Periodic orbits as the skeleton of classical and quantum chaos

Predrag Cvitanović

Niels Bohr Institute, Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark

A description of a low-dimensional deterministic chaotic system in terms of unstable periodic orbits (cycles) is a powerful tool for theoretical and experimental analysis of both classical and quantum deterministic chaos, comparable to the familiar perturbation expansions for nearly integrable systems. The infinity of orbits characteristic of a chaotic dynamical system can be resummed and brought to a Selberg product form, dominated by the short cycles, and the eigenvalue spectrum of operators associated with the dynamical flow can then be evaluated in terms of unstable periodic orbits. Methods for implementing this computation for finite subshift dynamics are introduced.

1. Introduction

Non-linear 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. This was already fully appreciated by Poincaré, who, describing his discovery of homoclinic tangles, mused that "the complexity of this figure will be striking, and I shall not even try to draw it" [1]. Today such drawings are cheap and plentiful; but Poincaré went a step further and noting that hidden in this apparent chaos is a rigid skeleton, a tree of *cycles* (periodic orbits) of increasing lengths and self-similar structure suggested that the cycles should be the key to chaotic dynamics:

"Étant données des équations...et une solution particulière quelconque de ces équations, on peut toujours trouver une solution périodique (dont la période peut, il est vrai, ètre trés longue), telle que la différence entre les deux solutions soit aussi petite qu'on le veut, pendant un temps aussi long qu'on le veut. D'ailleurs, ce qui nous rend ces solutions périodiques si précieuses, c'est qu'elles sont, pour ansi dire, la seule brèche par où nous puissions esseyer de pénétrer dans une place jusqu'ici réputée inabordable."

Periodic orbits have been at core of much of the mathematical work on the theory of the classical and quantum [2–4] dynamical systems ever since. We refer the reader to the reprint selection in ref. [5] and to the classic text of Ruelle [6] for a summary and overview of some of that literature.

Here we shall outline the general strategy of analyzing chaotic dynamical systems in terms of periodic orbits by working through a "generic" example, a repeller associated with the Hénon-type mappings. The strategy is: (1) describe the topology of the dynamical system; (2) convert this topology into a convergent computation of dynamical averages. We refer the reader to refs. [7, 8] for an introduction to the cycle expansions and their application to evaluation of entropies, dimensions, escape rates and other dynamical averages. Here we describe in more detail how a topology of a dynamical system is converted into symbolic dynamics and we introduce a new class of cycle expansions, Selberg product expansions, and discuss their convergence.

The strategy outlined here is perhaps too laborious if only a rough estimate of a dimension or a Lyapunov exponent is desired; a simple computer

generated average might suffice. However, the labor pays off if classical correlation or quantum resonance spectra are called for. We also hope that the high numerical precision attainable by cycle expansions might enable us to attack higher-dimensional problems where crude averaging is computationally infeasible.

The paper is organized as follows: in section 2 we sketch the construction of "pruned" Smale horseshoes, and in section 3 we outline the finite automata method for converting this topology into a grammar of allowed symbol sequences. In section 4 we sketch the conversion of automata graphs into determinants and apply them to evaluation of the topological entropy. In section 5 we introduce the transfer operators associated with deterministic dynamics, and relate their eigenvalues to periodic orbit expansions. In section 6 we discuss the form the cycle expansions take for the Selberg products and in section 7 we discuss their convergence.

2. Topology of hyperbolic flows

That the crucial ingredient in description of chaotic (unstable, hyperbolic) flows is the topology of the non-wandering set was clearly understood by Smale [9], who has provided us with the simplest visualization of such sets as intersections of Smale horseshoes.

Consider a three-dimensional invertible flow which returns an area of a Poincaré section of the flow stretched and folded into a "horseshoe", such that the initial area is intersected at most twice (see fig. 1). Run backwards, the flow generates the backward horseshoe which intersects the forward horseshoe at most four times, and so forth. We shall call such maps (flows) with at most 2^n transverse intersections the *once-folding* maps; they are the two-dimensional siblings of the unimodal maps in one dimension. The *non-wandering set* – the set of all points that do not escape to infinity (in particular, the set of all periodic orbits) – is contained in the intersections
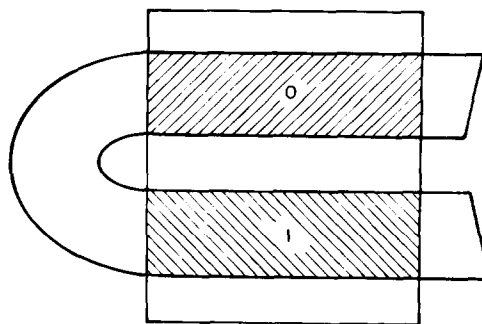

