WHAT ARE LYAPUNOV EXPONENTS, AND WHY ARE THEY INTERESTING?

AMIE WILKINSON

INTRODUCTION

At the 2014 International Congress of Mathematicians in Seoul, South Korea, Franco-Brazilian mathematician Artur Avila was awarded the Fields Medal for "his profound contributions to dynamical systems theory, which have changed the face of the field, using the powerful idea of renormalization as a unifying principle."[1] Although it is not explicitly mentioned in this citation, there is a second unifying concept in Avila's work that is closely tied with renormalization: Lyapunov (or characteristic) exponents. Lyapunov exponents play a key role in three areas of Avila's research: smooth ergodic theory, billiards and translation surfaces, and the spectral theory of 1-dimensional Schrödinger operators. Here we take the opportunity to explore these areas and reveal some underlying themes connecting exponents, chaotic dynamics and renormalization.

But first, what are Lyapunov exponents? Let's begin by viewing them in one of their natural habitats: the iterated barycentric subdivision of a triangle.

When the midpoint of each side of a triangle is connected to its opposite vertex by a line segment, the three resulting segments meet in a point in the interior of the triangle. The barycentric subdivision of a triangle is the collection of 6 smaller triangles determined by these segments and the edges of the original triangle:
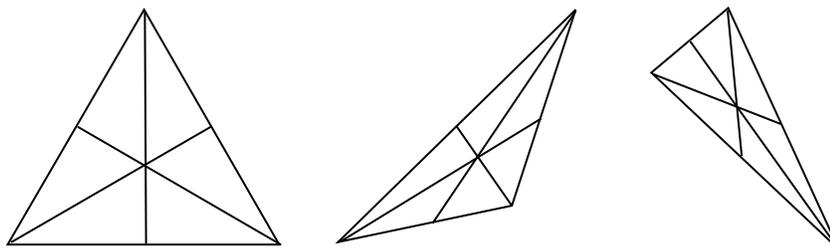
FIGURE 1. Barycentric subdivision.
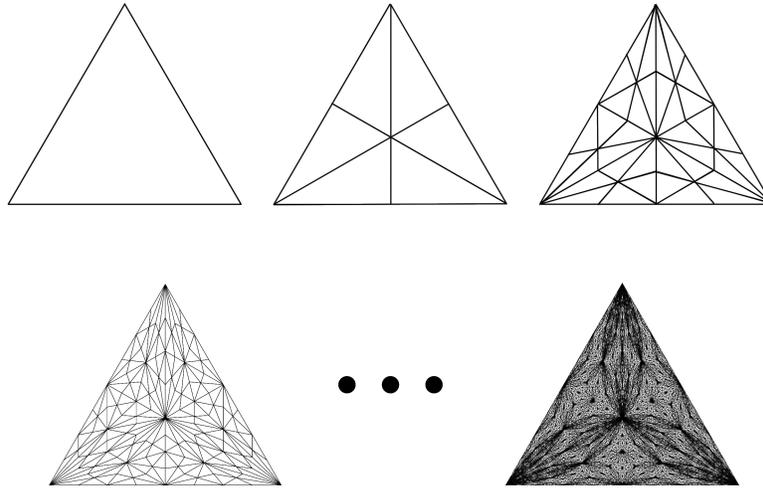
FIGURE 2. Iterating barycentric subdivision (courtesy of C. Mc-Mullen, [54], used with permission).

As the process of barycentric subdivision starts with a triangle and produces triangles, it is natural to iterate the process, barycentrically subdividing the 6 triangles obtained at the first step, obtaining 36 triangles, and so on, as in Figure 2.

Notice that as the subdivision gets successively finer, many of the triangles produced by subdivision get increasingly eccentric and needle-like. We can measure the skinniness of a triangle T via the *aspect ratio* $\alpha(T) = \operatorname{area}(T)/L(T)^2$, where $L(T)$ is the maximum of the side lengths; observe that similar triangles have the same aspect ratio. Suppose we fix a rule for labeling the triangles in a possible subdivision 1 through 6, roll a 6-sided fair die and at each stage choose a triangle to subdivide. The sequence of triangles $T_1 \supset T_2 \supset \cdots$ obtained have aspect ratios $\alpha_1, \alpha_2, \ldots$, where $\alpha_n = \alpha(T_n)$.

Theorem 0.1 ([54], see also [16]). *There exists a real number $\chi \approx 0.0446945 > 0$ such that almost surely,*

$$\lim_{n \to \infty} \frac{1}{n} \log \alpha_n = -2\chi.$$

In other words, if triangles are chosen successively by a random toss of the die, then with probability 1, their aspect ratios will tend to 0 in the nth toss at an exponential rate governed by $\exp(-2n\chi)$. The same conclusion holds with the same value of χ if the initial equilateral triangle is replaced by any marked triangle. This magical number χ is a Lyapunov exponent. We return to this example at the end of the next section.

Lyapunov exponents make multiple appearances in the analysis of dynamical systems. After defining basic concepts and explaining examples in Section 1, we describe in Sections 2–4 a sampling of Avila's results in smooth ergodic theory, Teichmüller theory and spectral theory, all of them tied to Lyapunov exponents in a fundamental way. We explore some commonalities of these results in Section 5. Section 6 is devoted to a discussion of some themes that arise in connection with Lyapunov exponents.

1. Cocycles, exponents, and hyperbolicity

Formally, Lyapunov exponents are quantities associated to a cocycle over a measure-preserving dynamical system. A measure-preserving dynamical system is a triple (Ω, μ, f), where (Ω, μ) is a probability space, and $f \colon \Omega \to \Omega$ is a measurable map that preserves the measure μ, meaning that $\mu(f^{-1}(X)) = \mu(X)$ for every measurable set $X \subset \Omega$. We say that (Ω, μ, f) is *ergodic* if the only f-invariant measurable sets have μ-measure 0 or 1. Equivalently, (Ω, μ, f) is ergodic if the only functions $\phi \in L^2(\Omega, \mu)$ satisfying $\phi \circ f = \phi$ are the constant functions. Any f-invariant measure μ can be canonically decomposed into ergodic invariant measures, a fact that allows us to restrict our attention to ergodic measures in some contexts, simplifying statements. The measures in such a decomposition are called *ergodic components,* and there can be uncountably many of them. The process of ergodic decomposition is a bit technical to describe; we refer the reader to [50] for more details.

1.1. Examples of measure-preserving systems. Here is a list of three types of measure-preserving systems that we will refer to again in the sections that follow.

Rotations on the circle. On the circle $\Omega = \mathbb{R}/\mathbb{Z}$, let $f_\alpha(x) = x + \alpha \, (\mathrm{mod}\, 1)$, where $\alpha \in \mathbb{R}$ is fixed. The map f_α preserves the Lebesgue–Haar measure μ (that assigns to an interval $I \in \mathbb{R}/\mathbb{Z}$ its length $|I|$). The map f_α is ergodic with respect to μ if and only if α is irrational. This has a straightforward proof: consider the equation $\phi \circ f_\alpha = \phi$, for some $\phi \in L^2(\mathbb{R}/\mathbb{Z}, \mu)$, and solve for the Fourier coefficients of ϕ.

When $\alpha = p/q$ is rational, every point $\omega \in \Omega$ is periodic, satisfying $f^q(\omega) = \omega$. Each ω then determines an ergodic f_α-invariant probability measure ν_ω obtained by averaging the Dirac masses along the orbit of ω:

$$\nu_\omega := \frac{1}{q} \left(\delta_\omega + \delta_{f_\alpha(\omega)} + \cdots + \delta_{f_\alpha^{q-1}(\omega)} \right).$$

Each ν_ω is an ergodic component of the measure μ, and hence there are uncountably many such components.

When α is irrational, μ is the *unique* f_α-invariant Borel probability measure. A homeomorphism of a compact metric space that has a unique invariant measure is called *uniquely ergodic*—a property that implies ergodicity and more. Unique ergodicity is mentioned again in Section 2 and is especially relevant to the discussion of Schrödinger operators with quasiperiodic potentials in Section 4.

There is nothing particularly special about the circle, and the properties of circle rotations listed here generalize easily to rotations on compact abelian groups.

Toral automorphisms. Let $\Omega = \mathbb{T}^2 := \mathbb{R}^2/\mathbb{Z}^2$, the 2-torus. Fix a matrix $A \in SL(2, \mathbb{Z})$. Then A acts linearly on the plane by multiplication and preserves the lattice $\mathbb{Z}^2$, by virtue of having integer entries and determinant 1. It therefore induces a map $f_A \colon \mathbb{T}^2 \to \mathbb{T}^2$ of the 2-torus, a group automorphism. The area μ is preserved because $\det(A) = 1$. Such an automorphism is ergodic with respect to μ if and only if the eigenvalues of A are λ and λ^{-1}, with $|\lambda| > 1$. This can be proved by examining the Fourier coefficients of $\phi \in L^2(\mathbb{T}^2, \mu)$ satisfying $\phi \circ f_A = \phi$: composing with f_A permutes the Fourier coefficients of ϕ by the adjoint action of A, and the assumption on the eigenvalues of A implies that if ϕ is not constant

there must be infinitely many identical coefficients, which violates the assumption that $\phi \in L^2(\mathbb{T}^2, \mu)$.[2]

In contrast to the irrational rotation f_α, the map f_A has many ergodic invariant Borel probability measures, even when f_A is ergodic with respect to the area μ. For example, as we have just seen, every periodic point of f_A determines an ergodic invariant measure, and f_A has infinitely many periodic points. This is a simple consequence of the Pigeonhole Principle, using the fact that $A \in SL(2, \mathbb{Z})$: for every natural number q, the finite collection of points $\{(\frac{p_1}{q}, \frac{p_2}{q}) : p_1, p_2 \in \{0, \ldots, q-1\}\} \subset \mathbb{T}^2$ is permuted by f_A, and so each element of this set is fixed by some power of f_A.

Bernoulli shifts. Let $\Omega = \{1, \ldots, k\}^{\mathbb{N}}$ be the set of all infinite, one sided strings $\omega = (\omega_1, \omega_2, \ldots)$ on the alphabet $\{1, \ldots, k\}$. Endowed with the product topology, the space Ω is compact, metrizable, and homeomorphic to a Cantor set. The shift map $\sigma \colon \Omega \to \Omega$ is defined by $\sigma(\omega)_k = \omega_{k+1}$. In other words, the image of the sequence $\omega = (\omega_1, \omega_2, \ldots)$ is the shifted sequence $\sigma(\omega) = (\omega_2, \omega_3, \ldots)$. Any nontrivial probability vector $p = (p_1, \ldots, p_k)$ (i.e., with $p_i \in (0, 1)$, and $\sum_i p_i = 1$) defines a product measure $\mu = p^{\mathbb{N}}$ supported on Ω.[3] The triple $\left(\Omega, \sigma, \mu = (p_1, \ldots, p_k)^{\mathbb{N}}\right)$ is called a *Bernoulli shift*, and μ is called a *Bernoulli measure*. It is not hard to see that the shift σ preserves μ and is ergodic with respect to μ.

The shift map $\sigma \colon \Omega \to \Omega$ manifestly has uncountably many ergodic invariant Borel probability measures, in particular the Bernoulli measures $(p_1, \ldots, p_k)^{\mathbb{N}}$, but the list does not end there. In addition to periodic measures (supported on orbits of periodic strings $(\omega_1, \ldots, \omega_q, \omega_1, \ldots, \omega_q, \ldots)$), there are σ-invariant probability measures on Ω encoding every measure-preserving continuous dynamical system[4]—in this sense the shift is a type of universal dynamical system.

