LARGE TIME ASYMPTOTICS OF FUNCTIONS TRANSPORTED ALONG CHAOTIC CONTACT FLOWS

DANIEL B. DIX

Contents

1. Introduction

Ergodic theory originated from physical applications, and yet some of the original physically motivated mathematics problems are still not well understood. Certain new applications—such as protein folding, RNA folding, and the assembly of and interactions within macromolecular complexes—have refocused attention on these old problems, but with a new twist. This proposal is about the classical problem of "decay of correlations" or "rate of mixing" in dynamical systems of a special but very important type. We propose to address appropriately formulated classical analytical problems, whose form has been motivated and guided by physical applications such as those mentioned above.

Suppose X (called the phase space) is a compact smooth manifold of dimension $2n-1$, $n \geq 1$, and ω is a closed two-form (called the *contact form*) on X such that for all $\mathbf{x} \in X$ we have $\{v \in T_{\mathbf{x}}X \mid \omega(\mathbf{x})(v,\cdot) = 0\}$ is one-dimensional. Suppose $\{\phi_t\}_{t\in\mathbb{R}}$ is the flow of a smooth complete never vanishing vector field F on X such that for all $\mathbf{x} \in X$ we have $\omega(\mathbf{x})(F(\mathbf{x}),\cdot) = 0$. X is therefore equipped with a canonical volume form; it is given locally by $\theta \wedge \omega^{n-1}$ in terms of any local 1-form θ such that $d\theta = \omega$ and $\theta(F(\cdot)) = 1$ [1],[14]. Let the measure μ associated to this volume form be normalized to be a probability measure. The flow $\{\phi_t\}_{t\in\mathbb{R}}$ will preserve the measure μ, since it preserves the volume form. Such flows, called *contact flows*, are central in physical applications; see examples below. Points $\mathbf{x}$ of X contain enough information to uniquely determine the trajectory for all positive and negative time. Much of this information is not observable or even of interest. Suppose Y is another smooth manifold and $\rho\colon X \to Y$ is a smooth mapping (which we will call a *fibration*). We think of points $\mathbf{y}$ of Y as encoding the observable

Date: October 14, 2002.

Key words and phrases. Protein folding, Decay of correlations, Special flows, Hyperbolic Toral automorphisms, Billiards, Geodesic flows on surfaces of constant negative curvature.

or potentially interesting information, and the fibration ρ as projecting out the uninteresting variables. The measure μ on X pushes forward via ρ to a probability measure ν on Y. Let $\mathcal{A}_\rho$ denote the σ-algebra of all subsets of X of the form $\rho^{-1}(V)$ where V is a Borel subset of Y. Suppose $f_0\colon X \to [0,\infty)$ is a smooth $\mathcal{A}_\rho$-measurable function such that $\int_X f_0\, d\mu = 1$. We can think of f_0 as determining an initial probability density on the phase space X which only depends on the interesting variables. The probability distribution of the phase point at a later time t is given by the density $f(\mathbf{x},t) = f_0(\phi_{-t}(\mathbf{x}))$. The central problem of this proposal is to determine the large time asymptotic behavior of $f(\cdot,t)$ in the sense of distributions on X. It should be obvious that $f(\mathbf{x},t)$ can never tend pointwise μ-almost everywhere to a limit function $\bar{f}(\mathbf{x})$ as $t \to \infty$ unless f_0 itself is constant along trajectories. However $E(f(\cdot,t) \mid \mathcal{A}_\rho)$ can be considered to be a function on Y, and we also wish to understand its large time asymptotic behavior as $t \to \infty$ in a suitable Banach space of functions on Y.

$f(\cdot,t)$ satisfies the following initial value problem:

$$f_t + \mathcal{X}_F f = 0, \qquad f(\cdot,0) = f_0,$$

where $\mathcal{X}_F$ is the derivation (first order partial differential operator) associated to the vector field F. This hyperbolic linear partial differential equation is of course solvable by the method of characteristics, but being linear these characteristics do not cross and there is no tendency to form shock waves in finite time as in nonlinear hyperbolic partial differential equations. However very steep gradients can form and the solution can become intensely oscillatory as $t \to \infty$. The investigation of the large-time behavior of such linear initial value problems is a natural new direction for the proposer given his previous work on the detailed large time asymptotics of linear dispersive equations [8]. However very different mathematical tools are required, again a major part of the project's appeal.

This problem is very subtle, but before we get deeply into these subtleties let us give some motivating examples. If (Q,g) is an n-dimensional Riemannian manifold and $V\colon Q \to \mathbb{R}$ is a smooth coercive function (i.e. for every $M > 1$ there is a compact set $K \subset Q$ such that $V(q) > M$ whenever $q \in Q \setminus K$) then $H(\alpha_q) = \frac{1}{2}\langle \alpha_q, g(q)^{-1}\alpha_q\rangle + V(q)$ defines a smooth Hamiltonian function on T^*Q, where $g(q)^{-1}\colon T_q^*Q \to T_qQ$ is the linear isomorphism induced by the Riemannian metric and $q \in Q$. Define $\theta_0\colon T^*Q \to T^*(T^*Q)$ such that for every $\alpha_q \in T_q^*Q$ we have $\theta_0(\alpha_q) = \alpha_q \circ T\pi_Q^* \in T_{\alpha_q}^*(T^*Q)$, where $\pi_Q^*\colon T^*Q \to Q$ is the projection, and $T\pi_Q^*\colon T_{\alpha_q}(T^*Q) \to T_qQ$ is its tangent mapping at the point α_q. This is the canonical one-form on a cotangent bundle, and the canonical symplectic structure is $\omega_0 = -d\theta_0$. H determines a smooth vector field F on T^*Q such that $dH(v_{\alpha_q}) = \omega_0(F(\alpha_q), v_{\alpha_q})$ for all $v_{\alpha_q} \in T_{\alpha_q}(T^*Q)$, where $\alpha_q \in T_q^*Q$ and $q \in Q$. If E is a regular value of H and $X = \{\mathbf{x} \in T^*Q \mid H(\mathbf{x}) = E\}$ then F is everywhere tangent to X and is complete, since X is a compact manifold. Let $\iota\colon X \to T^*Q$ be the inclusion mapping. Then the contact form ω on X is defined to be the pullback $\iota^*\omega_0$ of ω_0 over ι, i.e. the restriction of ω_0 to pairs of vectors tangent to X. Let $\theta = dh + \iota^*\theta_0$, where h is determined locally so that $\theta(F(\cdot)) = 1$; h can be arranged to solve $dh(F(\mathbf{x})) = 1 - \theta_0(F(\mathbf{x}))$ and to be zero on a transversal submanifold of X to the vector field F. The canonical volume form $\theta \wedge \omega^{n-1}$ on X determined by ω and F is a constant multiple of the symplectic volume form $\Omega = (-1)^{[\![\frac{n}{2}]\!]}\omega_0^n/n!$ on T^*Q "divided by dH". So for all $\mathbf{x}$ in the domain of h and for all linearly independent sets $\{v_1, \ldots, v_{2n-1}\} \subset T_{\mathbf{x}}X$ we have that

$$\frac{\Omega(\mathbf{x})(v, v_1, \ldots, v_{2n-1})}{dH(\mathbf{x})(v)} = \frac{(-1)^{[\![\frac{n}{2}]\!]+1}}{(n-1)!}(\theta \wedge \omega^{n-1})(\mathbf{x})(v_1, \ldots, v_{2n-1})$$

for all $v \in T_{\mathbf{x}}(T^*Q)$, v is transverse to $T_{\mathbf{x}}X$. Thus contact flows arise naturally from such "simple mechanical" Hamiltonian systems.

A particular Hamiltonian system of interest arises from the "rate of protein folding" problem. Proteins are large molecules which spontaneously fold into a definite shape in water under certain conditions (temperature, pressure, pH, etc.). One can describe the atoms of the protein and the solvent as classical particles subject to classical forces (the approximate nature of which are known). The motion of these particles can be described by a Hamiltonian system of the type we have discussed above. The phase point $\mathbf{x}$ specifies the positions and momenta of all the particles in the system, including both the protein atoms and the atoms in the water molecules. However the primary variables of interest are the N torsion angles specifying the conformation of the protein [19],[9]. Hence let $Y = (S^1)^N$ and $\rho\colon X \to Y$ be the map defining these torsion angles. One can define f_0 to be concentrated on a set of torsion variables corresponding to unfolded conformations of the protein molecule. Then one would like to show that $f(\cdot, t)$ tends in a distributional sense to $\bar{f}$ as $t \to \infty$, where $E(\bar{f} \mid \mathcal{A}_\rho)$ is a function on Y concentrated around the folded conformations. Moreover we wish to discover mathematical parameters governing the rate at which this "relaxation" process occurs.