Fig. 1. A Smale horseshoe.

of the forward and backward folds (fig. 2a) and can be labelled in the obvious manner [10] by bi-infinite binary sequences $\ldots s_{-3}s_{-2}s_{-1}s_0s_1s_2 s_3\ldots$, $s_i \in \{0,1\}$. For better visualization of the non-wandering set, fatten the intersection regions until they completely cover a unit square, as in fig. 2b. This *symbol plane* [11, 12] is a topologically accurate representation of the non-wandering set and serves as a street map for labelling its pieces. In the symbol plane, the dynamics is simply the decimal point shift; the _01.01_ square gets mapped into the _010.1_ rectangle (see fig. 2), and so on.

A generic once-folding map does not yield a complete horseshoe; some of the horseshoe pieces might be *pruned*, i.e. not realized by the particular mapping. In one dimension, the criterion for whether a symbolic sequence is allowed is easily formulated; symbolic sequences are topologically ordered by an "alternating binary" reordering of the binary symbols [13–15], and any orbit that strays to the right of the *kneading sequence* (the orbit of the critical point) is pruned [16, 17]. That does not mean that the symbolic dynamics is simple – as a matter of fact, already in one dimension its grammar can be arbitrarily complicated [18]. However, the grammar is finite if the critical point is in the basin of attraction of an attractive periodic point [18]; for example, if the critical point is attracted to a stable three-cycle, the repeller consists of an isolated $\bar{0}$ fixed point and all orbits built from the two letters $\{1,10\}$.
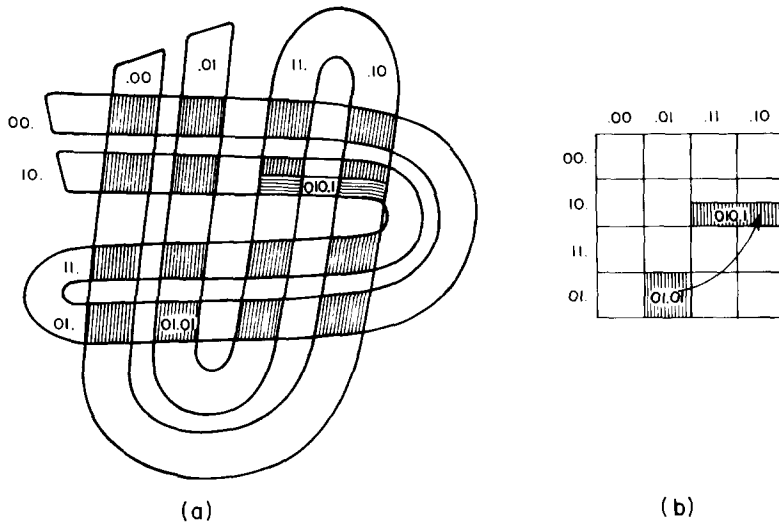
Fig. 2. (a) A complete Smale horseshoe twice iterated forwards and backwards in time and its (b) symbol plane representation (the intersections fattened into a unit square). In the symbol plane the dynamics maps rectangles into rectangles by a decimal point shift.

The analogous situation can be attained for a two-dimensional once-folding map if the parameters of the map adjusted so that the intersection of the backward and forward folds is still transverse, but no longer complete, as in fig. 3a. In this particular example the intersections _10.1_, _11.1_ have been lost, and consequently any trajectory containing substrings _101_, _111_ is pruned. We refer to the left border of this primary pruned region as the *pruning front*; an example of a pruning front is drawn in fig. 4. The topology puts two obvious constraints on the form of a pruning
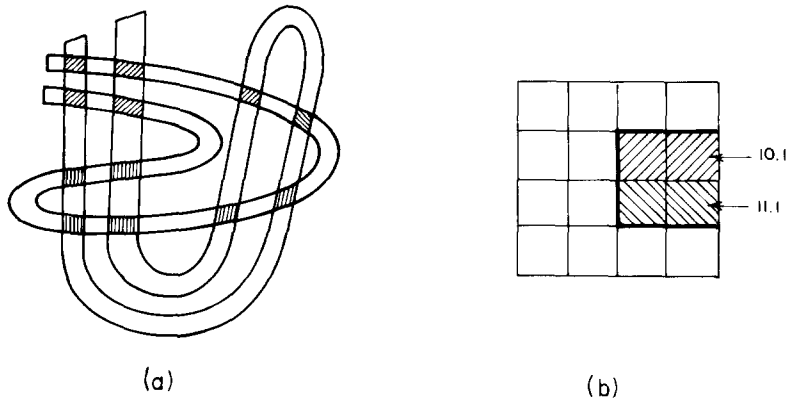

