Dynamical Systems Evolving

Lai-Sang Young[1]

ABSTRACT. This is an expanded version of a presentation given at ICM2018. It discusses a number of results taken from a cross-section of the author's work in Dynamical Systems. The topics include relation between entropy, Lyapunov exponents and fractal dimension, statistical properties of chaotic dynamical systems, physically relevant invariant measures, strange attractors arising from shear-induced chaos, random maps and random attractors. The last section contains two applications of Dynamical Systems ideas to Biology, one to epidemics control and the other to computational neuroscience.

Keywords: Lyapunov exponents, correlation decay, SRB measures, shear-induced chaos, random dynamical systems, epidemics, computational neuroscience

MSC2010 Classification: 37Dxx, 37Hxx, 92Bxx

I am grateful for the opportunity to share my ideas with the mathematics community. My field is *Dynamical Systems*, an area of mathematics concerned with time evolutions of processes natural and engineered, of iterative schemes and algorithms. It is the study of change, of moving objects and evolving situations, and the goal is to describe, analyze, explain, and ultimately predict.

I will be discussing mostly my own work, but to provide context – mostly for the benefit of readers not working in Dynamical Systems – I would like to start with a brief introduction to the field, where it has been, and where we are today. Dynamical Systems is now a little over 100 years old. As with any $2-3$ page overview of such a vast body of ideas, my account will necessarily be a gross simplification, a biased and intensely personal one at that, but still I hope it will help put the field in perspective.

[1]Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA. Email: lsy@cims.nyu.edu. This research is partially supported by NSF Grant DMS-1363161

Dynamical systems: a very brief overview

To facilitate referencing, I will divide the growth of the subject into four distinct phases, with the understanding that these are, for the most part, artificial boundaries.

I. A field was born. It is generally accepted that the origin of Dynamical Systems goes back to two sets of truly great ideas before and around the turn of the 20th century. One is the introduction of qualitative analysis into the study of differential equations by Poincaré in his work on celestial mechanics; the other consists of ideas surrounding the Ergodic Hypothesis.

Prior to the time of Poincaré, understanding a system defined by a set of (ordinary) differential equations meant solving analytically those equations, or approximating them with infinite series expansions. As we know now, most differential equations are not analytically solvable, and series expansions do not always shed light. Poincaré showed that a wealth of information about the qualitative behavior of a system, local and global, can be deduced without solving the equations explicitly, thereby introducing the world to the idea of geometric or qualitative analysis.

The Ergodic Hypothesis came from statistical mechanics. It asserts, roughly speaking, that because trajectories explore the phase space, time averages of an observable along individual trajectories should reflect the spatial average. In part because the needed language did not exist at the time, there is little consensus on how exactly this Hypothesis was first formulated or to whom to attribute it, though Boltzmann clearly had some of the ideas. However it came about, the impact that this profound idea has had on Dynamical Systems is undisputed.

II. Laying of foundation. Though Dynamical Systems was not recognized as a branch of mathematics until later, most of its foundation was laid in the first part of the twentieth century through the 1960s.

The framework for Ergodic Theory as the study of measure-preserving transformations was put on firm footing by the Ergodic Theorems of Birkhoff and von Neumann in the 1930s, and through the introduction of important invariants such as entropy later on. Ergodic theory can be seen as a probabilistic approach to Dynamical Systems, an approach in which one is not concerned with every initial condition but focuses instead on averages and almost-sure behaviors.

On the geometric-analytic side, the subject saw an explosion of creative activity in the middle of the 20th century for two distinct types of dynamical systems. I think of them as occupying the two ends of an "ordered-disordered" spectrum:

At the "ordered" end are quasi-periodic systems. In integrable Hamiltonian systems, orbits move about in an orderly fashion on highly constrained surfaces. The theory of Kolmogorov, Arnold and Moser (KAM) guarantees the persistence of this order on large parts of the phase space when such systems are perturbed.

At the "disordered" end are chaotic dynamical systems (though the word "chaos" was not yet in use at that time). Smale led a bold attempt to axiomatize chaotic behavior. Globalizing the idea of dynamics near a fixed point of saddle type, he invented as models of chaotic behavior the idea of hyperbolic invariant sets, characterized by exponential separation of nearby orbits on the set. His seminal work led to what eventually became the hyperbolic theory of dynamical systems.

III. Maturation and diversification. In the next few decades, Dynamical Systems gained formal recognition as an area of mathematics. The ideas outlined above continued to blossom, both conceptually and technically. New topics of research opened up, and the subject became more diverse. I mention below a sample of these developments:

Smooth ergodic theory, the use of ergodic theory techniques to study differentiable dynamical systems, was pioneered in the then Soviet Union in the early 1970s. The result was a nonuniform hyperbolic theory, which generalizes the uniformly hyperbolic invariant sets of Smale to a version where hyperbolicity occurs only *almost everywhere* and its onset is nonuniform in time.

KAM theory gave rise to two major topics: one studies structures that are remnants of KAM tori as one moves farther from integrable systems, and the other, pioneered by Arnold, studies the diffusion of orbits in near-integrable systems having more than two degrees of freedom, where motion is not blocked by invariant tori.

In the 1980s and 90s, a large number of researchers flocked to the study of one dimensional maps, real and complex, attracted by the fact that (i) the low dimensionality of the phase space made these systems more tractable, yet (ii) such "simple" systems already exhibited rich dynamical behaviors. The subject flourished; a number of deep and very striking theorems were proved. One dimension, however, is very special, and implications of 1D results for higher dimensions remain to be explored.

There was also growing interest in the analysis of concrete systems. Well known examples include the Lorenz system (3-mode truncation of convection in 2D), the periodically forced van der Pol Equation (a slow-fast system the forcing of which produces complex dynamics), and a class of dynamical systems known as "billiards" and "hard balls". Billiard systems are models of uniform motion of point particles in bounded domains, making elastic collisions with the walls of the domain.

Finally, smooth actions of groups other than $\mathbb{Z}$ and $\mathbb{R}$ on manifolds extending the theory of diffeomorphisms and flows became an active area. Many results concern "rigidity", as relations on group elements can severely limit the actions that can exist.

IV. Making connections. Starting from its inception, Dynamical Systems has always had contact with other fields. These contacts have intensified in recent years for multiple reasons: While it makes sense to have a general theory, what can be said about general systems without more specific context is limited. At the same time, the fact that many parts of mathematics and science can benefit from Dynamical Systems thinking became clearer than ever before.

Though I don't know enough to write about them individually, I know that Dynamical Systems has interactions with many branches of pure mathematics, including Number Theory, Geometric Group Theorey, Differential Geometry, Analysis and Probability, some through problems of mutual interest, or analogous phenomena, and also because many problems are naturally studied through iterative procedures.

Opportunities of collaborative interaction with applied mathematics, engineering, the physical and biological sciences abound, as almost all systems evolve with time. Many such collaborations are already in place or ongoing, but here I think we have some distance to go to achieve the full potential of Dynamical Systems' usefulness. I also think that to do that, some retooling may be necessary on our part.

This concludes the brief overview.

Turning now to my own research, I entered the field during Phase III in the Overview. In the sections to follow, I want to share with the reader some highlights of my work, the bulk of which lies in the ergodic theory of chaotic dynamical systems, the topic that descended from what I described as the "disordered" end of the spectrum in Part II — though my interests diversified with time. I have chosen to present several snapshots taken from a cross-section of my work, as opposed to an in-depth discussion of one or two results, because I believe that will give you a more balanced view of what I do, and of what my field is about.

This article is written with the aim of communicating with the broader mathematics community. I hope you will find it readable except for a few things here and there that I have to ask you to accept on faith.

1 Entropy, Lyapunov exponents, and fractal dimension

The middle third Cantor set can be seen as the invariant set

$$\Lambda := \left\{ x \in [0,1] : f^n x \in \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right] \forall n \in \mathbb{Z}^+ \right\}$$

of the map $f(x) = 3x \bmod 1$, $x \in [0,1]$. When iterated, the complexity of f restricted to Λ can be thought of as growing like 2^n, as in each iterate there is the possibility of getting sent to two distinct intervals. Clearly, $(f^n)' = 3^n$, and an elementary exercise shows that the Hausdorff dimension of Λ, $\mathrm{HD}(\Lambda) = \log 2 / \log 3$.

The main result of this section is a relation among three invariants for general dynamical systems: entropy, which measures the average growth in complexity, Lyapunov exponents, which measures average derivative growth, and the dimension of an invariant measure. The results presented will hold in all dimensions. In one dimension, it generalizes the Cantor set computation above, illustrating a concept at the heart of Ergodic Theory, namely that quantities that vary from point to point in the phase space can be represented by time averages along orbits.

For the benefit of readers not familiar with the subject, I have included in the first half of Sect. 1.1 some background material in smooth ergodic theory. The idea of Lyapunov exponents, in particular, will appear many times in Sections 1–4.

1.1 Setting and background information

Let M be a C^∞ compact Riemannian manifold without boundary. We consider a pair (f, μ), where $f : M \circlearrowleft$ is a diffeomorphism of M onto itself, of differentiability class $C^{1+\alpha}$ for some $\alpha > 0$, and μ is a Borel probability measure preserved by f, i..e., for every Borel subset $E \subset M$, $\mu(f^{-1}(E)) = \mu(E)$.

Given (f, μ), there are two important, and conceptually different, ways to quantify the dynamical complexity of the system. One is *entropy*, and the other is *Lyapunov exponents*. The entropy of f with respect to μ, denoted $h_\mu(f)$, is defined to be

$$h_\mu(f) = \sup_\eta h_\mu(f; \eta)$$

where the supremum is taken of all finite measurable partitions η of M and

$$h_\mu(f; \eta) = \lim_{n\to\infty} \frac{1}{n} H(\bigvee_{i=0}^{n-1} f^{-i}\eta) .$$

Here $H(\eta) = \int I(\eta) d\mu$ where $I(\eta)(x) = -\log \mu(\eta(x))$, so $h_\mu(f; \eta)$ has the interpretation of average information gain, or average uncertainty removed, per iteration of the map. (See [43] for an introductory text.) The Lyapunov exponent of f at x in the direction of a tangent vector v is the growth rate of $\|Df^n_x(v)\|$, i.e.,

$$\lambda(x, v) = \lim_{n\to\infty} \frac{1}{n} \log \|Df^n_x(v)\|$$

if this limit exists. By the well known Multiplicative Ergodic Theorem [33], given (f, μ), at μ-a.e. x, there is a set of numbers

$$\lambda_1(x) > \lambda_2(x) > \cdots \lambda_{r(x)}(x)$$

and a splitting of the tangent space T_xM at x into a direct sum of subspaces

$$T_xM = E_1(x) \oplus E_2(x) \oplus \cdots \oplus E_{r(x)}(x)$$

such that for all $v \in E_i(x)$, $v \neq 0$, we have $\lambda(x, v) = \lambda_i(x)$. The functions $x \mapsto r(x), \lambda_i(x), E_i(x)$ are Borel measurable, and since $\lambda_i(x) = \lambda_i(fx)$, we have that if (f, μ) is ergodic, then the Lyapunov exponents are given by a finite set of numbers $\lambda_1 > \cdots > \lambda_r$ with multiplicities $m_1, \cdots, m_r$ where $m_i = \dim(E_i)$.

