RECYCLING CHAOS

PREDRAG CVITANOVIĆ [1]
Institute for Mathematical Sciences
SUNY at Stony Brook, Stony Brook, NY 11794-3660
and
Niels Bohr Institute [2],
Blegdamsvej 17, DK-2100 Copenhagen Ø

1 INTRODUCTION

Nonlinear physics presents us with a perplexing variety of complicated fractal objects and strange sets. Notable examples include strange attractors for chaotic dynamical systems, regions of high vorticity in fully developed turbulence and fractal growth processes. The difficulty that these problems pose us was already noted in the Scriptures, which, following in the footsteps of the pioneering work of B.B. Mandelbrot, say: *That which is crooked cannot be made straight: and that which is wanting cannot be counted* (Ecclesiastes 1.15).

The goal of the present set of lectures is to introduce you to a theory of the low-dimensional deterministic chaos of predictive quality comparable to that of the traditional perturbation expansions for nearly integrable systems. In the traditional approach the integrable motions are used as zeroth-order approximations to physical systems, and weak non-linearities are then accounted for perturbatively. For strongly non-linear, non-integrable systems such expansions fail completely; the asymptotic states show an amazingly rich structure which is not at all apparent in the integrable approximations. However, hidden in this apparent chaos is a rigid skeleton, a tree of *cycles* (periodic orbits) of increasing lengths and self-similar structure. The insight of the modern dynamical systems theory is that the zeroth-order approximations to the harshly chaotic dynamics should be very different from those for the nearly integrable systems: a good starting approximation here is the linear stretching and folding of a baker's map, rather than the winding of a harmonic oscillator.

The lectures present a physics application of the dynamical systems theory developed by Smale[1], Sinai[2], Bowen[3], Ruelle[4, 5] and others; we refer the reader to the above literature for a survey of rigorous results. Perhaps I should hint at what the above sentence really means by quoting first the few starting sentences of Smale's review[1]:

---

[1] to appear in *Nonlinear Physical Phenomena, Brasilia 1989 Winter School*, À. Ferraz, F. Oliveira and R. Osorio, eds. (World Scientific, Singapore 1990)

[2] permanent address

" **I.1. Introduction to conjugacy problems for diffeomorphisms.** This is a survey article on the area of global analysis defined by differentiable dynamical systems or equivalently the action (differentiable) of a Lie group G on a manifold M. Here Diff(M) is the group of all diffeomorphisms of M and a diffeomorphism is a differentiable map with a differentiable inverse. ..."

In retrospect, much of the material that we shall cover is in this and related papers; but a physicist who has run into a chaotic looking time series in his laboratory has no clue that he is investigating the action (differentiable) of a Lie group G on a manifold M and that the Lefschetz trace formula is the way to go. In these lectures we shall try to teach you how to use such tools when confronted with a generic chaotic dynamical system. The theory developed here is based on the observation that the motion in dynamical systems of a few degrees of freedom is often organized around a few *fundamental* cycles. The strategy will be to express averages over chaotic phase space regions in terms of short unstable periodic orbits, with the small expansion parameter being the non-uniformity of the flow (here referred to as *curvature*) across neighborhoods of periodic points.

The lectures are organized as follows: in sect. 2 we introduce zeta functions by considering a physically measurable quantity, the escape rate from a repeller. In sect. 3 we introduce the *cycle expansions*[6, 7, 8] of the dynamical zeta functions[4]:

$$1/\zeta = 1 - \sum_f t_f z^{n_f} - \sum_n c_n z^n \ .$$

and motivate their division into the fundamental cycles t_f and the curvature corrections c_n.

In sect. 4 we discuss the symbolic dynamics prerequisites for implementing the cycle expansions. Symbolic dynamics of a generic dynamic system is arbitrarily complex; even for the humble parabola the grammar is finite only for special parameter values. Our strategy is akin to bounding a real number by a sequence of rational approximants; we converge toward the strange set under investigation by a sequence of self-similar Cantor sets. A "self-similar" Cantor set (in the sense in which we use the word here) is a Cantor set equipped with a *subshift of finite type*[1, 9], *ie.* the corresponding grammar can be stated as a finite number of pruning rules, each forbidding a finite symbol subsequence. The structure of the cycle expansions is more transparent if the pruning rules are implemented by redefining the alphabet; we shall give an example in sect. 4. The main idea will be to implement the pruning rules recursively, by replacing the initial alphabet by a new alphabet for which the rule is automatically obeyed. Finally, in sect. 5 we apply the techniques developed above to a concrete example, calculation of topological entropies.

Much of the other material covered in this course has been written up elsewhere - we refer the reader to ref. [10] for an introduction to the renormalization theory, ref. [11] for a discussion of the phase transitions on strange sets, and to refs. [7, 12] for further discussions of cycle expansions. The lectures concentrate on the general properties of cycle expansions; those have been applied in ref. [8] to a series of examples of low-dimensional chaos, such as 1-d strange attractors, the period-doubling repeller, the Hénon-type maps and the mode locking scaling for circle maps. Cycle expansions have also been applied to the irrational windings set of the critical circle maps[13], to the Hamiltonian period-doubling repeller[14], to a Hamiltonian three-disk pinball[15], to the three-disk quantum scattering resonances[16, 17, 18], to the non-leading eigenvalues[19] of transfer operators and to the correlation exponents[20]. Feasibility of analysis of experimental strange sets in terms of cycle expansions is discussed in refs. [6, 21].

Acknowledgements: These lectures are a presentation of extended recycling project[7, 8, 19] with R. Artuso and E. Aurell. The derivation of the Selberg product of sect. 2 is based on teachings of F. Tangerman and H.H. Rugh. Support by the Carlsberg Fundation is gratefully acknowledged.

