Chapter 20

Averaging

Why think when you can compute?
—Maciej Zworski

We discuss first the necessity of studying the averages of observables in chaotic dynamics. A time average of an observable is computed by integrating its value along a trajectory. The integral along trajectory can be split into a sum of over integrals evaluated on trajectory segments; if the observable is exponentiated, this yields a *multiplicative* weight for successive trajectory segments. This elementary observation will enable us to recast the formulas for averages in a multiplicative form that motivates the introduction of evolution operators and further formal developments to come. The main result is that any *dynamical* average measurable in a chaotic system can be extracted from the spectrum of an appropriately constructed evolution operator. In order to keep our toes closer to the ground, in sect. 20.5 we try out the formalism on the first quantitative diagnosis whether a system is chaotic, the Lyapunov exponent.

20.1 Dynamical averaging

In chaotic dynamics detailed prediction is impossible, as any finitely specified initial condition, no matter how precise, will fill out the entire accessible state space after a finite Lyapunov time (1.1). Hence for chaotic dynamics one cannot follow individual trajectories for a long time; what is attainable, however, is a description of the geometry of the set of possible outcomes, and the evaluation of long-time averages. Examples of such averages are transport coefficients for chaotic dynamical flows, such as escape rates, mean drifts and diffusion rates; power spectra; and a host of mathematical constructs such as generalized dimensions, entropies, and Lyapunov exponents. Here we outline how such averages are evaluated within the evolution operator framework. The key idea is to replace the expectation values of observables by the expectation values of exponential generating functionals. This associates an evolution operator with a given observable, and relates the expectation value of the observable to the leading eigenvalue of the evolution operator.

20.1.1 Time averages

Let $a = a(x)$ be any *observable*, a function that associates to each point in state space a number, a vector, or a tensor. The observable reports on a property of the dynamical system. The observable is a device, such as a thermometer or laser Doppler velocitometer. The device itself does not change during the measurement. The velocity field $a_i(x) = v_i(x)$ is an example of a vector observable; the speed $|v(x)|$ (the length of this vector), or perhaps a temperature measured in an experiment at instant τ are examples of scalar observable. We define the *integrated observable* A as the time integral of the observable a evaluated along the trajectory of the initial point x_0,

$$A(x_0, t) = \int_0^t d\tau\, a(x(\tau)) , \qquad x(t) = f^t(x_0) . \tag{20.1}$$

If the dynamics are given by an iterated mapping and the time is discrete, the integrated observable after n iterations is given by

$$A(x_0, n) = \sum_{k=0}^{n-1} a(x_k) , \qquad x_k = f^k(x_0)) \tag{20.2}$$

(we suppress vectorial indices for the time being).

example 20.1
p. 379

The *time average* of the observable along an orbit is defined by

$$\overline{a(x_0)} = \lim_{t\to\infty} \frac{1}{t} A(x_0, t) . \tag{20.3}$$

If a does not behave too wildly as a function of time –for example, if $a(x)$ is the Chicago temperature, bounded between $-80^o F$ and $+130^o F$ for all times– $A(x_0, t)$ is expected to grow no faster than t, and the limit (20.3) exists. For an example of a time average –the Lyapunov exponent– see sect. 20.5.

The time average is a property of the orbit, independent of the initial point on that orbit: if we start at a later state space point $f^T(x_0)$ we get a couple of extra finite contributions that vanish in the $t \to \infty$ limit:

$$\begin{aligned}
\overline{a(f^T(x_0))} &= \lim_{t\to\infty} \frac{1}{t} \int_T^{t+T} d\tau\, a(f^\tau(x_0)) \\
&= \overline{a(x_0)} - \lim_{t\to\infty} \frac{1}{t} \left(\int_0^T d\tau\, a(f^\tau(x_0)) - \int_t^{t+T} d\tau\, a(f^\tau(x_0)) \right) \\
&= \overline{a(x_0)} .
\end{aligned}$$

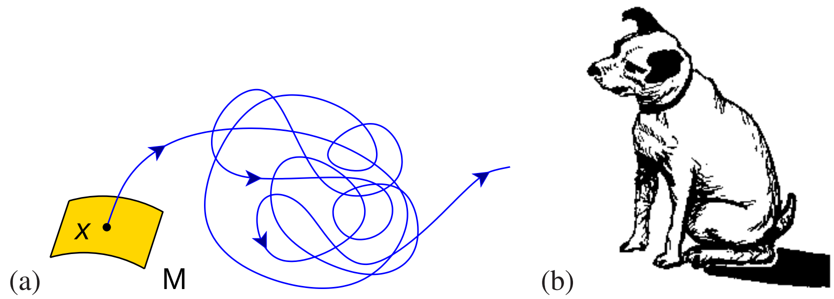


Figure 20.1: (a) A typical chaotic trajectory explores the state space with the long time visitation frequency building up the natural measure $\rho_0(x)$. (b) time average evaluated along an atypical trajectory such as a periodic orbit fails to explore the entire accessible state space. (A. Johansen)

The integrated observable $A(x_0, t)$ and the time average $\overline{a(x_0)}$ take a particularly simple form when evaluated on a periodic orbit. Define

exercise 4.6

$$A_p = \begin{cases} a_p T_p &= \int_0^{T_p} d\tau\, a(x(\tau)) \quad \text{for a flow} \\ a_p n_p &= \sum_{i=1}^{n_p} a(x_i) \quad \text{for a map} \end{cases}, \qquad x \in \mathcal{M}_p\,, \tag{20.4}$$

where p is a prime cycle, T_p is its period, and n_p is its discrete time period in the case of iterated map dynamics. The quantity A_p is a loop integral of the observable along a single traversal of a prime cycle p, so it is an intrinsic property of the cycle, independent of the starting point $x_0 \in \mathcal{M}_p$. If the trajectory retraces itself r times, we just obtain A_p repeated r times. Evaluation of the asymptotic time average (20.3) therefore requires only a single traversal of the cycle:

$$a_p = A_p/T_p\,. \tag{20.5}$$

Innocent as this seems, it implies that $\overline{a(x_0)}$ is in general a wild function of x_0; for a hyperbolic system it takes the same value $\langle a \rangle$ for almost all initial x_0, but a different value (20.5) on (almost) every periodic orbit (figure 20.1 (b)).

example 20.2
p. 379
section 24.1

20.1.2 Spatial averages

The *space average* of a quantity a evaluated over all state space trajectories $x(t)$ at time t is given by the d-dimensional integral over all initial points x_0 at time $t = 0$:

$$\begin{aligned} \langle a \rangle(t) &= \frac{1}{|\mathcal{M}|}\int_{\mathcal{M}} dx_0\, a(x(t))\,, \qquad x(t) = f^t(x_0) \\ |\mathcal{M}| &= \int_{\mathcal{M}} dx \;= \text{volume of } \mathcal{M}\,. \end{aligned} \tag{20.6}$$

