Chapter 19

Why cycle?

> "Progress was a labyrinth ... people plunging blindly in and then rushing wildly back, shouting that they had found it ... the invisible king the élan vital the principle of evolution ... writing a book, starting a war, founding a school...."
>
> —F. Scott Fitzgerald, *This Side of Paradise*

In the preceding chapters we have moved rather briskly through the evolution operator formalism. Here we slow down in order to develop some fingertip feeling for the traces of evolution operators.

19.1 Escape rates

We start by verifying the claim (15.11) that for a nice hyperbolic flow the trace of the evolution operator grows exponentially with time. Consider again the game of pinball of figure 1.3. Designate by $\mathcal{M}$ a state space region that encloses the three disks, say the surface of the table $\times$ all pinball directions. The fraction of initial points whose trajectories start out within the state space region $\mathcal{M}$ and recur within that region at the time t is given by

$$\hat{\Gamma}_{\mathcal{M}}(t) = \frac{1}{|\mathcal{M}|} \int\int_{\mathcal{M}} dxdy\, \delta\big(y - f^t(x)\big) \,. \tag{19.1}$$

This quantity is eminently measurable and physically interesting in a variety of problems spanning nuclear physics to celestial mechanics. The integral over x takes care of all possible initial pinballs; the integral over y checks whether they are still within $\mathcal{M}$ by the time t. If the dynamics is bounded, and $\mathcal{M}$ envelops the entire accessible state space, $\hat{\Gamma}_{\mathcal{M}}(t) = 1$ for all t. However, if trajectories exit $\mathcal{M}$ the recurrence fraction decreases with time. For example, any trajectory that falls off the pinball table in figure 1.3 is gone for good.

These observations can be made more concrete by examining the pinball phase space of figure 1.10. With each pinball bounce the initial conditions that survive get thinned out, each strip yielding two thinner strips within it. The total fraction of survivors (1.2) after n bounces is given by

$$\hat{\Gamma}_n = \frac{1}{|\mathcal{M}|} \sum_i^{(n)} |\mathcal{M}_i| \,, \tag{19.2}$$

where i is a binary label of the ith strip, and $|\mathcal{M}_i|$ is the area of the ith strip. The phase space volume is preserved by the flow, so the strips of survivors are contracted along the stable eigendirections, and ejected along the unstable eigendirections. As a crude estimate of the number of survivors in the ith strip, assume that the spreading of a ray of trajectories per bounce is given by a factor Λ, the mean value of the expanding eigenvalue of the corresponding fundamental matrix of the flow, and replace $|\mathcal{M}_i|$ by the phase space strip width estimate $|\mathcal{M}_i|/|\mathcal{M}| \sim 1/\Lambda_i$. This estimate of a size of a neighborhood (given already on p. 89) is right in spirit, but not without drawbacks. One problem is that in general the eigenvalues of a fundamental matrix for a finite segment of a trajectory have no invariant meaning; they depend on the choice of coordinates. However, we saw in chapter 16 that the sizes of neighborhoods are determined by stability eigenvalues of periodic points, and those are invariant under smooth coordinate transformations.

In the approximation $\hat{\Gamma}_n$ receives 2^n contributions of equal size

$$\hat{\Gamma}_1 \sim \frac{1}{\Lambda} + \frac{1}{\Lambda}, \cdots \qquad , \hat{\Gamma}_n \sim \frac{2^n}{\Lambda^n} = e^{-n(\lambda - h)} = e^{-n\gamma} \,, \tag{19.3}$$

up to pre-exponential factors. We see here the interplay of the two key ingredients of chaos first alluded to in sect. 1.3.1: the escape rate γ equals local expansion rate (the Lyapunov exponent $\lambda = \ln \Lambda$), minus the rate of global reinjection back into the system (the topological entropy $h = \ln 2$). As we shall see in (21.16), with correctly defined "entropy" this result is exact.