2 ZETA FUNCTIONS

The dynamical zeta functions[4] provide a powerful formalism for evaluating averages over chaotic dynamical systems. Perhaps the simplest example of such average is the escape rate from a repeller. The escape rate γ from a finite enclosure V around a d-dimensional repeller is defined[22] as the $n \to \infty$ limit of γ_n, where γ_n is given by the fraction of the initial volume that has not escaped by time n:

$$e^{-n\gamma_n} = \frac{\int_V dx dy \delta(y - f^{(n)}(x))}{\int_V dx} \tag{1}$$

More generally, consider the effect of a d-dimensional deterministic map $f(x)$ (or a Poincaré section of a $d+1$ dimensional flow) on a distribution $\phi_\alpha(x)$ on V given by a transfer operator $(\mathcal{L} \circ \phi)_\beta(y)$,

$$\mathcal{L}_{\beta\alpha}(y, x) = \omega_{\beta\alpha}(x)\delta(y - f(x)) \; . \tag{2}$$

Here $\omega_{\beta\alpha}(x)$ is any weight factor multiplicative along the trajectory (in the escape rate example above simply 1), and the indices refer to possible extra matrix structure (for example, group elements associated with discrete symmetries[15, 16]). Such transfer operators $\mathcal{L}$ appear in the literature under a variety of names: the Perron-Frobenious operator[23], scaling functions[24], and so forth. The integral (1) is proportional to $\mathcal{L}^n$ and is dominated by the leading eigenvalue of $\mathcal{L}$,

$$e^{-n\gamma_n} \propto \lambda_0^n \; ,$$

so the $n \to \infty$ limit of γ_n is $\gamma = -\log \lambda_0$. The eigenvalues λ_k of $\mathcal{L}$ are given by the zeros of

$$\det(1 - z\mathcal{L}) = \prod_{k=0}^{\infty} (1 - z\lambda_k) \; . \tag{3}$$

If $\mathcal{L}$ is real valued, the eigenvalues λ_k are either real, or come in complex pairs λ_k, λ_k^*. The determinant (3) can be evaluated in terms of traces by the identity

$$\det(1 - z\mathcal{L}) = e^{tr \log(1-z\mathcal{L})} = \exp\left(-\sum_{n=1}^{\infty} \frac{z^n}{n} tr\mathcal{L}^n\right) \tag{4}$$

$\mathcal{L}^n$ is simply the sum over all periodic points x_i of period n:

$$\begin{aligned} tr\mathcal{L}^n &= \sum_{\alpha} \int dx dy \delta(x-y) \mathcal{L}^n_{\alpha\alpha}(y,x) \\ &= \sum_{i}^{(n)} \frac{tr\ \omega^n(x_i)}{|\det[\mathbf{1} - \mathbf{J}^{(n)}(x_i)]|}\ , \end{aligned} \tag{5}$$

where

$$\mathbf{J}^{(n)}(x_i) = \prod_{j=0}^{n-1} \mathbf{J}(f^{(j)}(x_i)), \quad J_{kl}(x) = \frac{\partial}{\partial x_l} f_k(x) \tag{6}$$

is the i-cycle $[d \times d]$ Jacobian matrix, with eigenvalues $\Lambda_{i,1}$, $\Lambda_{i,2}$, …, $\Lambda_{i,d}$. We shall refer to the closure of the union of periodic points as the *strange set.*

In evaluating (5), we have assumed *hyperbolicity*, *i.e.* that no eigenvalue is marginal: $|\Lambda_{i,a}| \neq 1$. This is the key assumption underpinning the zeta function formalism; if there are cycles with marginal stability, evaluation of (5) requires inclusion of second order, nonlinear terms, and special care has to be exercised - the exponential scalings will be modified by power laws, typical of intermittency effects[25, 8]. These effects, though beyond the scope of present lectures, are generic to mixtures of order and chaos in nonintegrable dynamical systems and physically very important.

The denominators $\det[\mathbf{1} - \mathbf{J}]$ in (5) are inconvenient for our purposes, as they are not multiplicative (what that means will be clear in a moment). To bring (5) to the desired form, we factorize the denominator determinants into products of expanding eigenvalues $\Lambda_{i,1}$, $\Lambda_{i,2}$, …, $\Lambda_{i,e}$ and contracting eigenvalues $\Lambda_{i,e+1}$, …, $\Lambda_{i,d-1}$, $\Lambda_{i,d}$:

$$\frac{1}{|\det[\mathbf{1} - \mathbf{J}^{(n)}(x_i)]|} = \frac{1}{|\Lambda_i|} \prod_{a=1}^{e} \frac{1}{1 - 1/\Lambda_{i,a}} \prod_{b=e+1}^{d} \frac{1}{1 - \Lambda_{i,b}}\ . \tag{7}$$

Here $\Lambda_i = \prod_a^e \Lambda_{i,a}$ is the product of the expanding eigenvalues. Expanding $1/(1 - 1/\Lambda_{i,a})$, $1/(1 - \Lambda_{i,b})$ as geometric series, and substituting (5) back into (4), we find that $\det(1 - z\mathcal{L})$ is given by the infinite product

$$\det(1 - z\mathcal{L}) = \prod_{k_1 \cdots k_e} \prod_{l_{e+1} \cdots l_d} \exp\left(-\sum_{n=1}^{\infty} \frac{z^n}{n} \sum_{i}^{(n)} \frac{tr\ \omega_i}{|\Lambda_i|} \frac{\Lambda_{i,e+1}^{l_{e+1}} \Lambda_{i,e+2}^{l_{e+2}} \cdots \Lambda_{i,d}^{l_d}}{\Lambda_{i,1}^{k_1} \Lambda_{i,2}^{k_2} \cdots \Lambda_{i,e}^{k_e}}\right) . \tag{8}$$

Here tr refers to the $\omega_{\beta\alpha}$ indices, and $\omega_i = \prod_{j=0}^{n-1} \omega(f^{(j)}(x_i))$. Now the weights are multiplicative along the trajectory in the following sense; if t_i is the weight associated with a periodic point x_i of a shorter prime period n_p, $x_i = f^{(n_p)}(x_i)$, then $t_i = t_p^r$, where $r = n/n_p$. The eigenvalues of the nth iterate of the Jacobian $\mathbf{J}^{(n)}(x_i) = \left[\mathbf{J}^{(n_p)}(x_i)\right]^r$ satisfy this property. Furthermore, the stability of a cycle is, by the chain rule, the same everywhere along the orbit, so each prime cycle of length n_p contributes $\sum_{i\in p} t_i = n_p t_p^{n/n_p}$ terms, and the exponents of (8) can be rearranged as sums over *prime* cycles:

$$
\begin{aligned}
1/\zeta(z) &= \exp\left(-tr\sum_{n=1}^{\infty}\frac{z^n}{n}\sum_i^{(n)} t_i\right) = \exp\left(-tr\sum_p\sum_{r=1}^{\infty}\frac{(z^{n_p}t_p)^r}{r}\right) \\
&= e^{tr\sum_p \log(1-z^{n_p}t_p)} = \prod_p \det(1-z^{n_p}t_p)\ . \qquad (9)
\end{aligned}
$$

