3-23-2015

Thirty Years of Turnstiles and Transport

James Meiss  
*University of Colorado Boulder, jdm@Colorado.EDU*

Follow this and additional works at: [http://scholar.colorado.edu/appm_facpapers](http://scholar.colorado.edu/appm_facpapers)  
Part of the [Dynamical Systems Commons](http://scholar.colorado.edu/appm_facpapers), and the [Non-linear Dynamics Commons](http://scholar.colorado.edu/appm_facpapers)

Recommended Citation  
[http://scholar.colorado.edu/appm_facpapers/2](http://scholar.colorado.edu/appm_facpapers/2)

This Article is brought to you for free and open access by Applied Mathematics at CU Scholar. It has been accepted for inclusion in Applied Mathematics Faculty Contributions by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
Thirty years of turnstiles and transport

J. D. Meiss

Citation: Chaos: An Interdisciplinary Journal of Nonlinear Science 25, 097602 (2015); doi: 10.1063/1.4915831
View online: http://dx.doi.org/10.1063/1.4915831
View Table of Contents: http://scitation.aip.org/content/aip/journal/chaos/25/9?ver=pdfcov
Published by the AIP Publishing

Articles you may be interested in
A microscopic model for noise induced transport: Heat-bath nonlinearly driven by external white noise
Chaos 21, 013117 (2011); 10.1063/1.3545373

Transport in time-dependent dynamical systems: Finite-time coherent sets
Chaos 20, 043116 (2010); 10.1063/1.3502450

Weak mixing and anomalous kinetics along filamented surfaces
Chaos 11, 295 (2001); 10.1063/1.1355358

Numerical simulation of multicomponent ion beams in transport lines
AIP Conf. Proc. 480, 189 (1999); 10.1063/1.59499

Entropy balance, time reversibility, and mass transport in dynamical systems
Chaos 8, 396 (1998); 10.1063/1.166322
Thirty years of turnstiles and transport

J. D. Meiss
Department of Applied Mathematics, University of Colorado, Boulder, Colorado 80309-0526, USA

(Received 18 January 2015; accepted 2 March 2015; published online 23 March 2015)

To characterize transport in a deterministic dynamical system is to compute exit time distributions from regions or transition time distributions between regions in phase space. This paper surveys the considerable progress on this problem over the past thirty years. Primary measures of transport for volume-preserving maps include the exiting and incoming fluxes to a region. For area-preserving maps, transport is impeded by curves formed from invariant manifolds that form partial barriers, e.g., stable and unstable manifolds bounding a resonance zone or cantori, the remnants of destroyed invariant tori. When the map is exact volume preserving, a Lagrangian differential form can be used to reduce the computation of fluxes to finding a difference between the actions of certain key orbits, such as homoclinic orbits to a saddle or to a cantorus. Given a partition of phase space into regions bounded by partial barriers, a Markov tree model of transport explains key observations, such as the algebraic decay of exit and recurrence distributions. © 2015 AIP Publishing LLC.

[http://dx.doi.org/10.1063/1.4915831]

The problem of transport in dynamical systems is to quantify the motion of collections of trajectories between regions of phase space with physical significance, for example, to determine chemical reaction rates, mixing rates in a fluid, or particle confinement times in an accelerator or fusion plasma device. For Hamiltonian systems with a phase space containing regular (elliptic islands and tori) and irregular (chaotic) components, transport is impeded by partial barriers bounding resonance zones or by remnants of invariant tori. In the former case, the barriers are formed from the broken separatrices of an unstable periodic orbit, and in the latter, they are formed from similar stable and unstable manifolds of a remnant torus—a cantorus. Thirty years ago, MacKay, Meiss, and Percival discovered that cantori form robust partial barriers for area-preserving maps: in any region of phase space, the most resistant barriers are the cantori. In this review, I discuss this work and survey subsequent developments. While much has been done, there are still many open questions.

I. INTRODUCTION

In this article, I attempt to survey the history of the theory of transport in deterministic, conservative dynamics, as well as to discuss some of the many questions that remain. My interest in this subject arose from a collaboration with MacKay and Percival that began just over thirty years ago with the publication of two papers on “Transport in Hamiltonian Systems.” Like many fruitful ideas, there was some synchronicity to this discovery: we learned at the “Dynamics Days” meeting in June of 1983 [this was the sixth Dynamics Days and the fourth in Enschede, organized Robert Helleman that David Bensimon and Leo Kadanoff were working on a similar approach. I remember that when Robert, Ian, and I celebrated the acceptance of our papers, we thought that they contained ideas for at least ten years of additional research. Some thirty years later, I believe there are still many fruitful, open problems in this field. Much has been done: our articles have been cited 697 times, according to the Web of Science. The work of Vered Rom-Kedar and Stephen Wiggins, which introduced the term lobe dynamics and (with Anthony Leonard) applied these ideas to fluid mixing, has also been extensively cited (443 times).

The problem of deterministic transport also has been studied frequently in the first 24 volumes of Chaos: An Interdisciplinary Journal of Nonlinear Science. Indeed, the terms turnstile or lobe have been mentioned in 114 articles, beginning with an article on transport in a four-dimensional map in the very first issue. It would be impractical to try to review all of this work. Since I wrote a review in 1992, I will concentrate here on describing some of the more recent advances as well as on posing some questions that I hope will be answered before the 50th anniversary of Chaos. However, to make this paper coherent, it is necessary to tell some of the early story.

A theory of transport seeks to characterize the motion of groups of trajectories from one region of phase space to another. Typical dynamical systems have both regular and chaotic (irregular) trajectories. The question “Where does this trajectory go?” is a particularly difficult one to answer for chaotic trajectories since they—by definition—exhibit sensitive dependence upon initial conditions. The intermixtures of regular and irregular dynamics is particularly evident in the Hamiltonian case, where stable periodic orbits are surrounded by elliptic islands that are embedded in the chaotic sea as a fractal.

So how does one study transport? Our original paper set out a method, as follows:

In this paper, we study transport in the irregular components and initiate a theory of the organization inherent in the apparent chaos. This is achieved by...
recognizing a natural division of the irregular components into regions separated by partial barriers formed by joining the gaps in invariant Cantor sets. We discuss the mechanism of transport between these regions.

There were many antecedents to these ideas. For example, in his paper, introducing the quadratic area-preserving map, Hénon computed the set of orbits that do not escape to infinity, a kind of anti-transport result. More direct antecedents include the development of transition state theory in chemistry and especially the notions of Wigner for finding surfaces of minimal flux. For a chemical reaction that has no activation energy and for which classical mechanics is a good approximation, Wigner thought that there should be a surface in phase space that distinguishes between dissociated and reacted states. Unless the surface has “wrinkles,” that is, unless there are trajectories that cross it twice during one collision (similar to what MacKay later called “sneaky returns”), then the one-way flux across this surface gives a bound on the reaction rate.

For the purposes of this exposition, consider a map \( f : M \to M \) on some \( n \)-dimensional manifold \( M \). Most of the ideas here can be applied directly to flows, either directly using a Poincaré or stroboscopic section or by extension to continuous time. For a map, a trajectory is a sequence \( \{ z_t \} = \{ z_0, z_1, z_2, \ldots \} \), such that \( z_t \in M \) and

\[
\tag{3}
z_{t+1} = f(z_t).
\]

We will always assume that \( f \) is a homeomorphism, and typically that it is a diffeomorphism. In addition, we will suppose that \( f \) preserves a volume measure \( \mu \): for any measurable set \( A \)

\[
\mu(f(A)) = \mu(A). \tag{1}
\]

When the map is a diffeomorphism, volume can be computed using a differential form

\[
\Omega = \rho(z) dz^1 \wedge dz^2 \wedge \ldots \wedge dz^n, \tag{2}
\]

with density \( \rho \), i.e., \( \mu(A) = \int_A \Omega \). Then, \( f \) preserves volume when

\[
\tag{3} f^* \Omega = \Omega.
\]

Here, the pullback, \( f^* \), is the local action of the map \( f \) on differentials. An easy mnemonic for how this works is to denote the image as a function \( f'(z) = f(z) \), and then \( f^* dz = df' \) is the differential of this function, i.e., \( f^* dz = \sum Df_i(z) dz^i \). Here, \( Df_{ij} = \partial f_i / \partial z_j \) is the Jacobian matrix of \( f \). The implication is that (3) for the form (2) implies that

\[
\det(Df(z)) \rho(f(z)) = \rho(z).
\]

For the common case that \( \rho(z) = 1 \), this implies that the Jacobian of \( f \) has determinant one.