The ability to predict the rate of folding of a particular polypeptide chain is a crucial part of protein design, with wide ramifications and impact on health and biotechnology. Contact flows arise throughout classical statistical mechanics; thus protein folding is just one example of many that could be given of an application. Understanding general mathematical principles governing the rate of mixing in these types of dynamical systems is one of the most important contributions mathematical analysis can make to these applied problems. To achieve understanding of general principles we must first analyze a multiplicity of specific examples.

The relaxation to equilibrium in focus is of a subtle sort, since there is no dissipation of any kind in the equations of motion. It is difficult to justify from first principles the various *ad hoc* dissipative mechanisms that have been proposed over the years to model this irreversible trend toward equilibrium. We prefer to deal with it directly in its time reversible setting. In fact, on the Hamiltonian systems of the type we have discussed there is an involutive diffeomorphism $\mathcal{I}\colon T^*Q \to T^*Q$ which "reverses the direction of all the momenta", i.e. $\mathcal{I}(\alpha_q) = -\alpha_q$. The Hamiltonian flow then has the property that $\phi_t \circ \mathcal{I} = \mathcal{I} \circ \phi_{-t}$. The diffeomorphism $\mathcal{I}$ restricts to a diffeomorphism of X, and $\mathcal{I}^*\omega = -\omega$. We will assume that all our contact manifolds have a mapping $\mathcal{I}$ that is involutive (i.e. $\mathcal{I} \circ \mathcal{I} = \mathrm{id}_X$), contact structure reversing (i.e. $\mathcal{I}^*\omega = -\omega$), and such that $\phi_t \circ \mathcal{I} = \mathcal{I} \circ \phi_{-t}$ for all $t \in \mathbb{R}$. Our assumptions are intended to exclude manefestly dissipative and phase volume contracting models. But we cannot yet answer the main question we have posed in the generality that we have posed it; certainly systems of the size and complexity of our protein folding example are presently out of reach. But in our choice of simplified models we wish to preserve as much as possible these general features.

In the following sections we begin our discussion of a variety of simplified systems. Our primary purpose is to provide motivation and perspective on the main model to which we propose to devote our energies, namely special flows over hyperbolic toral automorphisms. However, this sequence of examples could also be used in reverse as a roadmap toward realistic protein folding models. These simplified models include Hamiltonian systems such as the Hènon–Heiles system ($n = 2$) and the classical water molecule ($n = 3$), where the phase space is divided between invariant tori and chaotic regions. We discuss the interaction of the ergodic components and the fibration $\rho\colon X \to Y$, the introduction of Poincaré sections, and the representation of the flow on X as a special flow over a (usually discontinuous) Poincaré return mapping. Then we discuss billiards models, which can be thought of as limits of Hamiltonian systems where the potential energy function V assumes only the values 0 or ∞. These systems can be designed to have strong mixing properties, and can also be equipped with natural fibrations. They also have special flow representations, but the Poincaré return mappings are in the most interesting cases discontinuous. Then we discuss in detail special flows over hyperbolic toral automorphisms, and describe our approach to the rate of mixing problem for these systems. Geodesic flows on the unit tangent bundle of compact surfaces of constant negative curvature are the examples in which the most detailed information is known, and these results set a standard to strive for elsewhere. Our short account is about how applicable those results are to our main problem. Billiards and geodesic flows (as well as special flows over hyperbolic toral automorphisms) can be addressed via symbolic dynamics as special flows over subshifts of finite type. We briefly summarize this strategy and indicate why we need to seek other tools in our inquiry.

2. Interacting Particle Hamiltonians

Consideration of simple examples such as a rotational flow on $X = S^1$ shows that $f(\cdot, t)$ can oscillate with no attenuation as $t \to \infty$. A case with $n = 2$ is the three particle Toda lattice [16]. These are *integrable* (hence not generic) Hamiltonian systems, but it is clearly necessary to regularize $f(\cdot, t)$ to $\frac{1}{t}\int_0^t f(\cdot, s)\,ds$ before a trend as $t \to \infty$ can be observed in general. Since the flow $\{\phi_t\}$ preserves the probability measure μ the ergodic theorems [7] imply that there is a nonnegative function $\bar{f} \in L^1(X, \mu)$ such that $\frac{1}{t}\int_0^t f(\cdot, s)\,ds \to \bar{f}$ as $t \to \infty$ in L^1 and pointwise μ-almost everywhere. This $\bar{f}$ is the conditional expectation $E(f_0 \mid \mathcal{U})$ of f_0 with respect to the sub-σ algebra $\mathcal{U}$ of the Borel σ-algebra $\mathcal{B}$ on X consisting of the sets invariant under the flow. If $P(X, \{\phi_t\})$ denotes the space of all probability measures on $\mathcal{B}$ which are preserved by the flow $\{\phi_t\}$ then the *ergodic decomposition theorem* [15], [14], [27], implies the existence of a $\mathcal{U}$-measurable mapping $m\colon X \to P(X, \{\phi_t\})\colon \mathbf{x} \mapsto m_{\mathbf{x}}$ such that for μ-almost all $\mathbf{x} \in X$ the set $m^{-1}(\{m_{\mathbf{x}}\})$ is invariant under the flow and contains the support of $m_{\mathbf{x}}$, and the measure $m_{\mathbf{x}}$ is ergodic (i.e. $m_{\mathbf{x}}(A)m_{\mathbf{x}}(X \setminus A) = 0$ for all $A \in \mathcal{U}$). Furthermore we have that the Choquet decomposition $\mu = \int_X m_{\mathbf{x}}\,d\mu(\mathbf{x})$ holds (the ergodic members of the convex set $P(X, \{\phi_t\})$ are its extreme points). Thus $\bar{f}(\mathbf{x}) = \int_X f_0\,dm_{\mathbf{x}}$. The set $m^{-1}(\{m_{\mathbf{x}}\})$ is called the *ergodic component* containing $\mathbf{x}$. An example where the ergodic components can be identified is the three particle Toda lattice mentioned above, whose Hamiltonian is given by

$$H_T(x, y, p_x, p_y) = \frac{1}{2}(p_x^2 + p_y^2) + \frac{1}{24}[e^{2y+2\sqrt{3}x} + e^{2y-2\sqrt{3}x} + e^{-4y}] - \frac{1}{8}.$$

Each ergodic component is a smooth submanifold of X (X is the subset of $\mathbb{R}^4$ where $H_T = E$, $E > 0$) diffeomorphic to a two-dimensional torus. These turn out to be level sets of the conserved quantity

$$8p_x(p_x^2 - 3p_y^2) + (p_x + \sqrt{3}p_y)e^{2y-2\sqrt{3}x} + (p_x - \sqrt{3}p_y)e^{2y+2\sqrt{3}x} - 2p_x e^{-4y}.$$

This invariant function is not derived from general physical conserved quantities, but is an accident of the integrable structure. But one generally tries to set up systems such that all the known physical conserved quantities are already constant on X, with the hope that this will mean that X is itself the only ergodic component. But this hope is in vain, as can be seen by truncating a Taylor expansion of the above Toda Hamiltonian after the cubic terms:

$$H_{HH}(x, y, p_x, p_y) = \frac{1}{2}(p_x^2 + p_y^2 + x^2 + y^2) + x^2y - \frac{1}{3}y^3.$$

This is the Hènon–Heiles Hamiltonian, and we make the analogous definition of X except that now $0 < E < \frac{1}{6}$. For very small E the KAM theorem implies the existence of a Cantor set of invariant tori, each of which would be an ergodic component. But in the regions of phase space between these invariant tori corresponding to vibrational frequencies which are nearly *resonant* there appear to be ergodic components which are three dimensional. These chaotic regions can be seen more easily at larger values of E [16]. The ergodic component in focus is not simply the entire region between two invariant tori, because such a region contains additional knotted invariant tori, whose smaller radii are again taken from a Cantor set. The gaps between these knotted tori are probably also chaotic in the same way. Thus the decomposition into ergodic components is in general very complicated.

Ergodicity of the flow $\{\phi_t\}$ on X is important mainly to precisely identify $\bar{f}$ given f_0, since in that case $\bar{f} = \int_X f_0 \, d\mu = 1$. In our protein folding example $\bar{f}$, or more precisely $\bar{f}\, d\mu$, is the idealized mathematical representation of the folded state of the protein, which is a nice thing to have in light of the herculean efforts being made to compute it by biochemists and biophysicists. However, as we hinted earlier, the probability $E(\bar{f} \mid \mathcal{A}_\rho)\, d\nu$ on Y more realistically represents the folded state of the protein. Thus we really want to know the function $E(\bar{f} \mid \mathcal{A}_\rho)$, which could be very nearly equal to 1 without it being true that $\bar{f} = 1$. This would depend on how the fibres of the map ρ intersect the ergodic components (i.e the level sets of $\bar{f}$). It would be very nice to find an alternate approach to that requiring a detailed knowledge of the ergodic components, and perhaps (although this is not our main proposal) this could be studied in the context of simplified models like Hènon–Heiles.