Here the index p runs through all distinct prime cycles. A prime cycle is a single traversal of the orbit; its label is a non-repeating symbol string. Returning back to (8) we obtain

$$
\begin{aligned}
\det(1-z\mathcal{L}) &= \prod_{k_1\cdots k_e}\prod_{l_{e+1}\cdots l_d}\frac{1}{\zeta_{k_1\cdots l_d}(z)} \\
1/\zeta_{k_1\cdots k_e l_{e+1}\cdots l_d}(z) &= \prod_p \det\left(1-\frac{\omega_p}{|\Lambda_p|}\frac{\Lambda_{p,e+1}^{l_{e+1}}\Lambda_{p,e+2}^{l_{e+2}}\cdots\Lambda_{p,d}^{l_d}}{\Lambda_{p,1}^{k_1}\Lambda_{p,2}^{k_2}\cdots\Lambda_{p,e}^{k_e}}z^{n_p}\right)\ . \qquad (10)
\end{aligned}
$$

The above expression for the eigenvalues of the transfer operator in terms of prime periodic orbits is the main result of this section; it should work for any hyperbolic average over any strange set for which the weight associated with a cycle is multiplicative along the trajectory. In the above, z is a formal parameter which keeps track of the topological cycle lengths; in more general settings, ω factors hide variables such as continuous time or energy. In what follows we shall often absorb z into the transfer operator: $z\mathcal{L} \to \mathcal{L}$, $z^{n_p}t_p \to t_p$. Equation (9) is an example of a dynamical ζ function[4], and we shall refer to the infinite product (10) as the Selberg product $Z(z) = \det(1-z\mathcal{L})$. The names are motivated by the (purely formal) similarity of these infinite products to the Euler product representation of the Riemann ζ function and to the Selberg Zeta function[26].

A variety of physically interesting averages are determined by the eigenvalues of the transfer operator, *ie.* the zeros of $\det(1-z\mathcal{L})$. In the escape rates example considered above, $\omega_{\alpha\beta} = 1$, and the leading eigenvalue λ_0 can be extracted from the first, $|\Lambda_p|$ "volume" weighted factor $\zeta_0 = \zeta_{0\cdots 0}$ in the Selberg product (10):

$$
1/\zeta_0(z) = \prod_p\left(1-\frac{z^{n_p}}{|\Lambda_p|}\right)\ . \qquad (11)
$$

The escape rate (1) is then given by $\gamma = -\log\lambda_0$. The other $1/\zeta_k$ factors in (10) affect the non-leading, anisotropy dependent eigenvalues of $\mathcal{L}$. In particular (see for ex. ref. [20]), the ratio λ_1/λ_0 yields the correlation exponent for chaotic flows.

To simplify the discussion of the non-leading eigenvalues, in what follows we shall assume that $f(x)$ is a 1-d repelling map, $|f'(x)| > 1$, monotone on two non-overlapping intervals

$$\begin{aligned} f(x) &= f_0(x), \quad x \in I_0 \\ &= f_1(x), \quad x \in I_1 \end{aligned} \tag{12}$$

with fixed points $f_0(x_0) = x_0$, $f_1(x_1) = x_1$. For 1-d repellers the Selberg Zeta (10) is simply

$$Z(z) = \prod_{k=0}^{\infty} \frac{1}{\zeta_k(z)}, \quad \frac{1}{\zeta_k} = \prod_p \left(1 - \frac{t_p}{\Lambda_p^k}\right) , \tag{13}$$

where, for the escape rate example, $t_p = z^{n_p}/|\Lambda_p|$. Note that in the above Λ_p, not $|\Lambda_p|$ enters, hence ζ_k depend on the eigenvalue signs and are *not*[20] the integer exponent cases of the expressions used in evaluation of generalized Lyapunov exponents[27, 28]. Neither do the zeros of $1/\zeta_k$ yield the correlation exponents; those are determined[20] by the zeros of $Z(z)$, not the individual factors $1/\zeta_k$.

3 CYCLE EXPANSIONS

How are formulas such as (10) used? We start by computing the weights ω_p and the eigenvalues $\Lambda_{p,1}, \ldots, \Lambda_{p,d}$ of the shortest prime cycles. This usually requires some numerical work, such as the Newton's method searches for periodic solutions; we shall assume that the numerics are under control, and that *all* short prime cycles up to given length have been found. It is very important not to miss any short cycles, as in this approach the consequences are catastrophic. The result is a finite list of prime cycles p (examples can be found in ref. [8]).

Now we formally expand the Euler product (9)

$$\begin{aligned} 1/\zeta(z) &= \prod_p (1 - t_p) = 1 - \sum_{p_1 p_2 \ldots p_k} t_{p_1 + p_2 \ldots + p_k}, \\ t_{p_1 + p_2 \ldots + p_k} &= (-1)^{k+1} t_{p_1} t_{p_2} \ldots t_{p_k} \quad , \end{aligned} \tag{14}$$

where the sum goes over all distinct non-repeating combinations of prime cycles. For $k > 1$, $t_{p_1 + p_2 \ldots + p_k}$ are "pseudo" orbits; they are sequences of shorter orbits that shadow the orbit with the symbol sequence $p_1 p_2 \ldots p_k$ along segments $p_1, p_2, \ldots, p_k$. For sufficiently small z (we have absorbed z into the weights by $z^{n_p} t_p \to t_p$ substitution) the sum makes sense as a power series in z. For the binary dynamics (for example, the map (12))

$$\begin{aligned} 1/\zeta &= (1 - t_0)(1 - t_1)(1 - t_{10})(1 - t_{100})(1 - t_{101})(1 - t_{1000}) \\ &\quad (1 - t_{1001})(1 - t_{1011})(1 - t_{10000})(1 - t_{10001}) \\ &\quad (1 - t_{10010})(1 - t_{10011})(1 - t_{10101})(1 - t_{10111}) \ldots \end{aligned} \tag{15}$$

the first few terms of the expansion are:

$$
\begin{aligned}
1/\zeta &= 1 - t_0 - t_1 - t_{01} - t_{001} - t_{011} - t_{0001} - t_{0011} - t_{0111} - \dots \\
&\quad - t_{0+1} - t_{0+01} - t_{01+1} - t_{0+001} - t_{0+011} - t_{001+1} - t_{011+1} \\
&\quad - t_{0+01+1} - \dots
\end{aligned}
\tag{16}
$$