As at each bounce one loses routinely the same fraction of trajectories, one expects the sum (19.2) to fall off exponentially with n. More precisely, by the hyperbolicity assumption of sect. 16.1.1 the expanding eigenvalue of the fundamental matrix of the flow is exponentially bounded from both above and below,

$$1 < |\Lambda_{min}| \le |\Lambda(x)| \le |\Lambda_{max}| \,, \tag{19.4}$$

and the area of each strip in (19.2) is bounded by $|\Lambda_{max}^{-n}| \le |\mathcal{M}_i| \le |\Lambda_{min}^{-n}|$. Replacing $|\mathcal{M}_i|$ in (19.2) by its over (under) estimates in terms of $|\Lambda_{max}|$, $|\Lambda_{min}|$ immediately leads to exponential bounds $(2/|\Lambda_{max}|)^n \leq \hat{\Gamma}_n \leq (2/|\Lambda_{min}|)^n$, i.e.,

$$\ln|\Lambda_{max}| - \ln 2 \geq \; -\frac{1}{n}\ln\hat{\Gamma}_n \; \geq \ln|\Lambda_{min}| - \ln 2 \,. \tag{19.5}$$

The argument based on (19.5) establishes only that the sequence $\gamma_n = -\frac{1}{n}\ln\Gamma_n$ has a lower and an upper bound for any n. In order to prove that γ_n converge to the limit γ, we first show that for hyperbolic systems the sum over survivor intervals (19.2) can be replaced by the sum over periodic orbit stabilities. By (19.4) the size of $\mathcal{M}_i$ strip can be bounded by the stability Λ_i of ith periodic point:

$$\mathcal{C}_1\frac{1}{|\Lambda_i|} < \frac{|\mathcal{M}_i|}{|\mathcal{M}|} < \mathcal{C}_2\frac{1}{|\Lambda_i|}\,, \tag{19.6}$$

for any periodic point i of period n, with constants $\mathcal{C}_j$ dependent on the dynamical system but independent of n. The meaning of these bounds is that for longer and longer cycles in a system of bounded hyperbolicity, the shrinking of the ith strip is better and better approximated by the derivatives evaluated on the periodic point within the strip. Hence the survival probability can be bounded close to the cycle point stability sum

$$\hat{\mathcal{C}}_1\,\Gamma_n < \sum_i^{(n)}\frac{|\mathcal{M}_i|}{|\mathcal{M}|} < \hat{\mathcal{C}}_2\,\Gamma_n \quad , \tag{19.7}$$

where $\Gamma_n = \sum_i^{(n)} 1/|\Lambda_i|$ is the asymptotic trace sum (16.26). In this way we have established that for hyperbolic systems the survival probability sum (19.2) can be replaced by the periodic orbit sum (16.26).

We conclude that for hyperbolic, locally unstable flows the fraction (19.1) of initial x whose trajectories remain trapped within $\mathcal{M}$ up to time t is expected to decay exponentially,

$$\Gamma_{\mathcal{M}}(t) \;\propto\; e^{-\gamma t}\,,$$

where γ is the asymptotic *escape rate* defined by

$$\gamma = -\lim_{t\to\infty}\frac{1}{t}\ln\Gamma_{\mathcal{M}}(t)\,. \tag{19.8}$$

19.1
page 373
14.4
page 282

19.2 Natural measure in terms of periodic orbits

We now refine the reasoning of sect. 19.1. Consider the trace (16.7) in the asymptotic limit (16.25):

$$\mathrm{tr}\,\mathcal{L}^n = \int dx\,\delta(x - f^n(x))\; e^{\beta A^n(x)} \approx \sum_i^{(n)}\frac{e^{\beta A^n(x_i)}}{|\Lambda_i|}\,.$$

The factor $1/|\Lambda_i|$ was interpreted in (19.2) as the area of the ith phase space strip. Hence $\mathrm{tr}\,\mathcal{L}^n$ is a discretization of the *integral* $\int dx e^{\beta A^n(x)}$ approximated by a tessellation into strips centered on periodic points x_i, figure 1.12, with the volume of the ith neighborhood given by estimate $|\mathcal{M}_i| \sim 1/|\Lambda_i|$, and $e^{\beta A^n(x)}$ estimated by $e^{\beta A^n(x_i)}$, its value at the ith periodic point. If the symbolic dynamics is a complete, any rectangle $[s_{-m}\cdots s_0.s_1s_2\cdots s_n]$ of sect. 11.4.1 always contains the cycle point $\overline{s_{-m}\cdots s_0s_1s_2\cdots s_n}$; hence even though the periodic points are of measure zero (just like rationals in the unit interval), they are dense on the non–wandering set. Equipped with a measure for the associated rectangle, periodic orbits suffice to cover the entire non–wandering set. The average of $e^{\beta A^n}$ evaluated on the non–wandering set is therefore given by the trace, properly normalized so $\langle 1\rangle = 1$:

$$\left\langle e^{\beta A^n}\right\rangle_n \approx \frac{\sum_i^{(n)} e^{\beta A^n(x_i)}/|\Lambda_i|}{\sum_i^{(n)} 1/|\Lambda_i|} = \sum_i^{(n)}\mu_i\, e^{\beta A^n(x_i)}\,. \tag{19.9}$$

Here μ_i is the *normalized natural measure*

$$\sum_i^{(n)}\mu_i = 1\,, \qquad\qquad \mu_i = e^{n\gamma}/|\Lambda_i|\,, \tag{19.10}$$

correct both for the closed systems as well as the open systems of sect. 15.1.3.

Unlike brute numerical slicing of the integration space into an arbitrary lattice (for a critique, see sect. 14.3), the periodic orbit theory is smart, as it automatically partitions integrals by the intrinsic topology of the flow, and assigns to each tile the invariant natural measure μ_i.

19.2.1 Unstable periodic orbits are dense

(L. Rondoni and P. Cvitanović)

Our goal in sect. 15.1 was to evaluate the space and time averaged expectation value (15.9). An average over all periodic orbits can accomplish the job only if the periodic orbits fully explore the asymptotically accessible state space.

Why should the unstable periodic points end up being dense? The cycles are intuitively expected to be *dense* because on a connected chaotic set a typical trajectory is expected to behave ergodically, and pass infinitely many times arbitrarily close to any point on the set, including the initial point of the trajectory itself. The argument is more or less the following. Take a partition of $\mathcal{M}$ in arbitrarily small regions, and consider particles that start out in region $\mathcal{M}_i$, and return to it in n steps after some peregrination in state space. In particular, a particle might return a little to the left of its original position, while a close neighbor might return a little to the right of its original position. By assumption, the flow is continuous, so generically one expects to be able to gently move the initial point in such a way that the trajectory returns precisely to the initial point, i.e., one expects a periodic point of period n in cell i. As we diminish the size of regions $\mathcal{M}_i$, aiming a trajectory that returns to $\mathcal{M}_i$ becomes increasingly difficult. Therefore, we are guaranteed that unstable orbits of larger and larger period are densely interspersed in the asymptotic non–wandering set.

The above argument is heuristic, by no means guaranteed to work, and it must be checked for the particular system at hand. A variety of ergodic but insufficiently mixing counter-examples can be constructed - the most familiar being a quasiperiodic motion on a torus.

19.3 Flow conservation sum rules

If the dynamical system is bounded, all trajectories remain confined for all times, escape rate (19.8) vanishes $\gamma = -s_0 = 0$, and the leading eigenvalue of the Perron-Frobenius operator (14.10) is simply $\exp(-t\gamma) = 1$. Conservation of material flow thus implies that for bound flows cycle expansions of dynamical zeta functions and spectral determinants satisfy exact *flow conservation* sum rules:

$$
\begin{aligned}
1/\zeta(0,0) &= 1 + \sum_\pi{}' \frac{(-1)^k}{|\Lambda_{p_1}\cdots\Lambda_{p_k}|} = 0 \\
F(0,0) &= 1 - \sum_{n=1}^{\infty} c_n(0,0) = 0
\end{aligned}
\tag{19.11}
$$

obtained by setting $s = 0$ in (18.15), (18.16) cycle weights $t_p = e^{-sT_p}/|\Lambda_p| \to 1/|\Lambda_p|$. These sum rules depend neither on the cycle periods T_p nor on the observable $a(x)$ under investigation, but only on the cycle stabilities $\Lambda_{p,1}$, $\Lambda_{p,2}$, $\cdots$, $\Lambda_{p,d}$, and their significance is purely geometric: they are a measure of how well periodic orbits tessellate the state space. Conservation of material flow provides the first and very useful test of the quality of finite cycle length truncations, and is something that you should always check first when constructing a cycle expansion for a bounded flow.