II. AREA-PRESERVING MAPS

Most of the theory of transport has been developed for the area-preserving case. For an area-preserving map, there are coordinates \( z = (x, y) \) such that the preserved volume form is \( \Omega = dy \wedge dx \). The well-studied preserved volume form on the cylinder \( M = S \times \mathbb{R} \)

\[
(x', y') = \left( x + y \mod 1, \ y - \frac{k}{2\pi} \sin(2\pi x) \right) \tag{4}
\]

is one such family. A typical phase portrait for this map is shown in Fig. 1. The standard map is an example of a twist map on the cylinder: a map for which the image of a vertical line is a graph over the angle. The powerful Aubry-Mather theory applies to this class of maps. Another example is Hénon’s family of area-preserving maps

\[
(x', y') = (-y + 2(a - x^2), y). \tag{5}
\]

Fig. 1. Phase portrait of the standard map (4) for \( k = 0.975 \) on the domain \([-0.5, 0.5] \times [-0.5, 0.5] \). Black trajectories show island chains associated with various elliptic periodic orbits. There are no rotational invariant circles for this value of \( k \). The red trajectory, started near the hyperbolic fixed point at (0.5, 0.0), is iterated \( 10^6 \) times and appears to be confined by rotational circles; however, it will escape after \( 2.3 \times 10^6 \) steps. The blue and yellow trajectories are similarly only temporarily confined. This figure was generated with the StdMap application.
Aref’s blinking vortex\(^5\) for which the area-preserving theory is directly applicable.

### A. Periodic orbits and resonance zones

Periodic orbits for area-preserving maps are typically born in pairs, one a saddle (hyperbolic) and the other elliptic. For example, the standard map \((4)\) for \(k = 0\) becomes the trivial shear \((x, y) \rightarrow (x + y, y)\); it is integrable with action variable \(y\). For this map, each circle on which \(y\) is rational is composed of period-\(n\) orbits since

\[
(x_n, y_n) = f^n(x_0, y_0) = (x_0 + ny_0 \mod 1, y_0) = (x_0, y_0),
\]

when \(y_0 = \frac{m}{n}\) with \(m \in \mathbb{Z}\) and \(n \in \mathbb{N}\). The minimal period of the orbit is \(n\) if \(\gcd(m, n) = 1\). These periodic orbits then have rotation number \(\alpha = \frac{m}{n}\), and we call them \((m, n)\)-orbits.

When the integrable shear is perturbed, at least two orbits survive for any rational rotation number; these are called the Birkhoff orbits, since they are predicted by the famous Poincaré-Birkhoff theorem.\(^{74}\) In Fig. 1, such orbits can be seen at the centers of island chains. For the case of a twist map, like \((4)\), Aubry-Mather theory shows that these orbits always exist; they are the so-called minimizing and minimax orbits because they can be obtained from a variational principle based on the action.\(^{9,69}\) The minimizing orbit is hyperbolic whenever it is a nondegenerate minimum of the action, and the minimax orbit is elliptic when the perturbation is small enough.\(^{81}\)

An elliptic-hyperbolic pair of \((m, n)\)-orbits forms the skeleton of a resonance zone consisting of \(n\) islands surrounding the elliptic points with boundaries formed from the stable and unstable manifolds of the saddle.\(^{31,87}\) One can view an island chain as a region of phase space in which the local rotation number is, roughly speaking, \(\frac{m}{n}\). Computational methods for delineating resonances include Laskar’s frequency map\(^{60}\) and the width of an orbit relative to a given rotation number.\(^{74}\)

Geometrically, one can define a region of phase space that is bounded by broken separatrices of the saddle, an example is shown in Fig. 2. The boundary is formed using points on primary homoclinic orbits. Recall that a homoclinic point

\[
m \in W^s(h) \cap W^u(h)
\]

is primary if there are segments of stable, \(W^s(h)\), and unstable, \(W^u(h)\), manifold from the saddle \(h\) to \(m\) (that is, initial segments) that intersect only at their endpoints\(^{30,99}\) (see Ref. 61 for a generalization to higher dimensions). The resonance boundary in Fig. 2 consists of two initial segments of unstable manifold—for the “upper” and “lower” boundary—connected to two critical islands of stable manifold at primary homoclinic points. The region enclosed by the pair of broken separatrices is the resonance zone. The switch from unstable to stable manifold can occur at any point on the homoclinic orbit \(\{m_i\}\); even though the resulting geometry depends on this choice, the area of the resonance does not. In order for an orbit to leave the resonance zone, it must pass through either the upper, \(E_u\) or lower, \(E_l\), exit set (we will say, more about these in Sec. IIIA). Similarly, an orbit can enter the zone only by first landing in the preimage of one of the incoming sets, \(f^{-1}(I_u)\) or \(f^{-1}(I_l)\). This construction can be done for any pair of \((m, n)\)-orbits; a \((1, 3)\) resonance zone for \((4)\) is shown in Fig. 3.
Each exit and incoming set is a lobe for the resonance, and the set, $E_u \cup f^{-1}(I_u)$, is a turnstile: it acts like a rotating door [The name “turnstile” was suggested to us by Carl Murray]. For each iteration of the map, trajectories escape through $E_u$ and arrive from $f^{-1}(I_u)$.

### B. Cantori

The existence of invariant tori with sufficiently irrational frequency vectors in Hamiltonian systems that are sufficiently close to integrable is guaranteed by KAM theory. However, KAM theory says nothing about what happens to these structures far from integrability. The existence of remnants of these tori after they have been destroyed was first deduced by Aubry and Percival using variational arguments. Percival named the remnants cantori because—for an area-preserving map—they are invariant Cantor sets. Every orbit on a cantorus has the same irrational rotation number.

Figure 4 shows the invariant circle of (4) with golden mean rotation number, $\omega = \phi$, at the threshold of its destruction, $k = k_{cr}(\phi) \approx 0.97163540631$, and at seven larger values of $k$ where it has become a cantorus. For twist maps, the projection of a cantorus onto the angle coordinate, $x$, is a Cantor set on the circle, so it has a countably infinite set of gaps. The orbit of each gap is a family of related gaps. For the standard map, each cantorus appears to have one family, and each gap grows continuously from zero as $k$ increases from $k_{cr}(\omega)$. More generally, there may be more than one family of gaps and invariant circles may reform.

Mather and Aubry showed that for twist maps there is an invariant set for each irrational rotation number that is either an invariant circle or a cantorus. The same analysis applies to optical Hamiltonian flows. It is not clear, however, whether maps that do not satisfy a twist condition have cantori. Cantori do exist for a large class of maps close enough to an anti-integrable limit.

**Question I** (Nontwist Cantori). What is the fate of invariant tori for area-preserving maps that do not satisfy a twist condition? Are there “shearless” cantori for nontwist maps like those studied by Ref. 41?

The anti-integrable limit also can be used to show that certain higher-dimensional symplectic maps and higher-order recurrence relations modeling crystal lattices have analogues of cantori. However, Aubry-Mather theory has not been generalized to these cases.

**Question II** (Higher-Dimensional Cantori). When an invariant torus in a multi-dimensional symplectic map is destroyed, does it immediately become a cantorus, or are there other topological configurations such as Sierpinski curves? The mechanism for destruction of two-tori in volume-preserving maps seems analogous to the symplectic case, but are there remnant cantori?