An elementary example which is a toy model of protein folding and also illustrates the role of the fibration is the Hamiltonian system

$$H(x, y, p_x, p_y) = \frac{1}{2}(p_x^2 + p_y^2) + V(x, y),$$

where the contours of the potential function $0 \le V(x, y) \le 1$ are shown in Figure 1. As usual $X \subset \mathbb{R}^4$ is a level set of H, say where $H = 1$. Let $\rho\colon X \to \mathbb{R}\colon (x, y, p_x, p_y) \mapsto x$. The shallow potential well near $(x, y) = (1, 0)$ represents the unfolded state and the deep well near $(x, y) = (0, 0)$ represents the folded state. We introduce local coordinates (x, y, ϑ) where $V(x, y) < 1$ and $-\pi < \vartheta \le \pi$ by the rule: $(p_x, p_y) = \sqrt{2(1 - V(x, y))}(\cos\vartheta, \sin\vartheta)$. The canonical volume form on X is given by $dx \wedge dy \wedge d\vartheta$ in this coordinate chart, which covers everything in X except

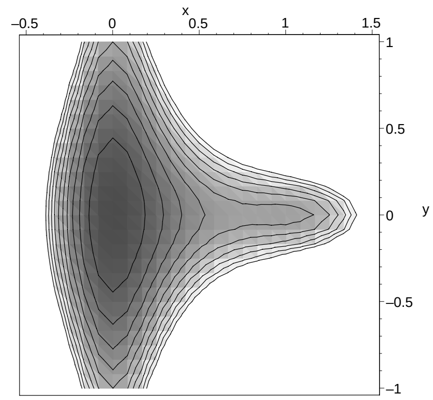

FIGURE 1

for a set of measure zero. The projected measure ν on $Y = \mathbb{R}$ is $\nu(x)\,dx$, where the density $\nu(x)$ is proportional to the width of the interval $\{y \in \mathbb{R} \mid V(x,y) \le 1\}$. The conditional expectation $E(\cdot \mid \mathcal{A}_\rho)$ can be computed by integrating in y and ϑ. The Hamiltonian vector field in this coordinate chart is as follows:

$$\begin{aligned}\dot{x} &= \sqrt{2(1 - V(x,y))}\cos\vartheta\\ \dot{y} &= \sqrt{2(1 - V(x,y))}\sin\vartheta\\ \dot{\vartheta} &= \frac{V_x(x,y)\sin\vartheta - V_y(x,y)\cos\vartheta}{\sqrt{2(1-V(x,y))}}.\end{aligned}$$

We can define a smooth function $\tilde{f}_0(x) \ge 0$ with support near $x = 1$ and such that $\int_{\mathbb{R}} \tilde{f}_0(x)\nu(x)\,dx = 1$. Then $f_0 = \tilde{f}_0 \circ \rho$ is an $\mathcal{A}_\rho$-measurable function on X. We clearly expect $E(f_0 \circ \phi_{-t} \mid \mathcal{A}_\rho)$, thought of as a function on Y, to converge as $t \to \infty$ to 1, or a function nearly equal to 1. Thus the initial density $\tilde{f}_0(x)\nu(x)$ would converge irreversibly to $\nu(x)$ as $t \to \infty$. This assertion can be checked computationally. It is proposed that numerical experiments be performed, by undergraduate or graduate students, to test this convergence. Another attractive feature of this example is the possibility of fully visualizing the flow in phase space. It is proposed that these flow images be developed using the CAVE visualization technology already available through the Industrial Mathematics Institute at the University of South Carolina. They will be a wonderful teaching aid in the subject of chaotic dynamics as well as possibly yielding theoretical insights. Unfortunately, at present without these images, rigorously proved assertions still seem far away.

The standard means of visualization of these flows is by means of a Poincaré surface of section $\mathcal{S} \subset X$ transverse to the flow $\{\phi_t\}$ and a Poincaré return mapping $\mathcal{P}\colon \mathcal{S} \to \mathcal{S}$. For example in the Toda or Hènon–Heiles systems above we could define $\mathcal{S} = \{\mathbf{x} \in X \mid x = 0\}$. It is natural to introduce coordinates (x, y, p_y) in X where we choose the branch $p_x > 0$. Then (y, p_y) are natural coordinates on $\mathcal{S}$. In general we take $\mathcal{S}$ to be a codimension one submanifold of the contact manifold X. If $\mathbf{x} \in \mathcal{S}$ then let $\tilde{\tau}(\mathbf{x}) > 0$ be the smallest positive time t such that $\phi_t(\mathbf{x}) \in \mathcal{S}$. Define $\mathcal{P}(\mathbf{x}) = \phi_{\tilde{\tau}(\mathbf{x})}(\mathbf{x})$ whenever $\tilde{\tau}(\mathbf{x}) < \infty$. The contact form ω on X restricts to a closed two-form which is nondegenerate at every $\mathbf{x} \in \mathcal{S}$ such that $F(\mathbf{x})$ is transverse to $T_{\mathbf{x}}\mathcal{S}$. The mapping $\mathcal{P}$ preserves this symplectic structure at each point where it is defined and smooth. Except for a restricted set of cases it is not possible to choose $\mathcal{S}$ everywhere transverse to the flow [4]. Some study [11] has been devoted to the question of what sections might exist with good properties, such as being a smooth manifold and $\tilde{\tau}$ being everywhere finite. Another desirable property $\mathcal{S}$ should have is invariance under the "momentum inversion" mapping $\mathcal{I}\colon X \to X$, i.e. $\mathcal{I}$ should restrict to a diffeomorphism of $\mathcal{S}$ onto itself. Define $\mathcal{J} = \mathcal{I} \circ \mathcal{P}$. We require that $\tilde{\tau} \circ \mathcal{J} = \tilde{\tau}$ so that $\mathcal{J}$ is involutive, i.e. $\mathcal{J} \circ \mathcal{J} = \mathrm{id}_{\mathcal{S}}$. Note that $\mathcal{P} = \mathcal{I} \circ \mathcal{J}$ and $\mathcal{P}^{-1} = \mathcal{J} \circ \mathcal{I}$. An interesting and natural example of such a choice of Poincaré section for Hamiltonian systems is the *equipotential* section [11]. This defined by the rule: $\mathcal{S} = \{\mathbf{x} \in X \mid \mathbf{x} = \alpha_q \in T_q^*Q, dV(q)(T\pi_Q^*(F(\alpha_q))) = 0\}$. In our toy model for protein folding the points from $\mathcal{S}$ in the coordinate chart (x, y, ϑ) are those where $V_x(x, y)\cos\vartheta + V_y(x, y)\sin\vartheta = 0$, i.e. the momentum is parallel to an equipotential curve. In these coordinates the momentum inversion mapping $\mathcal{I}$ is given by $(x, y, \vartheta) \mapsto (x, y, \vartheta + \pi \mod 2\pi)$, and clearly this maps $\mathcal{S}$ into itself. Also in this case $\tilde{\tau}$ seems to be a bounded function. Because a trajectory can pass through $\mathcal{S}$ non-transversely, the mappings $\tilde{\tau}$ and $\mathcal{P}$ are usually discontinuous, and this makes difficult the study of the discrete dynamical system $(\mathcal{S}, \mathcal{P})$. The singularities of the symplectic structure on $\mathcal{S}$ and the discontinuities of the mapping $\mathcal{P}$ are highly structured, reflecting their smooth origins, so there is hope generalizations of current results on discrete dynamical systems can be made in this direction.

Another interesting example with $n = 2$ which has been much studied is the system of one particle on $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ interacting with a short range radially symmetric potential centered at the origin [17]. The periodic boundary conditions break the rotational symmetry giving rise to a three dimensional constant energy manifold X. By adjusting the properties of the potential one can obtain ergodicity, indeed the mixing property of the dynamical system on X, or nonergodicity as the case may be. *Mixing* is the condition on the flow that for any $f_0, g \in L^2(X, \mu)$ we have

$$\lim_{t\to\infty} \int_X f(\cdot, t) g \, d\mu = \int_X f_0 \, d\mu \int_X g \, d\mu.$$

This certainly implies that $f(\cdot, t) \to \bar{f} = \int_X f_0 \, d\mu$ in the sense of distributions on X as $t \to \infty$, and is stronger than ergodicity. There is a natural fibration in this example, namely the mapping $\rho\colon X \to S^1$ induced by projecting onto the first space variable. If (r, ϑ) are polar coordinates on the region of the torus near the origin then the potential has support in the disk $r \le R < 1$. The equation $r = R$ defines a Poincaré section $\mathcal{S}$. In this case the reduction of the flow on X to the study of the discrete dynamical system $(\mathcal{S}, \mathcal{P})$ is a very effective technique. When the flow on X fails to be ergodic it will be interesting to study $E(\bar{f} \mid \mathcal{A}_\rho)$ and the natural of the convergence (or lack thereof) of $E(f(\cdot, t) \mid \mathcal{A}_\rho)$ to $E(\bar{f} \mid \mathcal{A}_\rho)$ as $t \to \infty$.