The next step is the key step in our approach: we observe that the expansion (14) allows a regrouping of terms into dominant *fundamental* contributions t_f and decreasing *curvature* corrections c_n:

$$
1/\zeta = 1 - \sum_f t_f z^{n_f} - \sum_n c_n z^n \ . \tag{17}
$$

We shall refer to such series as the *cycle expansions*. For the binary case (12) the cycle expansion is obtained by grouping together the terms of the same total symbol string length

$$
\begin{aligned}
1/\zeta &= 1 - t_0 - t_1 - [t_{10} - t_1 t_0] - [(t_{100} - t_{10} t_0) + (t_{101} - t_{10} t_1)] \\
&\quad - [(t_{1000} - t_0 t_{100}) + (t_{1110} - t_1 t_{110}) \\
&\quad + (t_{1001} - t_1 t_{001} - t_{101} t_0 + t_{10} t_0 t_1)] - \dots
\end{aligned}
\tag{18}
$$

The fundamental cycles t_0, t_1 have no shorter approximants; they are the "building blocks" of the dynamics in the sense that all longer orbits can be approximately pieced together from them (note that such expansion makes sense only if the long trajectories are shadowed by shorter trajectories with coincident symbol substrings). We call the sum of all terms of the same total length n (grouped in brackets above) the nth curvature correction c_n, for geometrical reasons that are discussed in ref. [7]. If all orbits are weighted equally ($t_p = z^{n_p}$), such combinations cancel exactly; if orbits of similar symbolic dynamics have similar weights, the weights in such combinations will almost cancel.

For example, consider a piecewise linear map of type (12), the tent map

$$
\begin{aligned}
f_0(x) &= \Lambda_0 x, \qquad 0 \le x \le 1/\Lambda_0 \\
f_1(x) &= \Lambda_1 (1 - x), \quad 1 - 1/\Lambda_1 \le x \le 1 \ .
\end{aligned}
\tag{19}
$$

By the chain rule the stability of any n-cycle factorizes as $\Lambda_{\epsilon_1 \epsilon_2 \dots \epsilon_n} = \Lambda_0^m \Lambda_1^{n-m}$. Clearly the information carried by individual cycles is highly redundant; the cycle expansions eliminate the redundancy by resumming this infinity of cycles and extracting the generating scales of a strange set. In the example at hand, all curvature terms in (18) vanish, and for the escape rate calculation, the ζ_k functions (13) are given exactly by

$$
1/\zeta_k(z) = 1 - \left(\frac{\epsilon_0}{\Lambda_0^{k+1}} + \frac{\epsilon_1}{\Lambda_1^{k+1}} \right) z \ , \tag{20}
$$

where ϵ_p is the sign of Λ_p. With increasing k, the eigenvalues (3)

$$\lambda_k = \frac{\epsilon_0}{\Lambda_0^{k+1}} + \frac{\epsilon_1}{\Lambda_1^{k+1}} \tag{21}$$

tend to $\lambda_k \to \epsilon/\Lambda^{k+1}$, with $|\Lambda|$ the smaller of the two stability eigenvalues $|\Lambda_0|$, $|\Lambda_1|$. Hence for the piecewise linear map (19), (20) yields the entire spectrum of the Selberg product (13), with eigenvalues λ_k falling off exponentially as Λ^{-k}:

$$Z(z) \approx \hat{Z}(z) = \prod_{k=0}^{\infty} \left(1 - \frac{1}{\Lambda^k}\frac{z}{|\Lambda|}\right) . \tag{22}$$

The product $\hat{Z}(z)$ is also named after Euler; unlike the Euler product (14), $\hat{Z}(z)$ is an entire function, and the zeros of the factors are the true zeros of $Z(z)$.

The above spectrum of $\mathcal{L}$ for a piecewise linear map is indicative of the spectra for non-linear maps as well, but the details are subtler. The problem is that while for the piecewise linear mappings the $1/\zeta_k$ functions reduce to finite polynomials, in general the curvature corrections c_n in (17) do not vanish, but fall off exponentially[29, 30, 7] as $c_n \propto c^n$, and cycle expansions (17) have finite radia of convergence. This can be seen by considering a pair of transfer operators[29]: $\mathcal{L}_0(y,x) = \delta(y - f(x))$, $\mathcal{L}_1(y,x) = (f'(x))^{-1}\,\delta(y - f(x))$. The ratio of the associated Selberg products (13) yields $1/\zeta_0$:

$$\frac{\det(1 - z\mathcal{L}_0)}{\det(1 - z\mathcal{L}_1)} = \prod_{k=0}^{\infty}\prod_p \frac{1 - t_p/\Lambda_p^k}{1 - t_p/\Lambda_p^{k+1}} = \frac{1}{\zeta_0(z)} \tag{23}$$

For nonlinear $f(x)$ the leading eigenvalue $\lambda_0^{(1)}$ of $\det(1 - z\mathcal{L}_1)$ induces a pole in $1/\zeta_0(z)$ and limits the radius of convergence of its cycle expansion (17) to the disk $|z| < |1/\lambda_0^{(1)}|$. Proximity of this pole affects the convergence of the leading eigenvalue λ_0 estimates from $1/\zeta_0$ cycle expansions, and it is a good strategy[8, 20] to expand instead the truncated Selberg products $1/\zeta_0\zeta_1\cdots\zeta_k$. The simplest example is

$$\begin{aligned} \frac{1}{\zeta_0\zeta_1} &= \left(1 - \sum_f t_f^{(0)} z^{n_f} - \sum_n c_n^{(0)} z^n\right)\left(1 - \sum_f t_f^{(1)} z^{n_f} - \sum_n c_n^{(1)} z^n\right) \\ &= 1 - \sum_f (t_f^{(0)} + t_f^{(1)}) z^{n_f} + \sum_{ff'} t_f^{(0)} t_{f'}^{(1)} z^{n_f + n_{f'}} - \sum_n \hat{c}_n z^n , \end{aligned} \tag{24}$$