### III. THE PROBLEM OF TRANSPORT

The inherent loss of predictability in chaotic dynamical systems implies that computation of individual trajectories is neither practical nor is it especially useful. A theory of transport seeks to make predictions about the longtime evolution of ensembles of trajectories and answer questions such as “How long does it take, on average, to get from one region to another?” or “What is the probability distribution of escape times from a region?” Thus, for the chemical reaction $AB + C \rightarrow A + BC$, the first region might correspond to the separated reactants $AB$ and $C$, and the second to the products $A$ and $BC$. In plasma physics, the first region may correspond to the interior of a confinement device such as a tokamak and escape to the divertor.

The first step in a theory of transport is the computation of the flux across codimension-one surfaces. For chaotic maps, this idea was used by Channon and Lebowitz who computed escape times for Hénon’s quadratic map. They used a broken separatrix of a period-five orbit to construct a simple, closed curve through which orbits can escape. In this section, I review the notions of flux and net flux for volume-preserving and exact volume-preserving maps.

It is important to note that one of the fundamental measures of chaos, the Lyapunov exponent is not usually a good measure for transport. In Ref. 86, we said:

The Lyapunov exponents give the rate at which a nice cat in a region of phase space turns into a mixed-up cat. The turnstile rate constants give the mean rate at which bits of the cat are transported to regions of the phase space.

### A. Exit sets and flux

A first step in a global theory of transport is to quantify the motion of an ensemble for short times by computing the flux of trajectories across codimension-one surfaces in phase space. Suppose $R$ is a region with a piecewise smooth
boundary $\partial R$. When $R$ is invariant, $f(R) = R$, and $\partial R$ is a “complete” barrier. When $R$ is “almost invariant,” its boundary is a partial barrier: though trajectories may escape, the leakage is slow. For a flow, the flux through a surface is computed by integrating the normal component of the velocity over the surface.\textsuperscript{68,82} For a map, the flux is the volume that escapes from the region upon each iterate, i.e., the volume of points that have just entered $I$ for a region $R$. Similarly, the incoming set $I$ for a region $R$ is the set of points that have just entered

$$E = \{ z \in R \mid f(z) \notin R \} = \{ z \in R \mid z \notin f^{-1}(R) \} = R \setminus f^{-1}(R), \quad (6)$$

see Fig. 5(a). The flux of escaping orbits from $R$ is just the volume of the exit set

$$\Phi(R) = \mu(E) = \int_E \Omega \quad (7)$$

Similarly, the incoming set $I$ for a region $R$ is the set of points that have just entered

$$I = R \setminus f(R).$$

Formally, $R$ is almost invariant in the sense of measure if $\Phi(R) \ll \mu(R)$. One technique to compute almost invariant sets uses the Perron-Frobenius operator; this idea was pioneered by Delnitz and Jünger.\textsuperscript{28} In lieu of this method, we will construct regions using resonance zones and cantori, for which the flux can be localized to turnstiles, see Sec. IV.

A volume form (2) $\Omega$ is exact if

$$\Omega = dx \quad (8)$$

for some $(n-1)$-form $x$. For example, the standard map (4) is volume preserving with $\Omega = dy \wedge dx$. Since $x \in S^1$ is periodic, we cannot integrate $\Omega$ with respect to this variable, but the Liouville form $x = ydx$ is well-defined on the cylinder and satisfies (8). When (8) holds, Stokes’s theorem can be used to reduce volume integrals to surface integrals.

The boundary of the exit set, $\partial E$, is the union of two oriented subsets $\mathcal{U} \subset \partial R$ and $\mathcal{S} \subset f^{-1}(\partial R)$ as sketched in Fig. 5(a). For a resonance zone, these are pieces of unstable and stable manifolds, recall Fig. 2. Assuming that the boundaries are chosen with an orientation consistent with that of $R$, then $\partial E = \mathcal{U} - \mathcal{S}$, and we can use (8) to rewrite (7) as

$$\Phi(R) = \mu(E) = \int_E dx = \int_{\partial E} x = \int_{\mathcal{U}} x - \int_{\mathcal{S}} x. \quad (9)$$

For two-dimensional maps, $\mathcal{U}$ and $\mathcal{S}$ are curves, and the line integrals in (9) are more efficient to compute than the original two-dimensional integral over $E$.

In a sense, the complete transport problem is solved if one knows the orbits of $E$ and $I$ for each region of interest. For example, the minimum transit time from $A$ to $B$ is the first time that $f^t(E_A) \cap I_B \neq \emptyset$. Similarly, a complete description of transport through a region requires knowing only the future history of the entering trajectories, i.e., the sets $f^t(I)$. These ideas form the basis for the theory of lobe dynamics as formulated by Rom-Kedar and Wiggins.\textsuperscript{97,99} As we will recall in Sec. V, this is especially useful in the study of transport through resonance zones.

B. Exact maps and net flux

When a region $R$ has finite measure, it is easy to see that the net flux through $\partial R$ vanishes, so that the measure of the incoming and exit sets must be the same

$$\mu(I) = \mu(R) - \mu(R \cap f(R)) = \mu(R) - \mu(f^{-1}(R) \cap R) = \mu(R) - \mu(R \cap f^{-1}(R)) = \mu(E),$$

where we used (1) and (6).

However, when the phase space is not contractible there may exist closed surfaces that that do not bound a region. For example, the standard map (4) models a kicked rotor for which the canonical coordinate is an angle, and the phase space is the cylinder $S \times \mathbb{R}$. A rotational circle is a simple closed curve that is not contractible, for example, $C_0 = \{(x, 0) : x \in S \}$. An area-preserving map on the cylinder need not have zero net flux through rotational circles; a
simple example is the map \( f(x, y) = (x, y + 1) \) that shifts each rotational circle upward but still has \( \det Df = 1 \).

Zero net flux is equivalent to exactness: a map \( f \) is exact volume-preserving when there exists a Lagrangian form \( \lambda \) such that

\[
f^*\alpha - \alpha = d\lambda,
\]

where \( \alpha \) satisfies (8) (recall the discussion of pullback in Sec. I). The Lagrangian for a two-dimensional map is a zero-form—a function—and is the discrete-time analog of the continuous-time Lagrangian of mechanics. Indeed, for a flow, \( \lambda \) is the time-integral of the Lagrangian along a trajectory between sections.\(^{86,115}\)

For example, using \( \alpha = ydx \), the standard map (4) is exact with the Lagrangian

\[
\lambda(x, y) = \frac{1}{2}y^2 + \frac{k}{4\pi} \cos(2\pi x),
\]

where \( y'(x, y) = y - \frac{k}{2\pi} \sin(2\pi x) \). This follows because

\[
f^*\alpha - \alpha = y'dx - ydx = y'd(x + y') - ydx = y'dy + \left( y - \frac{k}{2\pi} \sin(2\pi x) \right) dx - ydx = d\left( \frac{1}{2}y^2 + \frac{k}{4\pi} \cos(2\pi x) \right),
\]

the differential of (11). In our original work for area-preserving twist maps, we wrote \( \lambda = F(x, x') \), which is one of the standard generating functions for canonical transformations.\(^{36}\) I first learned of the more general formulation from Bob Easton.\(^{31}\)

For \( n \)-dimensional maps \( \lambda \) is an \((n - 2)\)-form. For example, the ABC map\(^{36}\)

\[
\begin{pmatrix}
x'
\end{pmatrix} = \begin{pmatrix}
x + A \sin z + C \cos y \\
y + B \sin x' + A \cos z \\
z + C \sin y' + B \cos x'
\end{pmatrix},
\]

preserves the standard volume \( \Omega = dx \wedge dy \wedge dz \), a fact that is most easily seen by viewing it as the composition of three shears, \( f = s_z \circ s_y \circ s_x \), where, \( s_z(x, y, z) = (x + A \sin z + C \cos y, y, z) \), etc. Though it is natural to view (12) as a map on the three-torus, it can also be thought of as a map on the cover \( M = T^2 \times \mathbb{R} \). Now the volume form \( \Omega \) is exact with \( \alpha = zdx \wedge dy \), and (12) is also exact with Lagrangian one-form

\[
\lambda = A[\sin(z)(dx' + zdy) + \cos(z)(dy - zdx')] + B \sin(x')dy' + C \cos(y')dx'.
\]

Here, the primed variables are to be thought of as the functions defined in (12). Such a representation is especially pertinent when \( B, C \ll 1 \), since then \( z \) is an action-like variable (invariant when \( B = C = 0 \)), and one can use it to compute fluxes that correspond to vertical motion.