I started to learn about these ideas when a number of exciting results had just been proved. Among them are the following relations between entropy and Lyapunov exponents: Most basic are Ruelle's inequality [38], which asserts that

$$h_\mu(f) \le \int \sum_i \lambda_i^+(x) m_i(x) d\mu , \qquad a^+ = \max\{a, 0\} , \tag{1}$$

and Pesin's entropy formula [34], which asserts that if μ is volume on M, then the inequality in (1) is an equality. These results can be interpreted as saying that all uncertainty is created by expansion, and that in conservative systems, all expansion is used to create entropy. The entropy formula was subsequently generalized to hold for SRB measures [26], i.e., μ is required only to have conditional densities on unstable manifolds; it does not have to be volume.

My contribution to this topic is to explain the gap in (1) in terms of the dimension of μ. For an arbitrary finite Borel measure ν on a metric space, we define the *dimension* of ν at x, $\dim(\nu, x)$, to be δ if $\nu(B(x,\rho)) \sim \rho^\delta$ for small ρ where $B(x,\rho)$ is the ball of radius ρ centered at x. More precisely,

$$\dim(\nu, x) := \lim_{\rho\to 0} \frac{\log \nu(B(x,\rho))}{\log \rho} = \delta$$

assuming this limit exists, and write $\dim(\nu)$ if $\dim(\nu, x)$ is constant ν-a.e. So Lebesgue measure on $\mathbb{R}^d$ has dimension d, but δ in general does not have to be an integer; it is a notion of fractal dimension.

1.2 Results

My first result in this topic was for surface diffeomorphisms. The notation is as above.

Theorem 1 [51] *Let dim$(M) = 2$, and let f be an arbitrary $C^{1+\alpha}$ diffeomorphism of M. We assume (f, μ) is ergodic with $\lambda_1 > 0 > \lambda_2$. Then dim$(\mu)$ is well defined and is given by*

$$\mathit{dim}(\mu) = h_\mu(f)\left[\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right] .$$

The idea, roughly speaking, is that the dimension of μ "in the unstable direction" is equal to $h_\mu(f)/\lambda_1$, it is equal to $h_\mu(f)/(-\lambda_2)$ "in the stable direction", and $\dim(\mu)$ is the sum of these two contributions. The number $h_\mu(f)/\lambda_1$ is the direct analog of $\log 2/\log 3$ in the middle third Cantor set example at the beginning of this section.

When there are multiple positive Lyapunov exponents, the nonconformality makes the situation considerably more complicated. The next result, proved in a joint work with Ledrappier, deals with that.

Theorem 2 [22] *Let f be a $C^{1+\alpha}$ diffeomorphism of a compact Riemannian manifold of arbitrary dimension, and let (f, μ) be ergodic, with positive Lyapunov exponents $\lambda_1 > \cdots > \lambda_r$ and multiplicities $m_1, \ldots, m_r$ respectively. Then $dim(\mu|W^u)$, the dimension of the conditional measures of μ on unstable manifolds, is well defined, and there are numbers $\delta_i \in [0, m_i], i = 1, \cdots, r$, such that*

$$h_\mu(f) = \sum_{i=1}^{r} \lambda_i \delta_i \qquad \text{and} \qquad \dim(\mu|W^u) = \sum_{i=1}^{r} \delta_i \ .$$

The numbers δ_i have the interpretation of being the "partial dimension" of μ in the direction of E_i, the subspace corresponding to the Lyapunov exponent λ_i. Interpreting $\lambda_i \delta_i$ as the entropy of (f, μ) in the direction of E_i, the first equality in Theorem 2 asserts that $h_\mu(f)$ is the sum of the "partial entropies" in the different expanding directions, while the second equality asserts that the dimension of $\mu|W^u$ is the sum of the partial dimensions. To be technically correct, I should say that the system (f, μ) has a hierarchy of unstable manifolds $W^1 \subset W^2 \subset \cdots \subset W^r = W^u$ defined μ-a.e., with W^k tangent to $E_1 \oplus \cdots \oplus E_k$, but there need not be invariant manifolds tangent to each E_i; and that in the actual proof, the δ_i are the dimensions of quotient measures on W^i/W^{i-1}, and $\lambda_i \delta_i = h_i - h_{i-1}$ where h_i can be made precise as a notion of entropy along the invariant foliation W^i.

Clearly, one can apply Theorem 2 to f^{-1}, obtaining an analogous result for $\dim(\mu|W^s)$. The result of [2], which states that in the absence of zero Lyapunov exponents,

$$\dim(\mu) = \dim(\mu|W^u) + \dim(\mu|W^s) \ ,$$

completes this circle of ideas.

I want to stress again the generality of the results in this section: they hold for *all diffeomorphisms and all invariant Borel probability measures*, with no restriction whatsoever on the geometry or the defining equations of the maps, or on the invariant measure.

2 Statistical properties of chaotic dynamical systems via Markov extensions

In Paragraph II of the Overview in the Introduction, I portrayed, impressionistically, all dynamical systems as lying on an "ordered-disordered" spectrum.

Let (f, μ) be as in Section 1, and think of it as being on the "disordered" side of this spectrum. Given an observable $\varphi : M \to \mathbb{R}$, how random can sequences of the form

$$\varphi(x),\ \varphi(fx),\ \varphi(f^2x),\ \ldots,\ \varphi(f^n x), \ldots$$

be for randomly chosen initial condition $x \in M$? Can these sequences be, for example, as random as functions of coin flips? This is a tricky question, for dynamical systems generated by maps or flows are *deterministic* in the sense that given an initial condition, the entire future trajectory is fully determined and nothing is left to chance. At the same time, if f has chaotic dynamics, then it may be hard to predict $\varphi(f^n x)$ from approximate knowledge of x. Indeed uniformly hyperbolic systems have been shown to produce statistics that obey some the same probabilistic limit laws as genuinely random stochastic processes.

A fairly complete theory of the statistical properties of uniformly hyperbolic or Axiom A systems (à la Smale) was developed in the 1970s; see [39, 7]. After that, the community began to move beyond Axiom A, to confront the challenges of examples such as the Lorenz attractor, Hénon maps and various billiard systems to which previously developed analytical techniques did not apply. It was against this backdrop that the work presented in this section was carried out.

2.1 A unified view for predominantly hyperbolic systems

First, a basic definition: Let $\Lambda \subset M$ be a compact invariant set of a diffeomorphism $f : M \circlearrowleft$. We say $f|_\Lambda$ ("f restricted to Λ") is *uniformly hyperbolic* if the tangent bundle $T_\Lambda M$ over Λ can be split into the direct sum of two Df-invariant subbundles

$$T_\Lambda M = E^u \oplus E^s$$

with the property that $Df^n|_{E^u}$ is uniformly expanding and $Df^n|_{E^s}$ is uniformly contracting. As mentioned in the Overview, this idea was introduced by Smale in his ground-breaking work in [42]. It is slightly inaccurate, but I will use the term "uniformly hyperbolic" and "Axiom A" interchangeably in this article.

The existence of hyperbolic invariant sets in a dynamical system is often seen as the presence of chaotic behavior. Many of the examples that challenged the community in the 1980s are not uniformly hyperbolic, but their dynamics are dominated by large hyperbolic invariant sets. This prompted me to propose the following construction as an attempt to provide a unified view for a class of dynamical systems with weaker forms of hyperbolicity than Axiom A. The material in the rest of this subsection is taken from [53].

Given $f : M \circlearrowleft$, my proposal was to construct, where possible, a *countable Markov extension*, i.e., a map $F : \Delta \circlearrowleft$ with

$$\begin{tikzcd} \Delta \arrow[r, "F"] \arrow[d, "\pi"'] & \Delta \arrow[d, "\pi"] \\ M \arrow[r, "f"] & M \end{tikzcd} \tag{2}$$

where $F : \Delta \circlearrowleft$ has the structure of a countable state Markov chain. To be clear, F is not a Markov chain, or isomorphic to one in any sense, only that there is a countable partition on Δ with respect to which the action of F has the flavor of a Markov chain.

Finite Markov partitions were constructed for Axiom A systems and were instrumental in the study of their statistical properties, see e.g. [41, 7, 39]. What I proposed was a generalization of these ideas to systems with weaker hyperbolicity. The usefulness of this proposal will, of course, depend on what one can do with it, and I will discuss that in Sects. 2.2 and 2.3.

It's a little technical, but I will give more detail on $F : \Delta \circlearrowleft$. We look for a set Λ_0 with a nice hyperbolic structure, and study "hyperbolic returns" to this set. More precisely, we seek Λ_0 of the form

$$\Lambda_0 \ = \ \left(\bigcup_{\gamma^u \in \Gamma^u} \gamma^u \right) \bigcap \left(\bigcup_{\gamma^s \in \Gamma^s} \gamma^s \right)$$

where Γ^s and Γ^u are two families of local stable and unstable manifolds, each element of $\gamma^s \in \Gamma^s$ intersecting each $\gamma^u \in \Gamma^u$ transversally in a unique point. We call a subset $\Theta \subset \Lambda_0$ a *u-subset* if $\Theta = \Lambda_0 \cap (\cup_{\gamma^u \in \tilde{\Gamma}^u} \gamma^u)$ for some subset $\tilde{\Gamma}^u \subset \Gamma^u$; *s-subsets* of Λ_0 are defined analogously. Pictorially, if one thinks of Γ^u as a stack of roughly "horizontal" disks and Γ^s a stack of roughly "vertical" disks, then Λ_0 is the lattice of intersection points of the disks in the two families, a u-subset of Λ_0 is a sublattice that runs from left to right, and an s-subset is a sublattice that runs from top to bottom.

Now suppose there is a decomposition of Λ_0 into

$$\Lambda_0 \;=\; \bigcup_{i=1}^{\infty} \Lambda^s_{0,i} \qquad \text{(disjoint union)}$$

with the property that each $\Lambda^s_{0,i}$ is a s-subset, and for each i, there exists $r_i \in \mathbb{Z}^+$ such that $f^{r_i}(\Lambda^s_{0,i}) = \Lambda^u_{0,i}$ where $\Lambda^u_{0,i}$ is an u-subset of Λ_0. This is what I meant by a "hyperbolic return". The function $R : \Lambda_0 \to \mathbb{Z}^+$ given by $R(x) = r_i$ for $x \in \Lambda^s_{0,i}$ is called the *return time* function. (We do not require R to be the first return time, and the $\Lambda^u_{0,i}$ do not have to be pairwise disjoint.)

For $f : M \circlearrowleft$ admitting the construction above, we let $\Delta = \bigcup_{i=0}^{\infty} \Delta_i$, where each Δ_i is a copy of $f^i(\{x \in \Lambda_0 : R(x) > i\})$, let π is the identification map, and define $F : \Delta \circlearrowleft$ so the diagram in (2) commutes. This in our Markov extension.