1.2. Cocycles. Let $M_{d \times d}$ be the d^2-dimensional vector space of $d \times d$ matrices (real or complex). A *cocycle* is a pair (f, A), where $f \colon \Omega \to \Omega$ and $A \colon \Omega \to M_{d \times d}$ are measurable maps. We also say that A is a *cocycle over f*. For each $n > 0$ and $\omega \in \Omega$, we write

$$A^{(n)}(\omega) = A(f^{n-1}(\omega))A(f^{n-2}(\omega)) \cdots A(f(\omega))A(\omega),$$

where f^n denotes the n-fold composition of f with itself. For $n = 0$, we set $A^{(n)}(\omega) = I$, and if both f and the values of the cocycle A are invertible, we also define, for $n \geq 1$,

$$A^{(-n)}(\omega) = (A^{(n)}(f^{-n+1}(\omega)))^{-1} = A^{-1}(f^{-n+1}(\omega)) \cdots A^{-1}(\omega).$$

One comment about the terminology "cocycle": while A is colloquially referred to as a cocycle over f, to fit this definition into a proper cohomology theory, one should reserve the term cocycle for the function $(n, \omega) \to A^{(n)}(\omega)$ and call A the *generator* of this (1-)cocycle. See [11] for a more thorough discussion of this point.

---

[2]In higher dimensions, a matrix $A \in SL(d, \mathbb{Z})$ similarly induces an automorphism f_A of $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$. The same argument using Fourier series shows that f_A is ergodic if and only if A does not have a root of 1 as an eigenvalue.

[3]The product measure has a simple description in this context: setting $C_i(j) := \{\omega \in \Omega : \omega_i = j\}$, the measure $\mu = (p_1, \ldots, p_k)^{\mathbb{N}}$ is defined by the properties $\mu(C_i(j)) = p_j$, and $\mu(C_i(j) \cap C_{i'}(j')) = p_j p_{j'}$, for any $i \neq i' \in \mathbb{N}$ and $j, j' \in \{1, \ldots, k\}$.

[4]Subject to some constraints involving invertibility and entropy.

A fruitful way of viewing a cocycle A over f is as a hybrid dynamical system $(f, A) \colon \Omega \times M_{d \times d} \to \Omega \times M_{d \times d}$ defined by

$$(f, A)(\omega, B) = (f(\omega), A(\omega)B).$$

Note that the nth iterate $(f, A)^n$ of this hybrid map is the hybrid map $(f^n, A^{(n)})$. The vector bundle $\Omega \times M_{d \times d}$ can be reduced in various ways to obtain associated hybrid systems, for example, the map $(f, A) \colon \Omega \times \mathbb{R}^d \to \Omega \times \mathbb{R}^d$ defined by $(f, A)(\omega, v) = (f(\omega), A(\omega)v)$. Thus a natural generalization of a cocycle over f is a map $F \colon \mathcal{B} \to \mathcal{B}$, where $\pi \colon \mathcal{B} \to \Omega$ is a vector bundle, and F acts linearly on fibers, with $\pi \circ F = f \circ \pi$. We will use this extended definition of cocycle to define the derivative cocycle in Subsection 1.4.

1.3. **Lyapunov exponents.** Let $f \colon \Omega \to \Omega$ be a measurable map (not necessarily preserving a probability measure). We say that a real number χ is a *Lyapunov exponent for the cocycle $A \colon \Omega \to M_{d \times d}$ over f at the point $\omega \in \Omega$* if there exists a nonzero vector $v \in \mathbb{R}^d$, such that

$$(1) \qquad \lim_{n \to \infty} \frac{1}{n} \log \|A^{(n)}(\omega)v\| = \chi.$$

Here $\| \cdot \|$ is a fixed norm on the vector space space $M_{d \times d}$. The limit in (1), when it exists, does not depend on the choice of such a norm (exercise).

Oseledets proved in 1968 [55] that if (f, Ω, μ) is a measure-preserving system and A is a cocycle over f satisfying the integrability condition $\log \|A\| + \log \|A^{-1}\| \in L^1(\Omega, \mu)$, then for μ-almost every $\omega \in \Omega$ and for every nonzero $v \in \mathbb{R}^d$ the limit in (1) exists. This limit assumes at most d distinct values $\chi_1(\omega) > \chi_2(\omega) > \cdots > \chi_{k(\omega)}(\omega)$. Each exponent $\chi_i(\omega)$ is achieved with a multiplicity $d_i(\omega)$ equal to the dimension of the space of vectors v satisying (1) with $\chi = \chi_i$, and these multiplicities satisfy $\sum_{i=1}^{k(\omega)} d_i(\omega) = d$.

If the cocycle A takes values in $SL(d, \mathbb{R})$, then, since $\log \det(A(\omega)) \equiv \log(1) = 0$, one also has that $\sum_{i=1}^{k(\omega)} d_i(\omega) \chi_i(\omega) = 0$. Thus if A takes values in $SL(2, \mathbb{R})$, then the exponents are of the form $-\chi(\omega) \leq 0 \leq \chi(\omega)$.

If (Ω, μ, f) is ergodic, then the functions $k(\omega)$, $\chi_i(\omega)$ and $d_i(\omega)$ are constant μ-almost everywhere. In this case, the essential values $\chi_1, \ldots, \chi_k \in \mathbb{R}$ are called the *Lyapunov exponents of A with respect to the ergodic measure μ.*

1.4. **Two important classes of cocycles.** *Random matrix cocycles* encode the behavior of a random product of matrices. Let $\{A_1, \ldots, A_k\} \subset M_{d \times d}$ be a finite collection of matrices. Suppose we take a k-sided die and roll it repeatedly. If the die comes up with the number j, we choose the matrix A_j, thus creating a sequence $A_{\omega_1}, A_{\omega_2}, \ldots$, where $\omega = (\omega_1, \omega_2, \ldots) \in \{1, \ldots, k\}^{\mathbb{N}}$. This process can be packaged in a cocycle A over a measure-preserving system (Ω, μ, σ) by setting

FIGURE 3. A 30-sided die, a.k.a. d30.

$\Omega = \{1, \ldots, k\}^{\mathbb{N}}$, $\mu = (p_1, \ldots, p_k)^{\mathbb{N}}$, where p_j is the probability that the die shows j on a roll, and setting σ to be the shift map. The cocycle is defined by $A(\omega) = A_{\omega_1}$. Then $A^{(n)}(\omega)$ is simply the product of the first n matrices produced by this process.

More generally, suppose that η is a probability measure on the set of matrices $M_{d \times d}$. The space $\Omega = M_{d \times d}^{\mathbb{N}}$ of sequences $(M_1, M_2, \ldots)$ carries the product measure $\eta^{\mathbb{N}}$, which is invariant under the shift map σ, where as above $\sigma(M_1, M_2, \ldots) = (M_2, M_3, \ldots)$. There is a natural cocycle $A \colon \Omega \to M_{d \times d}$ given by $A((M_1, M_2, \ldots)) = M_1$. The matrices $A^{(n)}(\omega)$, for $\omega \in \Omega$ are just n-fold random products of matrices chosen independently with respect to the measure η.

In the study of smooth dynamical systems the *derivative cocycle* is a central player. Let $f \colon M \to M$ be a C^1 map on a compact d-manifold M. Suppose for simplicity that the tangent bundle is trivial: $TM = M \times \mathbb{R}^d$. Then for each $x \in M$, the derivative $D_x f \colon T_x M \to T_{fx} M$ can be written as a matrix $D_x f \in M_{d \times d}$. The map $x \mapsto D_x f$ is called the derivative cocycle. The Chain Rule implies that if $A = Df$ is a derivative cocycle, then $D_x f^n = A^{(n)}(x)$.

The case where TM is not trivializable is easily treated: either one trivializes TM over a suitable subset of M or one expands the definition of cocycle as described at the end of Subsection 1.2: the map $Df \colon TM \to TM$ is an automorphism of the vector bundle TM, covering the map f. Lyapunov exponents for the derivative cocycle are defined analogously to (1). We fix a continuous choice of norms $\{\|\cdot\|_x \colon T_x M \to \mathbb{R}_{\geq 0} : x \in M\}$, for example the norms given by a Riemannian metric (more generally, such a family of norms is called a *Finsler*). Then χ is a Lyapunov exponent for Df at $x \in M$ if there exists $v \in T_x M$ such that

$$\text{(2)} \qquad \lim_{n \to \infty} \frac{1}{n} \log \|D_x f^n v\|_{f^n(x)} = \chi.$$

Since M is compact, the Lyapunov exponents of Df do not depend on the choice of Finsler. The conclusions of Oseledets's theorem hold analogously for derivative cocycles with respect to any f-invariant measure on M.

A simple example of a derivative cocycle is provided by the toral automorphism $f_A \colon \mathbb{T}^2 \to \mathbb{T}^2$ described above. Conveniently, the tangent bundle to $\mathbb{T}^2$ is trivial, and the derivative cocycle is the constant cocycle $D_x f_A = A$.

1.5. **Uniformly hyperbolic cocycles.** A special class of cocycles whose Lyapunov exponents are nonzero with respect to *any* invariant probability measure are the uniformly hyperbolic cocycles.

Definition 1.1. A continuous cocycle $A \colon \Omega \to Gl(d)$ over a homeomorphism $f \colon \Omega \to \Omega$ of a compact metric space Ω is *uniformly hyperbolic* if there exists an integer $n \geq 1$, and for every $\omega \in \Omega$, there is a splitting $\mathbb{R}^d = E^u(\omega) \oplus E^s(\omega)$ into subspaces that depend continuously on ω, such that for every $\omega \in \Omega$:

   (i) $A(\omega)E^u(\omega) = E^u(f(\omega))$, and $A(\omega)E^s(\omega) = E^s(f(\omega))$,
   (ii) $v \in E^u(\omega) \implies \|A^{(n)}(\omega)v\| \geq 2\|v\|$, and
   (iii) $v \in E^s(\omega) \implies \|A^{(-n)}(\omega)v\| \geq 2\|v\|$.

The definition is independent of choice of norm $\|\cdot\|$; changing the norm on $M_{d \times d}$ simply changes the value of n. The number 2 in conditions (ii) and (iii) can be replaced by any fixed real number greater than 1; again this only changes the value of n. Notice that measure plays no role in the definition of uniform

hyperbolicity, as it is a topological property of the cocycle. For short, we say that (f, A) is uniformly hyperbolic.

A trivial example of a uniformly hyperbolic cocycle is the constant cocycle $A(\omega) \equiv A_0 \in SL(2, \mathbb{R})$, where A_0 is any matrix with eigenvalues $\lambda > \lambda^{-1}$ satisfying $\lambda > 1$. Here the splitting $\mathbb{R}^2 = E^u(\omega) \oplus E^s(\omega)$ is the constant splitting into the sum of the λ and λ^{-1} eigenspaces of A_0, respectively. For a constant cocycle, the Lyapunov exponents are defined everywhere and are also constant; for this $SL(2, \mathbb{R})$ cocycle, the exponents are $\pm \log \lambda$.

A nontrivial example of a uniformly hyperbolic cocycle is any nonconstant, continuous $A \colon \Omega \to SL(2, \mathbb{R})$ with the property that the entries of $A(\omega)$ are all positive, for any $\omega \in \Omega$. In this case the splitting is given by

$$E^u(\omega) := \bigcap_{n \geq 0} A^{(n)}(f^{-n}(\omega)) \, (\mathcal{C}_+) \quad \text{and} \quad E^s(\omega) := \bigcap_{n \geq 0} A^{(-n)}(f^n(\omega)) \, (\mathcal{C}_-),$$

where $\mathcal{C}_+$ denotes the set of $(x, y) \in \mathbb{R}^2$ with $xy \geq 0$, and $\mathcal{C}_-$ is the set of (x, y) with $xy \leq 0$. For an example of this type, the Lyapunov exponents might not be everywhere defined, and their exact values with respect to a particular invariant measure are not easily determined, although they will always be nonzero where they exist. In this example and the previous one, the base dynamics $f \colon \Omega \to \Omega$ are irrelevant as far as uniform hyperbolicity of the cocycle is concerned.