The trace formula version of the flow conservation flow sum rule comes

in two varieties, one for the maps, and another for the flows. By flow conservation the leading eigenvalue is $s_0 = 0$, and for maps (18.14) yields

$$
\mathrm{tr}\,\mathcal{L}^n = \sum_{i\in \mathrm{Fix} f^n} \frac{1}{|\det\left(\mathbf{1} - M^n(x_i)\right)|} = 1 + e^{s_1 n} + \ldots\,. \tag{19.12}
$$

For flows one can apply this rule by grouping together cycles from $t = T$ to $t = T + \Delta T$

$$
\begin{aligned}
\frac{1}{\Delta T} \sum_{p,r}^{T\le rT_p\le T+\Delta T} \frac{T_p}{\left|\det\left(\mathbf{1}-M_p^r\right)\right|} &= \frac{1}{\Delta T}\int_T^{T+\Delta T} dt\,\left(1 + e^{s_1 t} + \ldots\right) \\
&= 1 + \frac{1}{\Delta T}\sum_{\alpha=1}^{\infty} \frac{e^{s_\alpha T}}{s_\alpha}\left(e^{s_\alpha \Delta T} - 1\right) \approx 1 + e^{s_1 T} + \cdots
\end{aligned}
\tag{19.13}
$$

As is usual for the the fixed level trace sums, the convergence of (19.12) is controled by the gap between the leading and the next-to-leading eigenvalues of the evolution operator.

19.4 Correlation functions

The *time correlation function* $C_{AB}(t)$ of two observables A and B along the trajectory $x(t) = f^t(x_0)$ is defined as

$$
C_{AB}(t;x_0) = \lim_{T\to\infty}\frac{1}{T}\int_0^T d\tau A(x(\tau+t))B(x(\tau))\,, \qquad x_0 = x(0) \tag{19.14}
$$

If the system is ergodic, with invariant continuous measure $\rho_0(x)dx$, then correlation functions do not depend on x_0 (apart from a set of zero measure), and may be computed by a state space average as well

$$
C_{AB}(t) = \int_{\mathcal{M}} dx_0\,\rho_0(x_0) A(f^t(x_0))B(x_0)\,. \tag{19.15}
$$

For a chaotic system we expect that time evolution will loose the information contained in the initial conditions, so that $C_{AB}(t)$ will approach the *uncorrelated* limit $\langle A\rangle \cdot \langle B\rangle$. As a matter of fact the asymptotic decay of correlation functions

$$
\hat{C}_{AB} := C_{AB} - \langle A\rangle\,\langle B\rangle \tag{19.16}
$$

for any pair of observables coincides with the definition of *mixing*, a fundamental property in ergodic theory. We now assume $\langle B\rangle = 0$ (otherwise we may define a new observable by $B(x) - \langle B \rangle$). Our purpose is now to connect the asymptotic behavior of correlation functions with the spectrum of the Perron-Frobenius operator $\mathcal{L}$. We can write (19.15) as

$$\tilde{C}_{AB}(t) = \int_{\mathcal{M}} dx \int_{\mathcal{M}} dy\, A(y)B(x)\rho_0(x)\delta(y - f^t(x))\,,$$

and recover the evolution operator

$$\tilde{C}_{AB}(t) = \int_{\mathcal{M}} dx \int_{\mathcal{M}} dy\, A(y)\mathcal{L}^t(y,x)B(x)\rho_0(x)$$

We recall that in sect. 14.1 we showed that $\rho(x)$ is the eigenvector of $\mathcal{L}$ corresponding to probability conservation

$$\int_{\mathcal{M}} dy\, \mathcal{L}^t(x,y)\rho(y) = \rho(x)\,.$$

Now, we can expand the x dependent part in terms of the eigenbasis of $\mathcal{L}$:

$$B(x)\rho_0(x) = \sum_{\alpha=0}^{\infty} c_\alpha \rho_\alpha(x)\,,$$

where $\rho_0(x)$ is the natural measure. Since the average of the left hand side is zero the coefficient c_0 must vanish. The action of $\mathcal{L}$ then can be written as

$$\tilde{C}_{AB}(t) = \sum_{\alpha \neq 0} e^{-s_\alpha t} c_\alpha \int_{\mathcal{M}} dy\, A(y)\rho_\alpha(y). \qquad (19.17)$$