where $\hat{c}_n$ are the curvature terms, *ie.* terms that vanish for piecewise linear repellers. For example, for the binary repeller (12) we have

$$\begin{aligned} \frac{1}{\zeta_0\zeta_1} &= 1 - \left(t_0 + t_1 + \frac{t_0}{\Lambda_0} + \frac{t_1}{\Lambda_1}\right) z + (t_0 + t_1)\left(\frac{t_0}{\Lambda_0} + \frac{t_1}{\Lambda_1}\right) z^2 \\ &\quad -[(t_{10} - t_1 t_0) + \left(\frac{t_{10}}{\Lambda_{10}} - \frac{t_0}{\Lambda_0}\frac{t_1}{\Lambda_1}\right)]z^2 \\ &\quad -[(t_{100} - t_{10} t_0) + \left(\frac{t_{100}}{\Lambda_{100}} - \frac{t_{10}}{\Lambda_{10}}\frac{t_0}{\Lambda_0}\right) \end{aligned}$$

$$
\begin{aligned}
&-\frac{t_0}{\Lambda_0}(t_{10}-t_1t_0)-t_0\left(\frac{t_{10}}{\Lambda_{10}}-\frac{t_0}{\Lambda_0}\frac{t_1}{\Lambda_1}\right)\\
&+(t_{101}-t_{10}t_1)+\left(\frac{t_{101}}{\Lambda_{101}}-\frac{t_1}{\Lambda_1}\frac{t_{10}}{\Lambda_{10}}\right)\\
&-(t_{10}-t_1t_0)\frac{t_1}{\Lambda_1}-\left(\frac{t_{10}}{\Lambda_{10}}-\frac{t_0}{\Lambda_0}\frac{t_1}{\Lambda_1}\right)t_1]z^3-\ldots
\end{aligned}
\tag{25}
$$

The first line contains the fundamental contributions, *i.e.* the terms that contribute to the piecewise linear map (19). The rest is the beginning of the curvature corrections.

As the estimate (22) implies that the coefficients in the expansion $Z(z)=\sum_n C_n z^n$ fall off rapidly as $|C_n|\approx\Lambda^{-n^2}$, one might be tempted to simply expand the Selberg product (13) as a power series in z, truncating the series at the maximal cycle length used. However, note that if $1/\zeta_0\zeta_1\cdots\zeta_k$ is expanded as a power series in z, the curvature expansion starts only with cycle lengths $n>kn_F$, where n_F is the length of the longest fundamental cycle, hence the infinite Selberg product *does not* separate into the fundamental part and curvature corrections, and the rapid convergence of the $1/\zeta$ cycle expansions is lost.

The above discussion might make the evaluation of eigenvalues of $\mathcal{L}$ appear a rather complicated affair. We have gone in so much detail to clarify the separation into the fundamental and curvature terms, as this tells us what is the minimal number of terms needed in order that the corrections are of curvature type and small. Actual calculation is straightforward. We substitute the eigenvalues and lengths of prime cycles into the cycle expansion (18), and evaluate the coefficients in polynomial approximations $1/\zeta_k=1-\sum a_n^{(k)}z^n$, $1/\zeta_0\zeta_1\cdots\zeta_k=1-\sum a_n z^n$. The escape rate $\gamma=\ln z$ is determined by finding the leading root of the polynomial approximation; the zeros $1/\zeta_0\zeta_1\cdots\zeta_k=0$ can be easily determined by standard numerical methods. Examples of such calculations are carried out in refs. [8], [13]-[21].

4 SYMBOLIC DYNAMICS

The first prerequisite to effective application of zeta function techniques is a precise description of the topology of a chaotic dynamical system, or its *symbolic dynamics*. This is achieved by partitioning the phase space into topologically distinct regions, associating with each region a symbol from an *alphabet*, and using those symbols to uniquely label every possible trajectory. If all possible symbol sequences can be realized as physical trajectories, the symbolic dynamics is called *complete*; if some sequences are not allowed, the symbolic dynamics is pruned (the word is suggested by "pruning" of branches corresponding to forbidden sequences for symbol dynamics organized by a hierarchical tree).

In this section we illustrate extraction of a symbolic dynamics on a piecewise linear repeller for which the itinerary of a repeller point x is given by its binary expansion. The main result is the Algorithm 1 which converts recursively a given value of the "pruning point" x_p into the symbolic dynamics of the map. We shall apply this symbolic dynamics in the next section to construction of explicit examples of zeta functions.

The simplest example of a map with complete binary dynamics is the Bernoulli shift

$$x_{n+1} = 2x_n \quad mod\ 1\ . \tag{26}$$

The map acts by shifting the binary point to the right, so the itinerary of $x = x_0$ (if x_n<1/2, then s_n=0; if x_n>1/2, then s_n=1) is simply its binary expansion $x = .s_1s_2s_3\cdots$. The periodic points $\overline{s_1s_2\cdots s_n}$ correspond to rational x:

$$x_{s_1s_2\cdots s_n} = \sum_{k=1}^{n}\frac{s_k}{2^k}\sum_{m=0}^{\infty}\frac{1}{2^{nm}} = \frac{2^n}{2^n-1}.s_1s_2\cdots s_n = \frac{\sum_{k=1}^{n} s_k 2^{n-k}}{2^n-1}\ . \tag{27}$$

This is just the binary version of the familiar fact that the decimal expansion of a rational number is (eventually) periodic.

The *clipped Bernoulli shift* obtained by slicing off all $x \geq x_p$ is a two-branch linear map of form

$$\begin{array}{rclrcl} f_0(x) & = & 2x & 0 \leq & x & \leq 1/2 \\ f_1(x) & = & 2x-1 & 1/2 \leq & x & < x_p\ . \end{array} \tag{28}$$

We shall refer to x_p as the *pruning point*. All trajectories that land in the $x_p < x \leq 1$ interval escape; only those x whose binary expansion contains no subsequence $\_s_m s_{m+1}s_{m+2}\cdots s_{m+n}\_$ such that $.s_m s_{m+1}s_{m+2}\cdots s_{m+n} > x_p$ survive the pruning. The surviving trajectories are unstable (with the stability multiplier $\Lambda_c = 2^{n_c}$, where n_c is the length of the trajectory c), and form a repelling strange set. The symbolic dynamics of the clipped Bernoulli shift is specified by the binary expansion of x_p, in the same sense that the symbolic dynamics of a unimodal map is specified by the trajectory (the kneading invariant[31]) of its critical value $x_p = f(x_c)$.