It is also possible to generalize the idea of canonical generating functions to generating forms for volume-preserving transformations when a generalized twist condition is satisfied.\(^{62}\)

**Question III (Generating Forms). Can Lagrangian forms be used for perturbation theory and near-integrable normal forms, just as canonical generators are used for the symplectic case? Can they be used to build volume-preserving integrators, like those found for area-preserving maps in Ref. 115?**

When \( f \) is homotopic to the identity, the net-flux “across” a closed \((n - 1)\)-dimensional surface \( S \) has meaning: it is the signed volume contained in the region \( R \) bounded by \( S \) and \( f(S) \). If \( \partial R = f(S) - S \), Stokes’s theorem implies that

\[
\Phi_{Net}(S) = \int_{\partial R} \alpha = \int_S (f^*\alpha - \alpha) = \int_S d\lambda = 0,
\]

because the integral of a total differential (exact form) over a closed surface \( S \) vanishes. For example, in Fig. 5(b), the rotational circle \( C \) on the cylinder is the upper boundary of an unbounded set with \( y \to -\infty \). This set has exit set \( E \), the region below \( C \) whose image is above \( C \), and incoming set \( I \), the region below \( C \) whose preimage is above \( C \). The region contained between \( f(C) \) and \( C \) is \( R = f(E) \cup I \), and when the map has zero net flux (13) implies that \( \mu(f(E)) = \mu(E) = \mu(I) \).

Maps with nonzero net flux also have interesting transport properties.\(^{35}\)

**IV. ACTION AND FLUX**

Perhaps the most surprising aspect of the calculation of Ref. 86 is the connection between Mather’s action difference and the flux of trajectories across curves in phase space.

In the general context of an exact volume-preserving map (10), let \( S \) and \( U \) be a pair of codimension one surfaces that together bound a region \( E \), as sketched in Fig. 6. Suppose that

![FIG. 6. A three-dimensional lobe bounded by a pair of surfaces \( U \) (dark red) and \( S \) (light blue) that intersect along a loop \( B \) on their common boundaries.](image-url)
$B = U \cap S = \partial U = \partial S$

is the common $(n-2)$-dimensional boundary of the surfaces. The volume of $E$ is given by the difference between integrals over the surfaces as in (9). Each of these $(n-1)$-dimensional integrals can be partially integrated using (10) with Stokes’s theorem; for example,

$$\int_S \zeta = \int_S (-d\lambda + f^*\zeta) = -\int_B \lambda + \int_{f(S)} \zeta.$$ (14)

This iterative formula can be used to compute lobe volumes in a number of different contexts. For example, suppose that $U = f^n(S)$; then, we can iterate (14) $n$-times to obtain

$$\int_S \zeta = -\sum_{i=0}^{n-1} \int_{f^i(B)} \lambda + \int_{f^n(S)} \zeta.$$ (15)

Note that the final integral in (15) is, by assumption, exactly the integral over the surface $U$, and so from (9) we have

$$\mu(E) = \int_U \zeta - \int_S \zeta = \sum_{i=0}^{n-1} \int_{f^i(B)} \lambda.$$ (16)

The last sum represents a generalized action of the periodic boundary $B = f^n(B)$.

A second case in which (14) is useful is when $S$ is a patch of stable manifold of some lower-dimensional invariant set, i.e., $S \subset W^s(A)$, where $A$ is a normally hyperbolic invariant manifold. In this case, the limit of the remainder integral in (15) as $n \to \infty$ is zero, and thus

$$\int_S \zeta = -\sum_{i=0}^{\infty} \int_{f^i(B)} \lambda.$$ (17)

is the forward action sum. Similarly, if $U \subset W^u(B)$ is a patch of unstable manifold of an invariant set $B$, it is appropriate to reverse (14) to iterate backwards and obtain

$$\int_U \zeta = -\sum_{i=-\infty}^{-1} \int_{f^i(B)} \lambda.$$ (18)

Lomelí and I showed how to compute flux for a 3D volume-preserving map, and MacKay showed how to compute flux over a saddle in a Hamiltonian flow (a result implicit in Ref. 105, as well). Nevertheless, the general problem in higher dimensions still has many mysteries.

**Question IV (Multidimensional Flux).** Can one use the flux formulas to compute flux through 2D manifolds in a 3D map, see e.g., Ref. 80, and through resonance zones formed from 2D normally hyperbolic invariant sets in a 4D symplectic map, see, e.g., Ref. 42?

**A. Flux through periodic orbits**

When $n = 2$, $B$ in (14) degenerates to a set of points, and $\int_B \lambda$ reduces to evaluation at its upper and lower limits. For example, suppose that $U$ and $S$ are curve segments joined at their endpoints, points on a pair of period-$n$ orbits. An example would be, e.g., the pair $\{e_0\}$ and $\{h_0\}$ that form an island chain, recall Fig. 3. In this case, $\lambda$ is a function, and the boundary $B = e_0 - h_0$ is a pair of points, $e_0$ on the minimax orbit and $h_0$ on the hyperbolic orbit. Then, (16) becomes

$$\mu(E) = \sum_{i=0}^{n-1} [\lambda(e_i) - \lambda(h_i)] = \Delta W[e, h]$$ (periodic orbit flux) (19)

the difference between the actions of the two orbits.

Indeed, $\Delta W[e, h]$ is the flux “through” the pair of orbits, as can be seen by constructing (if possible) a region $R$ with boundary $C$ that satisfies two hypotheses.

(H1) Suppose $C = \partial R$ is a simple closed curve alternately joining successive points on a pair of period-$n$ orbits, $\{e_i\}$ and $\{h_i\}$.

(H2) Suppose that $f(C) \cap C = \{e_i\} \cup \{h_i\}$.

An example of such a curve is sketched in Fig. 7. The requirement (H2) means that the exit and incoming lobes are bounded by segments of $C$ and $f(C)$ that connect pairs of points on the two orbits. By definition (6), $f(E)$ is the region outside $C$ and inside $f(C)$. Each of the $n$-lobes of the image of the exit set is bounded by $U_t - S_t$, where $S_t(U_t)$ is the segment of $C (f(C))$ from $h_t$ to $e_t$. Since $f(S_t) = U_{t+1}$ and $f(S_{t-1}) = U_t$, (10) implies that the escaping flux is

$$\Phi(R) = \mu(f(E)) = \sum_{i=0}^{n-1} \int_{U_{t+S_t}} \zeta = \sum_{i=0}^{n-1} \int_{S_t} \lambda = \sum_{i=0}^{n-1} \int_{S_t} \lambda(e_i) - \lambda(h_i)) = \Delta W[e, h].$$

Thus, the flux crossing $C$ is exactly the action difference of (19). It is especially interesting, I think, that the flux is independent of the choice of curve $C$, given that it satisfies (H1).
and (H2). Such curves are easy to construct for twist map examples, where the \((m, n)\) orbits are ordered relative to \(x\), but it is not obvious that it is always possible to construct such curves for a pair of period \(n\) orbits.

**Question V** (Periodic Flux). Under what conditions is it possible to find a curve \(C\) that satisfies (H1) and (H2)?

### B. Separatrix lobe and resonance areas

The iterative formula (14) can also be used to compute lobe and resonance zone areas for 2D maps. Recall from Sec. II A that a resonance zone is bounded by segments of stable and unstable manifold of a hyperbolic periodic orbit joined at a primary homoclinic point. A separatrix lobe is a region bounded by segments of stable and unstable manifold that intersect on a neighboring pair of primary homoclinic points.