Systems with $n = 2$ are qualitatively different than those with $n \ge 3$ because in the latter case trajectories are not trapped in the region between invariant tori, and hence we might expect an ergodic component to be much larger, perhaps occupying most of X. The study of this has begun [28], [13], but there is still a great deal not understood. One representative example with $n = 3$ is the Hamiltonian system governing the internal vibrations of a classical water molecule with vanishing linear and angular momentum [18]. Haller (Theorem 4.2.1, [13]) concludes that this system is not integrable. For small energies E the five dimensional phase space X would be expected to contain some three-dimensional invariant tori except for the fact that the frequencies are nearly resonant—a consequence of the energetic characteristics of the equilibrium geometry of the water molecule. This fact then motivates a deeper study of our main problem first for the system of a single water molecule, as well as for systems of several water molecules.

One interesting aspect of the system of several water molecules is macroscopic indistinguishability—that is, the fact that in practice in samples of water the size of those involved in protein simulations there is no way to distinguish one water molecule from another. Thus we should restrict the σ-algebra to the one containing only sets which are invariant with respect to any permutation of the water molecules. It will be interesting to see if this has any effect on the rate of mixing, and this would be most easily tested on a system of two water molecules.

In conclusion, we have surveyed several well-known Hamiltonian systems with $n = 2$ or 3 where the rate of mixing problem is easy to state. But in none of these cases to the proposer's knowledge is the rate of mixing rigorously known. To make progress we need further simplifications.

3. Billiards

All these examples are complicated by the fact that the dynamics proceeds along curves which can only be found by solving the differential equations of motion. Thus *billiard* systems are appealing because the trajectories are straight lines except for reflections at boundaries. Consider a compact connected domain $\Omega \subset \mathbb{C}$ such that $\partial\Omega$ is a finite disjoint union of simple closed C^1 curves each of which is composed of finitely many linear segments and circular arcs. Let J denote a finite disjoint union of bounded half-closed intervals and $Z\colon J \to \partial\Omega$ be an arc length parameterization respecting the standard orientation of $\partial\Omega$. Let $S^1 \subset \mathbb{C}$ denote the unit circle. Let X denote $(\Omega \times S^1) \setminus \{(Z(s), e^{i\vartheta}) \mid s \in J, \Re[Z'(s)e^{-i\vartheta - i\pi/2}] > 0\}$. Consider for $\mathbf{x} = (z, e^{i\vartheta}) \in X$ the straight line $\tilde{z}(t) = z + te^{i\vartheta}$ and let $[0, t^*(\mathbf{x})]$ be the maximal time interval such that $\tilde{z}([0, t^*(\mathbf{x})]) \subset \Omega$. Then $\tilde{z}(t^*(\mathbf{x})) = Z(s)$ for some $s \in J$. If $0 \leq t < t^*(\mathbf{x})$ then define $\phi_t(\mathbf{x}) = (\tilde{z}(t), e^{i\vartheta})$. If $e^{i\vartheta} \neq \pm Z'(s)$ and $t \geq t^*(\mathbf{x})$ then use the recursive definition $\phi_t(\mathbf{x}) = \phi_{t-t^*(\mathbf{x})}((Z(s), Z'(s)^2 e^{-i\vartheta}))$. This means that the particle starting at z moves in a straight line until it hits the boundary, where it reflects using an equal angle rule. If $e^{i\vartheta} = \pm Z'(s)$ then let $\phi_t(\mathbf{x})$ remain undefined for all $t \geq t^*(\mathbf{x})$. Let X° denote the interior of X and $X^\circ_t = \{\mathbf{x} \in X^\circ \mid \phi_t(\mathbf{x}) \text{ is defined and in } X^\circ\}$ for any $t \in \mathbb{R}$. On X° define the one-form $\theta = \cos(\vartheta)\, dx + \sin(\vartheta)\, dy$, where $z = x + iy$, and the two-form $\omega = -\, d\theta = -\sin(\vartheta)\, dx \wedge d\vartheta + \cos(\vartheta)\, dy \wedge d\vartheta$. The real vector space $\{\mathbf{v} = (v_x, v_y, v_\vartheta) \in T_{\mathbf{x}}X^\circ \mid \omega(\mathbf{v}, \cdot) = 0\}$ is spanned by $F(\mathbf{x}) = (\cos\vartheta, \sin\vartheta, 0)$; note also that $\theta(F(\mathbf{x})) = 1$. Thus modulo the singularities this flow has an exact contact structure (exact because $\omega = -\, d\theta$ for a single one-form θ). The topological space X is compact. The canonical volume form $\theta \wedge \omega$ is $dx \wedge dy \wedge d\vartheta$, and the associated probability measure μ is obtained from this volume form after normalization (i.e dividing by $2\pi\, \text{area}(\Omega)$). We have $\mu(X^\circ_t) = 1$ for all $t \in \mathbb{R}$. If $f_0 \in L^1(X, \mu)$ then $f_0 \circ \phi_{-t}$ is a well-defined member of $L^1(X, \mu)$. It would be interesting to have a functional analytic characterization of the contact structure that is present here, since it does not fit into our earlier definitions. It is the conviction that such a generalized notion of contact structure exists that allows us to include billiards along with orthodox contact flows. Since Ω is a Riemannian manifold with boundary, $\{\phi_t\}$ can be thought of as a geodesic flow on X, which is a closed subset of the unit tangent

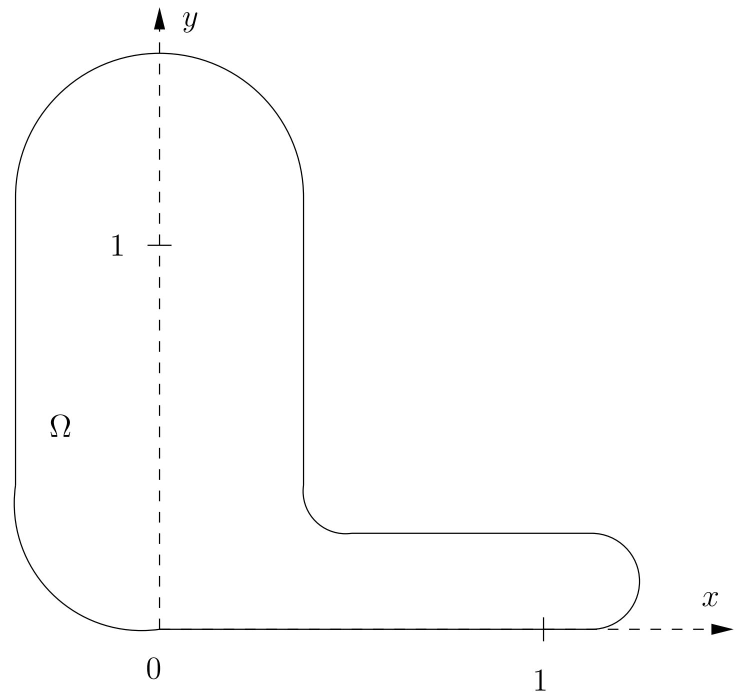

Figure 2

bundle. We could imagine the flow $\{\phi_t\}$ as being obtained as a limit of Hamiltonian flows associated to a sequence $V_n(x, y)$ of smooth potential functions vanishing on Ω but tending to infinity within a distance of $1/n$ of Ω.

Define $\mathcal{S} = \{(z, e^{i\vartheta}) \in X \mid z \in \partial\Omega\}$. The mapping $J \times [-1, 1] \to \mathcal{S}\colon (s, v) \mapsto (Z(s), Z'(s)e^{i\cos^{-1}(v)})$ is a bijective parameterization of $\mathcal{S}$. If we extend the form θ to the interior $\mathcal{S}^\circ$ of $\mathcal{S}$ and then pull it back onto $\mathcal{S}^\circ$, expressing it in the variables (s, v) we get $\theta = v\,ds$. Thus a similar extension and pullback of ω to $\mathcal{S}^\circ$ yields $\omega = -d\theta = ds \wedge dv$. The momentum inversion mapping $\mathcal{I}\colon X \to X$ is defined by the rule

$$(z, e^{i\vartheta}) \mapsto \begin{cases} (z, -e^{i\vartheta}) & z \in \Omega^\circ, \\ (z, -Z'(s)^2 e^{-i\vartheta}) & z = Z(s) \in \partial\Omega. \end{cases}$$

Clearly $\mathcal{I}$ maps $\mathcal{S}$ into itself; in the variables (s, v) we have $\mathcal{I}(s, v) = (s, -v)$. We have for $\mathbf{x} \in \mathcal{S}$ that $\tilde{\tau}(\mathbf{x}) = t^*(\mathbf{x})$, and this function is clearly bounded, although when Ω is not convex we have that $\tilde{\tau}$ is not continuous. This lack of continuity also applies to the Poincaré return mapping $\mathcal{P}$. If $\mathcal{J} = \mathcal{I} \circ \mathcal{P}$ notice that $\tilde{\tau} \circ \mathcal{J} = \tilde{\tau}$. Both $\tilde{\tau}$ and $\mathcal{J}$ are piecewise algebraic mappings, and the formula $\mathcal{P} = \mathcal{I} \circ \mathcal{J}$ gives a simple way to geometrically visualize this mapping.