We find this map pedagogically convenient, as the technique for extracting the symbolic dynamics is essentially the same as for the unimodal maps[32, 33, 31], but the conversion of the parameter x_p into symbolic dynamics is somewhat simpler (as the itineraries are ordered monotonically by the ordinary binary tree rather than the alternating binary tree).

Our strategy for converting x_p into symbolic dynamics is to check recursively whether the x_p falls into a window at successive levels of resolution. If it does, we obtain the exact alphabet; if it does not, the last letter in the approximate alphabet has to be refined, and the procedure repeated. We first phrase this recursive procedure as a general algorithm and then illustrate it by a few examples (the reader might prefer to glance at those first).

Algorithm 1: Pruning the symbolic dynamics from above

Given the pruning point x_p, the symbolic dynamics is determined recursively as follows:

1. expand $x_p = .s_1 s_2 s_3 \cdots$ in binary $s_i = \{0, 1\}$.

2. compactify the binary symbol sequences by rewriting them in the alphabet $\{a_1, a_2, a_3, a_4, \cdots\} = \{0, \underline{10}, \underline{110}, \underline{1110}, \cdots\}$ where a_n stands for a block of n-1 binary 1's followed by 0. At this level of resolution x_p is bracketed by

$$.a_n < x_p \leq .a_{n+1}$$

and the approximate alphabet is $\{a_1, a_2, a_3, a_4, \cdots, a_n\}$.

2. prohibition of a_{n+1} implies that the rightmost surviving point is $.\overline{a_n}$. If the pruning point x_p is in the window $\Delta_{a_n} = (.\overline{a_n}, .a_{n+1}]$, the exact alphabet is $\{a_1, a_2, a_3, \cdots, a_n\}$, and the algorithm stops.

3. If $x_p < .\overline{a_n}$, the pruning point falls somewhere within the $l_{a_n} = (.a_n, \overline{.a_n}]$ interval, and not all sequences starting with $.a_n$ are allowed. Subdivide the $.a_n$ interval into finer subintervals by appending to $.a_n$ all allowed basic blocks: $\{a_n\} = \{a_n a_1, a_n a_2, a_n a_3, \cdots, a_n a_m\}$. Here $m \leq n$ is determined by the value of x_p, $.a_n a_m < x_p \leq .a_n a_{m+1}$. The refined approximate alphabet is given by
$\{a_1, a_2, a_3, \cdots, a_{n-1}, b_1, b_2, , \cdots, b_m\} = \{a_1, \underline{a_2}, a_3, \cdots, a_{n-1}, \underline{a_n a_1}, \underline{a_n a_2}, \cdots, \underline{a_n a_m}\}$.

4. repeat step 2: if $x_p \in \Delta_{b_m} = (.\overline{b_m}, .b_{m+1}]$, the above alphabet is exact, otherwise continue refining the $l_{b_m} = (.b_m, \overline{.b_m}]$ interval by replacing b_m by new letters $\{c_1, c_2, , \cdots, c_k\}$, as in step 3.

Clearly any x_p whose binary expansion is finite yields a finite alphabet, and so does any x_p that falls into a Δ_{c_n} window. Otherwise the algorithm generates a monotone sequence of $l_{c_n} = (.c_n, \overline{.c_n}]$ covering intervals, $x_p \in l_{c_n}$, together with the associated approximate alphabets.

Example 1: the "golden mean" pruning $x_p = .11$

The simplest example of pruning for the clipped Bernoulli shift is given by the $x_p = .11$ pruning point value. As the substring _11_ is forbidden, 1 must always be followed by 0, so the allowed sequences can be built from any number of consecutive 0's and 10 blocks, and the alphabet is simply $\{0, 10\}$. Note that if x_p is set to .11, the rightmost surviving point of the repeller is not .11, but $.10101010\cdots$, *ie.* the periodic point $.\overline{10}$. Hence any x_p value in the window $\Delta_{10} = (.\overline{10}, .11] = (2/3, 3/4]$ leads to the same symbolic dynamics.

Example 2: The right ascending staircase Δ_0, Δ_{10}, $\Delta_{110}, \ldots$,

By the same argument as the above, $x_p = .11\ldots1$ (n binary "1"s, followed by a "0") pruning leads to the n letter alphabet

$\{a_1, a_2, a_3, \cdots, a_n\} = \{0, 10, 110, 1110, \ldots, 11\cdots10\}$,

and the symbolic dynamics unchanged over windows

$$\Delta_{11\cdots10} = (.\overline{11\cdots10}, .11\cdots1] = (\frac{2^{n-1}-2}{2^n-1}, \frac{2^n-1}{2^n}] \,, \tag{29}$$

whose width is shrinking as $|\Delta_{11\cdots 10}| = \frac{1}{2^n(2^n-1)}$.

Example 3: a “typical” pruning front value: $x_p = .11101101110$

1. At the first level of resolution x_p is bracketed by

$$.111 < x_p \leq .1111$$

so 1 can appear only within blocks 10, 110 and 1110. Rewrite x_p in the new alphabet $\{a_1, a_2, a_3, a_4\} = \{0, \underline{10}, \underline{110}, \underline{1110}\}$, where a_n stands for a block of $n-1$ binary 1's followed by 0: $x_p = .a_4a_3a_4a_2$

2. As in the above example, prohibition of $a_5 = 11110$ implies that the right-most surviving point is $.\overline{a_4} = .\overline{1110}$. Were $x_p \in \Delta_{a_4} = (.\overline{a_4}, .a_5]$, the alphabet $\{a_1, a_2, a_3, a_4\}$ would be exact, and we would be finished. However, as $x_p < .\overline{a_4}$, the pruning point falls somewhere within $.a_4 < x_p \leq .\overline{a_4}$, and not all sequences starting with a_4 are allowed.