The space $\mathcal{M}$ is assumed to have finite volume - open systems like the 3-disk game of pinball are discussed in sect. 20.4.

example 20.6
p. 381

What is it we *really* do in experiments? We cannot measure the time average (20.3), as there is no way to prepare a single initial condition with infinite precision. The best we can do is prepare an initial density $\rho(x)$, perhaps concentrated on some small (but always finite) neighborhood. Then we can abandon the uniform space average (20.6) and consider instead the weighted spatial average

$$\langle a\rangle_\rho(t) = \frac{1}{|\mathcal{M}_\rho|}\int_{\mathcal{M}} dx_0\, \rho(x_0)\, a(x(t))\,, \qquad |\mathcal{M}_\rho| = \int_{\mathcal{M}} dx\, \rho(x)\ . \tag{20.7}$$

For ergodic mixing systems, *any* smooth initial density will tend to the asymptotic natural measure in the $t \to \infty$ limit $\rho(x,t) \to \rho_0(x)$. This allows us to take any smooth initial $\rho(x)$ and define the *expectation value* $\langle a\rangle$ of an observable a as the asymptotic time and space average over the state space $\mathcal{M}$

$$\langle a\rangle = \frac{1}{|\mathcal{M}|}\int_{\mathcal{M}} dx\, \overline{a(x)} = \lim_{t\to\infty} \frac{1}{|\mathcal{M}|}\int_{\mathcal{M}} dx_0\, \frac{1}{t}\int_0^t d\tau\, a(x(\tau))\,. \tag{20.8}$$

We use the same $\langle\cdots\rangle$ notation as for the space average (20.6) and distinguish the two by the presence of the time variable in the argument: if the quantity $\langle a\rangle(t)$ being averaged depends on time, then it is a space average; if it is the infinite time limit, it is the expectation value $\langle a\rangle$.

The expectation value is a space average of time averages, with every $x \in \mathcal{M}$ used as a starting point of a time average. The advantage of averaging over space is that it smears the starting points which were problematic for the time average (such as periodic points). While easy to define, the expectation value $\langle a\rangle$ turns out not to be particularly tractable in practice.

Here comes a simple idea that is the basis of all that follows: Such averages are more conveniently studied by investigating instead of $\langle a\rangle$ the space averages of form

$$\langle e^{\beta\cdot A}\rangle = \frac{1}{|\mathcal{M}|}\int_{\mathcal{M}} dx\, e^{\beta\cdot A(x,t)}\,. \tag{20.9}$$

In the present context β is an auxiliary variable of no physical significance whose role is to enable us to recover the desired space average by differentiation,

$$\langle A_i\rangle = \frac{\partial}{\partial\beta_i}\langle e^{\beta\cdot A}\rangle\bigg|_{\beta=0}\ .$$

We write '$\beta\cdot A$' to indicate that if the observable is a d-dimensional vector $a(x) \in \mathbb{R}^d$, then $\beta \in \mathbb{R}^d$; if the observable is a $[d\times d]$ tensor, β is also a rank-2 tensor, and so on. Here we will mostly limit the considerations to scalar β and drop the dot in '$\beta\cdot A$'.

If the time average limit $a(x_0)$ (20.3) exists for 'almost all' initial x_0's and the system is ergodic and mixing (in the sense of sect. 1.3.1), we expect the time average along almost all trajectories to tend to the same value $\overline{a}$, and the integrated observable A to tend to $t\,\overline{a}$. The space average (20.9) is an integral over exponentials and hence also grows (or shrinks) exponentially with time. So as $t \to \infty$ we would expect the space average of $\exp(\beta A(x,t))$ to grow exponentially with time

$$\langle e^{\beta A}\rangle \to (\text{const})\, e^{ts(\beta)}\,,$$

and its rate of growth (or contraction) to be given by the limit

$$s(\beta) = \lim_{t\to\infty} \frac{1}{t} \ln\langle e^{\beta A}\rangle\,. \tag{20.10}$$

Now we understand one reason for why it is smarter to compute $\langle\exp(\beta A)\rangle$ rather than $\langle a\rangle$: the expectation value of the observable (20.8), the (generalized) diffusion tensor, and higher moments of the integrated observable (20.1) can be computed by evaluating the derivatives of $s(\beta)$

$$\begin{aligned} \left.\frac{\partial s}{\partial \beta_j}\right|_{\beta=0} &= \lim_{t\to\infty}\frac{1}{t}\langle A_j\rangle &&= \langle a_j\rangle\,, \\ \left.\frac{\partial^2 s}{\partial \beta_i \beta_j}\right|_{\beta=0} &= \lim_{t\to\infty}\frac{1}{t}\Big(\langle A_i A_j\rangle - \langle A_i\rangle\langle A_j\rangle\Big) && \\ &= \lim_{t\to\infty}\frac{1}{t}\langle (A_i - t\,\langle a_i\rangle)(A_j - t\,\langle a_j\rangle)\rangle &&= \Delta_{ij}\,, \end{aligned} \tag{20.11}$$

and so forth. We have explicitly written out the formulas for a scalar observable; exercise 20.1 the vector case is worked out in exercise 20.1 (we could have used full derivative notation $ds/d\beta$ in (20.11), but for vector observable we do need partial derivatives $\partial s/\partial\beta_i$). If we can compute the function $s(\beta)$, we have the desired expectation value without having to estimate any infinite time limits from finite time data.

Suppose we could evaluate $s(\beta)$ and its derivatives. What are such formulas good for? A typical application arises in the problem of determining transport coefficients from underlying deterministic dynamics.

example 20.3
p. 380

We turn to the problem of evaluating $\langle e^{\beta A}\rangle$ in sect. 20.3, but first we review some elementary notions of statistics that will be useful later on.

fast track:
sect. 20.3, p. 371

20.2 Moments, cumulants

Given a set of N data points, the unbiased empirical estimates for the *empirical mean* and the *unbiased sample variance* of observable a are

$$\hat{a} = \frac{1}{N}\sum_{i=1}^{N} a_i\,, \quad \hat{\sigma}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(a_i - \hat{a})^2\,. \tag{20.12}$$

(The $N - 1$ divisor in sample variance formula has to do with ensuring that $\hat{a}$ minimizes $\hat{\sigma}^2$.)