There is a natural fibration in such billiard systems, namely $Y = \mathbb{R}$ and $\rho\colon X \to Y\colon (z, e^{i\vartheta}) \mapsto x$. It is not difficult to design billiard systems which should have relaxation properties similar to protein folding; an example can be seen in Figure 2. Whereas in our earlier toy model for protein folding the folded state $(x, y) = (0, 0)$ actually had a lower potential energy than the unfolded state $(x, y) = (1, 0)$, in this model they have the same potential energy, namely 0. Rather the folded state $x = 0$ is only "entropically favored" over the unfolded state $x = 1$. That this should lead to relaxation from the unfolded to folded states is based on the assumption that this system is mixing, or at least that the fibres of ρ intersect the ergodic components equitably. Our design resembles the "stadium" of [6] which is known to be mixing. Design principles for mixing billiards were discussed in [29]. Techniques for establishing hyperbolicity, ergodicity and mixing are surveyed in [17].

The proposer has developed software to compute trajectories in an arbitrary billiard system of the above type. It is proposed that further software be written to provide a "proof of concept" for the ideas about relaxation propounded here. These computer experiments are suitable for undergraduate projects and because of their pictorial nature can lead to a much wider understanding of the trend toward equilibrium.

Given $h, g \in L^2(X, \mu)$ and $t \in \mathbb{R}$ we define the correlation $C_t(h, g)$ by the rule:

$$C_t(h, g) = \int_X (h \circ \phi_t) g \, d\mu - \int_X h \, d\mu \int_X g \, d\mu.$$

The *decay of correlations* problem concerns determining the exact rate that this quantity tends to 0 as $t \to \infty$. Since $\mathcal{I}$ preserves the measure μ on X we have that $C_{-t}(f_0, g) = C_t(f_0 \circ \mathcal{I}, g \circ \mathcal{I})$, so that the rate of mixing problem and the decay of correlations problem essentially coincide. If both f_0 and g are $\mathcal{A}_\rho$-measurable then strong enough estimates on the decay of $C_{-t}(f_0, g)$ can imply estimates on the convergence of $E(f_0 \circ \phi_{-t} \mid \mathcal{A}_\rho)$ to 1 in the norm of $L^2(Y, \nu)$ as $t \to \infty$. However such strong estimates are not attainable unless both f_0 and g_0 have some smoothness.

If λ denotes the probability measure on $\mathcal{S}$ induced by the invariant two-form $\omega = ds \wedge dv$, then we define the discrete time correlation $C_n(h, g) = \int_{\mathcal{S}} (h \circ \mathcal{P}^n) g \, d\lambda - \int_{\mathcal{S}} h \, d\lambda \int_{\mathcal{S}} g \, d\lambda$, where $h, g \in L^2(\mathcal{S}, \lambda)$ and $n \in \mathbb{Z}$. A comprehensive reference on the discrete time decay of correlations problem is [3]. The relevance of the discrete case to our continuous time problem is clearer once it is recognized that there is a metric isomorphism of the continuous time dynamical system we have defined above with a *special flow* over $\mathcal{P}$. Let $\mathcal{S}^* = \{\mathbf{x} \in \mathcal{S} \mid \mathcal{P}^n(\mathbf{x}) \in \mathcal{S} \text{ for all } n \in \mathbb{Z}\}$; it has λ-measure 1. Define the space $\mathcal{S}_{\tilde{\tau}}$ to be the set of orbits of an action of $\mathbb{Z}$ on $\mathcal{S}^* \times \mathbb{R}$ induced by the mapping $\mathcal{S}^* \times \mathbb{R} \to \mathcal{S}^* \times \mathbb{R} \colon (\mathbf{x}, \tau) \mapsto (\mathcal{P}(\mathbf{x}), \tau - \tilde{\tau}(\mathbf{x}))$. The flow $\tilde{\psi}_t(\mathbf{x}, \tau) = (\mathbf{x}, \tau + t)$ on $\mathcal{S}^* \times \mathbb{R}$ maps orbits into orbits, and hence determines a flow $\psi_t \colon \mathcal{S}_{\tilde{\tau}} \to \mathcal{S}_{\tilde{\tau}}$. The mapping $\mathcal{S}^* \times \mathbb{R} \to X \colon (\mathbf{x}, \tau) \mapsto \phi_\tau(\mathbf{x})$ is constant on orbits, and hence an injection $\xi \colon \mathcal{S}_{\tilde{\tau}} \to X$ is determined whose image has μ-measure 1. $\mathcal{S}_{\tilde{\tau}}$ has coordinates (s, v, τ) on the subset where $0 < \tau < \tilde{\tau}(s, v)$, and a local contact form $\zeta = ds \wedge dv = -d\eta$, where $\eta = d\tau + v \, ds$; $\eta \wedge \zeta = ds \wedge dv \wedge d\tau$ is a volume form, with associated probability measure κ. The degree to which $d\tau$ can be replaced by a continuous closed one-form on $\mathcal{S}_{\tilde{\tau}}$ depends on the continuity and other properties of $\tilde{\tau}$ and $\mathcal{P}$. Thus the flow $\{\psi_t\}$ is a contact flow on $\mathcal{S}_{\tilde{\tau}}$ only in the same (unspecified) generalized sense that $\{\phi_t\}$ is on X; however the flow $\{\psi_t\}$ does preserve the measure κ. The mapping ξ is a metric isomorphism between $(\mathcal{S}_{\tilde{\tau}}, \kappa)$ and (X, μ) which intertwines the flows, i.e. $\phi_t \circ \xi = \xi \circ \psi_t$ for all $t \in \mathbb{R}$.

The discontinuities of $\tilde{\tau}$ and $\mathcal{P}$ lead to difficulties analyzing the special flow $\{\psi_t\}$ on $\mathcal{S}_{\tilde{\tau}}$, and reducing continuous time decay of correlations to the discrete time case. However very good results have been obtained [26] for a different class of billiards (not including the example of Figure 2) using Markov partitions for flows [20] and refined spectral estimates of the Ruelle operator as in [10]. In particular exponential decay of $C_t(h,g)$ in time is proved for Hölder continuous functions h, g. The method used in this work does not respect the underlying "contact structure", and probably will not lead to accurate asymptotic expansions of $C_t(h,g)$ as $t \to \infty$. This is because the C^1 regularity of the original system is replaced by a model flow with Lipschitz regularity only. We are interested in exploring alternative approaches to the route through symbolic dynamics.

4. Special Flows over Hyperbolic Toral Automorphisms

Now we come to the main thrust of the proposal concerning the model which is sufficiently simple that we expect to make significant strides in understanding. Let $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ and define $A_{\mathcal{P}} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, $A_{\mathcal{I}} = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$, and $A_{\mathcal{J}} = \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix}$. Define the *hyperbolic toral automorphism* [14]

$$\mathcal{P}\colon \mathbb{T}^2 \to \mathbb{T}^2\colon (x,y) + \mathbb{Z}^2 \mapsto (2x+y, x+y) + \mathbb{Z}^2.$$

Actually any integer matrix leads to a mapping on $\mathbb{T}^2$ in the same way, so let mappings $\mathcal{I}, \mathcal{J}$ be associated to the matrices $A_{\mathcal{I}}$ and $A_{\mathcal{J}}$ respectively. Note that $\mathcal{I} \circ \mathcal{I} = \mathrm{id}_{\mathbb{T}^2}$, $\mathcal{J} \circ \mathcal{J} = \mathrm{id}_{\mathbb{T}^2}$, and $\mathcal{I} \circ \mathcal{J} = \mathcal{P}$. Let $\omega = dx \wedge dy$ be the symplectic form, inducing the probability measure λ on $\mathbb{T}^2$. Suppose $\tilde{\tau}\colon \mathbb{T}^2 \to (0,\infty)$ is a smooth function such that $\int_{\mathbb{T}^2} \tilde{\tau}\, d\lambda = 1$ and $\tilde{\tau} \circ \mathcal{J} = \tilde{\tau}$. Define X to be the set of orbits of the $\mathbb{Z}$-action on $\mathbb{T}^2 \times \mathbb{R}$ induced by the mapping $\mathbb{T}^2 \times \mathbb{R} \to \mathbb{T}^2 \times \mathbb{R}\colon (\mathbf{z}, \tau) \mapsto (\mathcal{P}(\mathbf{z}), \tau - \tilde{\tau}(\mathbf{z})$. X is a compact smooth manifold. We distinguish two coordinate charts:

(1) let V_1 be the image in X of $\{(\mathbf{z}, \tau_1) \in \mathbb{T}^2 \times \mathbb{R} \mid 0 < \tau_1 < \tilde{\tau}(\mathbf{z})\}$ with coordinates (x_1, y_1, τ_1), and
(2) let V_2 be the image in X of $\{(\mathbf{z}, \tau_2) \in \mathbb{T}^2 \times \mathbb{R} \mid -\frac{1}{2}\tilde{\tau}(\mathcal{I}(\mathbf{z})) < \tau_2 < \frac{1}{2}\tilde{\tau}(\mathbf{z})\}$ with coordinates (x_2, y_2, τ_2).