Fig. 3. (a) An incomplete Smale horseshoe: the inner forward fold does not intersect the rightmost backward folds. The backward folding in this figure and in fig. 4a is only schematic – in invertible mappings there are further missing intersections, all obtained by the forward and backward iterations of the primary pruned region. (b) The primary pruned region in the symbol plane and the corresponding forbidden binary substrings.
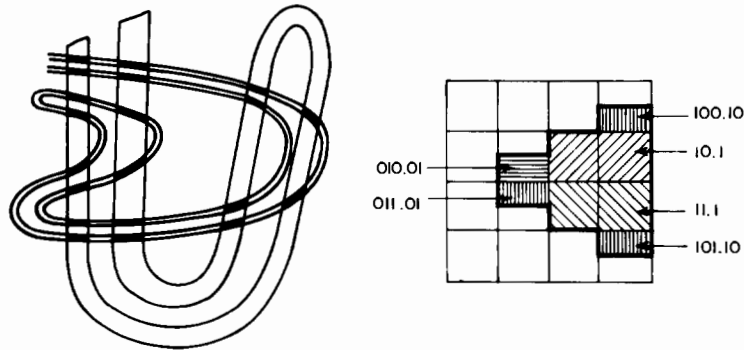
Fig. 4. An incomplete Smale horseshoe which illustrates the monotonicity of the pruning front, i.e. the thick line which delineates the left border of the primary pruned region.

front:

(1) The pruning front is symmetric across the horizontal $\frac{1}{2}$ line.

(2) The pruning front is *monotone* across either half of the symbol plane.

This is a consequence of the deterministic foliation; inner folds cannot pierce through the outer folds, and therefore have fewer transverse sections than the outer ones. The enumeration of all distinct allowed prunings is further considerably simplified by

(3) The *pruning front conjecture* [19][#1]: for once-folded maps the pruning front specifies *completely* the symbolic dynamics of the map, i.e. all pruned regions are obtained by forward and backward iterations of the primary pruned region, and there are no other independent pruning mechanisms.

The pruning front is a two-dimensional generalization of the one-dimensional kneading sequence; the location of each vertical step in the pruning front is the kneading sequence of the corresponding primary turnback of the unstable manifolds. If correct, the pruning front conjecture is a complete description of the symbolic dynamics of once-folding maps in the same sense in which the kneading sequence defines the symbolic dynamics of a one-dimensional unimodal map. The intuition behind this conjecture is that

the folding induced by a single iteration is the primary folding, and all other folds (turnbacks, homoclinic tangencies) are images or preimages of the primary ones.

3. Turning topology into symbolic dynamics

The pruning front fixes the topology of a once-folding map. If the pruning is a subshift of finite type (the symbolic dynamics can be presented as a finite alphabet together with a finite list of forbidden substrings), the topology can be converted into symbolic dynamics by means of a Markov diagram or a finite automaton [22, 18][#2].

For concreteness, take as an example the pruning of fig. 4. The pruned rectangles correspond to finite forbidden substrings _100.10_, _10.1_, _010.01_, _011.01_, _11.1_, _101.10_. Substrings _01101_, _10110_ contain the forbidden sub-block _101_, so they are redundant as pruning rules. Draw the *pruning tree* as a section of a binary tree with 0 and 1 branches and label each internal node by the sequence of 0's, 1's connecting it to the root of the tree (fig. 5a). These nodes are the potentially dangerous nodes – beginnings of substrings that might end up pruned. Add the side branches to those nodes (fig. 5b). As we continue down such branches we have to check whether the pruning tree imposes constraints on

---

[#1]Some numerical evidence for the correctness of the pruning front conjecture is given in refs. [12, 20].