Hyperbolicity is an *open* property of both the cocycle A and the dynamics f: if (f, A) is uniformly hyperbolic, and $\hat{f}$ and $\hat{A}$ are both uniformly close (i.e., in the C^0 metric) to f and A, then $(\hat{f}, \hat{A})$ is uniformly hyperbolic. The reason is that, as in the example just presented, uniform hyperbolicity is equivalent to the existence of continuously varying cone families $\mathcal{C}_+(\omega), \mathcal{C}_-(\omega) \subset \mathbb{R}^d$, jointly spanning $\mathbb{R}^d$ for each $\omega \in \Omega$, and an integer $n \geq 1$ with the following properties: $A^{(n)}(\omega) \, (\mathcal{C}_+(\omega)) \subset \mathcal{C}_+(f^n(\omega))$; $A^{(-n)}(\omega) \, (\mathcal{C}_-(\omega)) \subset \mathcal{C}_-(f^{-n}(\omega))$; vectors in $\mathcal{C}_+(\omega)$ are doubled in length by $A^{(n)}(\omega)$; and vectors in $\mathcal{C}_-(\omega)$ are doubled in length by $A^{(-n)}(\omega)$. The existence of such cone families is preserved under small perturbation.

1.6. **Anosov diffeomorphisms.** A diffeomorphism $f \colon M \to M$ whose derivative cocycle is uniformly hyperbolic is called *Anosov*. Again, one needs to modify this definition when the tangent bundle TM is nontrivial: the splitting of $\mathbb{R}^d$ in the definition is replaced by a splitting $TM = E^u \oplus E^s$ into subbundles—that is, a splitting $T_x M = E^u(x) \oplus E^s(x)$ into subspaces, for each $x \in M$, depending continuously on x. The norm $\| \cdot \|$ on the space $M_{d \times d}$ is replaced by a Finsler, as in the discussion at the end of Subsection 1.4. Since M is assumed to be compact, the Anosov property does not depend on the choice of Finsler.

Anosov diffeomorphisms remain Anosov after a C^1-small perturbation, by the openness of uniform hyperbolicity of cocycles. More precisely, the C^1 distance $d_{C^1}(f, g)$ between two diffeomorphisms is the sum of the C^0 distance between f and g and the C^0 distance between Df and Dg; thus if f is Anosov and $d_{C^1}(f, g)$ is sufficiently small, then Dg is hyperbolic, and so g is Anosov. Such a g is often called a C^1-*small perturbation of f*.

The toral automorphism $f_A \colon \mathbb{T}^2 \to \mathbb{T}^2$, with $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ is Anosov; since the derivative cocycle is constant, the splitting $\mathbb{R}^2 = E^u(x) \oplus E^s(x)$, for $x \in \mathbb{T}^2$ does not depend on x: as above, $E^u(x)$ is the expanding eigenspace for A corresponding to the larger eigenvalue $\lambda = (3 + \sqrt{5})/2 > 1$, and $E^s(x)$ is the contracting eigenspace

for A corresponding to the smaller eigenvalue $\lambda^{-1} = (3 - \sqrt{5})/2 < 1$. Since $\lambda > 2$, we can choose $n = 1$ to verify that uniform hyperbolicity holds in the definition. The Lyapunov exponents of this cocycle are $\pm \log \lambda$.

The map $g_\epsilon \colon \mathbb{T}^2 \to \mathbb{T}^2$ given by

$$(3) \qquad\qquad g_\epsilon(x, y) = (2x + y + \epsilon \sin(2\pi(x + y)), x + y)$$

is a C^1-small perturbation of f_A if ϵ is sufficiently small, and so g_ϵ is Anosov for small ϵ. Moreover, since $\det D_{(x,y)} g_\epsilon \equiv 1$, the map g_ϵ preserves the area μ on $\mathbb{T}^2$; we shall see in the next section that g_ϵ is ergodic with respect to μ. The two Lyapunov exponents of g_ϵ with respect to the ergodic measure μ are $\pm \log \lambda_\epsilon$, where $\lambda_\epsilon < \lambda$ for $\epsilon \neq 0$ small. There are several ways to see this: one way to prove it is to compute directly using the ergodic theorem (Theorem 2.1) that $\epsilon \mapsto \lambda_\epsilon$ is a smooth map whose local maximum is achieved at $\epsilon = 0$.

In contrast to f, the exponents of g_ϵ are not constant on $\mathbb{T}^2$ but depend on the invariant measure. For example, the averaged Dirac measures $\nu_{x_1} = \delta_{(0,0)}$ and $\nu_{x_2} = \frac{1}{3}\left(\delta_{(\frac{1}{2},\frac{1}{2})} + \delta_{(0,\frac{1}{2})} + \delta_{(\frac{1}{2},0)}\right)$ corresponding to the fixed point $x_1 = (0,0)$ and the periodic point $x_2 = (\frac{1}{2}, \frac{1}{2})$, respectively, are both invariant and ergodic under g_ϵ, for any ϵ. Direct computation with the eigenvalues of the matrices $D_{x_1} g_\epsilon$ and $D_{x_2} g_\epsilon^3$ shows that the Lyapunov exponents with respect to ν_{x_1} and ν_{x_2} are different for $\epsilon \neq 0$.

1.7. **Measurably (nonuniformly) hyperbolic cocycles.** We say that a cocycle A over (Ω, μ, f) is *measurably hyperbolic* if for μ-a.e. point $\omega \in \Omega$, the exponents $\chi_j(\omega)$ are all nonzero. Since the role played by the measure is important in this definition, we sometimes say that μ is a hyperbolic measure for the cocycle A, or A is hyperbolic with respect to μ.

Uniformly hyperbolic cocycles over a homeomorphism f are hyperbolic with respect to *any* f-invariant probability measure μ (exercise). An equivalent definition of measurable hyperbolicity that neatly parallels the uniformly hyperbolic condition follows: A is hyperbolic with respect to the f-invariant probability measure μ if there exist a set $X \subset \Omega$ with $\mu(X) = 1$ and splittings $\mathbb{R}^d = E^u(\omega) \oplus E^s(\omega)$ depending measurably on $\omega \in X$, such that for every $\omega \in X$ there exists an integer $n = n(\omega) \geq 1$ such that conditions (i)–(iii) in Definition 1.1 hold. The splitting into unstable and stable spaces is not necessarily continuous (or even globally defined on Ω), and the amount of time $n(\omega)$ to wait for doubling in length to occur depends on ω; for these reasons, measurably hyperbolic cocycles are often referred to as "nonuniformly hyperbolic".[5] In this nonuniform setting it is possible for a cocycle to be hyperbolic with respect to one invariant measure, but not another.

A measurably hyperbolic cocycle lurks behind random barycentric subdivision. The random process generating the triangles in iterated barycentric subdivision can be encoded in a cocycle as follows. First, we identify upper half-plane $\mathbb{H} \subset \mathbb{C}$ with the space of marked triangles (modulo Euclidean similarity) by sending a triangle T with vertices cyclically labeled a, b, c to a point $z \in \mathbb{H}$ by rescaling, rotating, and translating, sending a to 0, b to 1 and c to z; see Figure 4.

---

[5] The terminology is not consistent across fields. In smooth dynamics, a cocycle over a measurable system that is measurably hyperbolic is called nonuniformly hyperbolic, whether it is uniformly hyperbolic or not. In the spectral theory community, a cocycle is called nonuniformly hyperbolic if it is measurably hyperbolic but *not* uniformly hyperbolic.
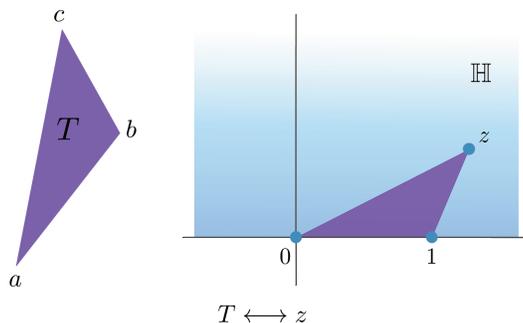
FIGURE 4. The upper half-plane $\mathbb{H}$ is the space of marked triangles, up to Euclidean similarity. The triangle T on the left corresponds to the point $z \in \mathbb{H}$ on the right.

Labelling cyclically the triangles $T^{(1)}, T^{(2)}, \ldots, T^{(6)}$ in the subdivision as in Figure 5, the Möbius transformation $B(z) = 2(z+1)/3$ sends the marked triangle T to the marked triangle $T^{(1)}$. The involutions pictured in Figure 6 generate the symmetric group S_3, whose nontrivial elements we label $P_1, \ldots, P_5$. A bit of thought shows that the transformations $B, BP_1, \ldots, BP_5$ are the 6 maps of the plane sending T to the rescaled triangles in the subdivision.

Fixing an identification of the lower half-plane with the upper half-plane via $z \mapsto \overline{z}$, the action of these 6 transformations on $\mathbb{H}$ are identified with the projective action of 6 elements $A_1, \ldots, A_6$ of $PGL(2, \mathbb{R})$, where B is identified with
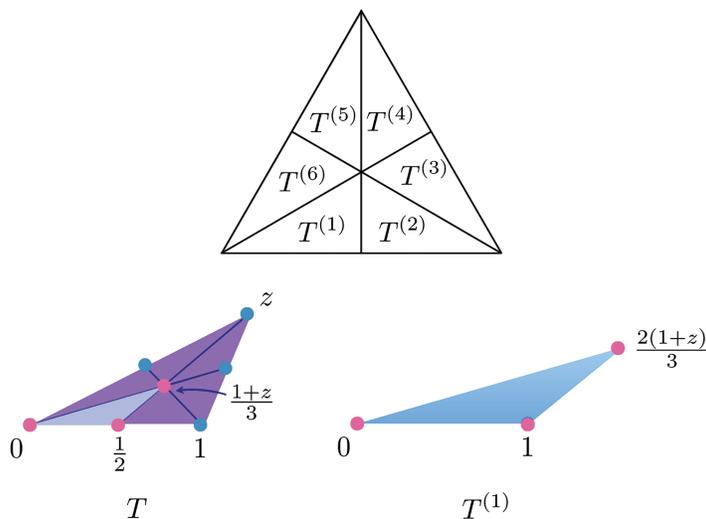
FIGURE 5. A Möbius transformation that selects the first triangle in barycentric subdivision.
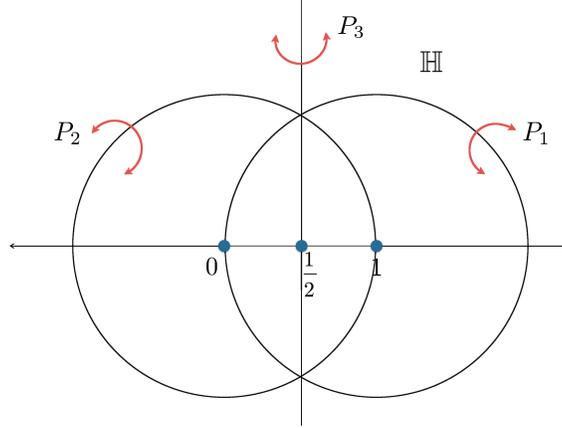
FIGURE 6. Three involutions generating the symmetric group S_3: inversion in the circles $|z| = 1$ and $|z - 1| = 1$, and reflection across the line $\text{Re}(z) = \frac{1}{2}$.

$\begin{pmatrix} 2/\sqrt{6} & 2/\sqrt{6} \\ 0 & 3/\sqrt{6} \end{pmatrix}$, P_1 is identified with $\begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$, P_2 is identified with $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, and P_3 is identified with $\begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$.