19.2
page 373

We see immediately that if the spectrum has a *gap*, i.e., if the second largest leading eigenvalue is isolated from the largest eigenvalue ($s_0 = 0$) then (19.17) implies *exponential* decay of correlations

$$\tilde{C}_{AB}(t) \sim e^{-\nu t}\,.$$

The correlation decay rate $\nu = s_1$ then depends only on intrinsic properties of the dynamical system (the position of the next-to-leading eigenvalue of the Perron-Frobenius operator), while the choice of a particular observable influences only the prefactor.

Correlation functions are often accessible from time series measurable in laboratory experiments and numerical simulations: moreover they are linked to transport exponents.

19.5 Trace formulas *vs.* level sums

Trace formulas (16.10) and (16.23) diverge precisely where one would like to use them, at s equal to eigenvalues s_α. Instead, one can proceed as follows; according to (16.27) the "level" sums (all symbol strings of length n) are asymptotically going like $e^{s_0 n}$

$$\sum_{i \in \mathrm{Fix} f^n} \frac{e^{\beta A^n(x_i)}}{|\Lambda_i|} \to e^{s_0 n}\,,$$

so an nth order estimate $s_{(n)}$ of the leading eigenvalue is given by

$$1 = \sum_{i \in \mathrm{Fix} f^n} \frac{e^{\beta A^n(x_i)} e^{-s_{(n)} n}}{|\Lambda_i|} \qquad (19.18)$$

which generates a "normalized measure". The difficulty with estimating this $n \to \infty$ limit is at least twofold:

1. due to the exponential growth in number of intervals, and the exponential decrease in attainable accuracy, the maximal n attainable experimentally or numerically is in practice of order of something between 5 to 20.

2. the pre-asymptotic sequence of finite estimates $s_{(n)}$ is not unique, because the sums Γ_n depend on how we define the escape region, and because in general the areas $\mathcal{M}_i$ in the sum (19.2) should be weighted by the density of initial conditions x_0. For example, an overall measuring unit rescaling $\mathcal{M}_i \to \alpha\mathcal{M}_i$ introduces $1/n$ corrections in $s_{(n)}$ defined by the log of the sum (19.8): $s_{(n)} \to s_{(n)} - \ln\alpha/n$. This can be partially fixed by defining a level average

$$\left\langle e^{\beta A(s)} \right\rangle_{(n)} := \sum_{i \in \mathrm{Fix} f^n} \frac{e^{\beta A^n(x_i)} e^{sn}}{|\Lambda_i|} \qquad (19.19)$$

and requiring that the ratios of successive levels satisfy

$$1 = \frac{\left\langle e^{\beta A(s_{(n)})} \right\rangle_{(n+1)}}{\left\langle e^{\beta A(s_{(n)})} \right\rangle_{(n)}}\,.$$

This avoids the worst problem with the formula (19.18), the inevitable $1/n$ corrections due to its lack of rescaling invariance. However, even though much published pondering of "chaos" relies on it, there is no need for such gymnastics: the dynamical zeta functions and spectral determinants are already invariant not only under linear rescalings, but under *all* smooth nonlinear conjugacies $x \to h(x)$, and require no $n \to \infty$ extrapolations to asymptotic times. Comparing with the cycle expansions (18.7) we see what the difference is; while in the level sum approach we keep increasing exponentially the number of terms with no reference to the fact that most are already known from shorter estimates, in the cycle expansions short terms dominate, longer ones enter only as exponentially small corrections.