2.2 Statistical properties of systems with Markov extensions

To address the question raised at the beginning of this section, we begin with the following definitions: Let (f, μ) be as in Section 1, and let $\mathcal{F}$ be a class of functions on M. We say (f, μ) has *exponential decay of time correlations* for test functions in $\mathcal{F}$ if there exists $\tau < 1$ such that for all $\varphi, \psi \in \mathcal{F}$, there exists $C = C(\varphi, \psi)$ such that

$$\left| \int (\varphi \circ f^n)\psi d\mu - \int \varphi d\mu \int \psi d\mu \right| \le C\tau^n \qquad \forall n \ge 1 \ .$$

Thus if $\varphi \circ f^n$ represents the observation on "Day n", then (f, μ) having exponential correlation decay means in particular that $\varphi \circ f^n$ and φ decorrelate exponentially fast in n. Informally, on Day n, there is still some memory of Day 0, but memory is fading exponentially fast. *Polynomial decay of correlations* is defined similarly.

In ergodic theory, a measure-preserving transformation (f, μ) is said to be *mixing* if for all measurable sets A, B, $\mu(f^{-n}A \cap B) \to \mu(A)\mu(B)$ as $n \to \infty$. The decay of time correlations to 0 is just a functional form of the same idea, i.e., exponential correlation decay means exponential mixing of sets in the phase space.

We say φ satisfies the *Central Limit Theorem* (CLT) with respect to (f, μ) if

$$\frac{1}{\sqrt{n}}\left(\sum_{i=0}^{n-1}\varphi\circ f^i - n\int\varphi d\mu\right) \ \to\ \mathcal{N}(0,\sigma)\ ,$$

the normal distribution with variance σ^2 for some $\sigma \geq 0$.

Consider now a map $f : M \circlearrowleft$ admitting a Markov extension. To this construction we now add a *reference measure*: Recall that Γ^s and Γ^u are the families of stable and unstable disks defining the hyperbolic product set Λ_0. We let $m = m_{\gamma^u}$ be the Riemannian measure on γ^u, and assume that

$$m(\gamma^u \cap \Lambda_0) > 0 \qquad \text{for some} \qquad \gamma^u \in \Gamma^u\ . \tag{3}$$

Furthermore, we require only that the return time function R be defined on m-a.e. $x \in \Lambda_0 \cap \gamma^u$. The reason for this focus on m is that we are primarily interested in SRB measures, the conditional measures on unstable manifolds of which are equivalent to Lebesgue. The importance of SRB measures is discussed in Sect. 3.1.

For systems satisfying the conditions above, I proved the following results:

Theorem 3 [53, 54] *Let $f : M \circlearrowleft$ be as above, i.e., there is a set $\Lambda_0 \subset M$ satisfying (3) with return time function $R : \Lambda_0 \to \mathbb{Z}^+$ and reference measure m on γ^u, and let $F : \Delta \circlearrowleft$ be its Markov extension. Then the following hold:*

(a) If $\int Rdm < \infty$, then f has an SRB measure μ.

(b) If $\int Rdm < \infty$ and $\gcd\{R\} = 1$, then (f, μ) is mixing.

Here gcd *= greatest common divisor. Assume below that the conditions in (b) hold. All results pertain to (f, μ), and all test functions are Hölder continuous.*

(c) If $m\{R > n\} \leq C\theta^n$ for some $\theta < 1$, then correlation decay is exponential.

(d) If $m\{R > n\} = \mathcal{O}(n^{-\alpha}), \alpha > 1$, then correlation decays at $\mathcal{O}(n^{-\alpha+1})$.

(e) If $m\{R > n\} = \mathcal{O}(n^{-\alpha}), \alpha > 2$, then the CLT holds.

The results in Theorem 3 were first proved for $F : \Delta \circlearrowleft$ and then passed to f, the idea being that proofs for F are simpler because $F : \Delta \circlearrowleft$ has the topological structure of a countable state Markov chain. Parts (a) and (b) have obvious analogs with Markov chain theory: finite expectation of return times is equivalent to positivity of recurrence, $\gcd\{R\} = 1$ means periodicity. Parts (c)–(e) establish the connection between statistical properties of the system and the tails of "renewal" times. Though I was not aware of it until later, results along similar lines for (real) Markov chains were obtained around that same general time frame [32]. My "Markov extensions" are not Markov chains as I have indicated earlier, but the ideas are similar.

2.3 Periodic Lorentz gas and other potential applications

The method discussed in Sects. 2.1 and 2.2 was used to study the statistical properties for a number of examples of dynamical systems that have hyperbolic properties but are not necessarily uniformly hyperbolic (see e.g. [49, 53, 54, 10, 36]). I will discuss in some detail one example, the 2-dimensional periodic Lorentz gas, and finish with some remarks on other applications.

The 2D periodic Lorentz gas. The Lorentz gas is a model for electron gases in metals. Mathematically, the 2D periodic case is represented by the motion of a point mass in $\mathbb{R}^2$ bouncing elastically off a fixed periodic configuration of convex scatterers. It was first studied by Sinai around 1970 [40], and is sometimes called the Sinai billiard. Putting the dynamics on the 2-torus $\mathbb{T}^2$, we obtain the billiard flow φ_t on $\Omega \times \mathbb{S}^1$ where $\Omega = \mathbb{T}^2 \setminus \Omega_i$ and the Ω_i's are disjoint convex regions with C^3 boundaries. Points in $\Omega \times \mathbb{S}^1$ are denoted (x, θ) where $x \in \Omega$ is the footpoint of the arrow pointing in direction θ, the direction of the motion. A section to the billiard flow is the collision manifold $M = \partial\Omega \times [-\frac{\pi}{2}, \frac{\pi}{2}]$. We consider the collision map f of the billiard flow, or the first return map from M to itself. It is straightforward to check that f leaves invariant the probability measure $\mu = c \cos\theta dx d\theta$ where θ is the angle the arrow makes with the normal pointing into Ω and c is the normalizing constant. For simplicity, we will assume the finite horizon condition, which requires that the time between collisions be uniformly bounded.

Following the strategy outlined in Sects. 2.1 and 2.2, i.e., by constructing for f a countable Markov extension and investigating the tail properties of the return time function R, I proved the following result. Let C^α denote the class of Hölder functions on M with Hölder exponent α.

Theorem 4 [53] *Let (f, μ) be as above. Then correlation decays exponentially fast for observables in C^α. More precisely, there exists $\beta = \beta(\alpha) > 0$ such that for every $\varphi, \psi \in C^\alpha$,*

$$\left| \int (\varphi \circ f^n)\psi d\mu - \int \varphi d\mu \int \psi d\mu \right| \leq Ce^{-\beta n} \qquad \forall n \geq 1$$

for some $C = C(\varphi, \psi)$.

For more information on statistical properties of billiards and hard balls, see [11]. A weaker version of this result showing "stretched exponential decay" was first proved for this class of billiards in [8] along with other statistical properties like the CLT. Markov extensions are, needless to say, not the only way to study correlation decay of dynamical systems. Stronger results on larger classes of billiard maps were obtained by Chernov [10], and exponential decay of the billiard *flow*, a much harder problem, was solved only recently by Baladi, Demers and Liverani [1].

Summary and other applications. The ideas of this section were intended for systems possessing a good amount of hyperbolicity but are not necessarily uniformly hyperbolic. The approach I proposed was to set aside individual characteristics of the dynamical system, focus on return times to a good reference set, and to deduce the statistical properties of the system from tail properties of these return times.

This method has been particularly effective for systems with a localized source of nonhyperbolicity, an identifiable set that spoils the hyperbolicity of orbits passing near it. The simplest examples are 1D maps of the form $f(x) = 1 - ax^2, x \in [-1, 1]$, $a \in [0, 2]$. Here the "bad set" is $\{x = 0\}$, where $f' = 0$. I showed in [49] that if the orbit of 0 does not approach the point 0 too fast, then the map has exponential correlation decay. In the Lorentz gas example above, the "bad set" is where particle trajectories graze the scatterers. In other examples, it can take the form of "traps" or "eddies", where orbits linger for long periods of times, thereby slowing down the speed with which different regions of the phase space are mixed. Two examples in this category are neutral fixed points of 1D maps [54] and billiard maps with parabolic regions such as the stadium [31], both of which have polynomial decay.

Yet another kind of "bad set" is where directions of expansions and contractions are switched, i.e., as the orbit passes near the bad set, tangent vectors that have been growing in length get rotated, causing them to shrink in subsequent iterates. Bad sets of this type are very challenging to deal with, as we will see. I want to mention that the results of this section have also been applied successfully to prove exponential correlation decay for the “good maps” in Theorem 7 [48].

3 Strange attractors from shear-induced chaos

This was one of my first attempts to connect the abstract theory of chaotic dynamical systems to concrete settings. An immediate question is: *Which invariant measure should one consider?* Except in the case of Hamiltonian or volume preserving flows, dissipative systems such as those with attractors do not come equipped with a natural invariant measure. In general, the number of invariant measures is very large, and not all of them are relevant for purposes of describing what one sees.

Sect. 3.1 discusses general mathematical issues associated with observable chaos. Sect. 3.2 describes some examples and Sect. 3.3 the analysis behind these examples.

3.1 Observable chaos and SRB measures

It is one thing for a dynamical system to have orbits that behave chaotically, another for this chaotic behavior to be observable. In finite-dimensional dynamics, one often equates positive Lebesgue measure sets with observable events. Adopting such a view, we say $f : M \circlearrowleft$ has *observable chaos* if $\lambda_{\max} > 0$ on at least a positive Lebesgue measure set, where

$$\lambda_{\max}(x) := \liminf_{n\to\infty} \frac{1}{n} \log \|Df_x^n\| ,$$

i.e. $\lambda_{\max}$ is the largest Lyapunov exponent at x when that is defined. In the rest of this section, I will write “$\lambda_{\max} > 0$” as abbreviation for “observable chaos”.

Hyperbolic invariant sets such as Smale’s horseshoes contain orbits with chaotic dynamics, but the presence of a horseshoe does not imply $\lambda_{\max} > 0$, for the horseshoe itself occupies a zero Lebesgue measure set, and its presence does not preclude the possibility that orbits starting from Leb-a.e. $x \in M$ may tend eventually to a stable equilibrium, called a “sink”. By contrast,

$\lambda_{\max} > 0$ implies that instability is not only observable (in the sense of Lebesgue measure), but it persists for all future times. It is a much stronger form of chaos than the presence of horseshoes alone.

The question, then, is: which systems have $\lambda_{\max} > 0$? Following [12], we call an invariant probability measure *physically relevant* if it reflects the properties of initial conditions on positive Lebesgue measure sets. For Hamiltonian or volume-preserving systems, Liouville measure or Riemannian measure are clearly physically relevant. For systems that are not conservative (no technical meaning intended), the situation is more subtle: In such systems, trajectories of most orbits tend toward *attractors*. For a set to attract, there is often volume contraction, and when there is volume contraction, all invariant probability measures are necessarily singular with respect to Lebesgue measure. *A priori*, it is not clear if the concept of a physically relevant invariant measure makes sense for attractors.

An important discovery in the 1970s by Sinai, Ruelle and Bowen was that every uniformly hyperbolic or Axiom A attractor Λ admits a special invariant probability measure μ which plays the role of Liouville measure for Hamiltonian systems, see [41, 37, 7]. This measure – when Λ is not a sink – is called an *SRB measure*. SRB measures are generally singular with respect to Lebesgue. Their physical relevance is derived from the fact that they have smooth conditional measures on unstable manifolds. A precise definition of physical relevance here is that for every continuous observable $\varphi : M \to \mathbb{R}$, Leb-a.e. x in the basin of the attractor has the property that

$$\frac{1}{n}\sum_{i=0}^{n-1} \varphi(f^i x) \;\to\; \int \varphi d\mu \; . \tag{4}$$

I stress that (4) is not a consequence of Birkhoff's Ergodic Theorem: In general, μ is supported on the attractor Λ, a zero Lebesgue measure set Λ; yet it governs the large-time dynamics of trajectories starting from Leb-a.e. point in the basin of attraction, a much larger open set containing Λ.