Thus random barycentric subdivision is governed by a random matrix cocycle over a Bernoulli shift. If repeated rolls of the die generate the sequence $\omega_1, \omega_2, \ldots$ with $\omega_i \in \{1, \ldots, 6\}$, then the nth triangle $T_n(\omega)$ generated is the projective image of $\begin{pmatrix} i \\ 1 \end{pmatrix}$ under $A^{(n)}(\omega)$, and an exercise shows that the aspect ratio of $T_n(\omega)$ is given by

$$\alpha(T_n(\omega)) = \|(0, 1) \cdot A^{(n)}(\omega)\|^{-2}.$$

We thus obtain the formula

$$\lim_{n \to \infty} \frac{1}{n} \log \alpha(T_n(\omega)) = -2 \lim_{n \to \infty} \frac{1}{n} \log \left\| (0, 1) \cdot A^{(n)}(\omega) \right\|,$$

and out pops the Lyapunov exponent $\chi = \lim_{n \to \infty} \frac{1}{n} \log \left\| (0, 1) \cdot A^{(n)}(\omega) \right\|$.[6]

Numerical simulation gives the value $\chi \approx 0.0446945$, but the fact that this number is positive follows from a foundational result of Furstenberg (stated precisely as a special case in Theorem 6.1) that underlies some of Avila's results as well. The upshot is that a random product of matrices in $SL(2, \mathbb{R})$ (or $PGL(2, \mathbb{R})$) can't have exponents equal to 0, except by design. In particular, if the matrices do not simultaneously preserve a collection of one or two lines, and the group generated by the matrices is not compact, then the exponents with respect to any nontrivial Bernoulli measure will be nonzero. These conditions are easily verified for the barycentric cocycle. Details of this argument about barycentric subdivision can be found in [54] and the related paper [16].

The barycentric cocycle is *not* uniformly hyperbolic, as can be seen by examining the sequence of triangles generated by subdivision on Figure 2: at each stage it

---

[6] As with eigenvalues, the Lyapunov exponents for left and right matrix multiplication coincide.

is always possible to choose some triangle with aspect ratio bounded below, even though most triangles in a subdivision will have smaller aspect ratio than the starting triangle. For products of matrices in $SL(d, \mathbb{R})$, uniform hyperbolicity must be carefully engineered, but for random products, measurable hyperbolicity almost goes without saying.

A long-standing problem in smooth dynamics is to understand which diffeomorphisms have hyperbolic derivative cocycle with respect to some natural invariant measure, such as volume (see [19]). Motivating this problem is the fact that measurable hyperbolicity produces interesting dynamics, as we explain in the next section.

2. Smooth ergodic theory

Smooth ergodic theory studies the dynamical properties of smooth maps from a statistical point of view. A natural object of study is a measure-preserving system (M, vol, f), where M is a smooth, compact manifold without boundary equipped with a Riemannian metric, vol is the volume measure of this metric, normalized so that $\mathrm{vol}(M) = 1$, and $f \colon M \to M$ is a diffeomorphism preserving vol. It was in this context that Boltzmann originally hypothesized ergodicity for ideal gases in the 1870s. Boltzmann's nonrigorous formulation of ergodicity was close in spirit to the following statement of the pointwise ergodic theorem for diffeomorphisms:

Theorem 2.1. *If f is ergodic with respect to volume, then its orbits are equidistributed in the following sense: for almost every $x \in M$ and any continuous function $\phi \colon M \to \mathbb{R}$,*

$$(4) \qquad \lim_{n \to \infty} \frac{1}{n} \left(\phi(x) + \phi(f(x)) + \cdots + \phi(f^{n-1}(x)) \right) = \int_M \phi \, d\mathrm{vol}.$$

On the left of equation (4) is the limit of the average value of the "observable" ϕ along the f-orbit of a typical point x; on the right is the average (or expected value) of ϕ over M. The ergodic theorem thus asserts that for an ergodic system, "time averages equal space averages."

As remarked previously, an example of an ergodic diffeomorphism is the rotation f_α on $\mathbb{R}/\mathbb{Z}$, for α irrational. In fact this transformation has a stronger property of unique ergodicity, which is equivalent to the property that the limit in (4) exists for *every* $x \in \mathbb{R}/\mathbb{Z}$.[7] While unique ergodicity is a strong property, the ergodicity of irrational rotations is fragile; the ergodic map f_α can be perturbed to obtain the nonergodic map $f_{p/q}$, where $\alpha \approx p/q$.

Another example of an ergodic diffeomorphism, at the opposite extreme of the rotations in more than one sense, is the automorphism f_A of the 2-torus induced by multiplication by the matrix $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ with respect to the area μ. In spirit, this example is closely related to the Bernoulli shift, and in fact its orbits can be measurably coded in such a way as to produce a measure-preserving isomorphism with a Bernoulli shift. As observed in the previous section, ergodicity of this map can be proved using Fourier analysis, but there is a much more robust proof, due to Anosov [1], who showed that any C^2 Anosov diffeomorphism that preserves volume is ergodic with respect to volume.

---

[7]This is a consequence of Weyl's equidistribution theorem and can be proved using elementary analysis; see, e.g., [37].

2.1. **Ergodicity of Anosov diffeomorphisms and Pesin theory.** Anosov's proof of ergodicity has both analytic and geometric aspects. For the map f_A, it follows several steps:

(1) The expanding and contracting subbundles E^u and E^s of the splitting $T\mathbb{T}^2 = E^u \oplus E^s$ are tangent to foliations $\mathcal{W}^u$ and $\mathcal{W}^s$ of $\mathbb{T}^2$ by immersed lines. These lines are parallel to the expanding and contracting eigendirections of A and wind densely around the torus, since they have irrational slope. The leaves of this pair of foliations are perpendicular to each other, since A is symmetric.

(2) A clever application of the pointwise ergodic theorem (presented here as a special case in Theorem 2.1) shows that any $\phi \in L^2(\mathbb{T}^2, \mu)$ satisfying $\phi \circ f = \phi$ is, up to a set of area 0, constant along leaves of the foliation $\mathcal{W}^u$ and (again, up to a set of area 0) constant along leaves of $\mathcal{W}^s$. This part of the argument goes back to Eberhard Hopf's study of geodesics in negatively curved surfaces in the 1930s.

(3) Locally, the pair of foliations $\mathcal{W}^u$ and $\mathcal{W}^s$ are just (rotated versions of) intervals parallel to the x and y axes. In these rotated coordinates, $\phi(x, y)$ is a measurable function constant a.e. in x and constant a.e. in y. Fubini's theorem then implies that such a ϕ must be constant a.e. This conclusion holds in local charts, but since $\mathbb{T}^2$ is connected, ϕ must be constant.

(4) Since any f-invariant function $\phi \in L^2(\mathbb{T}^2, \mu)$ is constant almost everywhere with respect to μ, we conclude that f_A is ergodic with respect to μ.

The same proof works for any smooth, volume-preserving Anosov diffeomorphism—in particular, for the maps g_ϵ defined in equation (3)—if one modifies the steps appropriately. The foliations by parallel lines in step (1) are replaced by foliations by smooth curves (or submanifolds diffeomorphic to some $\mathbb{R}^k$, in higher dimension). Step (2), the so-called Hopf argument, is almost the same, since it uses only volume preservation and the fact that the leaves of $\mathcal{W}^u$ and $\mathcal{W}^s$ are expanded and contracted, respectively. Step (3) is the most delicate to adapt and was Anosov's great accomplishment, and Step (4) is of course the same.

The ergodicity of Anosov diffeomorphisms is also a consequence of the much stronger property of being measurably encoded by a Bernoulli shift. This so-called *Bernoulli property* implies that Anosov diffeomorphisms are mixing with respect to volume, meaning that for any L^2 function ϕ, we have $\int_M \phi \cdot \phi \circ f^n \to (\int_M \phi)^2$ as $n \to \infty$. Visually, sets are mixed up by Anosov diffeomorphisms; see Figure 7, which explains why f_A is sometimes referred to as the "cat map". The map g_ϵ will similarly do quite a number on a cat. The proof of the Bernoulli property for Anosov diffeomorphisms builds the Anosov–Hopf proof of ergodicity. The Anosov–Hopf argument has the additional advantage that it can be adapted to prove ergodicity and mixing for systems that are not Bernoulli.

As explained in Subsection 1.6, any C^1-small perturbation of an Anosov diffeomorphism is Anosov, and C^2 Anosov diffeomorphisms preserving volume are ergodic. Hence volume-preserving C^2 Anosov diffeomorphisms are *stably ergodic:* the ergodicity can't be destroyed by a C^1-small perturbation, in marked contrast with the irrational rotation f_α.

The Anosov condition thus has powerful consequences in smooth ergodic theory. But, like uniformly hyperbolic matrix products, Anosov diffeomorphisms are
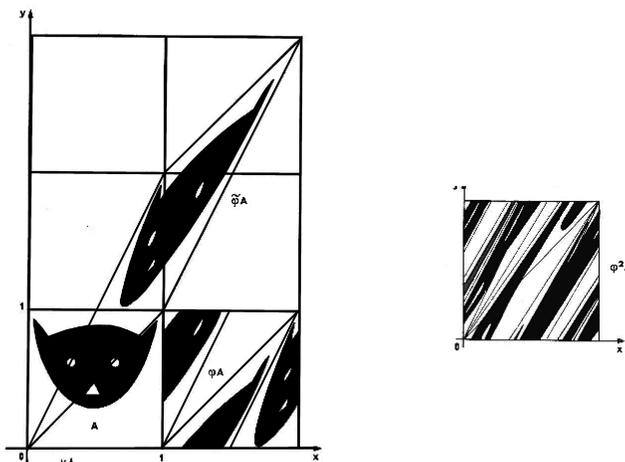
FIGURE 7. The action of f_A on a cat, from [2]. A cat is drawn in a square fundamental domain for $\mathbb{T}^2$ at the lower left. Its image under A is shown in the parallelogram, and it is reassembled into another fundamental domain to show its image under f_A. The image of the cat under f_A^2 is depicted at right.

necessarily contrived. In dimension 2, the only surface supporting an Anosov diffeomorphism is the torus $\mathbb{T}^2$, and conjecturally, the only manifolds supporting Anosov diffeomorphisms belong to a special class called the infra-nilmanifolds. On the other hand, *every* smooth manifold supports a volume-preserving diffeomorphism that is hyperbolic with respect to volume, as was shown by Dolgopyat and Pesin in 2002 [29].

In the 1970s Pesin [57] introduced a significant innovation in smooth ergodic theory: a nonuniform, measurable analogue of the Anosov–Hopf theory. Under the assumption that a volume-preserving diffeomorphism f is hyperbolic with respect to volume, Pesin showed that volume has at most countably many ergodic components with respect to f. Starting with Oseledets's theorem and repeatedly employing Lusin's theorem that every measurable function is continuous off of a set of small measure, Pesin developed an ergodic theory of smooth systems that has had numerous applications. Some limitations of Pesin theory are: first, that it begins with the hypothesis of measurable hyperbolicity, which is a condition that is very hard to verify except in special cases; and second, without additional input, measurable hyperbolicity does not imply ergodicity, as the situation of infinitely many ergodic components can and does occur [28].

2.2. Ergodicity of "typical" diffeomorphisms.

The question of whether ergodicity is a common property among volume-preserving diffeomorphisms of a compact manifold M is an old one, going back to Boltzmann's ergodic hypothesis of the late 19th century. We can formalize the question by fixing a differentiability class $r \in [1, \infty]$ and considering the set $\mathrm{Diff}^r_{\mathrm{vol}}(M)$ of C^r, volume-preserving diffeomorphisms of M. This is a topological space in the C^r topology, and we say that

$$\mathrm{Diff}^1_{\mathrm{vol}}(M)$$

■ diffeomorphisms with entropy 0

■ ergodic, measurably hyperbolic diffeomorphisms
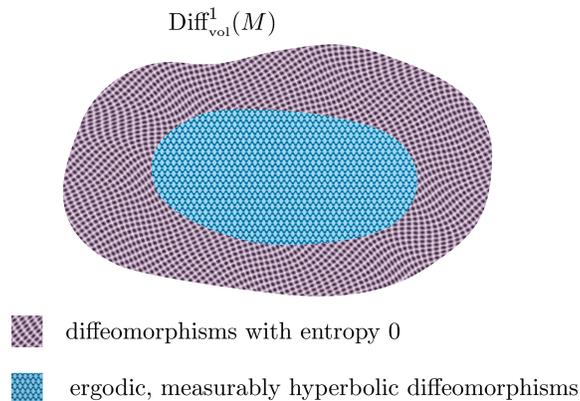
FIGURE 8. Generically, positive entropy implies ergodicity and measurable hyperbolicity.

a property holds *generically in* $\mathrm{Diff}^r_{\mathrm{vol}}(M)$ (or C^r *generically*, for short) if it holds for all f in a countable intersection of open and dense subsets of $\mathrm{Diff}^r_{\mathrm{vol}}(M)$.[8]

Oxtoby and Ulam [56] proved in 1939 that a generic volume-preserving *homeomorphism* of a compact manifold is ergodic. At the other extreme, KAM (Kolmogorov–Arnol'd–Moser) theory, introduced by Kolmogorov in the 1950s [43], implies that ergodicity is *not* a dense property, let alone a generic one, in $\mathrm{Diff}^\infty_{\mathrm{vol}}(M)$, if $\dim(M) \geq 2$.