Let $\mathcal{S}$ be defined to be the set of orbits of elements of the form $(\mathbf{z}, 0)$, so that $\mathcal{S} \subset X$ is a smooth submanifold. We consider $\mathcal{P}, \mathcal{I}, \mathcal{J}$ to map $\mathcal{S}$ into itself. The form $\omega = dx_j \wedge dy_j$ on V_j, $j = 1, 2$, defines a contact form on X extending ω on $\mathcal{S}$. For each $t \in \mathbb{R}$ the mapping $\mathbb{T}^2 \times \mathbb{R} \to \mathbb{T}^2 \times \mathbb{R}\colon (\mathbf{z}, \tau) \mapsto (\mathbf{z}, \tau + t)$ takes orbits to orbits, and so determines a mapping $\phi_t\colon X \to X$, which is a smooth contact flow; note that $F = \partial_{\tau_j}$ on V_j, $j = 1, 2$. The locally well-defined one-forms $\theta = d\tau_j + y_j\, dx_j$ on parts of V_j, $j = 1, 2$ ($y_j\, dx_j$ is not a globally defined one-form on $\mathbb{T}^2$), satisfy $\theta(F(\cdot)) = 1$, $\omega = -d\theta$, and the volume form $\theta \wedge \omega$ is given by $dx_j \wedge dy_j \wedge d\tau_j$, $j = 1, 2$. The mapping $\mathcal{I}\colon \mathcal{S} \to \mathcal{S}$ can be extended to a diffeomorphism (also denoted by $\mathcal{I}$) of X to itself: the mapping $\mathbb{T}^2 \times \mathbb{R} \to \mathbb{T}^2 \times \mathbb{R}\colon (\mathbf{z}, \tau) \mapsto (\mathcal{I}(\mathbf{z}), -\tau)$ maps orbits to orbits and so descends to a mapping of X to itself. Furthermore we have $\phi_t \circ \mathcal{I} = \mathcal{I} \circ \phi_{-t}$. We also have $\mathcal{I}^*\omega = -\omega$ on X. Therefore we obtain a contact flow with all the structure that we have cited as being characteristic of Hamiltonian systems.

However this system is better behaved than the Hamiltonian systems we mentioned earlier in that the Poincaré return mapping $\mathcal{P}$ is a very nice mapping. There is additional structure which makes the analysis of this example easier. Define

$$D(\mathbf{z}) = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ -\tilde{\tau}_x(\mathbf{z}) & -\tilde{\tau}_y(\mathbf{z}) & 1 \end{pmatrix}, \quad X(\mathbf{z}) = \begin{pmatrix} 1+\sqrt{5} & 1-\sqrt{5} & 0 \\ 2 & 2 & 0 \\ a_u(\mathbf{z}) & a_s(\mathbf{z}) & 1 \end{pmatrix},$$

and $\Lambda = \mathrm{diag}(\frac{3+\sqrt{5}}{2}, \frac{3-\sqrt{5}}{2}, 1)$. We choose the functions a_u, a_s such that $D(\mathbf{z})X(\mathbf{z}) = X(\mathcal{P}(\mathbf{z}))\Lambda$ for all $\mathbf{z} \in \mathbb{T}^2$. We find that the unique bounded solution [14] is $a_u(\mathbf{z}) = -\sum_{n=1}^{\infty}(\frac{3+\sqrt{5}}{2})^{-n}[(1+\sqrt{5})\tilde{\tau}_x(\mathcal{P}^{-n}(\mathbf{z})) + 2\tilde{\tau}_y(\mathcal{P}^{-n}(\mathbf{z}))]$, and $a_s(\mathbf{z}) = -a_u(\mathcal{I}(\mathbf{z}))$. These functions are only Hölder continuous even when $\tilde{\tau}$ is analytic (except when $\tilde{\tau} \equiv 1$). Define $b_1(\mathbf{z}, \tau_1) = 2\tau_1/\tilde{\tau}(\mathbf{z})$ on V_1 and

$$b_2(\mathbf{z}, \tau_2) = \begin{cases} 2\tau_2/\tilde{\tau}(\mathbf{z}) & \tau_2 \geq 0 \\ 2\tau_2/\tilde{\tau}(\mathcal{P}^{-1}(\mathbf{z})) & \tau_2 < 0 \end{cases}$$

on V_2. Define the Riemannian metric $g_j(\mathbf{z}, \tau_j) = X(\mathbf{z})^{-T}\Lambda^{b_j(\mathbf{z},\tau_j)}X(\mathbf{z})^{-1}$ on V_j, $j = 1, 2$. These together define a Hölder continuous Riemannian metric g on X. Let $\|\mathbf{v}\|_{\mathbf{x}}^2 = g(\mathbf{x})(\mathbf{v}, \mathbf{v})$ for all $\mathbf{v} \in T_{\mathbf{x}}X$. Define $\gamma(\mathbf{z}, t)$ to be the piecewise linear function passing through the points $\{(\alpha(n, \mathbf{z}), n)\}_{n\in\mathbb{Z}}$ where

$$\alpha(n, \mathbf{z}) = \begin{cases} \sum_{k=0}^{n-1} \tilde{\tau}(\mathcal{P}^k(\mathbf{z})) & n \geq 0, \\ -\sum_{k=1}^{-n} \tilde{\tau}(\mathcal{P}^{-k}(\mathbf{z})) & n < 0. \end{cases}$$

Then for all $\mathbf{x} = (\mathbf{z}, \tau_j) \in X$, $t \in \mathbb{R}$ we have

$$\begin{aligned} \|T\phi_t(\mathbf{x})[\mathbf{v}_{\pm,j}]\|_{\phi_t(\mathbf{x})} &= (\tfrac{3\pm\sqrt{5}}{2})^{\gamma(\mathbf{z},t+\tau_j)-\gamma(\mathbf{z},\tau_j)}\|\mathbf{v}_{\pm,j}\|_{\mathbf{x}}, \\ \mathbf{v}_{\pm,j} &= (1\pm\sqrt{5})\partial_{x_j} + 2\partial_{y_j} + a_u(\mathbf{z})\partial_{\tau_j} \end{aligned}$$

on V_j, $j = 1, 2$. $E_u(\mathbf{x}) = \mathrm{span}\{\mathbf{v}_{+,j}\}$ and $E_s(\mathbf{x}) = \mathrm{span}\{\mathbf{v}_{-,j}\}$ are called the *unstable* and *stable subspaces* respectively of $T_{\mathbf{x}}X$. If $E_0(\mathbf{x}) = \mathrm{span}\{F(\mathbf{x})\}$ then $T_{\mathbf{x}}X = E_u(\mathbf{x}) \oplus E_s(\mathbf{x}) \oplus E_0(\mathbf{x})$. The subbundles E_u, E_s, E_0 are invariant under the flow. Thus $\{\phi_t\}$ is an *Anosov flow.* Exact contact Anosov flows are discussed in [14]. Since hyperbolic toral automorphisms have many dense orbits, the flow $\{\phi_t\}$ has a trajectory dense in X.

Suppose $f_0\colon X \to \mathbb{R}$ is a (ideally smooth) function, where $\int_X f_0\, d\mu = 1$, and $f(\mathbf{x}, t) = f_0(\phi_{-t}(\mathbf{x}))$ for all $(\mathbf{x}, t) \in X \times \mathbb{R}$. Clearly if $\tilde{\tau} \equiv 1$ then the flow is not even weak mixing. However this does not seem to hinder the phenomenon we are studying. Let $\rho\colon V_1 \to \mathbb{R}/\mathbb{Z}\colon (x_1, y_1, \tau_1) \mapsto x_1 + \mathbb{Z}$ be a measurable fibration. Allow f_0 to be discontinuous and $\mathcal{A}_\rho$-measurable. Then by direct calculation one can show that $E(f(\cdot, t) \mid \mathcal{A}_\rho) = (1-t)f_0 + t$ for $0 \leq t \leq 1$ and $E(f(\cdot, t) \mid \mathcal{A}_\rho) = 1$ for all $t \geq 1$. This is "super-exponential" decay to the final state despite the limited regularity of f_0 and ρ. This behavior is unlikely to persist for nonconstant $\tilde{\tau}$.