More generally, following Ref. 31, suppose that \(R\) is a region bounded by \(K\) alternating segments of stable and unstable manifold, thus, \(\partial R = \sum_{j=1}^{K} (U_j + S_j)\), where the lengths of \(f^t(S_j)\) and \(f^{-t}(U_j)\) limit to zero as \(t \to \infty\), since they are stable and unstable segments, respectively. Denote the \(2K\) “corner points,” where these curves are joined, by \(m_j\) and \(h_j\), \(j = 1, 2, \ldots, K\), so that \(\partial U_j = m_j - h_j\) and \(\partial S_j = h_{j+1} - m_j\), with \(h_{K+1} = h_1\). Labeling the orbits of the corner points with subscripts, \(f^t(m_j) = m'_j\), as usual, then (17) and (18) become

\[
\int_{S_j} \alpha = -\sum_{t=0}^{\infty} (\lambda(h_j^{t+1}) - \lambda(m'_j)),
\]

\[
\int_{U_i} \alpha = -\sum_{t=-\infty}^{-1} (\lambda(m_j^t) - \lambda(h_j^t)).
\]

Finally, summing over the \(K\) alternating segments to obtain the full boundary of \(R\) implies, from (9)

\[
\mu(R) = \int_{\partial R} \alpha = \sum_{t=-\infty}^{\infty} \sum_{j=1}^{K} (\lambda(m_j^t) - \lambda(h_j^t)).
\]

(20)

For example, suppose that \(f\) and \(p\) are saddle fixed points and that \(U \subset W^u(f)\) and \(S \subset W^s(p)\) are segments that intersect on a pair of heteroclinic points \(h_0\) and \(m_0\) as shown in Fig. 8. Then, (20) applies with \(K = 1\): there are just two segments and two corner points. Thus, the area of the lobe is the action difference

\[
\mu(R) = \sum_{t=-\infty}^{\infty} (\lambda(m_t) - \lambda(h_t)) = \Delta W[m, h],
\]

(heteroclinic lobe area).

(21)

This applies to the computation of the lobe area shown in Fig. 2. The same formula applies when the saddles are periodic orbits.

The action principle can also be used to compute the area of a resonance zone as well, recall Sec. II A. Each island in the resonance is bounded by four segments of manifold and so the result (20) applies with \(K = 2\). There are four corner points: \(h'\) and \(h''\), the “left” and “right” saddle points of one island, and \(m',\), the “upper” and “lower” homoclinic points. For example, for the period-three resonance of Fig. 3, the leftmost island has corners \(h' = h_2, m_2, h'' = h_0,\) and \(m_1\). Since each of the islands have the same area, the total area of the resonance is \(n\) times that of one island

\[
\mu(R) = n \sum_{t=-\infty}^{\infty} (\lambda(m_t^+) - \lambda(h_t^+) + \lambda(m_t^-) - \lambda(h_t^-)),
\]

(period-\(n\) resonance area).

(22)

The result was first used to compute resonance areas in Ref. 87. Explicit results were obtained for the sawtooth map, a piecewise linear version of (4).

For a twist map like (4), the rotational periodic orbits are vertically ordered according to their rotation numbers, and there is a resonance zone for each \((m, n)\)-minimizing/ minmax pair of periodic orbits. A choice of homoclinic points can be made so that different zones intersect only in their turnstiles, and since the net flux vanishes, the areas of the two intersections must be equal. Thus, one can partition phase space into rotational resonance zones. An example is shown in Fig. 9. Chen showed that when there are no rotational invariant circles, the total area of the infinitely many resonance zones is the full area of phase space, that is, the partition is “complete.” One way to visualize this is to plot the area below each rotational curve as a function of rotation number.

Rotational curves can be formed from invariant circles, cantori, or broken separatrices. This area function jumps by the value (22) at every rational rotation number. The partition is a complete devil’s staircase when the total rise of the staircase is the sum of the jumps.

**Question VI** (Resonance Partition). Is the resonance partition complete for nontwist area-preserving maps when there are no rotational invariant circles? Is there a similar (complete) partition for one-action, volume-preserving maps? How could one construct a resonance partition for 4D symplectic maps (some of the intriguing geometry has been recently explored in Ref. 65)?
This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 174.16.144.241 On: Mon, 23 Mar 2015 14:39:24

C. Cantori as partial barriers

During the early years of the study of chaos, it was commonly observed that transport in a mixed phase space is highly nonuniform and is particularly slow as an orbit wends its way through regions containing recently destroyed tori. Indeed, Chirikov’s famous numerical experiments for the standard map showed that the time to cross from \( y \approx 0 \) to \( y \approx 0.5 \) goes to infinity as \( k \downarrow 0.989 \),\(^{22}\) recall Fig. 1. Thus, it seemed reasonable to speculate that cantori, even though they are typically hyperbolic, could act as strong barriers to transport. The importance of cantori was confirmed in Ref. 86 when we discovered that the flux is locally minimal on cantori. We also were able to explain Chirikov's result, showing that the average crossing time for (4) has a power law singularity with exponent 3.05 at \( k_c(\phi) \approx 0.971 \).

For symplectic twist maps, every orbit on an invariant Lagrangian torus is a minimum of the action.\(^ {88}\) For the area-preserving case, as noted in Sec. II B, a destroyed invariant circle becomes a cantorus: the minimizing orbits persist and are typically hyperbolic. Suppose that a cantorus \( T_\omega \) with rotation number \( \omega \) has is a single family of gaps: there are two minimizing orbits \( \{l_i\}, \{r_i\} \subset T_\omega \) that correspond to the left and right endpoints of the gaps. When the cantorus is hyperbolic there exist \( c, \lambda > 0 \) such that \( \| r_i - l_i \|. < c e^{-\lambda |i|} \). Aubry-Mather theory implies that there is another distinguished orbit \( \{m_j\} \), the minimax orbit, that falls in the gaps of the cantorus and is homoclinic to the orbit of gap endpoints, i.e., \( m_j \to l_i \) as \( t \to \pm \infty \).

The flux through \( T_\omega \) can be computed using the general action principle (20). Suppose \( C \) is any curve that satisfies hypotheses (H1) and (H2): it is a simple closed curve through the cantorus that crosses each gap going through the minimax point, and \( f(C) \cap C = T_\omega \cup \{m_i\} \) is exactly the cantorus and the minimax orbit. The same analysis as before implies that the flux crossing \( C \) is the action difference

\[
\Phi(T_\omega) = \sum_{l=\infty}^{\infty} (\lambda(m_l) - \lambda(l_l)) = \Delta W[m, l],
\]

(cantorus flux).

This action difference is also prominent in Aubry-Mather theory: when it is zero \( T_\omega \) is an invariant circle, and when it is positive, \( T_\omega \) is a cantorus.\(^ {9,69}\)

Computing invariant tori is difficult since orbits with irrational rotation numbers are infinitely long. One algorithm uses a Newton-like method to compute the Fourier series for the conjugacy to rigid rotation of an invariant torus.\(^ {48}\) However, this method does not seem to permit the computation of cantori. Another algorithm approximates a torus by periodic orbits. For the 2D case, this can be done using the continued fraction expansion \( \omega = [a_0, a_1, \ldots] \), with \( a_0 \in \mathbb{Z}, a_i \in \mathbb{N} \), of the rotation number to give a sequence of rationals

\[
\frac{m_i}{n_i} = [a_0, a_1, a_2, \ldots a_j] \equiv a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \cdots + \frac{1}{a_j}}},
\]

the convergents of \( \omega \). A similar technique uses the Farey tree, which, beginning with an interval [\( m_0, m_1 \)] bounded by neighboring rationals, \( m_0n_1 - m_1n_0 = 1 \), constructs the daughter \( m_2 = m_0 + m_1 \), and \( n_2 = n_0 + n_1 \), with \( \frac{m_0}{n_0} < \frac{m_2}{n_2} < \frac{m_1}{n_1} \).