The exact *mean* (or expectation or expected value) is the integral of the random variable with respect to its probability measure ρ, commonly denoted $\langle \cdots \rangle$, $E[\cdots]$, or $\overline{\cdots}$,

$$\langle a \rangle = E[a] = \overline{a} = \int_{\mathcal{M}} dx\, \rho(x)\, a(x)\,. \tag{20.13}$$

In ChaosBook we use $\langle \cdots \rangle_\rho$ to denote an integral over state space weighted by ρ, and $\overline{\cdots}$ to denote a time average. If the average is over a (finite or infinite) set of states labeled by labels π, each state contributing with a weight t_π, the expectation is given by

$$\langle A \rangle = \sum_\pi A_\pi t_\pi\,. \tag{20.14}$$

The kth *moment* is the expectation $\langle A^k \rangle$. The moments about the mean, $\langle (A - \langle A \rangle)^k \rangle$, are called *central moments*. For a scalar observable the second central moment is the *variance*, $\sigma^2 = \langle (A - \langle A \rangle)^2 \rangle$, and its positive square root is the *standard deviation* σ. For a multi-component observable the second central moment is the *covariance* matrix $Q_{ij} = \langle (A_i - \langle A_i \rangle)(A_j - \langle A_j \rangle) \rangle$, whose singular values are its standard deviations σ_j. *Standardized moment* is the kth central moment divided by σ^k, $\langle (A - \langle A \rangle)^k \rangle / \sigma^k$, a dimensionless representation of the distribution, independent of translations and linear changes of scale, but meaningful only for scalar observables. Moments can be collected in the moment-generating function (exponential generating function)

$$\langle e^{\beta A} \rangle = 1 + \sum_{k=1}^{\infty} \frac{\beta^k}{k!} \langle A^k \rangle\,. \tag{20.15}$$

However, we do not really care about describing deviations centered around A. If $|A| < 1$, A^k gets very small very fast, and conversely If $|A| > 1$, A^k gets very big, and what is so special about $|A| = 1$? What we care about are fluctuations around the mean $\langle A \rangle$. With some hindsight (Helmholtz, Gibbs, (20.10) above, $\cdots$), the natural to use momenta is given by the cumulant-generating function

$$\ln \langle e^{\beta A} \rangle = \sum_{k=1}^{\infty} \frac{\beta^k}{n!} \langle A^k \rangle_c\,, \tag{20.16}$$

where the subscript c indicates a *cumulant*, or, in statistical mechanics and quantum field theory contexts, the 'connected Green's function'. Expanding $\log\langle e^{\cdots} \rangle$ it is easy to check that the first cumulant is the mean, the second is the variance,

$$\langle A^2 \rangle_c = \sigma^2 = \langle (A - \langle A \rangle)^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2\,, \tag{20.17}$$

and $\langle A^3 \rangle_c$ is the third central moment, or the *skewness*,

$$\langle A^3 \rangle_c = \langle (A - \langle A \rangle)^3 \rangle = \langle A^3 \rangle - 3\langle A^2 \rangle \langle A \rangle + 2\langle A \rangle^3\,. \tag{20.18}$$

The higher cumulants are neither moments nor central moments. The fourth cumulant,

$$\begin{aligned}\langle A^4\rangle_c &= \langle (A-\langle A\rangle)^4\rangle - 3\langle (A-\langle A\rangle)^2\rangle^2 \\ &= \langle A^4\rangle - 4\langle A^3\rangle\langle A\rangle - 3\langle A^2\rangle^2 + 12\langle A^2\rangle\langle A\rangle^2 - 6\langle A\rangle^4 \,. \end{aligned} \tag{20.19}$$

rewritten in terms of standardized moments, is known as the *kurtosis*:

$$\frac{1}{\sigma^4}\langle A^4\rangle_c = \frac{1}{\sigma^4}\langle (A-\langle A\rangle)^4\rangle - 3 \,. \tag{20.20}$$

One of the reasons why cumulants are preferable to moments is that for a normalized Gaussian distribution all cumulants beyond the second one vanish, so they are a measure of deviation of statistics from the Gaussian one (see example 24.3).

A scholarly aside, safely ignored: In statistical mechanics and field theory, the partition function and the Helmholtz free energy have form

$$Z(\beta) = \exp(-\beta F) \,, \qquad F(\beta) = -\frac{1}{\beta}\ln Z(E) \,, \tag{20.21}$$

so in that sense $\langle e^{\beta A}\rangle$ is a 'partition function', and $s(\beta)$ in (20.10) is the 'free energy'. For a 'free' or 'Gaussian' field theory the only non-vanishing cumulant is the second one; for field theories with interactions the derivatives of $s(\beta)$ with respect to β then yield cumulants, or the Burnett coefficients (24.22), or 'effective' n-point Green functions or correlations.

20.3 Evolution operators

For it, the mystic evolution;
Not the right only justified
– what we call evil also justified.
—Walt Whitman,
Leaves of Grass: Song of the Universal

The above simple shift of focus, from studying $\langle a\rangle$ to studying $\langle \exp(\beta A)\rangle$ is the key to everything that follows. Make the dependence on the flow explicit by rewriting this quantity as

$$\langle e^{\beta A}\rangle = \frac{1}{|\mathcal{M}|}\int_{\mathcal{M}} dx \int_{\mathcal{M}} dy\, \delta\big(y - f^t(x)\big)\, e^{\beta A(x,t)} \,. \tag{20.22}$$

Here $\delta(y - f^t(x))$ is the Dirac delta function: for a deterministic flow an initial point x maps into a unique point y at time t. Formally, all we have done above is to insert the identity

$$1 = \int_{\mathcal{M}} dy\, \delta\big(y - f^t(x)\big) \,, \tag{20.23}$$

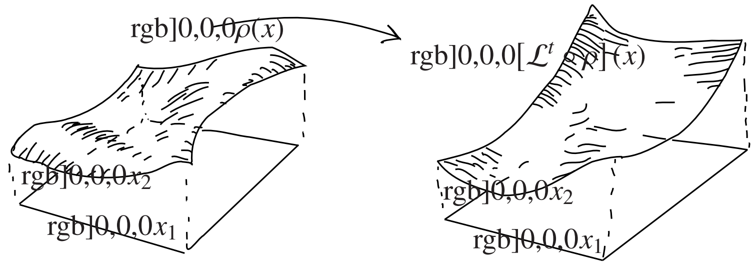


Figure 20.2: Space averaging pieces together the time average computed along the $t \to \infty$ orbit of figure 20.1 by a space average over infinitely many short t trajectory segments starting at all initial points at once.

into (20.9) to make explicit the fact that we are averaging only over the trajectories that remain in $\mathcal{M}$ for all times. However, having made this substitution we have replaced the study of individual trajectories $f^t(x)$ by studying the evolution of the density of *the totality* of initial conditions. Instead of trying to extract a temporal average from an arbitrarily long trajectory which explores the state space ergodically, we can now probe the entire state space with short (and controllable) finite time pieces of trajectories originating from every point in $\mathcal{M}$.