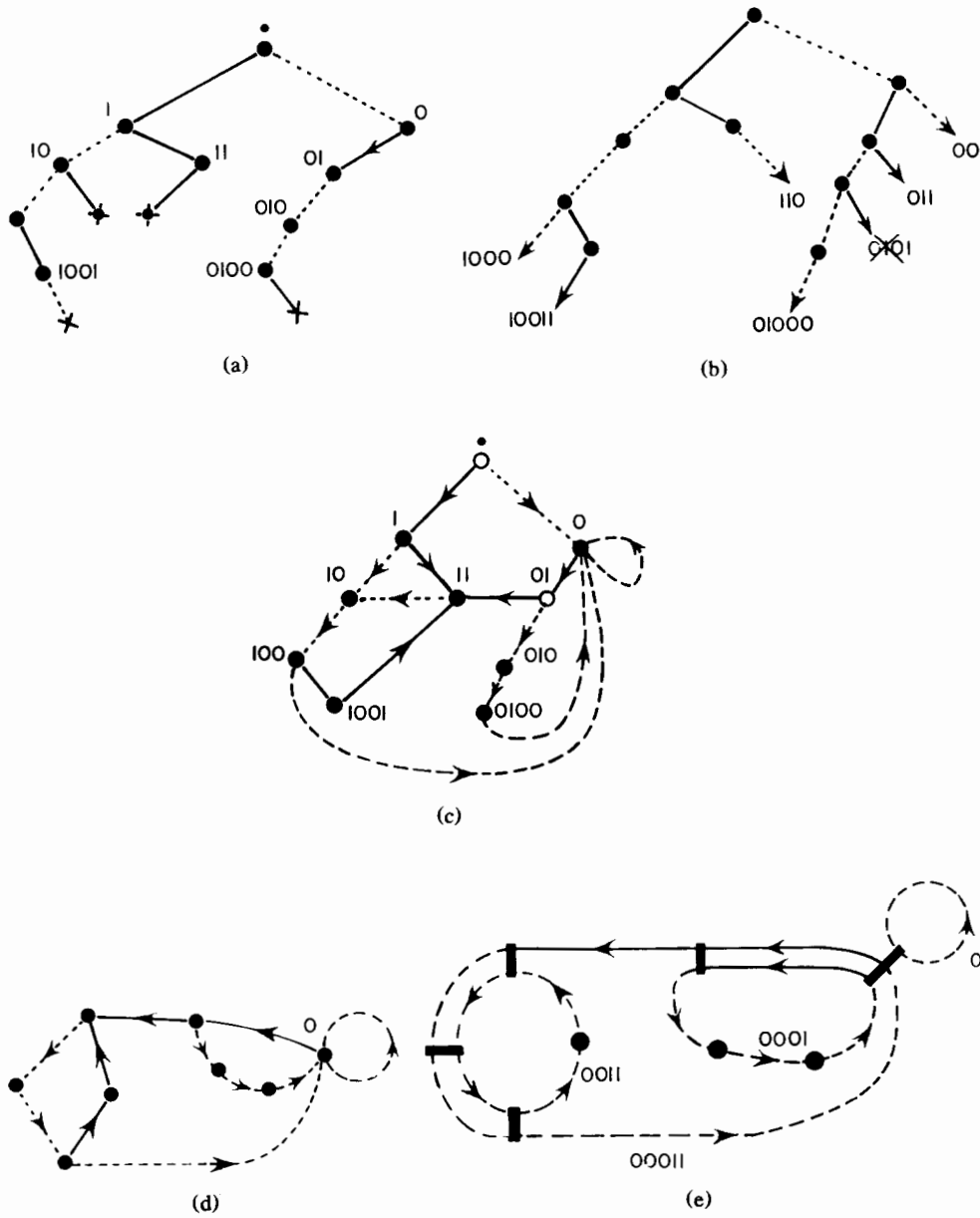[#2]We follow here Mats Nordahl's exposition [21].

Fig. 5. Conversion of the pruning front of fig. 4 into a finite automaton graph. (a) Starting with the start node ".", delineate all pruned substrings on the binary tree. Full line stands for "1" and the dashed line for "0". Ends of forbidden strings are marked with ×. Label all internal nodes by reading the bits connecting "." to the node. (b) Indicate all allowed starting substrings by arrows. (c) Drop recursively the leading bits in the allowed substrings; if the truncated string corresponds to an internal node in (a), connect them. (d) Delete all non-circulating nodes; all allowed sequences are generated as walks on this finite automaton graph. (e) Identify all distinct loops and construct the determinant (1).

the sequences so generated: we do this by knocking off the leading bits and checking whether the shortened strings coincide with any of the internal pruning tree nodes: $00 \rightarrow 0$; $110 \rightarrow 10$; $011 \rightarrow 11$; $011 \rightarrow 101$ (pruned); $1000 \rightarrow 00 \rightarrow 00 \rightarrow 0$; $10011 \rightarrow 0011 \rightarrow 011 \rightarrow 11$; $01000 \rightarrow 0$.