Repeating this construction on the two new intervals generates a binary tree that contains every rational in the original interval and for which every irrational is a limit of an infinite path.

The flux through a cantorus can be approximated as the flux through a sequence of approximating \((m, n)\)-orbit pairs, using (19). An example of the resulting flux function on the Farey tree is shown in Fig. 10 for \( k = k_\phi(\phi) \) when the golden mean invariant circle is at the threshold of destruction. What is remarkable is that this function appears to monotonically decrease: the flux through a Farey daughter is always smaller than that through each of its parents. For a Farey path that limits to the rotation number of an invariant circle, the flux goes to zero; this is the case in Fig. 10 for \( \omega = \phi^{-2} \), which has the path

\[
\begin{align*}
1 & 1 1 2 3 5 8 13 \\
2 & 3 5 8 13 21 34 & \\
\end{align*}
\]

When the Farey path limits to the rotation number of a cantorus, the periodic flux (19) limits to (23). Thus, the strongest barriers, corresponding to the smallest flux, are associated with cantori.

An alternative notion of minimal flux through cantori was formulated by Polterovich.\(^ {94}\)

**Question VII** (Cantori and Minimal Flux). Find a category of twist maps for which the flux is provably a monotone decreasing function on the Farey tree. For example, if the map is analytic, is the flux asymptotically monotone?

Though there is no completely satisfactory generalization of continued fractions to higher dimensions, Kim and
Olstund invented a binary tree that has most of the properties of the single frequency Farey tree (it does not have a unique path to each rational vector). Even though there have been many numerical studies of the break-up of invariant tori for symplectic maps and volume-preserving maps, there is no conclusive evidence for the formation of remnant tori that are Cantor sets.

**Question VIII** (Higher-Dimensional Cantori). Is there a multidimensional generalization of the Farey tree or continued fraction that is optimal for approximating tori and cantori? After destruction, is a torus replaced by a remnant invariant set with the same rotation vector; if so, what is its topology?

## V. EXIT AND RETURN TIMES

Though a region’s escaping and entering fluxes give one measure of transport, fluxes represent only the first step, both dynamically and theoretically. More global measures are obtained from distributions of exit, crossing, and return times. Early numerical experiments were done by Channon and Lebowitz, who showed that the exit time distribution from a resonance zone for the Hénon map (5) has a surprising power law form. This observation was greatly extended by Karney, who used integer arithmetic to eliminate floating point errors, and also by Chirikov and Shepelyansky. From the many subsequent observations of power law decays, it seems clear that this phenomenon is due to the presence of stable structures within the region of interest: these structures are extremely sticky. However, there is still much controversy about whether the asymptotic decay is truly a power law, and—if it is—if there are universal values for the exponent.

### A. Exit time distributions

Given a subset $A$ of phase space $M$, the (forward) exit time of a point $a \in A$ is the number of iterates for it to leave

$$
\tau^+_A(a) = \min_{t=0} \{ t : f^t(a) \notin A \}. 
$$

Similarly, the backward exit time is

$$
\tau^-_A(a) = \min_{t=0} \{ t : f^{-t}(a) \notin A \}. 
$$

In either case, the exit time is set to $\infty$ if the images remain in $A$ for all $t \in \mathbb{N}$. By definition (6), the exit set, $E$, of a region is the set with forward exit time one, and the incoming set, $I$, is the set with backward exit time one. The transit time is the time for an orbit to move from the incoming to the exit set of $A$; since this is constant along an orbit we can define

$$
\tau^+_A(a) = \tau^-_A(a) + \tau^-_A(a) - 1,
$$

subtracting the 1, to measure the time to get from $I$ to $E$. Finally, the return time for a point $a \in A$ is

$$
\tau^+_A(a) = \min_{t=0} \{ t : f^t(a) \in A \}. 
$$

Note that if $a$ is not in the exit set, its return time is 1 since the minimum is taken over positive $t$.

Chaos by its very definition precludes accurate long-time computations of (25)–(27). Nevertheless, one would like to obtain estimates of their distributions. Whenever $A$ is finite, nonzero measure, its exit time distribution

$$
\mu^+_A(j) = \frac{\mu(\{a \in A : \tau^+_A(a) = j\})}{\mu(A)} \tag{28}
$$

is the probability of exiting in time $j$. Another common measure of transport is the distribution of recurrence times: the probability that an orbit beginning in $A$ has first return time $j$

$$
\mu^-_A(j) = \frac{\mu(\{a \in A : \tau^-_A(a) = j\})}{\mu(A)} \tag{29}
$$

It is not hard to relate this distribution to others that we have defined. If a point in $A$ is not in the exit set, then its return time is 1, so we have

$$
\mu^-_A(1) = \frac{\mu(A\setminus E)}{\mu(A)} = 1 - \frac{\mu(E)}{\mu(A)}.
$$

Points that exit $A$ are in $E$, and $f(E)$ is the incoming set for $M \setminus A$. The first return to $A$ thus occurs when the orbit transits $M \setminus A$, so

$$
\mu^-_A(j) = \frac{\mu(E)}{\mu(A)} \mu^-_{M\setminus A}(j-1), \quad j > 1. \tag{30}
$$

Thus, the return time distribution to $A$ is essentially the transit time distribution through $M\setminus A$.

To compute the statistics of exit and return times, any invariant subset of $A$ should be excluded since it has infinite exit time. For example, the islands of stability that generically surround any elliptic periodic orbit should be excluded from exit time computations for the resonance zone of an area-preserving map. It is convenient to define the accessible
subset of \( A \) is the subset of points that can be reached from the outside, or equivalently, the set of points with finite backward exit time
\[
A_{acc} \equiv \{ a \in A : \tau^{-}(a) < \infty \}.
\]

This set has the same volume as the set with finite forward exit time. In order that the distribution (28) be normalized over \( \mathbb{N} \), it should be redefined using \( A_{acc} \) instead of \( A \).

In 1947, Mark Kac proved a result for stochastic processes that can be translated into a result for deterministic maps. Namely, suppose that \( M \) has finite measure and the invariant measure is normalized so that \( \mu(M) = 1 \). If \( M_{acc} \) is the subset of \( M \) that is accessible to orbits beginning in \( A \), then the average first return time to \( A \) is
\[
\langle \tau^{\text{return}} \rangle_{A} = \frac{1}{\mu(A)} \int_{A} \tau^{\text{return}} \, d\mu = \frac{\mu(M_{acc})}{\mu(A)}.
\]

Using the relation between transit time and return time distributions (29), gives the equivalent result:

**Lemma 1 (Average Exit Time).** If \( A \) is a measurable set with incoming set \( I \) and accessible set \( A_{acc} \) the average exit time for points in the incoming set is
\[
\langle \tau^{+} \rangle_{I} = \langle \tau^{\text{return}} \rangle_{I} = \frac{\mu(A_{acc})}{\mu(I)}.
\]

One interesting aspect of Lemma 1 is that it implies that the exit time distribution has finite mean, and by (29), that the mean return time to \( A \) is also finite when \( \mu(M) = 1 \). Higher moments of these distributions need not be finite; a counterexample is given by a simple shear map.

Since (29) is normalizable then one can define its cumulative distribution, the probability that the return is at least as long as \( k \)
\[
P^{\text{return}}_{A}(k) = \sum_{j=k}^{\infty} p^{\text{return}}_{A}(j).
\]

This is often called the **Poincaré recurrence distribution**. The Poincaré recurrence theorem states that when \( \mu(M) = 1 \), almost all trajectories return, i.e., \( P^{\text{return}}_{A}(1) = 1 \).

In most cases, there are no analytical formulae for exit and return time distributions. For a small perturbation to an integrable map with a homoclinic connection to a saddle, however, Melnikov theory can be used to estimate the exit set volume. Melnikov theory, originally developed for flows, was formulated for maps by Easton. Another case that can be treated is that of an adiabatic perturbation, where the lobes cover the region swept by the separatrices of the frozen time subsystems. If the structures of the homoclinic tangle or trellis can be given a topological classification, like that developed by Easton, then these methods can be extended to approximate additional intersection areas, and hence the exit time distributions.