The idea of SRB measures was generalized in the 1980s to a significantly broader context by Ledrappier and myself, meaning we identified a special class of invariant measures for general diffeomorphisms and flows that play a similar role as SRB measures for Axiom A attractors [26, 23, 52]. These measures continue to be characterized by their smooth conditional measures on unstable manifolds, though the picture is more complicated and less tidy than in the uniform hyperbolic case. Importantly, if a system admits an ergodic SRB measure μ with no zero Lyapunov exponents, then (4) holds for initial conditions from a positive Lebesgue measure set [35], implying $\lambda_{\max} > 0$.

Generalizing the idea of SRB measures to arbitrary dynamical systems does not, however, guarantee their *existence* in this larger context. Uniformly hyperbolic systems are special in that they have well aligned subspaces E^u consisting of vectors that are uniformly expanded from one iterate to the next. They satisfy what is called an "invariant cones" condition. For systems without continuous families of invariant cones – and that is the case for most dynamical systems – $\|Df^n_x(v)\|$ is likely to sometimes grow and sometimes shrink as n increases for most tangent vectors v, making it very challenging to identify aligned directions of exponential growth, a prerequisite for SRB measures.

Though SRB measures are thought to be prevalent among systems with chaotic attractors, current state of the art is that few instances of (genuinely) nonuniformly hyperbolic attractors have been shown rigorously to possess SRB measures. I will give some concrete examples in Sect. 3.2. All currently known examples in fact belong in the class of "rank one" attractors, which I will discuss in Sect. 3.3.

3.2 Shear-induced chaos in periodically kicked oscillators

A mechanism for producing $\lambda_{\max} > 0$ is shear-induced chaos, by which I mean the following: Start with a system with tame, nonchaotic dynamics, but some amount of shearing. The idea, in a nutshell, is that external forcing that magnifies the underlying shear can produce stretching and folding of the phase space, leading to $\lambda_{\max} > 0$.

A perfect setting for this mechanism is the periodic forcing of oscillators. The idea goes back nearly 100 years, to van der Pol and van der Mark, who observed in the course of their work on vacuum tube triode circuits that periodic forcing of relaxation oscillators could lead to "irregular noise". The problem was studied analytically by many authors: Cartwright, Littlewood, Levinson in the 1940s, Levi and others much later. The existence of horseshoes was proved analytically for a linearized system by Levi and for the original van der Pol equation by Haiduc with a computer assisted proof. See [15] for references on the above.

As explained in Sect. 3.1, the existence of horseshoes does not imply observable chaos. Noting that it is technically simpler to decouple the effects of the periodic drive from the dynamics of the unforced oscillator, my co-author Wang and I used pulsatile forcing, or *kicks*, with relatively long periods of relaxation in between. The phenomenology already manifests itself in the following very simple example:

Linear shear flow.[2] This model is given by

$$\begin{aligned} \dot{\theta} &= 1 + \sigma y \ , \\ \dot{y} &= -\lambda y \ + \ A \cdot \sin(2\pi\theta) \cdot \textstyle\sum_{n=-\infty}^{\infty} \delta(t - nT) \ , \end{aligned} \tag{5}$$

where $(\theta, y) \in \mathbb{S}^1 \times \mathbb{R}$, $\mathbb{S}^1 \equiv \mathbb{R}/\mathbb{Z}$, and σ, λ, A and T are constants with $\sigma, \lambda > 0$ and $T \gg 1$. Here, Eq. (5) with $A = 0$ is the unforced equation. Letting Φ_t denote the unforced flow, one sees that for all $z \in \mathbb{S}^1 \times \mathbb{R}$, $\Phi_t(z)$ tends to the limit cycle $\gamma = \{y = 0\}$ as $t \to \infty$. With T being the period of the forcing, it is natural to consider the time-T map $F_T = \Phi_T \circ \kappa$ of the forced system, where the effect of the instantaneous kick is given by $\kappa(\theta, y) = (\theta, y + A \sin(2\pi\theta))$.

We observed that for λT large, i.e., if the contraction between kicks is strong, the ratio

$$\frac{\sigma}{\lambda} A \ = \ \frac{\text{shear}}{\text{contraction}} \cdot \text{kick amplitude}$$

is key to determining whether the system is chaotic. See Fig 1. Our results can be summarized informally as follows:

Theorem 5 [45] *For each choice of σ, λ, A and T, the time-T map F_T has a (maximal) attracting set $\Lambda = \Lambda(\sigma, \lambda, A, T)$ with the following properties:*

(i) For $\frac{\sigma}{\lambda}A$ small, Λ is a closed invariant curve.

(ii) As $\frac{\sigma}{\lambda}A$ increases, the invariant curve breaks; the dynamics of F_T is initially of gradient type (with sinks and sources); then horseshoes start to develop.

(iii) In the case of large $\frac{\sigma}{\lambda}A$, regarding T as a parameter:

[2]After Wang and I had completed the study reported here, we learned that G. Zaslavsky, a physicist colleague of mine, had studied numerically a similar example 30 years earlier [55].

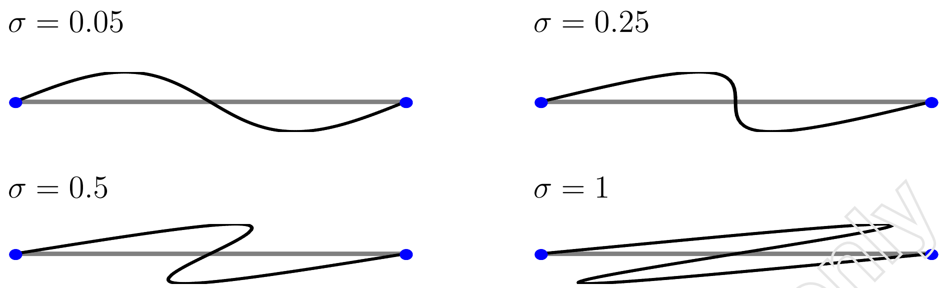

Figure 1: **Effect of increasing shear.** Images of $\Psi_\tau(\gamma)$ for $T = 10$ and $\lambda = A = 0.1$.

> *(a)* F_T *has horseshoes and sinks on an open set of* T*, and*
> *(b) provided that* $e^{-\lambda T}$ *is sufficiently small,* F_T *has an SRB measure, and* $\lambda_{\max} > 0$ *Leb-a.e. for a positive measure set of* T*.*

The ideas in this linear oscillator example are valid for general limit cycles in arbitrary dimensions, except that the effects of σ, λ and A cannot be separated. Instead, one has to interpret "shear" as the degree to which the kick "scrambles" the limit cycle, its variation measured with respect to the strong stable foliation of the unforced system. See the review article [29] for more detail.

The following is another illustration of the same mechanism at work.

Periodically forced Hopf bifurcations in ODEs and PDEs. Consider first the usual picture of a Hopf bifurcation in $2D$, for an equation $\dot{x} = F_\mu(x)$ where μ here is the bifurcation parameter. We assume $F_\mu(0) = 0$ for all μ, and that $x = 0$ undergoes a generic supercritical Hopf bifurcation at $\mu = 0$, so that a limit cycle of radius $\sim \sqrt{\mu}$ emerges as 0 destabilizes. Writing all this in normal form, we have

$$\dot{z} = k_0(\mu)z + k_1(\mu)z^2\bar{z} + k_2(\mu)z^3\bar{z}^2 + \text{h. o. t.}$$

We define the *twist* of the system to be

$$\tau \;=\; \frac{\text{Im } k_1(0)}{-\text{Re } k_1(0)} \;.$$

As we will see, this quantity is the analog of the shear in the previous example. The result below applies to dynamical systems defined by ODEs on phase spaces of any dimension $d \geq 2$ [46] as well as to systems defined by PDEs [30]. As illustration of the breadth of its applicability, I will state it for dissipative parabolic PDEs undergoing a Hopf bifurcation.

Consider, for example, a 1-parameter family of semilinear parabolic equations on a bounded domain $\Omega \subset \mathbb{R}^d$ with smooth boundary:

$$\begin{aligned} u_t &= D\Delta u + f_\mu(u), \qquad x \in \Omega,\ u \in \mathbb{R}^m, \\ u(x,t) &= 0, \qquad x \in \partial\Omega,\ t \geq 0. \end{aligned}$$

Here D is a diagonal matrix with positive entries, μ is a parameter, and $f_\mu : \mathbb{R}^d \to \mathbb{R}^m$ is a polynomial with $f_\mu(0) = 0$ for all μ. We assume, in a sense to be made precise, that the solution $u(x,t) \equiv 0$ is stable for $\mu < 0$, and that it loses its stability at $\mu = 0$. To this equation, we add a periodic forcing that is close to impulsive, i.e.,

$$u_t = D\Delta u + f_\mu(u) + \rho\varphi(x)p_T(t) \tag{6}$$

where $\rho \in \mathbb{R}$ is a constant, $\varphi : \Omega \to \mathbb{R}^m$ is a smooth function satisfying mild conditions, $p_T = \sum_{n=-\infty}^{\infty} \varepsilon^{-1} I_{[nT,nT+\varepsilon]}$ and I_A is the indicator function on A. As a *dynamical system* on $\left(H_0^1(\Omega)\right)^m$, (6) is a special case of an abstract system generated by an equation of the form

$$\dot{\mathbf{u}} = \mathcal{A}\mathbf{u} + \mathbf{f}_\mu(\mathbf{u}) + \rho\mathbf{\Phi}(\mathbf{u})p_T(t). \tag{7}$$

The result below applies to evolutionary equations of the form (7); we have motivated with a concrete PDE but the exact PDE that gives rise to it is immaterial.

Theorem 6 [30] *Let $\mathbb{H}$ be a Hilbert space. Our unforced equation has the form*

$$\dot{u} = \mathbf{A}u + \mathbf{F}(u,\mu), \qquad u \in \mathbb{H}, \quad \mu \in (-\mu_1, \mu_1) \subset \mathbb{R}, \tag{8}$$

where $-\mathbf{A}$ is a sectorial operator and $\mathbf{F} : \mathbb{H}^\sigma \times (-\mu_1, \mu_1) \to \mathbb{H}$ is C^5 for some $\sigma \in [0,1)$.[3] We assume $\mathbf{F}(0,\mu) = 0$ for all μ, and rewrite (8) to obtain

$$\dot{u} = \mathcal{A}_\mu u + f_\mu(u), \quad f_\mu(0) = 0, \quad \partial_u f_\mu(0) = 0. \tag{9}$$

This system is assumed to undergo a generic supercritical Hopf bifurcation at $\mu = 0$, with a limit cycle γ_μ emerging from $u = 0$ for $\mu > 0$.