Now connect the side branches into the corresponding nodes (fig. 5c). Nodes "." and 1 can only serve as starting nodes; no sequence returns to them, and as we are here interested only in cycles (infinitely recurrent sequences), we delete them. The result is the finite automaton of fig. 5d; the allowed sequences are generated as all possible walks along the graph. Equivalently, such a graph can be replaced by an alphabet. For example, starting with node 0, all possible non-self-intersecting return paths are $\{0, 1000, 1100(1100)^k 0;$ $\overline{1100}\}$, and any path can be composed of these "letters". As $k = 0, 1, 2, 3, \ldots,$ the alphabet is infinite, but complete: any allowed string is a sequence of letters from the alphabet, with no further pruning rules.

The procedure described above is one possible way of generating subshifts of finite type. One could have, for example, listed instead all non-vanishing elements of a $2^4 \times 2^4$ transition matrix; this yields a larger, non-minimal Markov graph. The finite automaton approach is more compact, but it is still not entirely natural in the context of the present problem: the graphs constructed by the algorithm outlined here are not necessarily minimal, the graphs associated with time reversed dynamics do not reflect this symmetry, and so on. Be it as it may, the method requires only simple string shifts and matches, and is easily implemented on a computer [23].

4. From diagrams to determinants

A finite automaton like the one given in fig. 5d is a compact encoding of the transition or the Markov matrix for a given subshift. It is a sparse matrix, and the associated determinant can be written down by inspection [24]: it is the sum of all possible partitions of the graph into products of non-intersecting loops, with each loop carrying a minus sign:

$$\det(1 - T) = 1 - t_0 - t_{0011} - t_{0001} - t_{00011}$$
$$+ t_0 t_{0011} + t_{0011} t_{0001}. \quad (1)$$

The simplest application of this determinant is to the evaluation of the topological entropy; if we set $t_p = z^{n_p}$, where n_p is the length of the p-cycle, the smallest root of

$$0 = \frac{1}{\zeta_{\text{top}}} = 1 - z - 2z^4 + z^8 \quad (2)$$

yields the topological entropy $h = -\ln z$, $z = 0.658779\ldots$, $h = 0.417367\ldots$.

In what follows we shall need all cycles up to length n, so it is handy to know their number. N_n, the number of periodic points of period n, is related to the topological polynomial $1/\zeta$ by [7]

$$\sum_{n=1}^{\infty} N_n z^n = -z \frac{\mathrm{d}}{\mathrm{d}z} \ln \frac{1}{\zeta}. \quad (3)$$

In the above example

$$\sum_{n=1}^{\infty} N_n z^n = \frac{z + 8z^4 - 8z^8}{1 - z - 2z^4 + z^8}.$$

This yields $N_1 = N_2 = N_3 = 1$, $N_n = 2n + 1$ for $n = 4, 5, 6, 7, 8$ and $N_n = N_{n-1} + 2N_{n-4} - N_{n-8}$ for $n > 8$.

The number of *prime* cycles follows by Möbius inversion:

$$M_n = n^{-1} \sum_{d|n} \mu\left(\frac{n}{d}\right) N_d,$$

where the Möbius function $\mu(1) = 1$, $\mu(n) = 0$ if n has a squared factor, and $\mu(p_1 p_2 \ldots p_k) = (-1)^k$ if all prime factors are different. In the above example $M_1 = 1$, $M_2 = M_3 = 0$, $M_4 = M_5 = M_6 = M_7 = 2$, $M_8 = 3$, $M_9 = 5, \ldots$.

The determinant (1) is exact for the finite graph in fig. 5e; for the associated (infinite dimensional) transfer operator, it is the beginning of the cycle expansion [7] of the corresponding dynamical zeta function [6]:

$$\frac{1}{\zeta} = \det(1 - T) = 1 - t_0 - t_{0011} - t_{0001} + t_{0001}t_{0011}$$

$$- (t_{00011} - t_0 t_{0011} + \ldots \text{curvatures}) \ldots . \quad (4)$$

The cycles 0, 0001 and 0011 are the *fundamental* cycles; they are not shadowed by any combinations of shorter cycles, and are the basic building blocks of the Cantor set generated by iterating the pruned region of fig. 4. All other cycles appear together with their shadows (for example, $t_{00011} - t_0 t_{0011}$ combination is of that type) and yield exponentially small corrections for hyperbolic systems. Parenthetically, the finite automaton construction sketched here is not the minimal one; for example, the graph of fig. 5 does not generate only the fundamental cycles, but also some shadowed cycles, t_{00011} in this example.