The general question of whether ergodicity is generic in $\mathrm{Diff}^r_{\mathrm{vol}}(M)$ remains open for $r \in [1, \infty)$, but we now have a complete answer for any manifold when $r = 1$ under the assumption of *positive entropy*. Entropy is a numerical invariant attached to a measure-preserving system that measures the complexity of orbits. The rotation f_α has entropy 0; the Anosov map f_A has positive entropy $\log(\lambda)$. By a theorem of Ruelle, positivity of entropy means that there is *some* positive volume subset of M on which the Lyapunov exponents are nonzero in *some* directions.

Theorem 2.2 (Avila, Crovisier, Wilkinson [7]). *Generically in* $\mathrm{Diff}^1_{\mathrm{vol}}(M)$, *positive entropy implies ergodicity and, moreover, measurable hyperbolicity with respect to volume.*

See Figure 8. This result was proved in dimension 2 by Mañé and Bochi [17, 51] and in dimension 3 by M. A. Rodriguez-Hertz [58]. Positive entropy is an *a priori* weak form of chaotic behavior that can be confined to an invariant set of very small measure, with trivial dynamics on the rest of the manifold. Measurable hyperbolicity, on the other hand, means that at almost every point *all* of the Lyapunov exponents of the derivative cocycle Df are nonzero. Conceptually, the proof of Theorem 2.2 divides into two parts:

(1) C^1 generically, positive entropy implies nonuniform hyperbolicity. One needs to go from some nonzero exponents on some of the manifold to all nonzero exponents on almost all of the manifold. Since the cocycle and the dynamics are intertwined, carrying this out is a delicate matter. This part

---

[8]Since $\mathrm{Diff}^r_{\mathrm{vol}}(M)$ is a Baire space, properties that hold generically hold for a dense set, and two properties that hold generically separately hold together generically.

of the argument relies on the fact that the C^1 topology is particularly well adapted to the problem. On the one hand, constructing C^1-small perturbations with a desired property is generally much easier than constructing C^2-small perturbations with the same property. On the other hand, many useful dynamical properties such as uniform hyperbolicity are C^1 open.

(2) C^1 generically, measurable hyperbolicity (with respect to volume) implies ergodicity. This argument uses Pesin theory but adds some missing input needed to establish ergodicity. This input holds C^1 generically. For example, [7] proves that the positive entropy assumption generically implies existence of a *dominated splitting*; this means that generically, a positive entropy diffeomorphism is something intermediate between an Anosov diffeomorphism and a nonuniformly hyperbolic one. There is a *continuous* splitting $TM = E^u \oplus E^s$, invariant under the derivative, such that for almost every $x \in M$, there exists an $n = n(x)$ such that for every $\xi^u \in E^u(x)$, $\|D_x f^n(\xi^u)\| \geq 2\|\xi^u\|$, and for every $\xi^s \in E^s(x)$, $\|D_x f^{-n}(\xi^s)\| \geq 2\|\xi^s\|$.

The proof incorporates techniques from several earlier results, most of which have been proved in the past 20 years [6, 18, 20, 59]. Also playing an essential technical role in the argument is a regularization theorem of Avila: every C^1 diffeomorphism that preserves volume can be C^1 approximated by a C^2 volume-preserving diffeomorphism [5]. The fact that this regularization theorem was not proved until recently highlights the difficulty in perturbing the derivative cocycle to have desired properties: you can't change Df without changing f too (and *vice versa*). This is why completely general results analogous to Furstenberg's theorem for random matrix products are few and far between for diffeomorphism cocycles.

3. Translation surfaces

A flat surface is any closed surface that can be obtained by gluing together finitely many parallelograms in $\mathbb{R}^2$ along coherently oriented parallel edges, as in Figure 9. Two flat surfaces are equivalent if one can be obtained from the other by cutting, translating, and rotating. A *translation* surface is a flat surface that comes equipped with a well-defined, distinguished vertical, "North" direction (or "South"
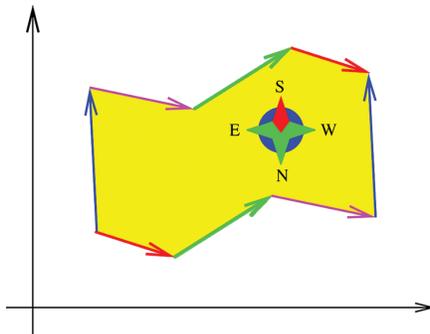
FIGURE 9. A flat surface (of genus 2) with a distinguished "South," also known as a translation surface. Parallel edges of the same color are identified (courtesy of Marcelo Viana, used with permission).

depending on your preferred orientation). Two translation surfaces are equivalent if one can be obtained from the other by cutting and translating (but *not* rotating).

Fix a translation surface Σ of genus $g > 0$. If one picks an angle θ and a point x on Σ, and follows the corresponding straight ray through Σ, there are two possibilities: either it terminates in a corner or it can be continued for all time. Clearly, for any θ, and almost every starting point (with respect to area), the ray will continue forever. If it continues forever, either it returns to the initial point and direction and produces a closed curve or it continues on a parallel course without returning. A version of the Pigeonhole Principle for area (Poincaré recurrence) implies that for almost every point and starting direction, the line will come back arbitrarily close to the starting point.
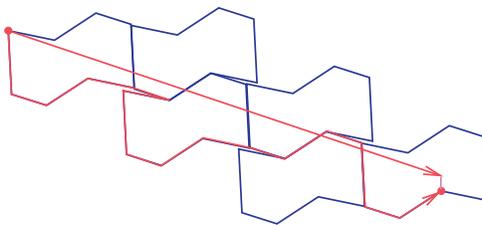
FIGURE 10. Closing up a ray that comes back close to itself (courtesy of Marcelo Viana, used with permission).

Kerckhoff, Masur, and Smillie [42] proved more: for a fixed Σ, and almost every θ, the ray through any point x is dense in Σ, and in fact is equidistributed with respect to area. Such a direction θ is called *uniquely ergodic*, as it is uniquely ergodic in the same sense that f_α is, for irrational α. Suppose we start with a uniquely ergodic direction and wait for the successive times that this ray returns closer and closer to itself. This produces a sequence of closed curves γ_n which produces a sequence of cycles $[\gamma_n]$ in homology $H_1(\Sigma, \mathbb{Z}) \simeq \mathbb{Z}^{2g}$.

Unique ergodicity of the direction θ implies that there is a unique $c_1 \in H_1(\Sigma, \mathbb{R})$ such that for any starting point x,

$$\lim_{n \to \infty} \frac{[\gamma_n]}{\ell(\gamma_n)} = c_1,$$

where $\ell(\gamma)$ denotes the length in Σ of the curve γ.

Theorem 3.1 (Forni [32], Avila and Viana [13], Zorich [65, 66]). *Fix a topological surface S of genus $g \geq 1$, and let Σ be almost any translation surface modeled on S.[9] Then there exist real numbers $1 > \nu_2 > \cdots > \nu_g > 0$ and a sequence of of subspaces $L_1 \subset L_2 \subset \cdots \subset L_g$ of $H_1(\Sigma, \mathbb{R})$ with $\dim(L_k) = k$ such that for almost every θ, for every x, and every γ in direction θ, the distance from $[\gamma]$ to L_g is bounded, and*

$$\limsup_{\ell(\gamma) \to \infty} \frac{\log \operatorname{dist}([\gamma], L_i)}{\log(\ell(\gamma))} = \nu_{i+1},$$

for all $i < g$.

---

[9] "Almost any" means with respect to the Lebesgue measure on possible choices of lengths and directions for the sides of the polygon. This statement can be made more precise in terms of Lebesgue measure restricted to various strata in the moduli space of translation surfaces.
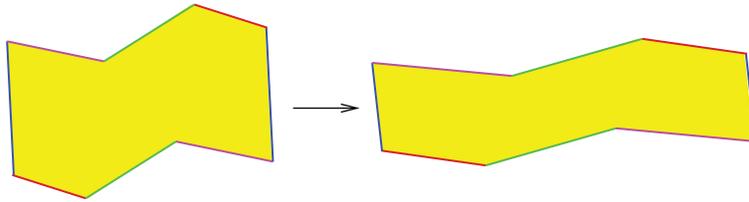
FIGURE 11. A local picture of the Teichmüller flow (courtesy of Marcelo Viana, used with permission).

This theorem gives precise information about the way the direction of $[\gamma_n]$ converges to its asymptotic cycle c_1: the convergence has a "directional nature" much in the way a vector $v \in \mathbb{R}^d$ converges to infinity under repeated application of a matrix

$$A = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ 0 & \cdots & \cdots & * \\ 0 & 0 & \cdots & \lambda_d \end{pmatrix},$$

with $\lambda_1 > \lambda_2 > \cdots \lambda_d > 1$.

The numbers ν_i are the Lyapunov exponents of the *Kontsevich–Zorich* (KZ) cocycle over the *Teichmüller flow*. The Teichmüller flow $\mathcal{F}_t$ acts on the moduli space $\mathcal{M}$ of translation surfaces (that is, translation surfaces modulo cutting and translation) by stretching in the East-West direction and contracting in the North-South direction. More precisely, if Σ is a translation surface, then $\mathcal{F}_t(\Sigma)$ is a new surface, obtained by transforming Σ by the linear map $\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$. Since a stretched surface can often be reassembled to obtain a more compact one, it is plausible that the Teichmüller flow has recurrent orbits (for example, periodic orbits). This is true and reflects the fact that the flow $\mathcal{F}_t$ preserves a natural volume that assigns finite measure to $\mathcal{M}$. The KZ cocycle takes values in the symplectic group $Sp(2g, \mathbb{R})$ and captures homological data about the cutting and translating equivalence on the surface.

Veech proved that $\nu_2 < 1$ [60], Forni proved that $\nu_g > 0$ [32], and Avila and Viana proved that the numbers $\nu_2, \nu_3, \ldots, \nu_{g-1}$ are all distinct [13]. Zorich established the connection between exponents and the so-called deviation spectrum in Theorem 3.1, a connection that holds in greater generality [65, 66]. Many more things have been proved about the Lyapunov exponents of the KZ cocycle, and some of their values have been calculated which are (until recently, conjecturally) rational numbers! See [25, 31].