[3] $\{\mathbb{H}^\sigma\}$ are a family of interpolating subspaces called fractional power spaces; see [16]

To (9) we add a forcing term, resulting in

$$\dot{u} = \mathcal{A}_\mu u + f_\mu(u) + \rho\Phi(u)p_T(t). \tag{10}$$

Here $\rho > 0$ is a constant, $\Phi : \mathbb{H}^\sigma \to \mathbb{H}^\sigma$ is C^5 with uniformly bounded C^5-norms, and p_T is as above. Let $\mathbb{H}^\sigma = E^c_\mu \oplus E^s_\mu$ be the decomposition into $\mathcal{A}_\mu$-invariant center and stable subspaces, E^c_μ corresponding to the two leading eigenvalues in the Hopf bifurcation. We assume $\Phi(0) \notin E^s_0$, and normalize to give $|P^c_0(\Phi(0))| = 1$ where P^c_0 is the projection of $\mathbb{H}^\sigma$ onto E^c_0.

Let F_T be the time-T map of the semiflow defined by (10). For $T >$ const $\cdot\mu^{-1}$, F_T has an attractor Λ_μ near the limit cycle γ_μ of (9), attracting all points in an open neighborhood $\mathcal{U}$ of Λ_μ in $\mathbb{H}^\sigma$. Assume further that

$$|\tau|\rho\mu^{-\frac{1}{2}} > L_0 \qquad \textit{for a certain } L_0 .$$

Then we have for each small enough $\mu > 0$ a roughly T-periodic positive measure set $\Delta_\mu \subset (M_\mu, \infty)$ for some $M_\mu \gg \mu^{-1}$ with the property that for every $T \in \Delta$,

(a) F_T has an ergodic SRB measure and

(b) $\lambda_{\max} > 0$ "almost everywhere" in $\mathcal{U}$.

The term "almost everywhere" above certainly requires justification, as $\mathcal{U}$ is an open set in an infinite dimensional function space. We discussed in Sect. 3.1 the idea of equating observable events with positive Lebesgue measure sets. There is, of course, no Lebesgue measure on Banach spaces, but one can use finite parameter families of initial conditions, and use Lebesgue as a reference measure on parameter space. In Theorem 6, "almost everywhere" refers to Lebesgue-a.e. initial condition in every 2-parameter family of initial conditions transversal to the strong stable foliation in $\mathcal{U}$. See [27, 30, 6] for more information.

3.3 A theory of rank-one attractors

The $\Lambda_{\max} > 0$ results in Sect. 3.2 were obtained by appealing to a general theorem on "rank-one attractors", the name my co-author Qiudong Wang and I coined for a class of attractors that we introduced and studied [44, 47, 48]. These attractors can live in phase spaces of any dimension. They are so called because they have only one direction of instability, with (strong) contraction in all other directions. Rank-one attractors are, in many ways, the

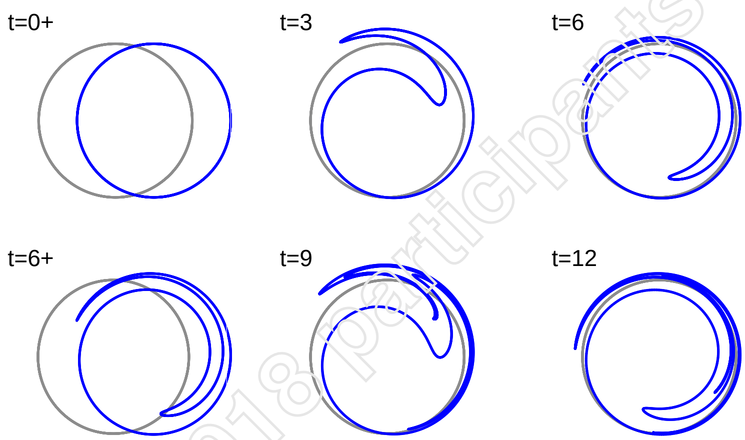

Figure 2: **Hopf attractor.** Grey circle is the limit cycle γ_μ of the unforced Hopf bifurcation. At time $t = 0$, the system receives a "kick", sending γ_μ to the blue circle. Between $t = 0$ and $t = 6$, the unforced flow brings the blue circle black to the grey, rotating counterclockwise with points farther from the center of the grey circle rotating at a higher speed due to a nonzero twist as explained in the text. At $t = 6$, the image of γ_μ has the shape shown, due to the differential in rotational speeds. The kick is repeated once every 6 units of time, i.e., $T = 6$. Here $\rho\mu^{-\frac{1}{2}} = \frac{1}{2}$ and $\tau = 10$.

simplest and least chaotic among chaotic attractors. They occur naturally, often following the loss of stability.

Of interest to us here are those rank-one attractors that possess SRB measures and exhibit observable chaos. The existence and abundance of such attractors is guaranteed by the theorem below. The precise statement of this result is technical, and since this article is not the right forum for technical details, I will suppress some of them, referring the reader to [47]. (A 2D version of this result was first proved in [44], but [47] is both more general and more readable.)

Theorem 7 [44, 47] (Informal version) *Let $M = I \times D_m$ where I is either an interval or a circle and D_m is an m-dimensional disk, $m \geq 1$. For each $\varepsilon > 0$, let $F_{a,\varepsilon} : M \circlearrowleft$ be a C^3 family of embeddings with $|\det(DF_{a,\varepsilon})| \sim \varepsilon^m$. Assume*

(a) as $\varepsilon \to 0$, $F_{a,\varepsilon} \to F_{a,0}$ in C^3 where $F_{a,0}$ is a family of maps from $M \to I \times \{0\}$;

(b) letting $f_a = F_{a,0}|_{I\times\{0\}}$, we obtain a family of 1D maps with
(i) nondegenerate critical points and
(ii) sufficiently strong expansion away from critical sets;

(c) the mappings $F_{a,0}$ satisfy certain nondegeneracy and transversality conditions.

Then for each sufficiently small $\varepsilon > 0$, there exists a positive measure set Δ_ε such that for all $a \in \Delta_\varepsilon$:

(i) $F_{a,\varepsilon}$ has an ergodic SRB measure;

(ii) $\lambda_{\max} > 0$ Leb-a.e. on M.

This is a perturbative result. The idea is to embed the systems of interest in a family that can be passed, in a meaningful way, to a *singular limit*, which is an object of a lower dimension (in this case 1D), the idea being that lower-dimensional objects are more tractable. The tricky part is to "unfold" the 1D results at $\varepsilon = 0$ to recover the dynamical picture for small $\varepsilon > 0$.

Theorem 7 is a generalization of the work of Benedicks and Carleson [3] and Benedicks and myself [4] on the Hénon maps; [3] and [4] can be seen as an extension of Jakobson's theorem [18] in 1D. Theorem 7 extends the core ideas in [3, 4], permitting the attractor to be embedded in a phase space of arbitrary dimension and replacing the formula-based arguments in [3] by geometric conditions that imply the existence of chaotic rank-one attractors with SRB measures.

The proof of Theorem 7 is technically quite involved, as it required a delicate parameter selection. This proof together with the study of dynamical properties in [48] for the "good maps" $F_{a,\varepsilon}$ in Theorem 7 occupy well over 150 pages. To avoid having to repeat a parallel analysis every time a similar situation is encountered, we have formulated Theorem 7 in such a way that the conclusion of SRB measure and $\lambda_{\max} > 0$ holds once certain conditions are met. The results in Sect. 3.2 were proved by checking these conditions. Another application was to slow-fast systems [14].

To summarize, Theorem 7 provides a general framework for producing rank-one attractors with SRB measures and $\lambda_{\max} > 0$, and all currently known examples of nonuniformly hyperbolic attractors with observable chaos belong in this class.

4 Random dynamical systems

Most realistic systems are governed by laws that are neither purely deterministic nor purely stochastic but a combination of the two. Noise terms are routinely added to differential equations to model uncontrolled fluctuations or forces not accounted for. Now it is known that solutions of stochastic differential equations have representations as *stochastic flows of diffeomorphisms*, i.e., for each ω corresponding to a realization of Brownian path, there is a 1-parameter family of diffeomorphisms $x \mapsto \varphi_t(x;\omega)$ satisfying $\varphi_{s+t}(x;\omega) = \varphi_t(\varphi_s(x;\omega);\sigma_s(\omega))$ where σ_s is time-shift along the path. See e.g. [20]. Thus systems modeled by SDEs can be seen as *i.i.d.* sequences of *random maps*, and as such, they have been studied a fair amount.

In Sect. 4.1, I will discuss two sets of results, both illustrating the fact that the averaging effect of randomness makes deterministic systems nicer and more tractable. An application of random dynamical systems is discussed in Sect. 4.2.

4.1 Extensions of deterministic theory to random maps

Consider first a *random maps system* $\mathcal{X}$ defined as follows. Fix a probability ν on Ω, the space of diffeomorphisms of a compact manifold M. We consider the composition

$$\cdots \circ f_n \circ \cdots \circ f_2 \circ f_1 \ , \quad n = 1, 2, \dots \ ,$$

where $f_1, f_2, \dots$ are chosen independently with law ν. Such a sequence defines a Markov chain on M, its transition probabilities being given by $P(A|x) = \nu\{f \in \Omega : f(x) \in A\}$ for Borel subsets $A \subset M$. A probability measure μ on M is called *stationary* for $\mathcal{X}$ if $\mu(A) = \int P(A|x) d\mu(x)$. Given a stationary measure μ, under the usual integrability conditions, Lyapunov exponents $\{\lambda_i\}$ are defined μ-a.e. for $\nu^{\mathbb{N}}$-almost every sequence $\underline{\mathbf{f}}^+ = (f_i)_{i=1}^\infty$ and are nonrandom. The (pathwise) entropy h_μ for $\mathcal{X}$ is as defined in the case of nonrandom maps, replacing the iteration of a single map by compositions of $(f_i)_{i=1}^\infty$; it is also nonrandom. See [19] for more information.

While μ is invariant under the Markov chain, i.e., when averaged over all random maps, there is the following pathwise notion of invariant measure: First we extend the sequence of maps $\underline{\mathbf{f}}^+ = (f_i)_{i=1}^\infty$to a bi-infinite sequence $\underline{\mathbf{f}} = (f_i)_{i=-\infty}^\infty$, chosen *i.i.d.* Viewing the system as having started from time $-\infty$, one obtains *sample measures* $\{\mu_{\underline{\mathbf{f}}}^-\}$ defined for a.e. $\underline{\mathbf{f}} = (f_i)_{i=-\infty}^\infty$ by conditioning μ on "the past". That is to say, $\mu_{\underline{\mathbf{f}}}^-$ describes the distribution at time 0 given that the maps $(f_i)_{i<0}$ have been applied. Equivalently,

$$\mu_{\underline{\mathbf{f}}}^- = \lim_{n\to\infty} (f_{-1} \circ f_{-2} \circ \cdots \circ f_{-n})_* \mu \ ;$$

the limit exists by martingale convergence. It is easy to see that $\mu_{\underline{\mathbf{f}}}^-$ depends only on $(f_i)_{i<0}$, that $\int \mu_{\underline{\mathbf{f}}}^- d\nu^{\mathbb{Z}}(\underline{\mathbf{f}}) = \mu$, and that $\mu_{\underline{\mathbf{f}}}^-$ is invariant in the sense that $(f_0)_* \mu_{\underline{\mathbf{f}}}^- = \mu_{\sigma\underline{\mathbf{f}}}^-$ where $\sigma\underline{\mathbf{f}}$ is the shifted sequence. See e.g. [19, 25].