As a matter of fact (and that is why we went to the trouble of defining the generator (19.24) of infinitesimal transformations of densities) *infinitesimally* short time evolution induced by the generator $\mathcal{A}$ of (19.24) suffices to determine the spectrum and eigenvalues of $\mathcal{L}^t$.

We shall refer to the kernel of the operation (20.22) as the *evolution operator*

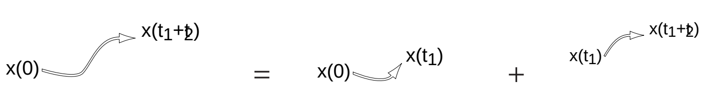
$$\mathcal{L}^t(y,x) = \delta\big(y - f^t(x)\big)\, e^{\beta A(x,t)} \,. \tag{20.24}$$

The simplest example is the $\beta = 0$ case, i.e., the Perron-Frobenius operator introduced in sect. 19.2. Another example - designed to deliver the Lyapunov exponent - will be the evolution operator (20.44) discussed below. The action of the evolution operator on a function ϕ is given by

$$\left[\mathcal{L}^t\phi\right](y) = \int_{\mathcal{M}} dx\, \delta\big(y - f^t(x)\big)\, e^{\beta A(x,t)} \phi(x) \,. \tag{20.25}$$

The evolution operator is different for different observables, as its definition depends on the choice of the integrated observable A in the exponential. Its job is to deliver the expectation value of a, but before showing that it accomplishes that, we need to verify the semigroup property of evolution operators.

By its definition, the integral over the observable a is additive along the trajectory



$$\begin{aligned} A(x_0, t_1 + t_2) &= \int_0^{t_1} d\tau\, a(f^\tau(x)) + \int_{t_1}^{t_1+t_2} d\tau\, a(f^\tau(x)) \\ &= A(x_0, t_1) + A(f^{t_1}(x_0), t_2) \,. \end{aligned}$$

As $A(x,t)$ is additive along the trajectory, the evolution operator generates a semigroup exercise 19.3 section 19.5

$$\mathcal{L}^{t_1+t_2}(y,x) = \int_{\mathcal{M}} dz\, \mathcal{L}^{t_2}(y,z)\mathcal{L}^{t_1}(z,x)\,, \tag{20.26}$$

as is easily checked by substitution

$$\left[\mathcal{L}^{t_2}\mathcal{L}^{t_1}a\right](y) = \int_{\mathcal{M}} dx\, \delta(y - f^{t_2}(x))e^{\beta A(x,t_2)}\left[\mathcal{L}^{t_1}a\right](x) = \left[\mathcal{L}^{t_1+t_2}a\right](y)\,.$$

This semigroup property is the main reason why (20.22) is preferable to (20.8) as a starting point for evaluation of dynamical averages: it recasts averaging in form of operators multiplicative along the flow.

In terms of the evolution operator, the space average of the moment-generating function (20.22) is given by

$$\langle e^{\beta A}\rangle = \frac{1}{|\mathcal{M}|}\int_{\mathcal{M}} dx \int_{\mathcal{M}} dy\, \phi(y)\mathcal{L}^t(y,x)\phi(x)\,.$$

where $\phi(x)$ is the constant function $\phi(x) = 1$. If the linear operator $\mathcal{L}^t$ can be thought of as a matrix, high powers of a matrix are dominated by its fastest growing matrix elements, and the limit (20.10)

$$s(\beta) = \lim_{t\to\infty}\frac{1}{t}\ln\langle\mathcal{L}^t\rangle\,. \tag{20.27}$$

yields the leading eigenvalue $s_0(\beta)$, and, through it, all desired expectation values (20.11).

In what follows we shall learn how to extract not only the leading eigenvalue of $\mathcal{L}^t$, but much of the dominant part of its spectrum. Clearly, we are not interested into the eigenvalues of $\mathcal{L}^t$ for any particular finite time t, but their behavior as $t \to \infty$. That is achieved *via* a Laplace transform, see sect. 20.3.3.

20.3.1 Spectrum of an evolution operator

> This operator is strange:
> it is not self-adjoint, so it is nothing good
> —Jean Bellissard

An exposition of a subject is of necessity sequential and one cannot explain everything at once. As we shall actually never use eigenfunctions of evolution operators, we postpone their discussion to sect. 28.6. For the time being we ask the reader to accept uncritically the following sketch:

Schematically, a linear operator has a spectrum of eigenvalues s_α and eigenfunctions $\varphi_\alpha(x)$

$$\left[\mathcal{L}^t\varphi_\alpha\right](x) = e^{s_\alpha t}\varphi_\alpha(x)\,, \qquad \alpha = 0,1,2,\ldots \tag{20.28}$$

ordered so that Re $s_\alpha \geq$ Re $s_{\alpha+1}$. For continuous time flow eigenvalues cannot depend on time, they are eigenvalues of the time-evolution generator (19.23) we always write the eigenvalues of an evolution operator in exponentiated form e^{s_α} rather than as multipliers λ_α We find it convenient to write them this way both for the continuous time $\mathcal{L}^t$ and the discrete time $\mathcal{L} = \mathcal{L}^1$ cases, and we shall assume that spectrum of $\mathcal{L}$ is discrete.

$\mathcal{L}^t$ is a linear operator acting on a density of initial conditions $\rho(x)$, $x \in \mathcal{M}$, so the $t \to \infty$ limit will be dominated by $s_0 = s(\beta)$, the leading eigenvalue of $\mathcal{L}^t$,

$$\left[\mathcal{L}^t \rho_\beta\right](y) := \int_\mathcal{M} dx\, \delta\!\left(y - f^t(x)\right) e^{\beta A(x,t)} \rho_\beta(x) = e^{t s(\beta)} \rho_\beta(y)\,, \qquad (20.29)$$

where $\rho_\beta(x)$ is the corresponding eigenfunction. For $\beta = 0$ the evolution operator (20.24) is the Perron-Frobenius operator (19.10), with $\rho_0(x)$ the natural measure.

From now on we have to be careful to distinguish the two kinds of linear operators. In chapter 5 we have characterized the evolution of the *local* linear neighborhood of a state space trajectory by eigenvalues and eigenvalues of the linearized flow Jacobian matrices. Evolution operators described in this chapter are *global*, and they act on densities of orbits, not on individual trajectories. As we shall see, one of the wonders of chaotic dynamics is that the more unstable individual trajectories, the nicer are the corresponding global density functions.