However if $\tilde{\tau}(\mathbf{z}) = 1 + \frac{1}{2}\cos(2\pi x)$, then the periodic orbits starting at $(0, 0)$ and $(\frac{1}{5}, \frac{2}{5})$ have periods $3/2$ and $\frac{7+\sqrt{5}}{4}$, implying that the flow is topologically mixing. This implies that it is mixing with respect to μ; also the metric entropy and the topological entropy of the flow coincide and are equal to $\frac{3+\sqrt{5}}{2}$ [20], [5], [22], and [14]. Note that this choice of $\tilde{\tau}$ also satisfies $\tilde{\tau} \circ \mathcal{J} = \tilde{\tau}$. We will henceforth restrict our discussion to this case. The manifold X can be smoothly embedded in $\mathbb{R}^N$ for some $N \leq 6$. Composing such an embedding with a coordinate function $\mathbb{R}^N \to \mathbb{R}$, we get a smooth fibration ρ. We propose to study such fibrations to find a natural explicit one suitable for our larger investigation.

Assume now that f_0 is smooth. Let $m(\mathbf{x}, z) = \int_0^\infty e^{-izt} f(\mathbf{x}, t)\, dt$ for $\Im z < 0$ be the Fourier-Laplace transform of f. If we use the coordinates $\mathbf{x} = (\mathbf{y}, \tau_1)$ on V_1 then from the equation $\partial_t f + \partial_{\tau_1} f = 0$ we have $izm(\mathbf{y}, \tau_1, z) + \partial_{\tau_1} m(\mathbf{y}, \tau_1, z) = f_0(\mathbf{y}, \tau_1)$, for all $\mathbf{y} \in \mathbb{T}^2$. Using an integrating factor we get

$$m(\mathbf{y}, \tau_1, z) = e^{-iz\tau_1}\left[r(\mathbf{y}, z) + \int_0^{\tau_1} e^{iz\tau'} f_0(\mathbf{y}, \tau')\, d\tau'\right],$$

where $r(\mathbf{y}, z) = m(\mathbf{y}, 0, z)$. Imposing the boundary condition $m(\mathbf{y}, \tilde{\tau}(\mathbf{y}), z) = m(\mathcal{P}(\mathbf{y}), 0, z)$ we obtain

$$(*) \qquad r(\mathcal{P}(\mathbf{y}), z) - e^{-iz\tilde{\tau}(\mathbf{y})} r(\mathbf{y}, z) = \int_0^{\tilde{\tau}(\mathbf{y})} e^{-iz[\tilde{\tau}(\mathbf{y}) - \tau']} f_0(\mathbf{y}, \tau')\, d\tau'.$$

This equation for a function r on $\mathbf{T}^2$ is the focus of our approach to the rate of mixing (or decay of correlations) problem for this system. Since $f(\mathbf{x}, t) = \frac{1}{2\pi}\int_{\mathcal{C}} e^{izt} m(\mathbf{x}, z)\, dz$, where the integration is along a contour $\mathcal{C}$ parameterized as $z = k - is$, $k \in \mathbb{R}$, $s > 0$, the decay in t of f is determined by the singularities in the z variable of m (the contour of integration can be shifted up with circuits around any poles spawning the dominant asymptotic contributions). The term involving f_0 on the right side of (*) is analytic in z so all the singularities come from problems with inverting the operator on the left side of (*). We know $f(\mathbf{x}, t)$ does not have any decay in t for fixed $\mathbf{x}$, so we expect a "wall of singularities" in $m(\mathbf{x}, z)$ on the real z axis. If $g\colon X \to \mathbb{R}$ is another smooth function then

$$\int_0^\infty e^{-izt} C_{-t}(f_0, g)\, dt = \int_X [m(\mathbf{x}, z) - \tfrac{1}{z}] g(\mathbf{x})\, d\mu(\mathbf{x}).$$

(Here we are assuming $\int_X f_0\, d\mu = 1$.) The singularities in the z behavior of this function is called the *correlation spectrum* of the flow. If there is to be exponential decay of correlations then this function should extend analytically to a half-plane $\Im z < \epsilon$, $\epsilon > 0$; it is known to be meromorphic in such a half-plane with no poles on the real z axis [22]. So the "wall of singularities" in m (or r) has disappeared after integration in X against a good function g. Instead of considering correlation we could look at the conditional expectation.

$$\int_0^\infty e^{-izt} E(f(\cdot, t) \mid \mathcal{A}_\rho)\, dt = E(m(\cdot, z) \mid A_\rho).$$

Thus the phenomenon of relaxation to equilibrium is related to analytically extending this function beyond its natural domain $\Im z < 0$. (Since $\bar{f} \equiv 1$ we have a simple pole at $z = 0$.) Again, the wall of singularities in r must disappear after integration against fibre measures. Roughly speaking, the part of the "distribution" $\partial_{\bar{z}} r$ concentrated on the real z-axis (except for the delta mass at $z = 0$) must give zero when applied essentially any smooth test function. This analytical subtlety is deserving of thorough investigation, which is what we propose to do.

One simple-minded approach is to introduce various finite dimensional approximations to the operators on the left side of (*) and to study how these approximations behave in the infinite dimensional limit. If we use a Fourier series representation for r then we find that multiplication by $e^{-iz\tilde{\tau}(\mathbf{y})}$ is not very simple to analyze in regard to its behavior in z. Another idea is to use the periodic orbit structure of $\mathcal{P}$. The 320 periodic orbits of least periods $1, 2, 3, 6$ are parameterized by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{40}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} j \\ k \end{pmatrix}, \qquad -10 < j \leq 10, -8 < k \leq 8.$$

Representing the function $r(\mathbf{y})$ by a column vector of its values at these points we can express $r(\mathcal{P}(\mathbf{y}))$ in terms of a 320×320 matrix C which is a direct sum of circulant matrices (cyclic permutations of the identity matrix). The term $e^{-iz\tilde{\tau}(\mathbf{y})}r(\mathbf{y})$ is expressed via a diagonal matrix $D(z)$. Thus the left side of (*) is represented by an invertible matrix $C - D(z)$, and the complete z dependence, including the poles and residues, of the inverse matrix $(I - C^{-1}D(z))^{-1}C^{-1}$ are explicitly computed. If $\mathbf{y} \in \mathbb{T}^2$ satisfies $\mathcal{P}^m\mathbf{y} = \mathbf{y}$, where m is its least period, then let $T(\mathbf{y}) = \sum_{j=0}^{m-1} \tilde{\tau}(\mathcal{P}^j(\mathbf{y}))$. The poles are located at the points $\frac{2\pi}{T(\mathbf{y})}\mathbb{Z}$, where $\mathbf{y}$ ranges over representatives of each periodic orbit. This is the beginning of the "wall of singularities". The residues are highly oscillatory functions of $\mathbf{y}$. Furthermore (and here things start getting really interesting) $\det(I - C^{-1}D(z))$ can be explicitly computed:

$$\det(I - C^{-1}D(z)) = \prod_{\mathbf{y}}(1 - e^{-izT(\mathbf{y})}),$$

where the product is over representatives $\mathbf{y}$ from each of the periodic orbits: one of period 1, two of period 2, 5 of period 3, and 50 of period 6. This is remarkable because the *unweighted dynamical zeta function* $\zeta^*(iz)$ [2], [20], [24] is given by

$$\zeta^*(s) = \prod_{\mathcal{O}}(1 - e^{-sT(\mathcal{O})})^{-1},$$

where the product is over all the periodic orbits $\mathcal{O}$ of the flow. Here we seem to have a direct connection between dynamical zeta functions and correlation spectra without the intervention of transfer operators.

But disturbing questions come in to mar this picture. The infinite product defining the zeta function will only converge if $\Re s > (\frac{3+\sqrt{5}}{2})$ (the topological entropy of the flow). The relation between truncations of the infinite product for a zeta function and the full zeta function is notoriously difficult. Thus transfer operators allow one (in some cases) to mediate between the zeta function singularities and the correlation spectrum (f_0, g in a Banach space of Hölder continuous functions) in a rigorous manner; the correlation spectra and the poles of $\zeta^*(iz)$ are offset by $\frac{3+\sqrt{5}}{2}$ in the imaginary direction (see references cited above). This suggests that our idea of representing a function by its values on periodic orbits is naive, and the infinite dimensional limit will be difficult to control.

However this is just the beginning of our search for finite dimensional representations of the operators on the left of (*) which yield insight. The operator $\mathcal{P}$ should be expressed using basis functions which have regularity which is dual in the unstable and stable directions [25], [12]. Although this is easy enough to arrange with Fourier series, the spatial nonlocality of that basis makes the multiplication operator complicated. One possibility is to develop a special "wavelet" style basis patterned after the lattice of periodic orbits, allowing spatial localization and yet permiting directional control of regularity [21]. Another more elementary approach is to develop finite dimensional norms, based on a piecewise linear approximation of r, which allow control of the regularity in the infinite dimensional limit. Perhaps best matrix approximations in these norms will explain the puzzle of the $i\frac{3+\sqrt{5}}{2}$ offset. We propose an investigation in all these directions.