5. Transfer operators

In the section above we have developed an elementary application of symbolic dynamics of a dynamical system; from a finite automaton we have computed the topological entropy, i.e. the average growth rate of the number of allowed orbits with increase in the symbolic string length. This is a special case of a general scheme for computing dynamical system averages in terms of periodic orbits [6]; as these methods are discussed at length elsewhere (see refs. [7, 8, 25]), here we only state the main result.

An average over a strange set, such as γ, the rate of escape from a d-dimensional repeller [26], given by the fraction of the initial volume that has not escaped by the time $n \to \infty$,

$$e^{-n\gamma} = \frac{\int_V dx \, dy \, \delta(y - f^{(n)}(x))}{\int_V dx}, \quad (5)$$

can be extracted from the eigenvalue spectrum of the *transfer operator*

$$\mathcal{L}(y, x) = \delta(y - f(x)) \quad (6)$$

(or its generalizations appropriate to other averages). Physically this spectrum is the correlation [27–29], resonances [30], quantum energy [31–36] spectrum, and so on. The spectrum is given by the zeros of the Fredholm determinant [6, 7] $\det(1 - z\mathcal{L})$ expressed in terms of the traces $\text{tr}\,\mathcal{L}^n$, i.e. the sums over all periodic points x_i of period n:

$$\text{tr}\,\mathcal{L}^n = \int dx \, dy \, \delta(x - y) \, \mathcal{L}^n(y, x)$$

$$= \sum_i^{(n)} \frac{1}{\left|\det\left[1 - J^{(n)}(x_i)\right]\right|}, \quad (7)$$

where

$$J^{(n)}(x_i) = \prod_{j=0}^{n-1} J(f^{(j)}(x_i)),$$

$$J_{kl}(x) = \frac{\partial}{\partial x_l} f_k(x) \quad (8)$$

is the i-cycle $d \times d$ Jacobian matrix, with eigenvalues $\Lambda_{i,1}, \Lambda_{i,2}, \ldots, \Lambda_{i,d}$. By factorizing the cycle determinants into products of expanding eigenvalues $\Lambda_{i,1}, \Lambda_{i,2}, \ldots, \Lambda_{i,e}$ and contracting eigenvalues $\Lambda_{i,e+1}, \ldots, \Lambda_{i,d-1}, \Lambda_{i,d}$,

$$\frac{1}{\left|\det\left[1 - J^{(n)}(x_i)\right]\right|} = \frac{1}{|\Lambda_i|} \prod_{a=1}^{d} \frac{1}{1 - u_{i,a}},$$

$$u_{i,a} = \Lambda_{i,a} \quad \text{if } |\Lambda_{i,a}| < 1,$$
$$= 1/\Lambda_{i,a} \quad \text{if } |\Lambda_{i,a}| > 1, \quad (9)$$

with $\Lambda_i = \prod_a^e \Lambda_{i,a}$ the product of the expanding eigenvalues, $\det(1 - z\mathcal{L})$ can be brought to a form of a generalized Selberg product

$$\det(1 - z\mathcal{L}) = Z(z)$$

$$= \prod_p \exp\left(-\sum_{r=1}^{\infty} \frac{t_p^r}{r d_{p,r}}\right)$$

$$= \prod_p \prod_{k_1 \dots k_d} \left(1 - t_p u_{p,1}^{k_1} u_{p,2}^{k_2} \dots u_{p,d}^{k_d}\right),$$

$$d_{p,r} = \prod_{a=1}^{d} \left(1 - u_{p,a}^r\right). \tag{10}$$

In the above, $t_p = z^{n_p}/|\Lambda_p|$ is a weight associated with a cycle, z is a formal parameter which keeps track of the topological cycle lengths, and 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. In what follows we shall often absorb z into the transfer operator: $z\mathcal{L} \to \mathcal{L}$, $z^{n_p} t_p \to t_p$.

What is called the "Selberg product" here could probably equally well be called the "Fredholm determinant" throughout this text. A minor distinction from the mathematical literature is that we do not restrict the determinants to the expanding eigendirections (as is done in refs. [37–39][*3]) and that we do not take recourse to the Fredholm theory [40], but manipulate the Selberg products as formal cycle expansions.