20.3.2 Evolution for infinitesimal times

For infinitesimal time δt, the evolution operator (20.6) acts as

$$\begin{aligned}
\rho(y, \delta t) &= \int dx\, e^{\beta A(x,\delta t)} \delta(y - f^{\delta t}(x))\, \rho(x, 0) \\
&= \int dx\, e^{\beta a(x) \delta t} \delta(y - x - \delta t\, v(x))\, \rho(x, 0) \\
&= (1 + \delta t\, \beta\, a(y))\, \frac{\rho(y, 0) - \delta t\, v \cdot \frac{\partial}{\partial x} \rho(y, 0)}{1 + \delta t \frac{\partial v}{\partial x}}\,,
\end{aligned}$$

(the denominator arises from the δt linearization of the jacobian) giving the continuity equation (19.22) a source term

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i}(v_i \rho) = \beta\, a\, \rho\,. \qquad (20.30)$$

The evolution generator (19.24) eigenfunctions now satisfy

$$(s(\beta) - \mathcal{A})\, \rho(x, \beta) = \beta\, a(x)\, \rho(x, \beta)\,. \qquad (20.31)$$

Differentiating with respect to β

$$\begin{aligned}
s'(\beta)\, \rho(x, \beta) \;&+\; s(\beta) \frac{\partial}{\partial \beta} \rho(x, \beta) + \frac{\partial}{\partial x}\left(v(x) \frac{\partial}{\partial \beta} \rho(x, \beta)\right) \\
&= a(x)\, \rho(x, \beta) + \beta\, a(x) \frac{\partial}{\partial \beta} \rho(x, \beta)
\end{aligned}$$

In the vanishing auxiliary parameter limit $\beta \to 0$, we have $s(0) = 0$, $\rho(x, 0) = \rho_0(x)$

$$s'(0)\,\rho_0(x) + \frac{\partial}{\partial x_i}\left(v_i(x)\,\frac{\partial}{\partial \beta}\rho(x, 0)\right) = a(x)\,\rho_0(x)\,.$$

By integrating, the second term vanishes by Gauss' theorem

$$s'(0) = \int dx\, a(x)\,\rho_0(x) = \langle a \rangle\,,$$

verifying equation (20.7): spatial average of the observable a is given by the derivative of the leading eigenvalue $s'(0)$.

fast track:
sect. 21, p. 384

20.3.3 Resolvent of $\mathcal{L}$

Here we limit ourselves to a brief remark about the notion of the 'spectrum' of a linear operator.

The Perron-Frobenius operator $\mathcal{L}$ acts multiplicatively in time, so it is reasonable to suppose that there exist constants $M > 0$, $s_0 \geq 0$ such that $\|\mathcal{L}^t\| \leq M e^{ts_0}$ for all $t \geq 0$. What does that mean? The operator norm is defined in the same spirit in which one defines matrix norms: We are assuming that no value of $\mathcal{L}^t \rho(x)$ grows faster than exponentially for any choice of function $\rho(x)$, so that the fastest possible growth can be bounded by e^{ts_0}, a reasonable expectation in the light of the simplest example studied so far, the escape rate (1.3). If that is so, multiplying $\mathcal{L}^t$ by e^{-ts_0} we construct a new operator $e^{-ts_0}\mathcal{L}^t = e^{t(\mathcal{A}-s_0)}$ which decays exponentially for large t, $\|e^{t(\mathcal{A}-s_0)}\| \leq M$. We say that $e^{-ts_0}\mathcal{L}^t$ is an element of a *bounded* semigroup with generator $\mathcal{A} - s_0 I$. Given this bound, it follows by the Laplace transform

$$\int_0^\infty dt\, e^{-st}\mathcal{L}^t = \frac{1}{s - \mathcal{A}}\,, \qquad \mathrm{Re}\, s > s_0\,, \tag{20.32}$$

that the *resolvent* operator $(s - \mathcal{A})^{-1}$ is bounded

$$\left\|\frac{1}{s - \mathcal{A}}\right\| \leq \int_0^\infty dt\, e^{-st} M e^{ts_0} = \frac{M}{s - s_0}\,. \tag{20.33}$$

If one is interested in the spectrum of $\mathcal{L}$, as we will be, the resolvent operator is a natural object to study; it has no time dependence, and it is bounded. It is called 'resolvent' because it separates the spectrum of $\mathcal{L}$ into individual constituents, one for each spectral 'line'. From (20.27), it is clear that the leading eigenvalue $s_0(\beta)$ corresponds to the pole in (20.33); as we shall see in chapter 21, the rest of the spectrum is similarly resolved into further poles of the Laplace transform.

The main lesson of this brief aside is that for continuous time flows, the Laplace transform is the tool that brings down the generator in (19.26) into the resolvent form (20.32) and enables us to study its spectrum.

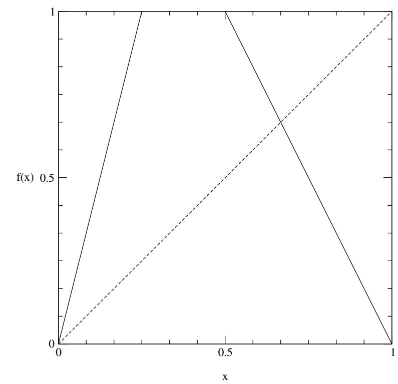


Figure 20.3: A piecewise-linear repeller (19.37): All trajectories that land in the gap between the f_0 and f_1 branches escape ($\Lambda_0 = 4$, $\Lambda_1 = -2$). See example 20.4.

20.4 Averaging in open systems

If $\mathcal{M}$ is a compact region or set of regions to which the dynamics is confined for all times, (20.8) is a sensible definition of the expectation value. However, if the trajectories can exit $\mathcal{M}$ without ever returning,

$$\int_{\mathcal{M}} dy\, \delta(y - f^t(x_0)) = 0 \qquad \text{for } t > t_{exit}\,, \qquad x_0 \in \mathcal{M}\,,$$

we might be in trouble. In particular, a *repeller* is a dynamical system for which the trajectory $f^t(x_0)$ eventually leaves the region $\mathcal{M}$, unless the initial point x_0 is on the repeller, so the identity

$$\int_{\mathcal{M}} dy\, \delta(y - f^t(x_0)) = 1\,, \quad t > 0\,, \qquad \text{iff } x_0 \in \text{non–wandering set} \tag{20.34}$$

might apply only to a fractal subset of initial points of zero Lebesgue measure (non–wandering set is defined in sect. 2.1.1). Clearly, for open systems we need to modify the definition of the expectation value to restrict it to the dynamics on the non–wandering set, the set of trajectories which are confined for all times.