Now for a bi-infinite sequence $\underline{\mathbf{f}} = (f_i)_{i=-\infty}^\infty$, unstable manifolds at time 0 depend also only on the past (while stable manifolds depend on the future). It therefore makes sense, given a random maps system $\mathcal{X}$, to define $\mu_{\underline{\mathbf{f}}}^-$ to be a *random SRB measure* if $\lambda_{\max} > 0$ $\mu_{\underline{\mathbf{f}}}^-$-a.e. and the conditional measures of $\mu_{\underline{\mathbf{f}}}^-$ on unstable manifolds have densities, following the definition in the deterministic case. The next result is roughly parallel to Theorem 2, to which we refer the reader for notation.

Theorem 8 *Let $\mathcal{X}$ be as above with an ergodic stationary measure μ. We assume that*

$$\int \log^+ \|f\|_{C^2} \nu(df), \quad \int \log^+ \|f^{-1}\|_{C^2} \nu(df) \ < \ \infty \ .$$

(1a) [21] *If $\lambda_{\max} < 0$, then $\mu^-_{\underline{\mathbf{f}}}$ is supported on a finite set of points for $\nu^{\mathbb{Z}}$-a.e. $\underline{\mathbf{f}}$.*

(1b) [25] *If μ has a density and $\lambda_{\max} > 0$, then*

$$h_\mu = \sum \lambda_i^+ m_i$$

and $\mu^-_{\underline{\mathbf{f}}}$ is a random SRB measure for $\nu^{\mathbb{Z}}$-a.e. $\underline{\mathbf{f}}$.

(2) [24] *Assume in addition to $\lambda_{\max} > 0$ that the backward derivative process associated with the Markov chain on the Grassmannian bundle of M has absolutely continuous transition probability kernels (see* [24]*). Then there is an i_0 such that the partial dimensions δ_i (see Sect. 1.2) satisfy*

$$\delta_i = 1 \quad \text{for all } i < i_0, \qquad \delta_i = 0 \quad \text{for all } i > i_0 \ .$$

Moreover, $\dim(\mu^-_{\underline{\mathbf{f}}}) = \sum_i \delta_i m_i$.

The condition that μ has a density is very natural for random maps; a sufficient (but not necessary) condition is that the transition probabilities $P(\cdot|x)$ have densities. Item (1) in Theorem 8 says that except when $\lambda_{\max} = 0$, $\mathcal{X}$ either has *random sinks*, i.e., almost all solutions coalesce in time into at most a finite number of (evolving) trajectories, or they have *random attractors* with *random SRB measures*, i.e., attracting sets having the geometric characteristics of attractors with SRB measures in the deterministic case – except that these attractors too evolve with time. Item (2) says that when randomly perturbed, sample measures align with the most expanding directions. With the configuration of δ_i in Theorem 8, the quantity $\sum_i \delta_i m_i$ is in fact another way to write the Kaplan-Yorke dimension [13]. Thus Theorem 8(2) proves that the Kaplan-Yorke conjecture holds for random maps.

Positivity of Lyapunov exponents via random perturbations. For deterministic maps, we discussed at the end of Sect. 3.1 the challenges in proving the existence of SRB measures. Similar challenges exist for proving $\lambda_{\max} > 0$ for volume-preserving maps. The *standard maps* family, a 1-parameter family of area-preserving maps $f_L : \mathbb{T}^2 \circlearrowleft$ with the property that $\|Df_L\| \sim L$ for $L > 1$, symbolizes the enormity of the challenge: In spite of considerable effort by leading researchers, no one has been able to prove – or disprove – the positivity of Lyapunov exponents for f_L for any L, however large.

The following result illustrates what the addition of random noise can do. Two points are of note: One is that the amount of noise needed is tiny; the other is that unlike Theorem 7, no parameter selection is needed.

Theorem 9 [5] *Let* $f_{L,a} : \mathbb{T}^2 \circlearrowleft$ *be given by*

$$f_{L,a}(x, y) = (L \sin(2\pi x) + a - y,\ x) \ .$$

To $f_{L,a}$*, we add a random perturbation of the form* $(x, y) \mapsto (x + \xi, y)$*,* $\xi \in [-\varepsilon, \varepsilon]$ *uniformly distributed. Then given any* $\alpha, \beta > 0$*, there exists* $C > 0$ *such that for all* L *sufficiently large, for all* a*, and for all* $\varepsilon \geq L^{-CL^{1-\beta}}$*, we have*

$$\lambda_{\max} > (1 - \alpha) \log L \qquad \textit{Lebesgue-a.e. on } \mathbb{T}^2 \ .$$

4.2 Interpretation as reliability of driven systems

Consider a continuous-time dynamical system defined on a manifold M. A signal $I(t) \in \mathbb{R}^n, t \in [0, \infty)$, is presented to the system at time 0. Think of it as an external input being switched on in an engineered system, or the onset of a stimulus in a biological system. The response of the system at time $t > 0$ is given by $F(t) = F(x_0, \{I(s)\}_{0 \leq s < t}; t)$, where $x_0 \in M$ is the internal state of the system at time 0. A system is called *reliable* with respect to a class of signals $\mathcal{I}$ if for almost all $I \in \mathcal{I}$, the dependence of $F(t)$ on x_0 vanishes with time. Initially, some dependence of $F(t)$ on x_0 is unavoidable. The idea is that a reliable system will, after a transient, entrain to the signal $I(t)$ and lose memory of its own initial state; see e.g. [28].

In the simple setting where $I(t)$ are (frozen) realizations of white noise, this setup can be described by a stochastic differential equation. For a typical sample Brownian path ω and $t_1 < t_2$, flow-maps $F_{t_1,t_2,\omega}$ from time t_1 to t_2 are well defined under mild regularity assumptions on the coefficients of the SDE [20]. This puts us in the setting of random maps. Let μ be the invariant probability measure, which we assume to be unique. As the sample measures μ_ω^- are given by $\mu_\omega^- = \lim_{t\to\infty}(F_{-t,0,\omega})_*\mu$ (see Sect. 4.1), we may assume that for large enough t, the observed distribution of x_t starting from μ at time 0 is approximately $\mu^-_{\sigma_t(\omega)}$ where σ_t is the time shift.

The interpretation therefore is as follows: In the case $\lambda_{\max} < 0$, μ_ω^- is a finite set for a.e. ω by Theorem 8(1)(a). Under suitable conditions, it consists of a single point. That is to say, the approximate location of x_t is largely independent of x_0 for large t, the definition of a reliable system. If

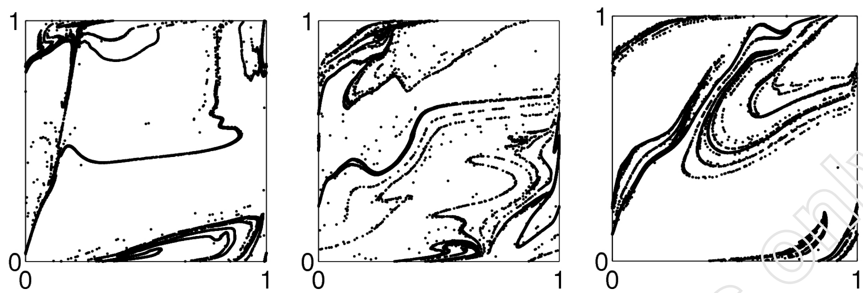

Figure 3: **Random attractors.** Snapshopts of distributions of x_t starting from μ for two coupled phase oscillators driven by a white-noise stimulus. These distributions approximate the sample measures $\mu^-_{\sigma_t(\omega)}$. The phase space is the 2-torus, and the system is unreliable. The curves seen are unstable manifolds of random strange attractors on which SRB measures are supported. The attractors evolve perpetually with time, retaining certain basic characteristics throughout.

$\lambda_{\max} > 0$, then by Theorem 8(1)(b), almost surely μ^-_ω has the characteristic geometry of a random SRB measure. That means x_t may be very different depending on x_0, and this dependence on x_0 will persist for all $t > 0$. The system is unreliable, and the effective dimension of the random SRB measure given by Theorem 8(2) describes the extent of its unreliability.

5 Applications to biology

In the last 5-10 years, I have taken an interest in the application of Dynamical Systems ideas to the biological sciences. I would like to report here on work in two different directions: Sect. 5.1 discusses a study on an idealized model of epidemics control, while Sect. 5.2 contains a glimpse into some work in computational neuroscience.

5.1 Control of epidemics via isolation of infected hosts

This work is on a simple model of infectious diseases. When an outbreak is unforeseen, the only available means of containment is the isolation of infected individuals. Isolation is very effective when implemented immediately and in full; one simply cuts off all contact between infected hosts and the rest of the population. But such perfect implementation is not feasible in reality:

facilities to house the infected, and medical personnel to care for them have to be available at a moment's notice, and infected hosts have to be identified as soon as they become infectious.

Below I describe the results of a theoretical study [50] the goals of which are to quantify minimum response capabilities needed to squash incipient outbreaks, and to predict the consequences when containment fails. It is an idealized model built on the well known susceptible-infected-susceptible (SIS) model of epidemics.

Consider, to begin with, a random network of N nodes; each node represents a host, and nodes that are linked by edges are neighbors. The mean degree at a node is given by $\langle k \rangle$, so the density of connections is denoted by $m = \langle k \rangle / N$. In the absence of any control mechanism, the situation is as follows. Each host is in one of three discrete states: healthy and susceptible (S), infectious (I) and incubating (E). Infectious hosts infect their neighbors at rate β; infected nodes incubate the disease for a period of σ units of time during which they are assumed to be neither symptomatic nor infectious. At the end of the incubation period they become infectious. Infectious nodes remain in that state until they recover and rejoin the susceptible group. The rate of recovery is γ. This, roughly speaking, is the SIS model.

To the setting above, we introduce the following isolation protocol: If a host remains infectious for τ units of time without having recovered, it enters a new state, Q (for *quarantine*), with probability p. The hosts that do not enter state Q at time τ remain infectious until they recover on their own. A host that enters state Q remains in this state for κ units of time, at the end of which it is discharged and rejoins the healthy and susceptible pool.

The quantities of interest in this model are $S(t)$, $E(t)$, $I(t)$, and $Q(t)$, representing the fractions of the population in the *susceptible*, *incubating*, *infectious* and *quarantine* states respectively at time t starting from some initial condition, and the new parameters of our model are $\tau, \kappa > 0$ and $p \in [0, 1]$: p is the probability that an infected individual will be isolated, τ is the time between becoming infectious and entering isolation, and κ is the duration of isolation.[4]

Assuming that links between the infectious and susceptible nodes are uncorrelated, we obtain, by moment closure, the following system of Delay

[4]We have not built into the model the idea of *immunity*, which is of course very relevant in the long run, but less so for shorter time scales, such as the time evolution following a single outbreak.

