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 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.

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 $a = a(x)$ b e a n *observable*, a function that associates a number to each point in state space. 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. A temperature measured in an experiment at instant τ is an example of an observables. 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).

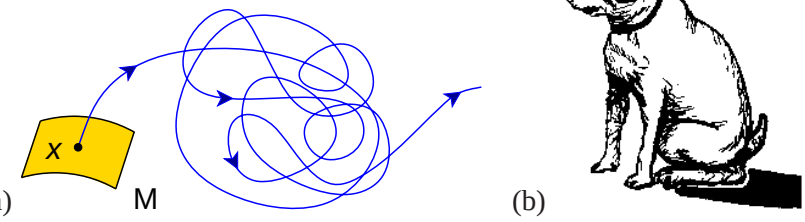


Figure 20.1: (a) A typical chaotic trajectory explores the state space with the long time visitation frequency building up the natural measure $\rho_0(x)$. (b) time average evaluated along an atypical trajectory such as a periodic orbit fails to explore the entire accessible state space. (A. Johansen)

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

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

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:

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

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.

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

$$\langle a\rangle_\rho(t) = \int_{\mathcal{M}} dx_0\, \rho(x_0)\, a(x(t))\,, \qquad \int \tag{20.7}$$

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 $\mathcal{M}$

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

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,

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

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 $s(\beta)$

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

and so forth.

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.

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

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

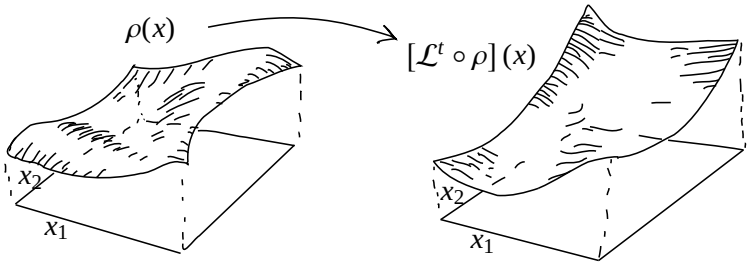


Figure 20.2: Space averaging pieces together the time average computed along the $t \to \infty$ orbit of figure 20.1 by a space average over infinitely many short t trajectory segments starting at all initial points at once.

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 $\mathcal{M}$

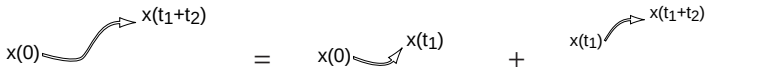
We shall refer to the kernel of the operation (20.22) as the *evolution operator*

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

The simplest example is the $\beta = 0$ case, i.e., the Perron-Frobenius operator introduced in sect. 19.2. The action of the evolution operator on a function ϕ is given by

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

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

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

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).

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 $\mathcal{L}^t$.

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 $\mathbf{L}$, 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 chapter18 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.

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* $[2\times 2]$ *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)

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

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.

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).