_R t) = \epsilon_L\, R_{R,0}(\epsilon_L t)$$

There are also scaling relations analagous to Eq.(18). The Laplace transform again obeys a quadratically nonlinear equation, whose solution in the neighborhood of s=0 has the form

$$f(s) = f_0(1 + as + \dots) + s^{\beta}\,(1 + bs + \dots) + \dots$$

where β obeys the trancendental equation[21]

$$w_L\, \epsilon_L^{-\beta} + w_R\, \epsilon_R^{-\beta} = 1 \qquad (24)$$

similar to (20). The root of (24) with the smallest real part occurs at $\mathrm{Im}(\beta) = 0$, and governs the asymptotic decay in time. Roots with $\mathrm{Im}(\beta) \neq 0$, of which there are infinitely many, provide an oscillatory correction to this decay, analogous to the periodic function in Eq.(19). Using the values of the parameters in Eqs. (16),(17), and (23) gives the decay rate

$$f(t) \to t^{-\beta-1} \qquad \beta = 1.6$$

The revised value of β can be used to compute the survival probability and correlation function from Eqs. (21) and (22). Note that since $\beta<2$ the correlation function falls off more slowly than t^{-1}. This means that the diffusion coefficient, Eq.(5), does not exist! This explains the secular anomalies that are seen in numerical computations of diffusion in the standard map. Of course, if some non-Hamiltonian effects (collisions, radiation, quantum fluctuations, etc.) were to destroy the correlations induced by the islands on a time scale short compared to that for becoming trapped, then the anomalies would be washed out, and diffusion recovered.

Our result can be compared with several numerical experiments. The best is that of Karney,[19] who computed the first return distribution for motion in the neighborhood of an island in the quadratic map. He iterates the map 10^{12} times, using fixed point arithmetic to make sure the map is exactly area preserving and reversible. The first return distribution is observed to decay with an overall algebraic form with $\beta \approx 1.4$ for $t < 10^8$, but has significant oscillatory structure which precludes a numerical evaluation of the exponent. Chirikov and Shepelyanski[20] perform a similar experiment for the standard map at k=0.9716 and obtain an exponent $\beta \approx 1.34$ for $t < 10^5$.

It is sensible that our theoretical result should decay more rapidly than the numerical experiments, since we have neglected some cantori that could trap the orbit and slow the decay. In particular we assumed there is exactly one class c+1 island between each pair of class c cantori, when in fact there are many, perhaps an infinity. If the Markov model is valid, the inclusion of more islands would result in a tree with more branches. The modification of Eq. (24) is straightforward, with an additional term for each branch.

Another possible problem with the theory is the assumption of strict scaling, which holds only in the case of noble boundary circles. I understand,[18] however, that the general boundary circle has a special type of frequency which "looks" noble from the chaotic side, and that the self-similarity is nearly true in this case as well.

CONCLUSIONS

The theory of transport in two-degree-of-freedom systems may allow the prediction of asymptotic escape rates that are all but impossible to obtain numerically. The state of the art in numerical experiments, as described by Karney,[19] involves 20 hours of Cray time for 10^{12} iterations of the simplest conceivable map. Even in this case the asymptotic behavior of the correlation function is not clear. Realistic maps, such as those describing the magnetic lattice of the

SSC for example, are much more complicated and one may need to iterate them more than 10^{12} times.

An advantage of the theoretical results is that the decay exponents predicted are universal. Any two-degree-of-freedom Hamiltonian is expected to give the same results.

Generalization of these results to higher degree-of-freedom systems is an important problem. In this case the dimension of the invariant surfaces is such that they do not divide the energy surface into regions. It is this fact that allows the leaking of trajectories around an invariant torus, or Arnold diffusion.[6] I am optimistic, however, that the ΔW formulation will generalize, and may even give a new picture of Arnold diffusion.

REFERENCES

1. E. Wigner, J. Chem. Phys. **5**, 720 (1937); R. Marcelin, Ann. Physique **3**, 120 (1915).
2. J.D. Meiss, R.S. MacKay, and I.C. Percival, Phys. Rev. Lett. **52**, 697 (1984); Physica **13D**, 55 (1984).
3. I.C. Percival, in Nonlinear Dynamics and the Beam-Beam Interation, M. Month and J.C. Herrera (eds.), American Inst. of Physics Conf. Proc. No. 57, p. 302 (1979).
4. J.D. Hanson, J.R. Cary, J.D. Meiss, J.Stat. Phys. **39**, 327 (1985); J.R. Greene, "A Simple Transport Model," GA Technologies Report #GA-A17511(1984).
5. B. V. Chirikov, Phys. Reports **52**, 265 (1979).
6. See for example M.V. Berry, in Topic in Nonlinear Dynamics, S. Jorna (ed)., American Inst. of Physics Conf. Proc. No. 46, p. 16 (1978); or A.N. Lichtenberg and M.V. Lieberman, Regular and Irregular Motion, (Springer Verlag, New York, 1982).
7. J.M. Greene, J. Math. Phys., **20**, 1183 (1979).

8. 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).
9. R.S. MacKay, "Renormalization in Area Preserving Maps," Ph.D. Thesis Princeton University (University Microfilms, Ann Arbor, Michigan, 1982).
10. J.N. Mather, "A Criterion for the Non-existence of Invariant Circles," preprint, Princeton University (1982).
11. J.R. Cary and J.D. Meiss, Physical Review **24A**, 2664 (1981).
12. A.B. Rechester, M.N. Rosenbluth, and R.B. White, Physical Review **23A**, 2664 (1981).
13. J. D. Meiss, J.R. Cary, C. Grebogi, J.D. Crawford, A.N. Kaufman, and H.D.I. Abarbanel, Physica **6D**, 360 (1983).
14. C.F.F. Karney, A.B. Rechester, and R.B.White, Physica **4D**, 425 (1982).
15. R.B. White, personal communication (1984).
16. J. D. Meiss, J. R. Cary, D. F. Escande, R. S. MacKay, I. C. Percival, and J. L. Tennyson, "Dynamical Theory of Anomalous Particle Transport," in Plasma Physics and Controlled Nuclear Fusion Research 1984, London (International Atomic Energy Agency, Vienna, 1985), in press.
17. R.S. MacKay, personal communication (1984).
18. J. Stark and R.S. MacKay, "Boundary Circles," in preparation.
19. C.F.F. Karney, Physica **8D**, 360 (1983).
20. B.V. Chirikov and D.L. Shepelyanski, Physica **13D**, 395 (1984).
21. J.D. Meiss and E. Ott, "Markov Tree Model of Transport in Area Preserving Maps," in preparation.