Differential Equations in the continuum limit as $N \to \infty$:

$$\begin{aligned}
\dot S(t) &= -\beta m S(t)I(t) + \gamma I(t) + \beta m \varepsilon S(t-\sigma-\tau-\kappa)I(t-\sigma-\tau-\kappa), \\
\dot E(t) &= \beta m[S(t)I(t) - S(t-\sigma)I(t-\sigma)] \\
\dot I(t) &= \beta m S(t-\sigma)\, I(t-\sigma) - \gamma I(t) - \beta m \varepsilon S(t-\sigma-\tau)\, I(t-\sigma-\tau)\,, \\
\dot Q(t) &= \beta m \varepsilon[S(t-\sigma-\tau)I(t-\sigma-\tau) - S(t-\sigma-\tau-\kappa)I(t-\sigma-\tau-\kappa)]\,,
\end{aligned}$$

where β and m are as above and $\varepsilon = pe^{-\gamma\tau}$.

Let $C := C\left([-\sigma-\tau-\kappa, 0], \mathbb{R}^4\right)$, the Banach space of continuous functions. Given an initial function $\phi \in C$, the solution $x(t,\phi) \in \mathbb{R}^4$, $t \geq 0$, to the initial value problem exists and is unique. Standard results imply, in fact, that the system above defines a C^1 semi-flow on C with the sup norm. Observe that by the conservation of mass property, if $\phi = (\phi_S, \phi_E, \phi_I, \phi_Q)$ and $x(t;\phi) = (S(t), E(t), I(t), Q(t))$, then $S(t)+E(t)+I(t)+Q(t) = \phi_S(0) + \phi_E(0) + \phi_I(0) + \phi_Q(0)$ for all $t \geq 0$. Therefore, the 3-D hyperplane $H^3 := \{S(t)+E(t)+I(t)+Q(t) = 1\} \subset \mathbb{R}^4$ is left invariant by the semi-flow. To obtain biologically relevant solutions we further restrict to the subset of H^3 on which $S(t), E(t), I(t), Q(t) \geq 0$.

In the SIS model, what determines whether the disease will propagate is the *disease reproductive number* $r := \beta m/\gamma$: that is, without an isolation protocol, a small outbreak is contained if $r < 1$, and it spreads if $r > 1$.

Theorem 10 [50] *(1) Given p, τ and κ, the* effective disease reproductive number

$$r_e = r(1-\varepsilon) = r(1 - pe^{-\gamma\tau})\ .$$

Starting from an initial condition $\phi = (\phi_S, \phi_E, \phi_I, \phi_Q)$ with $\phi_I > 0$ near the disease-free equilibrium $(1,0,0,0)$, the following response is needed to ensure $r_e < 1$:

(a) there is a minimum isolation probability $p_c = 1 - \frac{1}{r}$ *so that to contain the outbreak, one must have $p > p_c$; and*

(b) for $p > p_c$, there is a critical identification time

$$\tau_c(p) = \frac{1}{\gamma}\log\frac{p}{p_c}\ ,$$

so that for $\tau < \tau_c(p)$ the disease dies out and for $\tau > \tau_c(p)$ it spreads.

(2) Let p, τ, κ and ϕ be as above. If $r_e > 1$, then there is at most one possible endemic equilibrium, given by constant functions $(S_{\text{eq}}, E_{\text{eq}}, I_{\text{eq}}, Q_{\text{eq}})$ with

$$S_{\text{eq}} = \frac{1}{r_e} = \frac{1}{r(1-\varepsilon)} \qquad \textit{and} \qquad I_{\text{eq}} = \frac{(1-\varepsilon)}{\sigma\gamma + \kappa\gamma\varepsilon + (1-\varepsilon)} \cdot (1 - S_{\text{eq}}).$$

We conjecture that when $r_e > 1$, all solutions starting from ϕ in the theorem converge to the predicted endemic equilibrium. A proof is out of reach for now, but numerical evidence is strongly in favor of the conjecture for reasonable parameters.

5.2 Dynamics of the brain

First, why the brain? *The brain is a dynamical system*, a large and complex dynamical system, one of the most fascinating naturally occurring systems I have ever seen. It is made up of a large number of smaller subsystems, namely neurons, coupled together in a hierarchical network. Single neurons are themselves complicated biological entities. While one must eventually combine information across scales, it is simpler, to begin with, to first focus on either network level activity or cellular/subcellular properties of individual neurons. This section is about the former, that is to say, I will suppress detailed properties of individual neurons and focus on their dynamical interactions.

A crash course on neuronal dynamics. Neurons communicate with one another by electrical signals which they produce when they spike. When an Excitatory (E) neuron spikes, it brings the recipients of its signal, called its postsynaptic neurons, a little closer to their own spiking thresholds. Inhibitory (I) neurons do the opposite to their postsynaptic neurons.

Neglecting biochemical processes and keeping track of electrical properties only, the dynamics of a neuron are governed by its membrane potential and are described by the standard Leaky Integrate-and-Fire (LIF) equation

$$\frac{dv}{dt} = -g_R v - g_E(t)(v - V_E) - g_I(t)(v - V_I) \ . \tag{11}$$

Here v is membrane potential in normalized units: Without external input, v tends to 0 at rate g_R (a constant). The Excitatory conductance $g_E(t)$ of a neuron is elevated for a brief period of time ($10 - 20$ milliseconds) when the neuron receives excitatory input, such as a spike from a presynaptic E-neuron. The term $g_E(t)(v - V_E)$ is called the Excitatory current; it drives the membrane potential towards the value V_E ($= 14/3$ in the present normalization), called the excitatory reversal potential. But $v(t)$ never gets that high: by the time it reaches 1, the neuron spikes and v is reset to 0. The Inhibitory conductance $g_I(t)$ is elevated similarly when the neuron receives inhibitory input as a result of the spiking of a presynaptic I-neuron; this current drives the membrane potential towards $V_I = -2/3$.

To summarize, the membrane potential of a neuron swings up and down depending on the input it receives. When it reaches its spiking threshold 1, the neuron spikes, sending signals to other neurons thereby affecting their time evolutions, and its own membrane potential is reset to 0.

Network architecture, or wiring, is of course important. The primate cerebral cortex (which is what I know best) is organized into regions and subregions and layers, with neurons in local populations having similar preferences and functional roles. Within local circuitries, E- and I-neurons are relatively homogeneously though sparsely connected, whereas connections between distinct regions or layers are specific and tend to be Excitatory only.

Computational modeling of the visual cortex. Turning now to my own modeling work in Neuroscience, I have worked mostly with the visual cortex, the part of the brain responsible for the processing of visual information. The modeling I do is heavily data-driven, benchmarked by dozens if not hundreds of sets of experimental data from monkey, whose visual cortex is quite similar to our own. From point-to-point representation of visual images in the retina, cortex extracts, through multiple stages of processing, information such as edges, shapes, color, movement, and texture — features that enable the brain to make sense of visual scenes. This complex task is accomplished through the dynamical interaction of neurons. My immediate goal in this research is to unravel the dynamical processes responsible for feature extraction. My larger goals are to connect dynamical events on the neuronal level to cortical functions and ultimately to human perception and behavior.

Working with a small team of neuroscientists and postdocs, I have been involved in building a comprehensive model of the primary visual cortex, or V1. This is the area of cortex closest to sensory input. It is also the largest and most complex of all the visual cortical areas. We have worked mostly on the input layer of the magnocellular pathway, one of the two main pathways in V1. *Our challenge is to deduce underlying dynamical mechanisms from experimental data documenting how V1 responded to various stimuli,* i.e., to reverse-engineer the dynamical system from its outputs in response to stimuli.

This is less straightforward than one might think, because the brain does not simply reproduce a copy, pixel by pixel, of the stimulus presented. It extracts data, dissects and recombines, suppresses and enhances; it *processes* the information. Below I will illustrate this point with an example; the material is taken in part from [9].

To get a sense of size, think of the moon as projecting to a region of diameter $\sim 1/2$ degree in our retina, near the center of one's visual field. A somewhat surprising fact is that to cover a region that projects to $1/4$ deg $\times 1/4$ deg of the retina, there are, in the magno-pathway, only about 10 cells in the Lateral Geniculate Nucleus (LGN), the body that relays information from the retina to V1. Of these 10 cells in the LGN, about 5 are ON, and 5 are OFF. The ON-cells spike vigorously when luminance in its visual field goes from dark to light, and the OFF cells do the same when it goes from light to dark. All that is simple enough, an array of cells reporting the changes in luminance in their visual fields.

Now LGN neurons project directly to V1, and one of the most salient features of V1 is that its cells are *orientation selective*: most cells have a preference for edges of a particular orientation, and when such an edge passes through its receptive field, the cell gets excited and spikes. Orientation selectivity (OS) is a very important property; it helps us detect contours in the visual scene. The question is, how do V1 neurons acquire their OS, since LGN neurons have no such selectivity?

Half a century ago, Hubel and Wiesel, who eventually won the Nobel Prize, proposed an explanation for OS in terms of the alignment of LGN cells that converge on V1 cells [17]. Based on currently available data, our model confirmed that this idea is very likely the source of OS, but contrary to conventional wisdom, we found that the signal from LGN is weak — an idea that is only slowly gaining traction in the neuroscience community. In our model, the feedforward signal is substantially amplified (and mollified) through dynamical interactions among cortical neurons, which after all provide the bulk of the current received by neurons in V1. See Fig 4.

The modeling work above is computation based. Realistic models as complex as ours are not amenable to rigorous analysis at the present time, yet much of our model building has been guided by Dynamical Systems ideas, the only difference being that simulations played the role of proofs in the validation of conjectures.

6 Looking forward

Dynamical Systems was born a little over 100 years ago, inspired by curiosity about natural phenomena. The field has blossomed and matured; we have developed a rich collection of ideas and techniques that we can proudly call our own. Advances within the field have suggested new avenues of research and new questions, but it is also important to reconnect with original goals, and to build new connections.

Building connections will require developing new techniques and adopting new viewpoints. Let me give just one example, drawing from my own limited experience: Most naturally occurring dynamical systems are *large* – large in the sense of extended phase spaces, many degrees of freedom, a more complex dynamical landscape. The concerns are very different when studying "small" and "large" dynamical systems. For small systems, complexity means chaotic behavior, positive Lyapunov exponents, positive entropy. In "large" systems, it is not clear what chaotic behavior means. The focus is more naturally on *emergent* phenomena, behaviors that cannot be predicted from local rules and that emerge as a result of interaction among components. To gain insight, one looks for ways to reduce system complexity by identifying parameters, or modes of behavior, that are more important than others.

Since Dynamical Systems is naturally connected to many parts of mathematics and many scientific disciplines, I see a future full of possibilities, and I encourage my fellow dynamicists to embrace the challenge. Not only is Dynamical Systems a study of moving objects and evolving situations, the subject itself is also evolving. It *must* evolve, if it is to remain interesting, vibrant and relevant.

References

[1] Viviane Baladi, Mark F. Demers, and Carlangelo Liverani. Exponential decay of correlations for finite horizon Sinai billiard flows. *Invent. Math.*, 211(1):39–177, 2018.

[2] Luis Barreira, Yakov Pesin, and Jörg Schmeling. Dimension and product structure of hyperbolic measures. *Ann. of Math. (2)*, 149(3):755–783, 1999.

[3] Michael Benedicks and Lennart Carleson. The dynamics of the Hénon map. *Ann. of Math. (2)*, 133(1):73–169, 1991.

[4] Michael Benedicks and Lai-Sang Young. Sinaĭ-Bowen-Ruelle measures for certain Hénon maps. *Invent. Math.*, 112(3):541–576, 1993.