**Question IX (Higher-Dimensional Trellises).** Can one classify trellises in 3D volume-preserving and 4D symplectic maps to obtain information about transport properties?

**B. Stickiness and anomalous diffusion**

If a map is uniformly hyperbolic, then the number of periodic orbits of a given period grows exponentially at a rate given by the topological entropy. This results in an exponential form for exit time distributions
\[
p^{+}_{A}(j) \sim e^{-j/\gamma}.
\]

However, as was first observed by Ref. 24, such distributions for area-preserving maps with mixed regular and chaotic orbits appear to decay much more slowly
\[
p^{+}_{A}(j) \sim j^{-\gamma-1}.
\]

Such algebraic decay is a signal of the stickiness of stable structures in the phase space, see Sec. VI.

The existence of the average (31) implies that \( \gamma > 0 \). It can be shown that the exponent for \( f \) and \( f^{-1} \) are the same, thus (26) implies that the transit time distribution will have this exponent as well, and hence, by (29), so will the return distribution. Consequently, the cumulative recurrence distribution will decay as
\[
P^{\text{return}}_{A}(t) \sim t^{-\gamma}.
\]

Many studies have found \( \gamma \approx 1.5 \) though there are fluctuations that persist for large times, and it is still controversial whether the power law (33) is valid asymptotically, and if it is, whether there are classes of dynamical systems for which the exponent \( \gamma \) is universal.

**Question X (Power Law Decay).** Do exit and return-time distributions in typical volume-preserving maps with mixed regular and chaotic components have an asymptotic power law form? Is there a universal exponent \( \gamma \)?

The exponent \( \gamma \) is related to another measure of transport, the growth of the mean square displacement \( \langle \|x_t - x_0\|^2 \rangle \) with time. For Brownian motion, the squared displacement grows linearly in time and there exists a diffusion coefficient
\[
D = \lim_{T \to \infty} \frac{1}{2T} \langle \|x_T - x_0\|^2 \rangle.
\]

However, it is observed that maps like (4) sometime exhibits anomalous momentum diffusion with
\[
\langle (y_t - y_0)^2 \rangle \sim t^\beta.
\]

**Super-diffusion**, i.e., \( \beta > 1 \), is observed in the standard map when there are **accelerator modes**, stable orbits that advance linearly in the momentum direction. An accelerator island will drag nearby chaotic orbits along, so that even when the ensemble average does not include the accelerating island, the effect is that \( \beta > 1 \). The simplest accelerator mode forms at \( k = 2n \), where after one step, \( (x_1, y_1) = (x_0, y_0 + 1) \); however, there are probably high-period accelerator modes arbitrarily close to \( k_\gamma(\phi) \). More general maps that do not have the vertical periodicity of the standard map do not have accelerator modes; however, the effect of sticky islands is still important.
Karney showed that (33) and (34) can be related through the force correlation function, $C(t)$. Indeed, the effective diffusion after $T$ steps is

$$\frac{1}{2T} \langle (y_T - y_0)^2 \rangle = \frac{1}{2} C(0) + \sum_{t=1}^{T} C(t).$$

Karney’s result is that the power law (33) implies $C(t) \sim t^{-\gamma+1}$, which then, by (34), gives

$$\gamma + \beta = 3. \quad (35)$$

Consequently, the algebraic decay of the recurrence distribution and anomalous diffusion are two sides of one coin. Since numerical experiments estimate that $\gamma \approx 1.5$, then (35) implies that $\beta \approx 1.5$ as well. This has been confirmed by Venegeroles using the Perron-Frobenius operator to analytically estimate diffusion due to accelerator modes.

Stickiness can also be studied using finite time Lyapunov exponents: the distribution of exponents is bimodal due to orbits sticking near elliptic regions. This is also observed in higher-dimensional systems.

VI. MARKOV MODEL FOR TRANSPORT

The phase space of a typical area-preserving map is a complex mixture of islands and chaotic orbits, recall Fig. 1. Each island, a collection of nested invariant circles is also surrounded by periodic orbits that have rational rotation numbers relative to the elliptic periodic orbit at the island’s center. Each elliptic periodic orbit also generically forms islands. Numerically, there appear to be infinitely many islands comprising a fat fractal: a chaotic region has nonzero area but its boundary has non-integer dimension. Thus, embedded in the chaotic sea is an infinite hierarchy of islands around islands, see Fig. 11.

A connected chaotic component can be partitioned into tree of states. The nodes of the tree correspond to regions bounded by partial barriers. Each state is a portion of the chaotic component surrounding a given elliptic island, i.e., the accessible subset of a resonance zone. Transport between states is mediated by turnstiles. These could be lobes in resonance boundaries; however, at least for twist maps, there are also cantori between resonances (there are irrationalals between every pair of rationals). Since lowest flux is that through a cantorus, recall Sec. IV C, these form the most restrictive partial barriers. The elliptic orbits inside a resonance zone also form islands that are separated by cantori.

Transport through a chaotic region appears to be anomalous whenever it has a boundary component that is a stable invariant set. For area-preserving maps, these are typically boundary circles of islands, the outermost of the encircling loops. It is these that give rise to the power law (32).

The simplest model contains a single island so that the chaotic component is bounded by a single invariant circle. The stickiness of a boundary circle is caused by a sequence of cantori that limit on the boundary with ever-decreasing flux. As John Greene explained to us in 1983, such a scenario yields a power law (32); however, the exponent is much larger than that observed: $\gamma = 3.05$ (Ref. 47) or $\gamma = 3.27$. Alternatively, if only the hierarchy of islands around islands is kept, ignoring the cantori, the exponent is still too large: $\gamma = 2.25, 120$ The insufficiency of either model was also confirmed by the observations of Ref. 111: a trajectory following either one of these hierarchies has an increasing probability of becoming trapped by secondary islands or by smaller flux cantori as time increases. This phenomenon results in transient local density accumulations. Note also that systems not smooth enough to have cantori and with no hierarchy of islands, such as some piecewise linear maps, have a different exponent, $\gamma = 2$.

The simplest partition that keeps both phenomena is a binary Cayley tree (nodes have a parent and two daughters), see Fig. 12. Each node in the tree is labeled by a sequence $S = s_0 s_1 \ldots$, with $s_i \in \{0, 1\}$. The root of the tree, denoted $\emptyset$, can be taken to be an absorbing state: whenever a trajectory reaches this state, it is deemed to have escaped. For the Hénon map of Fig. 11, the null state is the region outside the resonance zone formed by the stable and unstable manifolds of the hyperbolic fixed point (blue and red curves).

Transitions to the right and left on the tree correspond to distinct topological structures in phase space. A transition to the right corresponds to motion towards a given island boundary. The generically irrational rotation number of a boundary circle has a continued fraction expansion, and each of its convergents, (24), corresponds to a periodic orbit. The resonance zones of every other convergent lie in the chaotic component, giving an infinite sequence of levels that limit on the boundary circle. A transition from one level to the next corresponds to adding a “1” to the state, $S \rightarrow S1$.

Each elliptic island within a state $S$ also has its own boundary circle and corresponding infinite sequence of levels of encircling periodic orbits. We denote the outermost of these new levels by $S0$; thus, a transition to the left on the tree, $S \rightarrow S0$, corresponds to being trapped around an island of higher class. For example, in Fig. 11, the outermost elliptic islands in the upper right pane form a $(1, 7)$ chain. This represents state $S = 1$ in Fig. 12. One island in this chain is enlarged in the upper left panel of Fig. 11; it is also encircled by a $(1, 7)$ chain that has rotation number $1/7$ relative to the seventh image, $f^7$, so that this orbit has period 49 relative to $f$. In the figures, this island is designated by its rotation number sequence $1/7$; these are the outermost elliptic islands of class one corresponding to the state $S = 10$.