Denote by $\mathcal{M}$ a state space region that encloses all interesting initial points, say the 3-disk Poincaré section constructed from the disk boundaries and all possible incidence angles, and denote by $|\mathcal{M}|$ the volume of $\mathcal{M}$. The volume of state space containing all trajectories, which start out within the state space region $\mathcal{M}$ and recur within that region at time t, is given by

$$|\mathcal{M}(t)| = \int_{\mathcal{M}} dx dy\, \delta\Big(y - f^t(x)\Big) \;\sim\; |\mathcal{M}| e^{-\gamma t}. \tag{20.35}$$

As we have already seen in sect. 1.4.3, this volume is expected to decrease exponentially, with the escape rate γ. The integral over x takes care of all possible initial points; the integral over y checks whether their trajectories are still within $\mathcal{M}$ by the time t. For example, any trajectory that falls off the pinball table in figure 1.1 is gone for good. section 27.1

If we expand an initial distribution $\rho(x)$ in (20.28), the eigenfunction basis $\rho(x) = \sum_\alpha a_\alpha \varphi_\alpha(x)$, we can also understand the rate of convergence of finite-time estimates to the asymptotic escape rate. For an open system the fraction of trapped trajectories decays as

section 20.4

$$\begin{aligned}\Gamma_{\mathcal{M}}(t) &= \frac{\int_{\mathcal{M}} dx\,[\mathcal{L}^t\rho]\,(x)}{\int_{\mathcal{M}} dx\,\rho(x)} = \sum_\alpha e^{s_\alpha t} a_\alpha \frac{\int_{\mathcal{M}} dx\,\varphi_\alpha(x)}{\int_{\mathcal{M}} dx\,\rho(x)} \\ &= e^{s_0 t}\left((\text{const.}) + O(e^{(s_1-s_0)t})\right) . \end{aligned} \tag{20.36}$$

The constant depends on the initial density $\rho(x)$ and the geometry of state space cutoff region $\mathcal{M}$, but the escape rate $\gamma = -s_0$ is an intrinsic property of the repelling set. We see, at least heuristically, that the leading eigenvalue of $\mathcal{L}^t$ dominates $\Gamma_{\mathcal{M}}(t)$ and yields the escape rate, a measurable property of a given repeller.

The non–wandering set can be very difficult to describe; but for any finite time we can construct a normalized measure from the finite-time covering volume (20.35), by redefining the space average (20.9) as

$$\langle e^{\beta A} \rangle = \int_{\mathcal{M}} dx\, \frac{1}{|\mathcal{M}(t)|} e^{\beta A(x,t)} \sim \frac{1}{|\mathcal{M}|} \int_{\mathcal{M}} dx\, e^{\beta A(x,t)+\gamma t} . \tag{20.37}$$

in order to compensate for the exponential decrease of the number of surviving trajectories in an open system with the exponentially growing factor $e^{\gamma t}$. What does this mean? Once we have computed γ we can replenish the density lost to escaping trajectories, by pumping in $e^{\gamma t}$ of new trajectories in such a way that the overall measure is correctly normalized at all times, $\langle 1 \rangle = 1$.

example 20.4
p. 380

20.5 Evolution operator evaluation of Lyapunov exponents

A solution to these problems was offered in sect. 20.3: replace time averaging along a single orbit by action of a multiplicative evolution operator on the entire state space, and extract the state space average of the Lyapunov exponent from its leading eigenvalue, computed from finite length cycles. The main idea - what is the Lyapunov ‘observable’ - can be illustrated by the dynamics of a 1-dimensional map.

example 20.5
p. 380

Here we have restricted our considerations to 1-d maps, as for higher-dimensional flows only the Jacobian matrices are multiplicative, not the individual eigenvalues. Construction of the evolution operator for evaluation of the Lyapunov spectra for a d-dimensional flow requires more skill than warranted at this stage in the narrative: an extension of the evolution equations to a flow in the tangent space.

If the chaotic motion fills the whole state space, we are indeed computing the asymptotic Lyapunov exponent. If the chaotic motion is transient, leading eventually to some long attractive cycle, our Lyapunov exponent, computed on a non–wandering set, will characterize the chaotic transient; this is actually what any experiment would measure, as even a very small amount of external noise suffices to destabilize a long stable cycle with a minute immediate basin of attraction.

All that remains is to determine the value of the Lyapunov exponent

$$\lambda = \langle \ln |f'(x)| \rangle = \left.\frac{\partial s(\beta)}{\partial \beta}\right|_{\beta=0} = s'(0) \qquad (20.38)$$

from (20.11), the derivative of the leading eigenvalue $s_0(\beta)$ of the evolution operator (20.44). example 23.1

The only question is: How? (By chapter 23 you will know.)

Résumé

The expectation value $\langle a \rangle$ of an observable $a(x)$ integrated, $A^t(x) = \int_0^t d\tau\, a(x(\tau))$, and time averaged, A^t/t, over the trajectory $x \to x(t)$ is given by the derivative

$$\langle a \rangle = \left.\frac{\partial s}{\partial \beta}\right|_{\beta=0}$$

of the leading eigenvalue $e^{ts(\beta)}$ of the evolution operator $\mathcal{L}^t$.

By computing the leading eigenfunction of the Perron-Frobenius operator (19.10), one obtains the expectation value (19.17) of any observable $a(x)$. Thus we can construct a specific, hand-tailored evolution operator $\mathcal{L}$ for each and every observable. The good news is that, by the time we arrive at chapter 23, the scaffolding will be removed, both $\mathcal{L}$'s and their eigenfunctions will be gone, and only the explicit and exact periodic orbit formulas for expectation values of observables will remain. chapter 23

The next question is: How do we evaluate the eigenvalues of $\mathcal{L}$? In example 20.4, we saw a piecewise-linear example where these operators reduce to finite matrices **L**, but for generic smooth flows, they are infinite-dimensional linear operators, and finding smart ways of computing their eigenvalues requires some thought. In chapter 14 we undertook the first step, and replaced the *ad hoc* partitioning (19.11) by the intrinsic, topologically invariant partitioning. In chapter 18 we applied this information to our first application of the evolution operator formalism, evaluation of the topological entropy, and the growth rate of the number of topologically distinct orbits. In chapters 21 and 22, this small victory will be refashioned into a systematic method for computing eigenvalues of evolution operators in terms of periodic orbits.