The simplest example of a Selberg product is associated with a 1D repelling map $f(x)$, monotone on two non-overlapping intervals,

$$f(x) = f_0(x), \quad x \in I_0,$$

$$= f_1(x), \quad x \in I_1, \tag{11}$$

with fixed points $f_0(x_0) = x_0$, $f_1(x_1) = x_1$ and $|f'_\epsilon(x)| > 1$. For 1D repellers the Selberg product

---

[*3]We are grateful to V. Baladi for explaining the distinction to us.

(10) can be written as

$$Z(z) = \prod_{k=0}^{\infty} \frac{1}{\zeta_k(z)}, \quad \frac{1}{\zeta_k} = \prod_p \left(1 - \frac{t_p}{\Lambda_p^k}\right), \tag{12}$$

where, for the escape rate example, $t_p = z^{n_p}/|\Lambda_p|$.

6. Cycle expansions

In numerical applications of the Selberg products (10) usually only a finite number of prime cycles is available. We expand [41] $Z(z)$ as a multinomial in prime cycle weights t_p (the functional form of the weight depends on the average under consideration; various particular cases are discussed in refs. [7, 8]):

$$Z(z) = \prod_p \sum_{k=0}^{\infty} C_{p^k} t_p^k = \sum_{k_1 k_2 k_3 \dots = 0}^{\infty} \tau_{p_1^{k_1} p_2^{k_2} p_3^{k_3} \dots}. \tag{13}$$

In the above we have defined

$$\tau_{p_1^{k_1} p_2^{k_2} p_3^{k_3} \dots} = \prod_{i=1}^{\infty} C_{p_i^{k_i}} t_{p_i}^{k_i}. \tag{14}$$

The coefficients C_{p^k} can be evaluated by expanding (10), $Z(z) = \prod_p Z_p$, where

$$Z_p = 1 - \left(\sum_{r=1}^{\infty} \frac{t_p^r}{r d_{p,r}}\right) + \frac{1}{2}\left(\sum_{r=1}^{\infty} \frac{t_p^r}{r d_{p,r}}\right)^2 - \dots.$$

Expanding and recollecting terms, and suppressing the p cycle label for the moment, we obtain

$$Z_p = \sum_{r=1}^{\infty} C_k t^k, \quad C_k = \frac{(-)^k c_k}{D_k},$$

$$D_k = \prod_{r=1}^{k} d_r = \prod_{a=1}^{d} \prod_{r=1}^{k} (1 - u_a^r), \tag{15}$$

where evaluation of c_k requires a certain amount

of unilluminating algebra:

$$c_0 = 1,$$
$$c_1 = 1,$$
$$c_2 = \frac{1}{2}\left(\frac{d_2}{d_1} - d_1\right)$$
$$= \frac{1}{2}\left(\prod_{a=1}^{d}(1 + u_a) - \prod_{a=1}^{d}(1 - u_a)\right),$$
$$c_3 = \frac{1}{3!}\left(\frac{d_2 d_3}{d_1^2} + 2d_1 d_2 - 3d_3\right),$$
$$= \frac{1}{6}\left(\prod_{a=1}^{d}\left(1 + 2u_a + 2u_a^2 + u_a^3\right)\right.$$
$$\left. + 2\prod_{a=1}^{d}\left(1 - u_a - u_a^2 + u_a^3\right) - 3\prod_{a=1}^{d}\left(1 - u_a^3\right)\right),$$

etc. For example, for a general two-dimensional map we have

$$Z_p = 1 - \frac{1}{D_1}t + \frac{u_1 + u_2}{D_2}t^2$$
$$- \frac{u_1 u_2(1 + u_1)(1 + u_2) + u_1^3 + u_2^3}{D_3}t^3 + \cdots.$$

$$(16)$$

The explicit form of the 1D and the Hamiltonian 2D cycle expansions is given in section 7, where we also discuss the convergence of such cycle expansions. The striking aspect of this cycle expansion is its resemblance to the factorization of natural numbers into primes. This is somewhat unexpected, as the cycle weights (for example, the stability eigenvalues (8)) factorize exactly with respect to r repetitions of a prime orbit, $t_{pp\ldots p} = t_p^r$, but only approximately (*shadowing*) with respect to subdividing a string into prime substrings, $t_{p_1 p_2} \approx t_{p_1} t_{p_2}$. However, with $\tau_{p_1^{k_1} p_2^{k_2} \ldots p_n^{k_n}}$ defined as above, the prime factorization of symbol strings is unique in the sense that *each symbol string can be written as a unique concatenation of prime strings*, up to a convention on ordering of primes.