Zorich's result reduces the proof of Theorem 3.1 to proving that the the exponents $\nu_1, \ldots, \nu_g$ are positive and distinct. In the $g = 1$ case where Σ is a torus, this fact has a simple explanation. The moduli space $\mathcal{M}$ is the set of all flat structures on the torus (up to homothety), equipped with a direction. This is the quotient $SL(2, \mathbb{R})/SL(2, \mathbb{Z})$, which is the unit tangent bundle of the modular surface $\mathbb{H}/SL(2, \mathbb{Z})$. The (continuous time) dynamical system $\mathcal{F}_t$ on Ω is given by left multiplication by the matrix $\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$. The cocycle is, in essence, the derivative cocycle for this flow (transverse to the direction of the flow). This flow

neighborhood
of a cusp
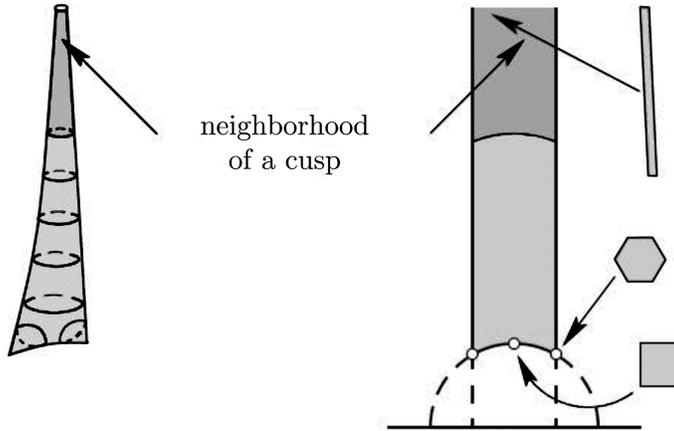
FIGURE 12. The moduli space $\mathbb{H}^2/PSL(2,\mathbb{Z})$ of flat structures on
the torus, which is a punctured sphere with two cone points, cor-
responding to the hexagonal and square lattices in $\mathbb{R}^2$. Pictured at
right is a fundamental domain in $\mathbb{H}^2$ (courtesy of Carlos Matheus,
used with permission).

is *uniformly hyperbolic* (i.e., Anosov), and its exponents are $\nu_1 = \log(e) = 1$ and
$-\nu_1 = \log(e^{-1}) = -1$.

The proof in [13] that the exponents $\nu_1, \ldots, \nu_g$ are positive and distinct for
general translation surfaces is considerably more involved. We can nonetheless boil
it down to some basic ideas.

(1) The Teichmüller flow itself is nonuniformly hyperbolic with respect to a
natural volume (Veech [60]), and it can be coded in a way that the dynamics
appear almost random.
(2) Cocycles over perfectly random systems (for example i.i.d. sequences of
matrices) have a tendency to have distinct, nonzero Lyapunov exponents.
This was first proved by Furstenberg in the 2×2 case [34] and later by
Guivarc'h and Raugi [36] in the $d \times d$ case (see also [35]).
(3) Cocycles over systems that are nonrandom, but sufficiently hyperbolic and
with good coding, also tend to have distinct, nonzero Lyapunov exponents.
This follows from series of results, beginning with Ledrappier in the 2×2
case [46], and in increasing generality by Bonatti and Viana [22], Viana
[61], and Avila and Viana [12].

4. HOFSTADTER'S BUTTERFLY

Pictured in Figure 13 is the spectrum of the operator $H_x^\alpha \colon \ell^2(\mathbb{Z},\mathbb{C}) \to \ell^2(\mathbb{Z},\mathbb{C})$
given by

$$[H_x^\alpha u](n) = u(n+1) + u(n-1) + 2\cos(2\pi(x+n\alpha))u(n),$$

where x is a fixed real number called the *phase*, and $\alpha \in [0,1]$ is a parameter called
the *frequency*. The vertical variable is α, and the horizontal variable is the spectral
energy parameter E, which ranges in $[-4,4]$. We can read off the spectrum of H_x^α
by taking a horizontal slice at height α; the black region is the spectrum.
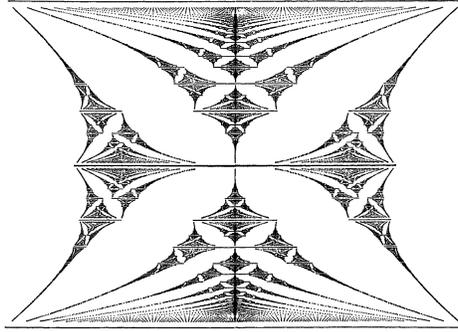
FIGURE 13. Hofstadter's butterfly. Reprinted figure with permission as follows: Douglas R. Hofstadter, *Energy levels and wave functions of Block electrons in rational and irrational magnetic fields*, Physical Review B **14** (1976), 2239–2249. Copyright 1976, American Physical Society.

In an influential 1976 paper, Douglas Hofstadter of *Gödel, Escher, Bach* fame discovered this fractal picture while modelling the behavior of electrons in a crystal lattice under the force of a magnetic field [38]. This operator plays a central role in the theory of the integer quantum Hall effect developed by Thouless et al., and, as predicted theoretically, the butterfly has indeed appeared in von Klitzing's QHE experiments. Similar operators are used in modeling graphene, and similar butterflies also appear in graphene related experiments (see, e.g., [27]).

Some properties of the butterfly have been established rigorously. For example, Avila and Krikorian proved:

Theorem 4.1 (Avila and Krikorian [9])**.** *For every irrational $\alpha \in [0,1]$, the α-horizontal slice of the butterfly has measure* 0*.*

Their proof complements and thus extends the earlier result of Last [45], who proved the same statement, but for a full measure set of α satisfying an arithmetic condition. In particular, we have:

Corollary 4.2. *The butterfly has measure* 0*.*

Other properties of the butterfly, for example its Hausdorff dimension, remain unknown.

The connection between the spectrum of this operator and cocycles is an interesting one. Recall the definition of the spectrum of H_x^α:

$$\sigma(H_x^\alpha) := \{E \in \mathbb{C} : H_x^\alpha - E \cdot \mathrm{Id} \text{ is not invertible}\}.$$

The eigenvalues are those E so that the eigenvalue equation $H_x^\alpha u = Eu$ admits $\ell^2(\mathbb{Z})$ solutions, i.e., those E such that $H_x^\alpha - E \cdot \mathrm{Id}$ is not injective.

The following simple observation is key. A sequence $(u_n : n \in \mathbb{Z}) \subset \mathbb{C}^{\mathbb{Z}}$ (not necessarily in $\ell^2(\mathbb{Z})$) solves $H_x^\alpha u = Eu$ if and only if

$$A_E(f_\alpha^n(x)) \begin{pmatrix} u_n \\ u_{n-1} \end{pmatrix} = \begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix}, \ n \in \mathbb{Z},$$

where $f_\alpha \colon \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ is the translation mentioned above, and

$$(5) \qquad\qquad A_E(x) = \begin{pmatrix} E - 2\cos(2\pi x) & -1 \\ 1 & 0 \end{pmatrix},$$

which defines an $SL(2, \mathbb{R})$-cocycle over the rotation f_α, an example of a *Schrödinger cocycle*. Using the cocycle notation, we have

$$A_E^{(n)}(x) \begin{pmatrix} u_0 \\ u_{-1} \end{pmatrix} = \begin{pmatrix} u_n \\ u_{n-1} \end{pmatrix}, \ n \in \mathbb{Z}.$$

Now let's connect the properties of this cocycle with the spectrum of H_x^α. Suppose for a moment that the cocycle A_E over f_α is uniformly hyperbolic, for some value of E. Then for every $x \in \mathbb{R}/\mathbb{Z}$ there is a splitting $\mathbb{R}^2 = E^u(x) \oplus E^s(x)$ invariant under cocycle, with vectors in $E^u(x)$ expanded under $A_E^{(mn)}(x)$, and vectors in $E^s(x)$ expanded under $A_E^{(-mn)}(x)$, both by a factor of 2^m. Thus no solution u to $H_x^\alpha u = Eu$ that can be polynomially bounded simultaneously in both directions, which implies E is not an ℓ^2 eigenvalue of H_x^α. It turns out that the converse is also true, and moreover:

Theorem 4.3 (R. Johnson [39]). *If α is irrational, then for every $x \in [0, 1]$,*

$$(6) \qquad\qquad \sigma(H_x^\alpha) = \{E : A_E \text{ is not } \textit{uniformly hyperbolic over } f_\alpha\}.$$

For irrational α, we denote by Σ_α the spectrum of $\sigma(H_x^\alpha)$, which by Theorem 4.3 does not depend on x. Thus for irrational α, the set Σ_α is the α-horizontal slice of the butterfly.

The butterfly is therefore both a dynamical picture and a spectral one. On the one hand it depicts the spectrum of a family of operators H_x^α parametrized by α, and on the other hand it depicts, within a 2-parameter family of cocycles $\{(f_\alpha, A_E) : (E, \alpha) \in [-4, 4] \times [0, 1]\}$, the set of parameters corresponding to dynamics that are *not* uniformly hyperbolic.

Returning to spectral theory, we continue to explore the relationship between spectrum and dynamics. If α is irrational, then f_α is ergodic, and Oseledets's theorem implies that the Lyapunov exponents for any cocycle over f_α take constant values over a full measure set. Thus the Lyapunov exponents of A_E over f_α take two essential values, $\chi_E^+ \geq 0$ and χ_E^-; the fact that $\det(A_E) = 1$ implies that $\chi_E^- = -\chi_E^+ \leq 0$. Then either A_E is nonuniformly hyperbolic (if $\chi_E^+ > 0$) or the exponents of A_E vanish.

Thus for fixed α irrational, the spectrum Σ_α splits, from a dynamical point of view, into two (measurable) sets: the set of E for which A_E is nonununiformly hyperbolic, and the set of E for which the exponents of A_E vanish. On the other hand, spectral analysis gives us a different decomposition of the spectrum:

$$\sigma(H_x^\alpha) = \sigma_{ac}(H_x^\alpha) \cup \sigma_{sc}(H_x^\alpha) \cup \sigma_{pp}(H_x^\alpha),$$

where $\sigma_{ac}(H_x^\alpha)$ is the absolutely continuous spectrum, $\sigma_{pp}(H_x)$ is the pure point spectrum (i.e., the closure of the eigenvalues), and $\sigma_{sc}(H_x^\alpha)$ is the singular continuous spectrum. All three types of spectra have meaningful physical interpretations. While the spectrum $\sigma(H_x^\alpha)$ does not depend in x (since α is irrational), the decomposition into subspectra can depend on x.[10] It turns out that the absolutely

---

[10]In fact, the decomposition is independent of a.e. x, just not all x.

continuous spectrum does not depend on x, so we can write $\Sigma_{ac,\alpha}$ for this common set.

The next deep relation between spectral theory and Lyapunov exponents is the following, which is due to Kotani:

Theorem 4.4 (Kotani [44]). *Fix α irrational. Let $\mathcal{Z}$ be the set of E such that the Lyapunov exponents of A_E over f_α vanish. Let $\overline{\mathcal{Z}^{ess}}$ denote the essential closure of $\mathcal{Z}$, i.e., the closure of the Lebesgue density points of $\mathcal{Z}$. Then*

$$\Sigma_{ac} = \overline{\mathcal{Z}^{ess}}.$$

Thus Lyapunov exponents of the cocycle are closely related to the spectral type of the operators H_x. For instance, Theorem 4.4 implies that if A_E is nonuniformly hyperbolic over f_α for almost every $E \in \Sigma_\alpha$, then $\Sigma_{ac,\alpha}$ is empty: H_x^α has no absolutely continuous spectrum.

We remark that Theorems 4.3 and 4.4 hold for much broader classes of Schrödinger operators over ergodic measure-preserving systems. For a short and self-contained proof of Theorem 4.3, see [64]. The spectral theory of one-dimensional Schrödinger operators is a rich subject, and we have only scratched the surface here; for further reading, see the recent surveys [40] and [26].

Avila's very recent work, some of it still unpublished, provides further fascinating connections of this type, linking the spectral properties of quasiperiodic operators with analytic potentials to properties of the Lyapunov exponents of their associated cycles [3, 4].

5. Spaces of dynamical systems and metadynamics

Sections 2, 3, and 4 are all about *families* of dynamical systems. In Section 2, the family is the space of all volume-preserving diffeomorphisms of a compact manifold M. This infinite-dimensional space is not locally compact, and we have thrown up our hands and depicted it in Figure 8 as a blob. Theorem 2.2 asserts that within a residual set of positive entropy systems (which turn out to be an open subset of the blob), measurable hyperbolicity (and ergodicity) is generic.

In Section 3, the moduli space $\mathcal{M}$ of translation surfaces can also be viewed as a space of dynamical systems, in particular the *billiard* flows on *rational* polygons, i.e., polygons whose corner angles are multiples of 2π. In a billiard system, one shoots a ball in a fixed direction and records the location of the bounces on the walls. By a process called unfolding, a billiard trajectory can be turned into a straight ray in a translation surface.[11] The process is illustrated in Figure 14 for the square torus billiard.