Commentary

Remark 20.1 'Pressure'. The quantity $\langle \exp(\beta A) \rangle$ is called a 'partition function' by Ruelle [A39.14]. Some authors decorate it with considerably more Greek and Gothic letters than is done in this treatise. Ruelle [20.1] and Bowen [A1.70] had given name 'pressure' or 'topological pressure' $P(a)$ to $s(\beta)$ (where a is the observable introduced in sect. 20.1.1), defined by the 'large system' limit (20.10). As we shall also apply the theory to computating the physical gas pressure exerted on the walls of a container by a bouncing particle , we refer to $s(\beta)$ as simply the leading eigenvalue of the evolution operator introduced in sect. 19.5. The 'convexity' properties such as $P(a) \leq P(|a|)$ will be pretty obvious consequences of the definition (20.10). In the case that $\mathcal{L}$ is the Perron-Frobenius operator (19.10), the eigenvalues $\{s_0(\beta), s_1(\beta), \cdots\}$ are called the *Ruelle-Pollicott resonances* [A1.61, A1.62, A1.63], with the leading one, $s(\beta) = s_0(\beta)$ being the one of main physical interest. In order to aid the reader in digesting the mathematics literature, we shall try to point out the notational correspondences whenever appropriate. The rigorous formalism is replete with lims, sups, infs, Ω-sets which are not really essential to understanding of the theory, and are avoided in this book.

Remark 20.2 State space discretization. Ref. [20.10] discusses numerical discretizatons of state space, and construction of Perron-Frobenius operators as stochastic matrices, or directed weighted graphs, as coarse-grained models of the global dynamics, with transport rates between state space partitions computed using this matrix of transition probabilities; a rigorous discussion of some of the former features is included in ref. [28.21].

20.6 Examples

Example 20.1 ***Integrated observables.*** *(a) If the observable is the velocity, $a_i(x) = v_i(x)$, its time integral $A(x_0, t_i)$ is the trajectory $A(x_0, t_i) = x_i(t)$.*

(b) For Hamiltonian flows the action *associated with a trajectory $x(t) = [q(t), p(t)]$ passing through a phase-space point $x_0 = [q(0), p(0)]$ is:*

$$A(x_0, t) = \int_0^t d\tau\, \dot{\mathbf{q}}(\tau) \cdot \mathbf{p}(\tau)\,. \tag{20.39}$$

click to return: p. ??

Example 20.2 ***Deterministic diffusion.*** *The phase space of an open system such as the Sinai gas (an infinite 2-dimensional periodic array of scattering disks, see sect. 24.1) is dense with initial points that correspond to periodic runaway trajectories. The mean distance squared traversed by any such trajectory grows as $x(t)^2 \sim t^2$, and its contribution to the diffusion rate $D \propto x(t)^2/t$, (20.3) evaluated with $a(x) = x(t)^2$, diverges. Seemingly there is a paradox; even though intuition says the typical motion should be diffusive, we have an infinity of ballistic trajectories.*

For chaotic dynamical systems, this paradox is resolved by also averaging over the initial x and worrying about the measure of the 'pathological' trajectories. (continued in example 20.3)

click to return: p. ??

Example 20.3 ***Deterministic diffusion.*** *(continued from example 20.2) Consider a point particle scattering elastically off a d-dimensional array of scatterers. If the scatterers are sufficiently large to block any infinite length free flights, the particle will diffuse chaotically, and the transport coefficient of interest is the diffusion constant $\langle x(t)^2 \rangle \approx 2dDt$. In contrast to D estimated numerically from trajectories $x(t)$ for finite but large t, the above formulas yield the asymptotic D without any extrapolations to the $t \to \infty$ limit. For example, for $a_i = v_i$ and zero mean drift $\langle v_i \rangle = 0$, in d dimensions the diffusion constant is given by the curvature of $s(\beta)$ at $\beta = 0$,* section 24.1

$$D = \lim_{t\to\infty} \frac{1}{2dt} \langle x(t)^2 \rangle = \frac{1}{2d} \sum_{i=1}^{d} \left. \frac{\partial^2 s}{\partial \beta_i^2} \right|_{\beta=0} , \tag{20.40}$$

so if we can evaluate derivatives of $s(\beta)$, we can compute transport coefficients that characterize deterministic diffusion. As we shall see in chapter 24, periodic orbit theory yields an exact *and explicit closed form expression for D.* click to return: p. ??

Example 20.4 ***Escape rate for a piecewise-linear repeller:*** *(continuation of example 19.1) What is gained by reformulating the dynamics in terms of 'operators'? We start by considering a simple example in which the operator is a $[2\times2]$ matrix. Assume the expanding 1-dimensional map $f(x)$ of figure 20.3, a piecewise-linear 2–branch repeller (19.37). Assume a piecewise constant density (19.38). There is no need to define $\rho(x)$ in the gap between $\mathcal{M}_0$ and $\mathcal{M}_1$, as any point that lands in the gap escapes.*

The physical motivation for studying this kind of mapping is the pinball game: f is the simplest model for the pinball escape, figure 1.8, with f_0 and f_1 modelling its two strips of survivors.

As can be easily checked using (19.9), the Perron-Frobenius operator acts on this piecewise constant function as a [2×2] 'transfer' matrix (19.39) exercise 19.1
exercise 19.5

$$\begin{pmatrix} \rho_0 \\ \rho_1 \end{pmatrix} \to \mathcal{L}\rho = \begin{bmatrix} \frac{1}{|\Lambda_0|} & \frac{1}{|\Lambda_1|} \\ \frac{1}{|\Lambda_0|} & \frac{1}{|\Lambda_1|} \end{bmatrix} \begin{pmatrix} \rho_0 \\ \rho_1 \end{pmatrix} ,$$

stretching both ρ_0 and ρ_1 over the whole unit interval Λ, and decreasing the density at every iteration. In this example the density is constant after one iteration, so $\mathcal{L}$ has only one non-zero eigenvalue $e^{s_0} = 1/|\Lambda_0| + 1/|\Lambda_1| \leq 1$, with constant density eigenvector $\rho_0 = \rho_1$. The quantities $1/|\Lambda_0|$, $1/|\Lambda_1|$ are, respectively, the sizes of the $|\mathcal{M}_0|$, $|\mathcal{M}_1|$ intervals, so the exact *escape rate (1.3) – the log of the fraction of survivors at each iteration for this linear repeller – is given by the sole eigenvalue of $\mathcal{L}$:*

$$\gamma = -s_0 = -\ln(1/|\Lambda_0| + 1/|\Lambda_1|) \,. \tag{20.41}$$

Voila! Here is the rationale for introducing operators – in one time step we have solved the problem of evaluating escape rates at infinite time. (continued in example 28.5) click to return: p. ??