If there is only one significant elliptic island in each state, then the tree is binary. More generally, some states in the tree can have more daughters when there are multiple islands in a given state, and some may have fewer if the periodic orbits are hyperbolic.

The simplest approximate model for transport on the tree of states is Markov. The assumption is that the chaos within a state $S$ causes rapid mixing, and so the probability of a trajectory leaving at any time is proportional to the flux, $\Delta W_{S,S'}$, of trajectories from $S$ to state $S'$, where $S'$ is the parent, $DS$, or a daughter, $S0$ or $S1$. This flux is the area of the turnstile of the most resistant cantorus that divides the states and can be computed as an action difference, recall Sec. IV. Note that the flux is symmetric, $\Delta W_{S,S'} = \Delta W_{S',S}$, when the
map is exact, recall Sec. III B. This model seems reasonable when the average exit time \((31)\) is large compared to a Lyapunov time—the time for significant memory loss in the dynamics. However, for a near-integrable system, the mixing approximation probably will not work, and there are strong correlations that result in oscillatory exit time distributions.97

**Question XI** (Markov Transport Models). Is there a limit in which the Markov model is valid? When can higher-order correlations be neglected?

Thus, a Markov model for transport on the tree is defined by transition probabilities \(p_{S \rightarrow S'}\) for each pair of connected nodes on the tree, recall Fig. 12. Under the assumption that a trajectory has equal probability to be anywhere in \(S\), the probability of such a transition is

\[
p_{S \rightarrow S'} = \frac{\Delta W_{SS'}}{A_S},
\]

where \(A_S\) is the accessible area of \(S\). By Lemma 1, \((36)\) is exactly the inverse of the average transit time through the region. Following Ref. 89, it is convenient to categorize the change in transition probabilities on the tree by two ratios

\[
w_i(S) = \frac{p_{S \rightarrow S_i}}{p_{S \rightarrow DS}}, \\
\varepsilon_i(S) = \frac{p_{S \rightarrow S_i}}{p_{S \rightarrow DS}},
\]

for \(i \in \{0, 1\}\). The first ratio measures asymmetry between motion “down” and “up” the tree, and the second measures the slowing of time scales deeper in the tree.
For example, renormalization theory implies that for a noble invariant circle, \( w_1 = 0.053112 \): it is 19 times more likely for an orbit to move away from the circle—to the parent state—than to move one level deeper.\(^{66}\) For the self-similar sequence of period-seven islands shown in Fig. 11, the class ratio is even smaller, \( w_2 = 0.014158 \), implying that trapping around an island is an even rarer event.\(^{73}\) For this reason, anomalous transport is observable only for extremely long trajectories. Accordingly, it is debatable whether an asymptotic regime is reached even for the longest computations (e.g., quadruple precision computations for orbits of length \( 10^{12} \), Ref. 111).

In the original Markov tree model, Ott and I assumed that the tree is self-similar.\(^{89}\) This would asymptotically hold if every boundary circle were noble and every island chain had the same period relative to its parent. In this case, the ratios (37) are independent of the state \( S \), though they depend on the choice of class, \( i = 0 \), or level, \( i = 1 \). More generally, the results of Ref. 43 imply that the level hierarchy is approximately self-similar, and that for almost all boundary circles

\[
   w_1 \approx e_1^{1.05}. \tag{38}
\]

Similarly, there are cases in which the class hierarchy is asymptotically self-similar.\(^{73}\) Numerical studies again show that the scaling coefficients are related

\[
   w_0 \approx e_0^{1.19}, \tag{39}
\]

even as the period of the successive islands varies.

If the binary tree is exactly self-similar, then there is an integral equation for the recurrence distribution.\(^{89}\) A solution to this gives the power law (32), with an exponent determined by the dispersion relation

\[
   w_0 e_0^{-\gamma} + w_1 e_1^{-\gamma} = 1.
\]

Given the relations (38)–(39), there are still two parameters needed to solve for \( \gamma \). Reasonable values for the \( e_i \) give \( \gamma = 1.96 \), which is closer to the observed value, \( \gamma \approx 1.5 \), than the models with no tree hierarchy. The analysis of Ref. 89 implies that each branch added to the tree decreases the value of \( \gamma \).

Note, however, that for a typical map, neither self-similar scenario will hold, and the scaling factors (37) will depend upon the state \( S \). Recent studies have modeled the lack of self-similarity statistically. The idea is that an orbit will sample different regions of phase space, and its long-time behavior will be equivalent to an ensemble of scalings. Effectively, the coefficients of the Markov tree model should be taken as a random draw from some probability distribution. Numerical studies do appear yield different values of \( \gamma \) in different regions of phase space.\(^{3,111}\) Using such a model, Cristadoro and Ketzmerick find \( \gamma = 1.57 \), in remarkable agreement with numerical experiments.\(^{23}\) A similar model, using a different ensemble was studied in Ref. 17. Indeed, it is not clear which random ensemble of scaling coefficients is valid, though there may be some sense in which there is a universal ensemble.\(^{3} \) The ensemble picture also provides an explanation for the difficulty of observing an asymptotic regime: the correlation function of the fluctuations in the mean decay exponent has a log-periodic behavior that decays very slowly in time.\(^{17}\)

The Markov model, then, not only gives a qualitative explanation for the power law decay of the exit time distribution (32) but also gives a value of \( \gamma \) close to that observed. Of course, this is only for the area-preserving case.

**Question XII** (Higher-Dimensional Transport). Are there Markov tree models that quantitatively explain transport in higher-dimensional volume-preserving and symplectic maps?

**VII. THE NEXT THIRTY YEARS**

I have tried to summarize both the state of the art in the theory of transport for conservative dynamical systems and to indicate some directions that seem to me to be still ripe for future research.

There are other fruitful techniques that I did not include in this survey, such as the characterization of mixing in two-dimensional flows by the braiding of trajectories,\(^{11,46}\) and the

---

**FIG. 12.** Tree for the Hénon map (5) at \( a = -0.17197997940 \) with the phase space shown in Fig. 11 where it has a period-seven self-similar hierarchy, from Ref. 3. Each node is labeled by the state \( S \) and the winding number sequence for the main island of that class, \( \xi = \xi_1 \xi_2 \ldots \). Several transition probabilities, \( p_{i \rightarrow j} \), are also indicated. Reproduced with permission from Phys. Rev. E 90(6), 062923 (2014). Copyright 2014 AIP Publishing LLC.
set theoretic methods based on the Perron-Frobenius or Koopman operator formalisms. I also did not discuss the continuous time and space approximation of anomalous transport models to obtain phenomenological, fractional kinetic equations. This theory, which is reviewed in Refs. 120 and 119, does not have as close a connection with the underlying topology of the dynamics as the discrete Markov tree. Finally, I did not discuss the application of the ideas of turnstiles to semi-classical and quantum mechanics. Indeed, in Ref. 86, we speculated that classical cantori whose turnstiles were small in units of Planck’s constant might play a significant role in the semiclassical mechanics of molecules, and MacKay and I later applied this idea to compute the ionization threshold for hydrogen in a strong microwave field. Several recent advances in this area confirm our original speculation.70

There is much that remains intriguing about this problem: the infinite complexity of phase space and the long-time correlations that result in apparent power law decays and anomalous diffusion. Even after thirty years, I still cannot resist the allure of phase space portraits for simple area-preserving maps.

In Ref. 86, we noted that “the theory is not yet complete, but we believe that the ideas introduced here are of central importance for a broad range of systems that are neither completely ordered nor completely chaotic”.

While there has been considerable progress, I think—as indicated by the series of questions posed in this paper—that there is still much to do in the study of transport.

ACKNOWLEDGMENTS

The work was partially supported NSF Grant No. DMS-1211350. I would like to thank Robert MacKay, Ian Percival, and Bob Easton for long and fruitful collaborations, many discussions, and a profusion of insights.