The moduli space $\mathcal{M}$ is not so easy to draw and not completely understood (except for genus $g = 1$). It is, however, a finite-dimensional manifold and carries some nice structures, which makes it easier to picture than Diff(M). Theorem 3.1 illustrates how dynamical properties of a *meta dynamical system*, i.e., the Teichmüller flow $\mathcal{F}_t \colon \mathcal{M} \to \mathcal{M}$, are tied to the dynamical properties of the elements of $\mathcal{M}$: The Lyapunov exponents of the KZ cocycle over $\mathcal{F}_t$ for a given billiard table with a given direction describe how well an infinite billiard ray can be approximated by closed, nearby billiard paths. In yet another example, Lyapunov exponents of the

---

[11]Not every translation surface comes from a billiard, since the billiards have extra symmetries. But the space of billiards embeds inside the space of translation surfaces, and the Teichmüller flow preserves the set of billiards.
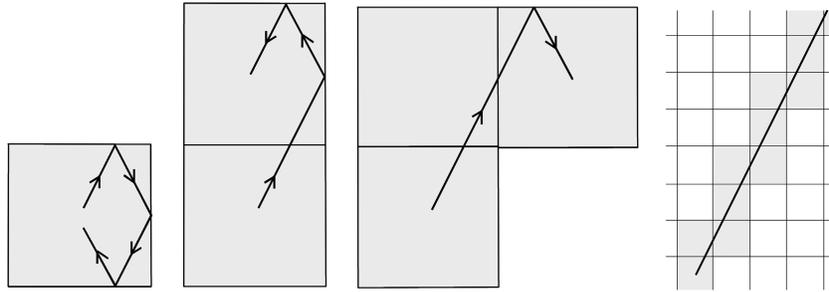
FIGURE 14. Unfolding billiards in a square to get lines in a torus
(courtesy of Diana Davis, used with permission).

Teichmüller flow control the rate of convergence of time averages to space averages
for the billiard flow.

In Section 4, we saw how the spectral properties of a family of operators
$\{H_x^\alpha : \alpha \in [0,1]\}$ are reflected in the dynamical properties of families of cocycles
$\{(f_\alpha, A_E) : (E, \alpha) \in [-4, 4] \times [0, 1]\}$. Theorems about spectral properties thus have
their dynamical counterparts. For example, Theorem 4.3 tells us that the butterfly
is the complement of those parameter values where the cocycle (f_α, A_E) is uniformly
hyperbolic. Since uniform hyperbolicity is an open property in both α and E, the
complement of the butterfly is open. Corollary 4.2 tells us that the butterfly has
measure 0. Thus the set of parameter values in the square that are hyperbolic form
an open and dense, full-measure subset. In fact, work of Bourgain and Jitomirskaya
[24] implies that *the butterfly is precisely the set of parameter values (E, α) where
the Lyapunov exponents of (f_α, A_E) vanish for some x.*[12] These results in some
ways echo Theorem 2.2, within a very special family of dynamics.
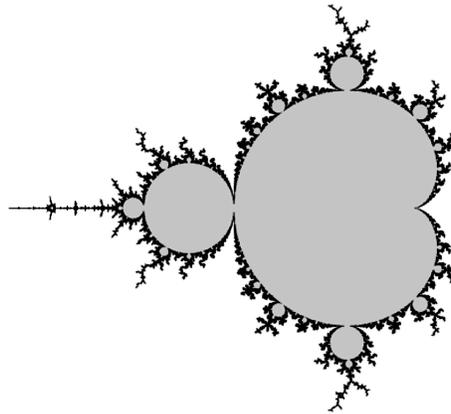
FIGURE 15. The Mandelbrot set.

---

[12]This automatically means for all x in the case of irrational α.

The Hofstadter butterfly is just one instance of a low-dimensional family of dynamical systems containing interesting dynamics and rich structure. A similar picture appears in complex dynamics,[13] in the one (complex) parameter family of dynamical systems $\{p_c(z) = z^2 + c : c \in \mathbb{C}\}$. The Mandelbrot set consists of parameters c for which the map p_c has a connected Julia set J_c (see Figure 15).

Note that in this noninvertible, conformal context, uniform hyperbolicity of the derivative cocycle of p_c on J_c just means that there exists an n such that $|(p_c^n(z))'| \geq 2$, for all $z \in J_c$. It is conjectured that the set of parameters c such that p_c is uniformly hyperbolic on J_c is (open and) dense in the Mandelbrot set.

6. THEMES

We conclude by summarizing a few themes, some of which have come up in our discussion.

Nonvanishing exponents sometimes produce chaotic behavior. The bedrock result in this regard is Anosov's proof [1] that smooth Anosov flows and diffeomorphisms are mixing (and in particular ergodic). Another notable result is Anatole Katok's proof [41] that measurable hyperbolicity of diffeomorphism with respect to some measure μ produces many periodic orbits—in particular, the number of orbits of period n grows exponentially in n. Pesin theory provides a sophisticated tool for exploiting measuable hyperbolicity to produce chaotic behavior such as mixing and even the Bernoulli property.

Exponents can carry geometric information. We have not discussed it here, but there are delicate relationships between entropy, exponents and Hausdorff dimension of invariant sets and measures, established in full generality by Ledrappier and Young [47, 48]. The expository article [63] contains a clear discussion of these relationships as well as some of the other themes discussed in this paper. The interplay between dimension, entropy and exponents has been fruitfully exploited in numerous contexts, notably in rigidity theory. Some examples are Ratner's theorem for unipotent flows, Elon Lindenstrauss's proof of Quantum Unique Ergodicity for arithmetic surfaces, and the Einsiedler–Katok–Lindenstrauss proof that the set of exceptions to the Littlewood conjecture has Hausdorff dimension 0; see [30, 49, 53].

Vanishing exponents sometimes present an exceptional situation that can be exploited. Both Furstenberg's theorem and Kotani theory illustrate this phenomenon. Here is Furstenberg's criterion, presented in a special case:

Theorem 6.1 (Furstenberg [34]). *Let $(A_1, \ldots, A_k) \subset SL(2, \mathbb{R})$, and let G be the smallest closed subgroup of $SL(2, \mathbb{R})$ containing $\{A_1, \ldots, A_k\}$. Assume that*

    (1) *G is not compact.*
    (2) *There is no finite collection of lines $\emptyset \neq L \subset \mathbb{R}^2$ such that $M(L) = L$, for all $M \in G$.*

Then for any probability vector $p = (p_1, \ldots, p_k)$ on $\{1, \ldots, k\}$ with $p_i > 0$, for all i, there exists $\chi^+(p) > 0$, such that for almost every $\omega \in \{1, \ldots, k\}^{\mathbb{N}}$ (with respect to

---

[13]Another field in which Avila has made significant contributions, which we do not touch upon here.

the Bernoulli measure $p^{\mathbb{N}}$),

$$\lim_{n \to \infty} \frac{1}{n} \log \|A^{(n)}(\omega)\| = \chi^+.$$

One way to view this result: if the exponent χ_+ vanishes, then the matrices either leave invariant a common line or pair of lines, or they generate a precompact group. Both possibilities are degenerate and are easily destroyed by perturbing the matrices. One proof of a generalization of this result [46] exploits the connections between entropy, dimension and exponents alluded to before. This result was formulated more completely in a dynamical setting by [21] as an "Invariance Principle," which has been further refined and applied in various works of Avila and others; see for example [10, 14, 15].

For general $SL(d, \mathbb{R})$ cocycles, vanishing of exponents is still an exceptional situation, but even more generally, the condition $d_i \geq 2$—that an exponent has multiplicity greater than 1—is also exceptional. This statement was rigorously established for random matrix products by Guivarc'h and Raugi [36] and undleries the Avila–Viana proof of simplicity of spectrum for the KZ cocycle. See the discussion at the end of Section 3. The same ideas play a role in Margulis's proof of superrigidity for higher rank lattices in semisimple Lie groups; see [52].

Continuity and regularity of exponents is a tricky business. In the family g_ϵ of Anosov diffeomorphisms considered in Subsection 2.6, the Lyapunov exponent $\chi_\epsilon = \log \lambda_\epsilon$ varies smoothly with the parameter ϵ. In general, however, Lyapunov exponents do not depend smoothly, or even continuously, on the cocycle. Understanding the exact relationship between exponents, cocycles, measures, and dynamics is an area still under exploration, and a few of Avila's deepest results, some of them still in preparation with Eskin and Viana, lie in this area. The book [62] is an excellent introduction to the subject.

Acknowledgments

About the author

Amie Wilkinson is professor of mathematics at the University of Chicago. She studies smooth dynamical systems and, in particular, the interplay between dynamics and other structures in pure mathematics—geometric, statistical, topological, and algebraic.

References

[1] D. V. Anosov, *Geodesic flows on closed Riemannian manifolds of negative curvature* (Russian), Trudy Mat. Inst. Steklov. **90** (1967), 209. MR0224110

[2] V. I. Arnol'd and A. Avez, *Ergodic problems of classical mechanics*, Translated from the French by A. Avez, W. A. Benjamin, Inc., New York-Amsterdam, 1968. MR0232910

[3] A. Avila, *Global theory of one-frequency Schrödinger operators*, Acta Math. **215** (2015), no. 1, 1–54, DOI 10.1007/s11511-015-0128-7. MR3413976

[4] A. Avila, *KAM, Lyapunov exponents and spectral dichotomy for one-frequency Schrödinger operators*. In preparation.

[5] A. Avila, *On the regularization of conservative maps*, Acta Math. **205** (2010), no. 1, 5–18, DOI 10.1007/s11511-010-0050-y. MR2736152

[6] A. Avila and J. Bochi, *Nonuniform hyperbolicity, global dominated splittings and generic properties of volume-preserving diffeomorphisms*, Trans. Amer. Math. Soc. **364** (2012), no. 6, 2883–2907, DOI 10.1090/S0002-9947-2012-05423-7. MR2888232

[7] A. Avila, S. Crovisier, A. Wilkinson, *Diffeomorphisms with positive metric entropy*, preprint.

[8] A. Avila and S. Jitomirskaya, *The Ten Martini Problem*, Ann. of Math. (2) **170** (2009), no. 1, 303–342, DOI 10.4007/annals.2009.170.303. MR2521117

[9] A. Avila and R. Krikorian, *Reducibility or nonuniform hyperbolicity for quasiperiodic Schrödinger cocycles*, Ann. of Math. (2) **164** (2006), no. 3, 911–940, DOI 10.4007/annals.2006.164.911. MR2259248

[10] A. Avila, J. Santamaria, and M. Viana, *Holonomy invariance: rough regularity and applications to Lyapunov exponents* (English, with English and French summaries), Astérisque **358** (2013), 13–74. MR3203216

[11] A. Avila, J. Santamaria, M. Viana, and A. Wilkinson, *Cocycles over partially hyperbolic maps*, Astérisque **358** (2013), 1–12. MR3203215

[12] A. Avila and M. Viana, *Simplicity of Lyapunov spectra: a sufficient criterion*, Port. Math. (N.S.) **64** (2007), no. 3, 311–376, DOI 10.4171/PM/1789. MR2350698

[13] A. Avila and M. Viana, *Simplicity of Lyapunov spectra: proof of the Zorich-Kontsevich conjecture*, Acta Math. **198** (2007), no. 1, 1–56, DOI 10.1007/s11511-007-0012-1. MR2316268

[14] A. Avila and M. Viana, *Extremal Lyapunov exponents: an invariance principle and applications*, Invent. Math. **181** (2010), no. 1, 115–189, DOI 10.1007/s00222-010-0243-1. MR2651382

[15] A. Avila, M. Viana, and A. Wilkinson, *Absolute continuity, Lyapunov exponents and rigidity I: geodesic flows*, J. Eur. Math. Soc. (JEMS) **17** (2015), no. 6, 1435–1462, DOI 10.4171/JEMS/534. MR3353805