The beauty of the trace formulas is that they are coordinatization independent: both $\left|\det\left(\mathbf{1} - M_p\right)\right| = |\det\left(\mathbf{1} - M^{T_p}(x)\right)|$ and $e^{\beta A_p} = e^{\beta A^{T_p}(x)}$ contribution to the cycle weight t_p are independent of the starting periodic point point x. For the fundamental matrix M_p this follows from the chain rule for derivatives, and for $e^{\beta A_p}$ from the fact that the integral over $e^{\beta A^t(x)}$ is evaluated along a closed loop. In addition, $\left|\det\left(\mathbf{1} - M_p\right)\right|$ is invariant under smooth coordinate transformations.

Résumé

We conclude this chapter by a general comment on the relation of the finite trace sums such as (19.2) to the spectral determinants and dynamical zeta functions. One might be tempted to believe that given a deterministic rule, a sum like (19.2) could be evaluated to any desired precision. For short finite times this is indeed true: every region $\mathcal{M}_i$ in (19.2) can be accurately delineated, and there is no need for fancy theory. However, if the dynamics is unstable, local variations in initial conditions grow exponentially and in finite time attain the size of the system. The difficulty with estimating the $n \to \infty$ limit from (19.2) is then at least twofold:

1. due to the exponential growth in number of intervals, and the exponential decrease in attainable accuracy, the maximal n attainable experimentally or numerically is in practice of order of something between 5 to 20;

2. the pre-asymptotic sequence of finite estimates γ_n is not unique, because the sums $\hat{\Gamma}_n$ depend on how we define the escape region, and because in general the areas $|\mathcal{M}_i|$ in the sum (19.2) should be weighted by the density of initial x_0.

In contrast, the dynamical zeta functions and spectral determinants are invariant under *all* smooth nonlinear conjugacies $x \to h(x)$, not only linear rescalings, and require no $n \to \infty$ extrapolations.

Commentary

Remark 19.1 Nonhyperbolic measures. $\mu_i = 1/|\Lambda_i|$ is the natural measure only for the strictly hyperbolic systems. For non-hyperbolic systems, the measure might develop cusps. For example, for Ulam type maps (unimodal maps with quadratic critical point mapped onto the "left" unstable fixed point x_0, discussed in more detail in chapter 22), the measure develops a square-root singularity on the $\overline{0}$ cycle:

$$\mu_0 = \frac{1}{|\Lambda_0|^{1/2}} . \tag{19.20}$$

The thermodynamics averages are still expected to converge in the "hyperbolic" phase where the positive entropy of unstable orbits dominates over the marginal orbits, but they fail in the "non-hyperbolic" phase. The general case remains unclear [11.22, G.18, G.14, G.11].

Remark 19.2 Trace formula periodic orbit averaging. The cycle averaging formulas are not the first thing that one would intuitively write down; the approximate trace formulas are more accessibly heuristically. The trace formula averaging (19.13) seems to have be discussed for the first time by Hannay and Ozorio de Almeida [G.1, 7.11]. Another novelty of the cycle averaging formulas and one of their main virtues, in contrast to the explicit analytic results such as those of ref. [18.4], is that their evaluation *does not* require any explicit construction of the (coordinate dependent) eigenfunctions of the Perron-Frobenius operator (i.e., the natural measure ρ_0).

Remark 19.3 Role of noise in dynamical systems. In any physical application the dynamics is always accompanied by additional external noise. The noise can be characterized by its strength σ and distribution. Lyapunov exponents, correlation decay and dynamo rate can be defined in this case the same way as in the deterministic case. You might fear that noise completely destroys the results derived here. However, one can show that the deterministic formulas remain valid to accuracy comparable with noise width if the noise level is small. A small level of noise even helps as it makes the dynamics more ergodic, with deterministically non-communicating parts of the state space now weakly connected due to the noise, making periodic orbit theory applicable to non-ergodic systems. For small amplitude noise one can expand

$$\overline{a} = \overline{a}_0 + \overline{a_1}\sigma^2 + \overline{a_2}\sigma^4 + \ldots ,$$

around the deterministic averages a_0. The expansion coefficients $\overline{a_1}, \overline{a_2}, \ldots$ can also be expressed via periodic orbit formulas. The calculation of these coefficients is one of the challenges facing periodic orbit theory, discussed in refs. [14.5, 14.6, 14.7].

References

[19.1] F. Christiansen, G. Paladin and H.H. Rugh, *Phys. Rev. Lett.* **65**, 2087 (1990).