[5] Alex Blumenthal, Jinxin Xue, and Lai-Sang Young. Lyapunov exponents for random perturbations of some area-preserving maps including the standard map. *Ann. of Math. (2)*, 185(1):285–310, 2017.

[6] Alex Blumenthal and Lai-Sang Young. Absolute continuity of stable foliations for mappings of Banach spaces. *Comm. Math. Phys.*, 354(2):591–619, 2017.

[7] Rufus Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*. Lecture Notes in Mathematics, Vol. 470. Springer-Verlag, Berlin-New York, 1975.

[8] L. A. Bunimovich, Ya. G. Sinaĭ, and N. I. Chernov. Statistical properties of two-dimensional hyperbolic billiards. *Uspekhi Mat. Nauk*, 46(4(280)):43–92, 192, 1991.

[9] Logan Chariker, Robert Shapley, and Lai-Sang Young. Orientation selectivity from very sparse lgn inputs in a comprehensive model of macaque v1 cortex. *Journal of Neuroscience*, 36(49):12368–12384, 2016.

[10] N. Chernov. Advanced statistical properties of dispersing billiards. *J. Stat. Phys.*, 122(6):1061–1094, 2006.

[11] N. Chernov and L. S. Young. Decay of correlations for Lorentz gases and hard balls. In *Hard ball systems and the Lorentz gas*, volume 101 of *Encyclopaedia Math. Sci.*, pages 89–120. Springer, Berlin, 2000.

[12] J.-P. Eckmann and D. Ruelle. Ergodic theory of chaos and strange attractors. *Rev. Modern Phys.*, 57(3, part 1):617–656, 1985.

[13] Paul Frederickson, James L. Kaplan, Ellen D. Yorke, and James A. Yorke. The Liapunov dimension of strange attractors. *J. Differential Equations*, 49(2):185–207, 1983.

[14] John Guckenheimer, Martin Wechselberger, and Lai-Sang Young. Chaotic attractors of relaxation oscillators. *Nonlinearity*, 19(3):701–720, 2006.

[15] Radu Haiduc. Horseshoes in the forced van der Pol system. *Nonlinearity*, 22(1):213–237, 2009.

[16] Daniel Henry. *Geometric theory of semilinear parabolic equations*, volume 840 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-New York, 1981.

[17] David H Hubel and Torsten N Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, 160(1):106–154, 1962.

[18] M. V. Jakobson. Absolutely continuous invariant measures for one-parameter families of one-dimensional maps. *Comm. Math. Phys.*, 81(1):39–88, 1981.

[19] Yuri Kifer. *Ergodic theory of random transformations*, volume 10 of *Progress in Probability and Statistics*. Birkhäuser Boston, Inc., Boston, MA, 1986.

[20] Hiroshi Kunita. *Stochastic flows and stochastic differential equations*, volume 24 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1990.

[21] Y. Le Jan. équilibre statistique pour les produits de difféomorphismes aléatoires indépendants. *Ann. Inst. H. Poincaré Probab. Statist.*, 23(1):111–120, 1987.

[22] F. Ledrappier and L.-S. Young. The metric entropy of diffeomorphisms. II. Relations between entropy, exponents and dimension. *Ann. of Math. (2)*, 122(3):540–574, 1985.

[23] F. Ledrappier and L.-S. Young. The metric entropy of diffeomorphisms. II. Relations between entropy, exponents and dimension. *Ann. of Math. (2)*, 122(3):540–574, 1985.

[24] F. Ledrappier and L.-S. Young. Dimension formula for random transformations. *Comm. Math. Phys.*, 117(4):529–548, 1988.

[25] F. Ledrappier and L.-S. Young. Entropy formula for random transformations. *Probab. Theory Related Fields*, 80(2):217–240, 1988.

[26] François Ledrappier and Jean-Marie Strelcyn. A proof of the estimation from below in Pesin's entropy formula. *Ergodic Theory Dynam. Systems*, 2(2):203–219 (1983), 1982.

[27] Zeng Lian, Lai-Sang Young, and Chongchun Zeng. Absolute continuity of stable foliations for systems on Banach spaces. *J. Differential Equations*, 254(1):283–308, 2013.

[28] Kevin K. Lin, Eric Shea-Brown, and Lai-Sang Young. Reliability of coupled oscillators. *J. Nonlinear Sci.*, 19(5):497–545, 2009.

[29] Kevin K. Lin and Lai-Sang Young. Dynamics of periodically kicked oscillators. *J. Fixed Point Theory Appl.*, 7(2):291–312, 2010.

[30] Kening Lu, Qiudong Wang, and Lai-Sang Young. Strange attractors for periodically forced parabolic equations. *Mem. Amer. Math. Soc.*, 224(1054):vi+85, 2013.

[31] Roberto Markarian. Billiards with polynomial decay of correlations. *Ergodic Theory Dynam. Systems*, 24(1):177–197, 2004.

[32] Sean Meyn and Richard L. Tweedie. *Markov chains and stochastic stability*. Cambridge University Press, Cambridge, second edition, 2009. With a prologue by Peter W. Glynn.

[33] V. I. Oseledec. A multiplicative ergodic theorem. Characteristic Ljapunov, exponents of dynamical systems. *Trudy Moskov. Mat. Obšč.*, 19:179–210, 1968.

[34] Ja. B. Pesin. Characteristic Ljapunov exponents, and smooth ergodic theory. *Uspehi Mat. Nauk*, 32(4 (196)):55–112, 287, 1977.

[35] Charles Pugh and Michael Shub. Ergodic attractors. *Trans. Amer. Math. Soc.*, 312(1):1–54, 1989.

[36] Luc Rey-Bellet and Lai-Sang Young. Large deviations in non-uniformly hyperbolic dynamical systems. *Ergodic Theory Dynam. Systems*, 28(2):587–612, 2008.

[37] David Ruelle. A measure associated with axiom-A attractors. *Amer. J. Math.*, 98(3):619–654, 1976.

[38] David Ruelle. An inequality for the entropy of differentiable maps. *Bol. Soc. Brasil. Mat.*, 9(1):83–87, 1978.

[39] David Ruelle. *Thermodynamic formalism*, volume 5 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Co., Reading, Mass., 1978. The mathematical structures of classical equilibrium statistical mechanics, With a foreword by Giovanni Gallavotti and Gian-Carlo Rota.

[40] Ja. G. Sinaĭ. Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards. *Uspehi Mat. Nauk*, 25(2 (152)):141–192, 1970.

[41] Ja. G. Sinaĭ. Gibbs measures in ergodic theory. *Uspehi Mat. Nauk*, 27(4(166)):21–64, 1972.

[42] S. Smale. Differentiable dynamical systems. *Bull. Amer. Math. Soc.*, 73:747–817, 1967.

[43] Peter Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982.

[44] Qiudong Wang and Lai-Sang Young. Strange attractors with one direction of instability. *Comm. Math. Phys.*, 218(1):1–97, 2001.

[45] Qiudong Wang and Lai-Sang Young. From invariant curves to strange attractors. *Comm. Math. Phys.*, 225(2):275–304, 2002.

[46] Qiudong Wang and Lai-Sang Young. Strange attractors in periodically-kicked limit cycles and Hopf bifurcations. *Comm. Math. Phys.*, 240(3):509–529, 2003.

[47] Qiudong Wang and Lai-Sang Young. Toward a theory of rank one attractors. *Ann. of Math. (2)*, 167(2):349–480, 2008.

[48] Qiudong Wang and Lai-Sang Young. Dynamical profile of a class of rank-one attractors. *Ergodic Theory Dynam. Systems*, 33(4):1221–1264, 2013.

[49] L.-S. Young. Decay of correlations for certain quadratic maps. *Comm. Math. Phys.*, 146(1):123–138, 1992.

[50] L.-S. Young, S. Ruschel, S. Yanchuk, and T. Pereira. Siq: control of infectious diseases via isolation. 2017.

[51] Lai Sang Young. Dimension, entropy and Lyapunov exponents. *Ergodic Theory Dynamical Systems*, 2(1):109–124, 1982.

[52] Lai-Sang Young. Bowen-Ruelle measures for certain piecewise hyperbolic maps. *Trans. Amer. Math. Soc.*, 287(1):41–48, 1985.

[53] Lai-Sang Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math. (2)*, 147(3):585–650, 1998.

[54] Lai-Sang Young. Recurrence times and rates of mixing. *Israel J. Math.*, 110:153–188, 1999.

[55] G. M. Zaslavsky. The simplest case of a strange attractor. *Phys. Lett. A*, 69(3):145–147, 1978/79.

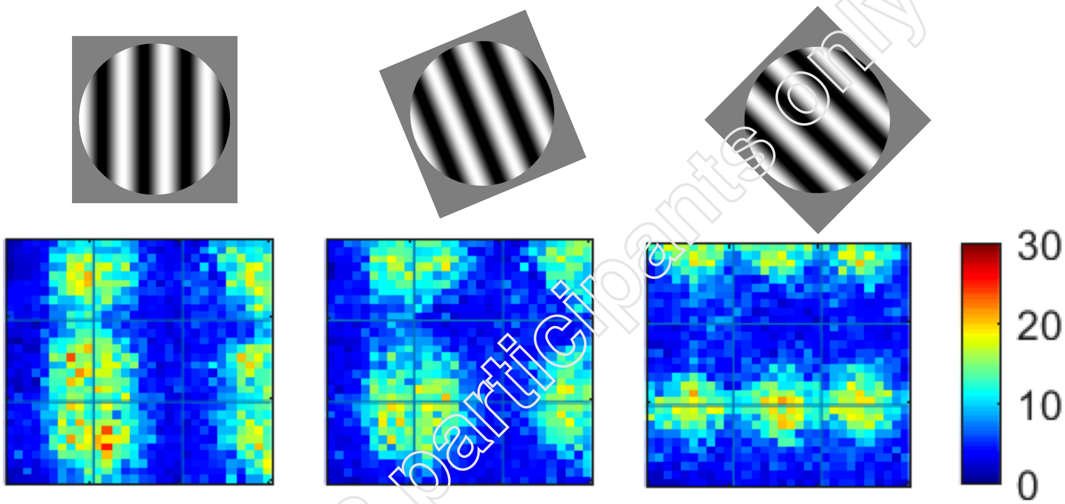

Figure 4: **Model response to drifting gratings.** In the top row are three drifting gratings oriented at 0°, 22.5° and 45° from the vertical; these stimuli are used a great deal in experiments in neuroscience, as they elicit a strong response from V1 neurons having a preference for orientations aligned with the gratings. The bottom three panels show a small piece of our model V1 cortex (layer 4Cα), about 1.5×1.5 mm^2 in actual size. Each tiny box corresponds to ~ 30 neurons, its color indicating the number of spikes fired per neuron over a 1 sec interval; see color bar. Visible in each panel is the group of neurons preferring the orientation of the grating above the panel. **Important:** The only input to the model are light intensity maps, i.e., functions $g(x, t)$ where x is location in visual space, t is time, and $g(x, t)$ is the luminance of the grating at location x at time t. LGN relays this information to cortex via the spiking of ON and OFF cells as explained in the text, and our model cortical neurons, the evolutions of which are governed by (11), interact dynamically to produce the response shown.