[16] I. Bárány, A. F. Beardon, and T. K. Carne, *Barycentric subdivision of triangles and semigroups of Möbius maps*, Mathematika **43** (1996), no. 1, 165–171, DOI 10.1112/S0025579300011669. MR1401715

[17] J. Bochi, *Genericity of zero Lyapunov exponents*, Ergodic Theory Dynam. Systems **22** (2002), no. 6, 1667–1696, DOI 10.1017/S0143385702001165. MR1944399

[18] J. Bochi, C^1-*generic symplectic diffeomorphisms: partial hyperbolicity and zero centre Lyapunov exponents*, J. Inst. Math. Jussieu **9** (2010), no. 1, 49–93, DOI 10.1017/S1474748009000061. MR2576798

[19] J. Bochi and M. Viana, *Lyapunov exponents: how frequently are dynamical systems hyperbolic?*, Modern dynamical systems and applications, Cambridge Univ. Press, Cambridge, 2004, pp. 271–297. MR2090775

[20] J. Bochi and M. Viana, *The Lyapunov exponents of generic volume-preserving and symplectic maps*, Ann. of Math. (2) **161** (2005), no. 3, 1423–1485, DOI 10.4007/annals.2005.161.1423. MR2180404

[21] C. Bonatti, X. Gómez-Mont, and M. Viana, *Généricité d'exposants de Lyapunov non-nuls pour des produits déterministes de matrices* (French, with English and French summaries), Ann. Inst. H. Poincaré Anal. Non Linéaire **20** (2003), no. 4, 579–624, DOI 10.1016/S0294-1449(02)00019-7. MR1981401

[22] C. Bonatti and M. Viana, *Lyapunov exponents with multiplicity 1 for deterministic products of matrices*, Ergodic Theory Dynam. Systems **24** (2004), no. 5, 1295–1330, DOI 10.1017/S0143385703000695. MR2104587

[23] J. Bourgain and M. Goldstein, *On nonperturbative localization with quasi-periodic potential*, Ann. of Math. (2) **152** (2000), no. 3, 835–879, DOI 10.2307/2661356. MR1815703

[24] J. Bourgain and S. Jitomirskaya, *Continuity of the Lyapunov exponent for quasiperiodic operators with analytic potential*, J. Statist. Phys. **108** (2002), no. 5-6, 1203–1218, DOI 10.1023/A:1019751801035. Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays. MR1933451

[25] J. Chaika and A. Eskin, *Every flat surface is Birkhoff and Oseledets generic in almost every direction*, J. Mod. Dyn. **9** (2015), 1–23, DOI 10.3934/jmd.2015.9.1. MR3395258

[26] D. Damanik, *Schrödinger operators with dynamically defined potentials: a survey*, to appear: Erg. Th. Dyn. Syst.

[27] C. Dean, L. Wang, P. Maher, C. Forsythe, F. Ghahari, Y. Gao, J. Katoch, M. Ishigami, P. Moon, M. Koshino, T. Taniguchi, K. Watanabe, K. L. Shepard, J. Hone, P. Kim, *Hofstadter's butterfly and the fractal quantum Hall effect in moiré superlattices*, Nature **497** (2013), 598–602.

[28] D. Dolgopyat, H. Hu, Y. Pesin, *An example of a smooth hyperbolic measure with countably many ergodic components*, Proceedings of Symposia in Pure Mathematics **69** (2001) 95–106.

[29] D. Dolgopyat and Y. Pesin, *Every compact manifold carries a completely hyperbolic diffeomorphism*, Ergodic Theory Dynam. Systems **22** (2002), no. 2, 409–435, DOI 10.1017/S0143385702000202. MR1898798

[30] M. Einsiedler, A. Katok, and E. Lindenstrauss, *Invariant measures and the set of exceptions to Littlewood's conjecture*, Ann. of Math. (2) **164** (2006), no. 2, 513–560, DOI 10.4007/annals.2006.164.513. MR2247967

[31] A. Eskin, M. Kontsevich, and A. Zorich, *Sum of Lyapunov exponents of the Hodge bundle with respect to the Teichmüller geodesic flow*, Publ. Math. Inst. Hautes Études Sci. **120** (2014), 207–333, DOI 10.1007/s10240-013-0060-3. MR3270590

[32] G. Forni, *Deviation of ergodic averages for area-preserving flows on surfaces of higher genus*, Ann. of Math. (2) **155** (2002), no. 1, 1–103, DOI 10.2307/3062150. MR1888794

[33] H. Furstenberg and H. Kesten, *Products of random matrices*, Ann. Math. Statist. **31** (1960), 457–469. MR0121828

[34] H. Furstenberg, *Noncommuting random products*, Trans. Amer. Math. Soc. **108** (1963), 377–428. MR0163345

[35] I. Ya. Gol′dsheĭd and G. A. Margulis, *Lyapunov exponents of a product of random matrices* (Russian), Uspekhi Mat. Nauk **44** (1989), no. 5(269), 13–60, DOI 10.1070/RM1989v044n05ABEH002214; English transl., Russian Math. Surveys **44** (1989), no. 5, 11–71. MR1040268

[36] Y. Guivarc'h and A. Raugi, *Propriétés de contraction d'un semi-groupe de matrices inversibles. Coefficients de Liapunoff d'un produit de matrices aléatoires indépendantes* (French, with English summary), Israel J. Math. **65** (1989), no. 2, 165–196, DOI 10.1007/BF02764859. MR998669

[37] H. Helson, *Harmonic analysis*, 2nd ed., Texts and Readings in Mathematics, vol. 7, Hindustan Book Agency, New Delhi, 2010. MR2656971

[38] D. Hofstadter, *Energy levels and wavefunctions of Bloch electrons in rational and irrational magnetic fields*, Physical Review B **14** (1976) 2239–2249.

[39] R. A. Johnson, *Exponential dichotomy, rotation number, and linear differential operators with bounded coefficients*, J. Differential Equations **61** (1986), no. 1, 54–78, DOI 10.1016/0022-0396(86)90125-7. MR818861

[40] S. Jitomirskaya, C. A. Marx, *Dynamics and spectral theory of quasi-periodic Schrödinger-type operators*, to appear: Erg. Th. Dyn. Syst.

[41] A. Katok, *Lyapunov exponents, entropy and periodic orbits for diffeomorphisms*, Inst. Hautes Études Sci. Publ. Math. **51** (1980), 137–173. MR573822

[42] S. Kerckhoff, H. Masur, and J. Smillie, *Ergodicity of billiard flows and quadratic differentials*, Ann. of Math. (2) **124** (1986), no. 2, 293–311, DOI 10.2307/1971280. MR855297

[43] A. N. Kolmogorov, *Théorie générale des systèmes dynamiques et mécanique classique* (French), Proceedings of the International Congress of Mathematicians, Amsterdam, 1954, Vol. 1, Erven P. Noordhoff N.V., Groningen; North-Holland Publishing Co., Amsterdam, 1957, pp. 315–333. MR0097598

[44] S. Kotani, *Ljapunov indices determine absolutely continuous spectra of stationary random one-dimensional Schrödinger operators*, Stochastic analysis (Katata/Kyoto, 1982), North-Holland Math. Library, vol. 32, North-Holland, Amsterdam, 1984, pp. 225–247, DOI 10.1016/S0924-6509(08)70395-7. MR780760

[45] Y. Last, *Zero measure spectrum for the almost Mathieu operator*, Comm. Math. Phys. **164** (1994), no. 2, 421–432. MR1289331

[46] F. Ledrappier, *Positivity of the exponent for stationary sequences of matrices*, Lyapunov exponents (Bremen, 1984), Lecture Notes in Math., vol. 1186, Springer, Berlin, 1986, pp. 56–73, DOI 10.1007/BFb0076833. MR850070

[47] F. Ledrappier and L.-S. Young, *The metric entropy of diffeomorphisms. I. Characterization of measures satisfying Pesin's entropy formula*, Ann. of Math. (2) **122** (1985), no. 3, 509–539, DOI 10.2307/1971328. MR819556

[48] F. Ledrappier and L.-S. Young, *The metric entropy of diffeomorphisms. II. Relations between entropy, exponents and dimension*, Ann. of Math. (2) **122** (1985), no. 3, 540–574, DOI 10.2307/1971329. MR819557

[49] E. Lindenstrauss, *Invariant measures and arithmetic quantum unique ergodicity*, Ann. of Math. (2) **163** (2006), no. 1, 165–219, DOI 10.4007/annals.2006.163.165. MR2195133

[50] R. Mañé, *Ergodic theory and differentiable dynamics*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 8, Springer-Verlag, Berlin, 1987. Translated from the Portuguese by Silvio Levy. MR889254

[51] R. Mañé, *Oseledets's theorem from the generic viewpoint*. Proc. Int. Congress of Mathematicians (Warszawa 1983) Vol. 2, 1259–76.

[52] G. A. Margulis, *Discrete subgroups of semisimple Lie groups*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 17, Springer-Verlag, Berlin, 1991. MR1090825

[53] D. W. Morris, *Ratner's theorems on unipotent flows*, Chicago Lectures in Mathematics, University of Chicago Press, Chicago, IL, 2005. MR2158954

[54] C. McMullen, *Barycentric subdivision, martingales and hyperbolic geometry*, preprint, 2011.

[55] V. I. Oseledec, *A multiplicative ergodic theorem. Characteristic Ljapunov, exponents of dynamical systems* (Russian), Trudy Moskov. Mat. Obšč. **19** (1968), 179–210. MR0240280

[56] J. C. Oxtoby and S. M. Ulam, *Measure-preserving homeomorphisms and metrical transitivity*, Ann. of Math. (2) **42** (1941), 874–920. MR0005803

[57] Ja. B. Pesin, *Characteristic Ljapunov exponents, and smooth ergodic theory* (Russian), Uspehi Mat. Nauk **32** (1977), no. 4 (196), 55–112, 287. MR0466791

[58] J. Rodriguez-Hertz, *Genericity of nonuniform hyperbolicity in dimension 3*, J. Mod. Dyn. **6** (2012), no. 1, 121–138, DOI 10.3934/jmd.2012.6.121. MR2929132

[59] M. Shub and A. Wilkinson, *Pathological foliations and removable zero exponents*, Invent. Math. **139** (2000), no. 3, 495–508, DOI 10.1007/s002229900035. MR1738057

[60] W. A. Veech, *The Teichmüller geodesic flow*, Ann. of Math. (2) **124** (1986), no. 3, 441–530, DOI 10.2307/2007091. MR866707

[61] M. Viana, *Almost all cocycles over any hyperbolic system have nonvanishing Lyapunov exponents*, Ann. of Math. (2) **167** (2008), no. 2, 643–680, DOI 10.4007/annals.2008.167.643. MR2415384

[62] M. Viana, *Lectures on Lyapunov exponents*, Cambridge Studies in Advanced Mathematics, vol. 145, Cambridge University Press, Cambridge, 2014. MR3289050

[63] L.-S. Young, *Ergodic theory of differentiable dynamical systems*, Real and complex dynamical systems (Hillerød, 1993), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., vol. 464, Kluwer Acad. Publ., Dordrecht, 1995, pp. 293–336. MR1351527

[64] Z. Zhang, *Resolvent set of Schrödinger operators and uniform hyperbolicity*, arXiv:1305.4226v2(2013).

[65] A. Zorich, *How do the leaves of a closed 1-form wind around a surface?*, Pseudoperiodic topology, Amer. Math. Soc. Transl. Ser. 2, vol. 197, Amer. Math. Soc., Providence, RI, 1999, pp. 135–178, DOI 10.1090/trans2/197/05. MR1733872

[66] A. Zorich, *Asymptotic flag of an orientable measured foliation on a surface*, Geometric study of foliations (Tokyo, 1993), World Sci. Publ., River Edge, NJ, 1994, pp. 479–498. MR1363744

Department of Mathematics, University of Chicago, 5734 S. University Avenue, Chicago, Illinois 60637