3. Therefore we subdivide the .1110 interval into finer subintervals by appending to $.a_4$ all allowed basic blocks: $\{a_4\} = \{a_4a_1, a_4a_2, a_4a_3, a_4a_4\}$. As $x_p = .a_4a_3a_4a_2 < .a_4a_4$, the $.a_4a_4$ interval is pruned. x_p can be rewritten in the new alphabet $\{a_1, a_2, a_3, b_1, b_2, b_3\} = \{a_1, a_2, a_3, \underline{a_4a_1}, \underline{a_4a_2}, \underline{a_4a_3}\}$ as $x_p = .b_3b_2$.

4. repeat step 2: is $x_p > .\overline{b_3}$? It is not, so

5. repeat step 3: subdivide $\{b_3\} = \{b_3a_1, b_3a_2, b_3a_3, b_3b_1, b_3b_2\}$. The b_3b_2 block is forbidden by the pruning point value, so we are done; the alphabet consists of 9 letters

$\{a_1, a_2, a_3, b_1, b_2, c_1, c_2, c_3, c_4\} = \{a_1, a_2, a_3, b_1, b_2, \underline{b_3a_1}, \underline{b_3a_2}, \underline{b_3a_3}, \underline{b_3b_1}\}$

We could have kept the binary notation throughout, but a two-letter alphabet makes for rather rather tedious reading; in the binary notation the fundamental blocks are

$\{0, 10, 110, 11100, 111010, 11101100, 111011010, 1110110110, 111011011100\}$.

This finishes the list of examples.

For the clipped Bernoulli map the fraction of the x_p parameter values for which the alphabet is finite can be estimated analytically. If c is a sequence corresponding to one of the windows unfolded recursively in the above, the symbolic dynamics is unchanged over the window

$$\Delta_c = (.\overline{c}, .c1] = (\frac{.c}{1-2^{-n_c}}, .c1] \tag{30}$$

whose width shrinks with n_c, the length of the binary string c, as

$$|\Delta_c| = .c1 - .\overline{c} = \frac{1 - .c_{n+1}}{2^{n_c}-1} = \frac{.c^+1}{2^{n_c}-1}\ , \tag{31}$$

where c^+ is the binary complement of c. This follows from summing the successive images of the the pruned interval $1 - .c1$ within the $(.c, .c1)$ interval.

The widths of the fatest $c = .1000 \cdots 0$ and the thinest $c = .11 \cdots 1$ steps in the devil staircase corresponding to strings of length n are, respectively,

$$\frac{2^{n-1}-1}{2^n(2^n-1)} \geq \Delta_c \geq \frac{1}{2^n(2^n-1)} \; .$$

The lower bound follows from (29) and the upper bound from $1-.c1 = 1-.10\ldots01 = 1/2 - 1/2^n$.

We leave the evaluation of the total measure $\sum \Delta_c$ taken up by finite alphabets as an exercise for the student. What is the measure taken up by the infinite alphabets? Is there a set of non-integer Hausdorff dimension, and what is its significance?

To summarize this section: the most important lesson of the pruning of the cycle expansions is that prohibition of a finite subsequence unbalances the head of a cycle expansion and increases the number of the fundamental cycles in (17). Hence the pruned expansions are expected to start converging only *after* all fundamental cycles have been incorporated. In this way the cycle expansions provide a crisp and clearcut definition of the fundamental set of scales.

5 TOPOLOGICAL ENTROPY

The simplest application of the cycle expansions is cycle counting. The growth of the number of cycles as a function of the symbol string length is characterized by the *topological entropy*:

$$h = \lim_{n\to\infty} \frac{\log N_n}{n} \; , \tag{32}$$

where the number of periodic points of length n is given by

$$N_n = \sum_i^{(n)} 1 = trT^n \; .$$

The growth of N_n with n is controlled by the leading eigenvalue of the transition matrix T_{ij}, the topological analogue of the transfer operator (2): $T_{ij} = 1$ if $x_i = f(x_j)$, otherwise $T_{ij} = 0$. As in (9), the leading eigenvalue of T is given by the leading zero $z = e^{-h}$ of the *topological* $1/\zeta$ function[34]

$$1/\zeta = \prod_p (1 - z^{n_p}) = 1 - \sum_f t_f z^{n_f} \; . \tag{33}$$

The periodic points are counted by setting $t_p = 1$ if cycle p exists, $t_p = 0$ if p is pruned. This expression for the entropy is *exact*; in contrast to the definition (32), no $n \to \infty$ extrapolations of $\ln N_n/n$ are required.

Note that for the cycle counting both t_{ab} and the pseudo-orbit combinations $t_{a+b} = t_a t_b$ in cycle expansions (17) have the same value $z^{n_a+n_b}$, so all curvature combinations $t_{ab} - t_a t_b$ vanish exactly. If the number of fundamental cycles is finite, we refer to (33) as the *topological polynomial*.

The symbolic dynamics considered in the preceeding section gives a class of rather simple topological polynomials. If the symbolic dynamics can be written as a complete (unrestricted, unpruned) alphabet in N symbols, then $t_f = 1$ if $f \in$ alphabet, $t_f = 0$ otherwise. According to the results of the preceeding section, for the clipped Bernoulli shift the symbolic dynamics is given by a finite (or infinite alphabet) built up of blocks of increasing binary length:

$\{a_1, a_2, a_3, \cdots, a_n, b_1, b_2, , \cdots, b_m, c_1, c_2, \cdots\}$

For a finite unrestricted alphabet, the topological entropy is given by the smallest root of the corresponding topological polynomial[35] (33):

$$\begin{aligned} 0 &= 1 - z^{n_{a_1}} - z^{n_{a_2}} - z^{n_{a_3}} - \cdots - z^{n_{a_n}} \\ &\quad - z^{n_{b_1}} - z^{n_{b_2}} - \cdots - z^{n_{b_m}} \\ &\quad - z^{n_{c_1}} - z^{n_{c_2}} - \cdots - z^{n_{c_k}} \ . \end{aligned} \tag{34}$$

The simplest example of a not entirely trivial topological polynomial follows from the 3-cycle pruning example 1 of sect. 4. The fundamental cycles 0, $\underline{01}$ are of length 1 and 2, so the topological polynomial is simply

$$\prod_p (1 - z^{n_p}) = 1 - z - z^2 \ , \tag{35}$$

and the topological entropy is $h = \log \frac{1+\sqrt{5}}{2}$.