Example 20.5 ***Lyapunov exponent, discrete time 1-dimensional dynamics.*** *Due to the chain rule (4.22) for the derivative of an iterated map, the stability of a 1-dimensional mapping is multiplicative along the flow, so the integral (20.1) of the observable $a(x) = \ln|f'(x)|$, the local trajectory divergence rate, evaluated along the trajectory of x_0, is additive:*

$$A(x_0, n) = \ln \left| f^{n\prime}(x_0) \right| = \sum_{k=0}^{n-1} \ln \left| f'(x_k) \right| \,. \tag{20.42}$$

For a 1-dimensional iterative mapping, the Lyapunov exponent is then the expectation value (20.8) given by a spatial integral (20.7) weighted by the natural measure

$$\lambda = \langle \ln |f'(x)| \rangle = \int_{\mathcal{M}} dx\, \rho_0(x) \ln |f'(x)| \,. \tag{20.43}$$

The associated one time step evolution operator (20.24) is

$$\mathcal{L}(y, x) \;=\; \delta(y - f(x))\; e^{\beta \ln |f'(x)|} \,. \tag{20.44}$$

click to return: p. ??

Example 20.6 ***Microcanonical ensemble.*** *In statistical mechanics the space average (20.6) performed over the Hamiltonian system constant energy surface invariant measure* $\rho(x)dx = dqdp\,\delta(H(q,p) - E)$ *of volume* $\omega(E) = \int_{\mathcal{M}} dqdp\,\delta(H(q,p) - E)$

$$\langle a(t) \rangle = \frac{1}{\omega(E)} \int_{\mathcal{M}} dqdp\,\delta(H(q,p) - E) a(q,p,t) \tag{20.45}$$

is called the microcanonical ensemble average.

click to return: p. ??

Exercises

20.1. **Expectation value of a vector observable.** Check and extend the expectation value formulas (20.11) by evaluating the derivatives of $s(\beta)$ up to 4-th order for the space average $\langle \exp(\beta \cdot A) \rangle$ with a_i a vector quantity:

(a)

$$\left.\frac{\partial s}{\partial \beta_i}\right|_{\beta=0} \;=\; \lim_{t\to\infty} \frac{1}{t} \langle A_i \rangle \;=\; \langle a_i \rangle \,, \tag{20.46}$$

(b)

$$\left.\frac{\partial^2 s}{\partial \beta_k \partial \beta_j}\right|_{\beta=0} \;=\; \lim_{t\to\infty} \frac{1}{t} \left(\langle A_k A_j \rangle - \langle A_k \rangle \langle A_j \rangle \right)$$
$$= \lim_{t\to\infty} \frac{1}{t} \langle (A_k - t\langle a_k \rangle)(A_j - t\langle a_j \rangle) \rangle \,.$$

Note that the formalism is smart: it automatically yields the *variance* from the mean, rather than simply the 2nd moment $\langle a^2 \rangle$.

(c) compute the third derivative of $s(\beta)$.

(d) compute the fourth derivative assuming that the mean in (20.46) vanishes, $\langle a_i \rangle = 0$. The 4-th order moment formula

$$K(t) = \frac{\langle x^4(t) \rangle}{\langle x^2(t) \rangle^2} - 3 \tag{20.47}$$

that you have derived is known as *kurtosis* (20.20): it measures a deviation from what the 4-th order moment would be were the distribution a pure Gaussian (see (24.22) for a concrete example). If the observable is a vector, the kurtosis $K(t)$ is given by

$$\frac{\sum_{kj} \left[\langle A_k A_k A_j A_j \rangle + 2 \left(\langle A_k A_j \rangle \langle A_j A_k \rangle - \langle A_k A_k \rangle \langle A_j A_j \rangle \right) \right]}{\left(\sum_k \langle A_k A_k \rangle \right)^2}$$

20.2. **Escape rate for a 1-dimensional repeller, numerically.** Consider the quadratic map

$$f(x) = Ax(1 - x) \tag{20.48}$$

on the unit interval. The trajectory of a point starting in the unit interval either stays in the interval forever or after some iterate leaves the interval and diverges to minus infinity. Estimate numerically the escape rate (27.8), the rate of exponential decay of the measure of points remaining in the unit interval, for either $A = 9/2$ or $A = 6$. Remember to compare your numerical estimate with the solution of the continuation of this exercise, exercise 23.2.

20.3. **Pinball escape rate from numerical simulation*.** Estimate the escape rate for $R : a = 6$ 3-disk pinball by shooting 100,000 randomly initiated pinballs into the 3-disk system and plotting the logarithm of the number of trapped orbits as function of time. For comparison, a numerical simulation of ref. [A1.40] yields $\gamma = .410\ldots$.

References

[20.1] D. Ruelle, *Bull. Amer. Math. Soc.* **78**, 988 (1972).

[20.2] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, *Springer Lect. Notes on Math.* **470** (1975).

[20.3] D. Ruelle, "The thermodynamical formalism for expanding maps," *J. Diff. Geo.* **25**, 117 (1987).

[20.4] M. Pollicott, "On the rate of mixing of Axiom A flows," *Invent. Math.* **81**, 413 (1985).

[20.5] D. Ruelle, *J. Diff. Geo.* **25**, 99 (1987).

[20.6] V. I. Oseledec, *Trans. Moscow Math. Soc.* **19**, 197 (1968).

[20.7] M. Pollicott, *Lectures on Ergodic Theory and Pesin Theory in Compact Manifolds*, (Cambridge Univ. Press, Cambridge 1993).

[20.8] Ya.B. Pesin, *Dynamical systems with generalized hyperbolic attractors: hyperbolic, ergodic and topological properties*, *Ergodic Theory and Dynamical Systems,* **12**, 123 (1992).

[20.9] Ya.B. Pesin, *Func. Anal. Applic.* **8**, 263 (1974).

[20.10] M. Dellnitz, O. Junge, W.S. Koon, F. Lekien, M.W. Lo, J.E. Marsden, K. Padberg, R. Preis, S.D. Ross, and B. Thiere, "Transport in Dynamical Astronomy and Multibody Problems," *Internat. J. Bifur. Chaos* **15**, 699 (2005); www.cds.caltech.edu/~koon/papers

[20.11] G. Froyland, "Computer-assisted bounds for the rate of decay of correlations," *Commun.Math.Phys.* **189**, 237 (1997); C. Liverani, "Rigorous numerical investigation of the statistical properties of piecewise expanding maps. A feasibility study," *Nonlinearity* **14**, 463 (2001).

[20.12] P. Cvitanović, "Chaos for cyclists," in E. Moss, ed., *Noise and chaos in nonlinear dynamical systems* (Cambridge Univ. Press, Cambridge 1989).

[20.13] P. Cvitanović, "The power of chaos," in J.H. Kim and J. Stringer, eds., *Applied Chaos*, (John Wiley & Sons, New York 1992).

[20.14] P. Cvitanović, ed., *Periodic Orbit Theory - theme issue*, *CHAOS* **2**, 1-158 (1992).

[20.15] P. Cvitanović, “Dynamical averaging in terms of periodic orbits,” *Physica* **D 83**, 109 (1995).