5. Geodesic Flows and Symbolic Dynamics

The most intensely studied example of a chaotic contact flow is that of the geodesic flow. If Σ is a compact Riemannian surface with constant negative curvature and $X = T_1\Sigma$ is a submanifold of the symplectic manifold $T^*\Sigma$, then the pullback of the symplectic form from $T^*\Sigma$ defines the contact form on X. The Hamiltonian is as described in the introduction where the potential energy function is identically zero. This example is amenable to detailed analysis because the phase space can be expressed as $\Gamma\backslash \mathrm{Sl}(2,\mathbb{R})$, where Γ is a discrete subgroup of $\mathrm{Sl}(2,\mathbb{R})$. The geodesic flow then has the explicit representation as $\phi_t(\Gamma g) = \Gamma g\,\mathrm{diag}(e^t, e^{-t})$, for all $g \in \mathrm{Sl}(2,\mathbb{R})$. [14] contains many details about geodesic flows; [23] contains a detailed proof of decay of correlations, [10] contains a proof of an analogous result for variable negative curvature. [22] contains a discussion of detailed large-time asymptotics of the correlation function. The exact exponential decay rate is computable in terms of the eigenvalues of the Laplace-Beltrami operator on Σ.

The question arises then as to whether our main example, i.e. special flows over hyperbolic toral automorphisms, is subsumed by the example of geodesic flows. The answer is no since the stable and unstable subbundles of geodesic flows are C^1 whereas the functions a_u, a_s defining those subbundles in our case are merely Hölder continuous.

Thus geodesic flows provide for us a source of inspiration as to what to prove, but not necessarily how to prove it. Also, after a clearer understanding is obtained for our problem these results might be combined with results on geodesic flows to suggest general principles about rates of mixing. Such general principles of course would be the most important broader impact of the proposed research.

Finally we mention a class of examples for which very complete information is known. These are the special flows over subshifts of finite type, c.f. [2] and references therein. The phase space X' of such a flow is constructed via a roof function $\tilde{\tau}'$ over a space $\mathcal{S}'$ equipped with a bilateral shift $\mathcal{P}'$. Unfortunately the space $\mathcal{S}'$ seems to have no discernable symplectic structure. $\mathcal{S}'$ is a metric space, so there is a well-defined notion of Lipschitz continuous function, but no notion of smoother functions. There is a fibration $\rho' \colon \mathcal{S}' \to \mathcal{S}'_+$, where the space $\mathcal{S}'_+$ supports a unilateral shift $\mathcal{P}'_+ \colon \mathcal{S}'_+ \to \mathcal{S}'_+$, where $\rho' \circ \mathcal{P}' = \mathcal{P}'_+ \circ \rho'$. If $\tilde{g} \colon \mathcal{S}'_+ \to \mathbb{R}$ is a continuous function then $g = \tilde{g} \circ \rho'$ is a continuous $\mathcal{A}_\rho$-measurable function on $\mathcal{S}$. The function $E(g \circ \mathcal{P}'^{-1} \mid \mathcal{A}_{\rho'})$ can be interpreted as a function $\mathcal{L}\tilde{g}$ on $\mathcal{S}'_+$. The operator $\mathcal{L}$ is called the *Frobenius-Perron operator*, or the *Ruelle transfer operator* [3]. The spectral properties of $\mathcal{L}$ mediate between the correlation spectra and the dynamical zeta function for these types of flows [2].

If one has another flow $\{\phi_t\}$ on X, say an Anosov (or Axiom A) flow, then by introducing a *Markov partition* one can obtain a mapping $\xi \colon X' \to X$ intertwining the special flow on X' with the flow $\{\phi_t\}$ [20]. This formalism is called *symbolic dynamics*. Unfortunately the mapping ξ does not faithfully represent all aspects of the structure of X; however it does remarkably well, and this approach is the basis of most of the strongest results about Anosov (or Axiom A) flows. For our example (i.e. special flows over hyperbolic toral automorphisms) a Markov partition for $\mathcal{P}$ is developed in [14]. It would be useful to compute the exact relation between the correlation spectrum for our example and the correlation spectrum for its symbolic dynamical model. Probably this has been done, but the proposer cannot find it in the literature. However a general summary of these sorts of calculations in [2] suggests that that the mapping ξ introduces "spurious poles", and sorting out the exact large-time behavior by this method could be very difficult. Thus we have given preference to more direct approaches.

References

[1] R. Abraham, J.E. Marsden, *Foundations of Mechanics*, second edition, Benjamin/Cummings, London, 1978.

[2] V. Baladi, *Periodic Orbits and Dynamical Spectra*, Ergod. Th. & Dynam. Sys., **18**, (1998), 255–292.

[3] V. Baladi, *Positive Transfer Operators and Decay of Correlations*, World Scientific, Singapore, 2000.

[4] A. Bolsinov, H.R. Dullin, A. Wittek, *Topology of Energy Surfaces and Existence of Transverse Poincaré Sections*, J. Phys. A: Math. Gen., **29**, (1996), 4977–4985.

[5] R. Bowen, D. Ruelle, *The Ergodic Theory of Axiom A Flows*, Invent. Math., **29**, (1975), 181–202.

[6] L.A. Bunimovich, *The Ergodic Properties of certain Billiards*, Funct. Anal. Appl., **8**, (1974), 268–269.

[7] I.P. Cornfeld, S.V. Fomin, Ya.G. Sinai, *Ergodic Theory*, Springer-Verlag, New York, 1982.

[8] D. B. Dix, *Large-Time Behavior of Solutions of Linear Dispersive Equations*, Lecture Notes in Math, **1668**, 1997.

[9] D. B. Dix, *Polyspherical Coordinate Systems on Orbit Spaces with Applications to Biomolecular Conformation*, preprint, 2002.

[10] D. Dolgopyat, *On Decay of Correlations in Anosov Flows*, Annals Math., **147**, (1998), 357–390.

[11] H.R. Dullin, A. Wittek, *Complete Poincaré Sections and Tangent Sets*, J. Phys. A: Math. Gen., **28**, (1995), 7157–7180.

[12] D. Fried, *Meromorphic Zeta Functions for Analytic Flows*, Commun. Math. Phys., **174**, (1995), 161–190.

[13] G. Haller, *Chaos Near Resonance*, Springer-Verlag, New York, 1999.

[14] A. Katok, B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, Cambridge, 1995.

[15] G. Keller, *Equilibrium States in Ergodic Theory*, Cambridge University Press, Cambridge, 1998.

[16] A.J. Lichtenberg, M.A. Lieberman, *Regular and Chaotic Dynamics*, Springer-Verlag, New York, 1992.

[17] C. Liverani, *Interacting Particles*, in *Hard Ball Systems and the Lorentz Gas*, 179–216, Encylopaedia Math. Sci., **101**, Springer, Berlin, 2000.

[18] J.E. Marsden, *Lectures on Mechanics*, London Mathematical Society Lecture Note Series, **174**, Cambridge University Press, Cambridge, 1992.

[19] A. Neumaier, *Molecular Modeling of Proteins and Mathematical Prediction of Protein Structure*, SIAM Rev., **39**, no. 3, (1997), 407–460.

[20] W. Parry, M. Pollicot, *Zeta Functions and the Periodic Orbit Structure of Hyperbolic Dynamics*, Astérisque, **187-188**, (1990).

[21] A.P. Petukhov, *Periodic Wavelets*, Sbornik: Math., **188**, (1997), 1481–1506.

[22] M. Pollicott, *On the Rate of Mixing of Axiom A Flows*, Invent. Math., **81**, (1985), 413–426.

[23] M. Ratner, *The Rate of Mixing for Geodesic and Horocycle Flows*, Ergod. Th. & Dynam. Sys., **7**, (1987), 267–288.

[24] D. Ruelle, *Dynamical Zeta Functions and Transfer Operators*, Notices A.M.S., **49**, no. 8, (2002), 887–895.

[25] H.H. Rugh, *The Correlation Spectrum for Hyperbolic Analytic Maps*, Nonlinearity, **5**, (1992), 1237–1263.

[26] L. Stoyanov, *Spectrum of the Ruelle Operator and Exponential Decay of Correlations for Open Billiard Flows*, Amer. J. Math., **123**, (2001), 715–759.

[27] D.W. Strook, *Probability Theory, An Analytic View*, Cambridge University Press, Cambridge, 1999.

[28] S. Wiggins, *Chaotic Transport in Dynamical Systems*, Springer-Verlag, New York, 1992.

[29] M. Wojtkowski, *Principles for the Design of Billiards with Nonvanishing Lyapunov Exponents*, Commun. Math. Phys., **105**, (1986), 391–414.

Department of Mathematics, University of South Carolina

E-mail address: `dix@math.sc.edu`