The topological polynomial for the example 2 of sect. 4 is given by

$$\prod_p (1 - z^{n_p}) = 1 - z - z^2 - \ldots - z^n = \frac{1 - 2z + z^{n+1}}{1 - z} \ . \tag{36}$$

The topological entropy is $h = \log \lambda_0$, where λ_0 is the leading eigenvalue $1 < \lambda_0 \leq 2$. The remaining roots of (36) lie (for large n) close to the unit circle in the complex plane and are of no physical interest.

The alphabet above was generated by resolving the longest fundamental string at a given level by a set of longer strings; so even if the grammar is not finite and the cycle expansion is not a polynomial, the convergence of the cycle expansion should be good, as the errors are bounded from below and above by truncating the expansions with terms z^n and z^{n+1}, where n is the length of the longest binary string in the alphabet. With increasing resolution n typically grows in leaps and bounds. The entropy is given by the isolated real zero $1/2 \leq z < 1$; the remaining zeros of the polynomial approximations to the entropy function for infinite grammar bunch on the unit circle. The rate of convergence depends on the separation of the leading, entropy eigenvalue from the non-leading eigenvalues; as long as there is a gap, the convergence will be exponential, though situations without gap also arise and are interesting (*cf.* period doubling $1/\zeta$ in ref. [31]).

The above considerations, in spite of the restriction to mere cycle counting, reveal a great deal about the spectra of more general transfer operators. For linear systems with a single scale Λ, $1/\zeta_0(z)$ is given by (34), simply by rescaling $z \to z/|\Lambda|$. For non-linear mappings, polynomial approximations to $1/\zeta_k$ have a rather similar structure; there is a physically significant $\lambda_0^{(k)}$, together with a family of unphysical poles in the complex plane, placed roughly on a circle of radius $|1/c|$, where c controls the asymptotic behavior of curvatures, $c_n \approx c^n$. The extraneous zeros delineate the boundary of the convergence of the cycle expansion of $1/\zeta_k$; and for longer and longer truncated Selberg products $1/\zeta_0\zeta_1\cdots\zeta_k$ this boundary is pushed further and further out, allowing determination of a finite number of leading eigenvalues of $\mathcal{L}$.

These are the tools - all that remains for you now is to return to your favorite chaotic system and recycle it.

References

[1] S. Smale, *Bull. Am. Math. Soc.* **73**, 747, (1967).

[2] Ya. G. Sinai, *Russ. Math. Surveys* **166**, 21, (1972).

[3] R. Bowen, *Equilibrium states and the ergodic theory of Anosov-diffeomorphisms*, Springer Lecture Notes in Math., **470**, (1975).

[4] D. Ruelle, *Statistical Mechanics, Thermodynamic Formalism*, (Addison-Wesley, Reading MA, 1978).

[5] D. Ruelle, *Elements of Differentiable Dynamics and Bifurcation Theory*, (Academic Press, San Diego, 1989).

[6] P. Cvitanović, *Phys. Rev. Lett.* **61**, 2729 (1988).

[7] R. Artuso, E. Aurell and P. Cvitanović, to appear in *Nonlinearity.*

[8] R. Artuso, E. Aurell and P. Cvitanović, to appear in *Nonlinearity.*

[9] see for example J. Guckenheimer and P. Holmes, *Non-linear Oscillations, Dynamical Systems and Bifurcations of Vector Fields* (Springer, New York, 1986).

[10] *Universality in Chaos, 2. edition*, P. Cvitanović, ed., (Adam Hilger, Bristol 1989).

[11] P. Cvitanović, in P. Zweifel, G. Gallavotti and M. Anile, eds., *Non-linear Evolution and Chaotic Phenomena* (Plenum, New York 1987).

[12] P. Cvitanović, in F. Moss ed., *Noise and Chaos in Nonlinear Dynamical Systems*, (Cambridge Univ. Press, Cambridge 1988).

[13] P. Cvitanović, G.H. Gunaratne and M.J. Vinson, submitted to *Nonlinearity.*

[14] E. Aurell, submitted to *Phys. Rev. A.*

[15] P. Cvitanović and B. Eckhardt (in preparation).

[16] P. Cvitanović and B. Eckhardt, *Phys. Rev. Lett.* **63**, 823 (1989).

[17] P. Cvitanović, B. Eckhardt and P. Scherer (in preparation).

[18] G. Russberg (in preparation).

[19] E. Aurell, to appear in *J. Stat. Phys.*.

[20] F. Christiansen, H.H. Rugh and G. Paladin, submitted to *Phys. Rev. A*.

[21] F. Christiansen, unpublished.

[22] L. Kadanoff and C. Tang, *Proc. Natl. Acad. Sci. USA,* **81**, 1276, (1984).

[23] see for ex. Y. Oono and Y. Takahashi, *Prog. Theor. Phys.* **63**, 1804 (1980); S.-J. Chang and J. Wright, *Phys. Rev.* **A 23**, 1419 (1981). Y. Takahashi and Y. Oono, *Prog. Theor. Phys.* **71**, 851 (1984).

[24] M.J. Feigenbaum, *J. Stat. Phys.* **52**, 527 (1988).

[25] Y. Pomeau and P. Manneville, *Commun. Math. Phys.* **74**, 189 (1980).

[26] A. Selberg, *J. Ind. Math. Soc.* **20**, 47 (1956).

[27] H. Fujisaka, *Progr. Theor. Phys.* **70**, 1264 (1983).

[28] H. Kantz and P. Grassberger, *Physica* **17D**, 75 (1985).

[29] D. Ruelle, *Inventiones math.* **34**, 231 (1976).

[30] D. Ruelle, "The thermodynamical formalism for expanding maps", IHES preprint P/89/08 (Jan 1989).

[31] J. Milnor and W. Thurston, "On iterated maps of the interval", in A. Dold and B. Eckmann, eds., *Dynamical Systems, Proceedings, U. of Maryland 1986-87, Springer Lec. Notes in Math.* **1342**, 465 (Springer, Berlin 1988).

[32] P. J. Myrberg, *Ann. Acad. Sc. Fenn., Ser. A,* **259**, 1 (1958).

[33] N. Metropolis, M.L. Stein and P.R. Stein, *J. Comb. Theo.* **A15**, 25 (1973).

[34] The ζ functions were originally introduced in this context by E. Artin and B. Mazur, *Annals. Math.* **81**, 82 (1965).

[35] P.A.P Moran, *Proc. Camb. Phil. Soc.* **42**,15 (1946).