To be more explicit, we illustrate the above by expressing binary strings as concatenations of prime factors. We start by computing N_n, the number of terms in the expansion (13) of the total cycle length n. Setting $C_k(\Lambda_p)t_p^k = z^{n_p k}$ in (13), we obtain

$$\sum_{n=0}^{\infty} N_n z^n = \prod_p \sum_{k=0}^{\infty} z^{n_p k} = \frac{1}{\prod_p(1 - z^{n_p})}.$$

The generating function for the number of terms in the Selberg product is the topological zeta function [42]. For the complete binary dynamics we have [7] $N_n = 2^n$ contributing terms of length n:

$$\zeta_{\text{top}} = \frac{1}{\prod_p(1 - z^{n_p})} = \frac{1}{1 - 2z} = \sum_{n=0}^{\infty} 2^n z^n.$$

Hence the number of distinct terms in the expansion (13) is the same as the number of binary strings, and conversely, the set of binary strings of length n suffices to label all terms of the total cycle length n in the expansion (13).

Next we implement the factorization by decomposing recusively binary strings into concatenations of prime strings. There are two strings of length 1, both prime: $p_1 = 0$, $p_2 = 1$. There are four strings of length 2: 00, 01, 11, 10. The first three are ordered concatenations of primes: $00 = p_1^2$, $01 = p_1 p_2$, $11 = p_2^2$; by ordered concatenations we mean that $p_1 p_2$ is legal, but $p_2 p_1$ is not. The remaining string is the only prime of length 2, $p_3 = 10$. Proceeding by discarding the strings which are concatenations of shorter primes $p_1^{k_1} p_2^{k_2} \ldots p_j^{k_j}$, with primes lexically ordered, we generate the standard list (table 1 of ref. [7]) of primes: 0, 1, 10, 101, 100, 1000, 1001, 1011, 10000, 10001, 10010, 10011, 10110, 10111, 100000, 100001, 100010, 100011, 100110, 100111, 101100, 101110, 101111, ... This factorization is illustrated in table 1.

How is this factorization used in practice? Suppose we have computed (or perhaps even measured in an experiment) all prime orbits up to length n, i.e. we have a list of t_p's and the corresponding Jacobian eigenvalues $\Lambda_{p,1}$,

Table 1
Factorization of all binary strings up to length 5 into ordered concatenations $p_1^{k_1} p_2^{k_2} \ldots p_n^{k_n}$ of prime strings $p_1 = 0$, $p_2 = 1$, $p_3 = 10$, $p_4 = 100, \ldots$.

Factors	String	Factors	String
p_1	0	p_1^5	00000
p_2	1	$p_1^4 p_2$	00001
p_1^2	00	$p_1^3 p_2^2$	00011
$p_1 p_2$	01	$p_1^2 p_2^3$	00111
p_2^2	11	$p_1 p_2^4$	01111
p_3	10	p_2^5	11111
p_1^3	000	$p_1^3 p_3$	00010
$p_1^2 p_2$	001	$p_1^2 p_2 p_3$	00110
$p_1 p_2^2$	011	$p_1 p_2^2 p_3$	01110
p_2^3	111	$p_2^3 p_3$	11110
$p_1 p_3$	010	$p_1 p_3^2$	01010
$p_2 p_3$	110	$p_2 p_3^2$	11010
p_4	100	$p_1^2 p_4$	00100
p_5	101	$p_1 p_2 p_4$	01100
		$p_2^2 p_4$	11100
		$p_3 p_4$	10100
p_1^4	0000	$p_1^2 p_5$	00101
$p_1^3 p_2$	0001	$p_1 p_2 p_5$	01101
$p_1^2 p_2^2$	0011	$p_2^2 p_5$	11101
$p_1 p_2^3$	0111	$p_3 p_5$	10101
p_2^4	1111	$p_1 p_6$	01000
$p_1^2 p_3$	0010	$p_2 p_6$	11000
$p_1 p_2 p_3$	0110	$p_1 p_7$	01001
$p_2^2 p_3$	1110	$p_2 p_7$	11001
p_3^2	1010	$p_1 p_8$	01011
$p_1 p_4$	0100	$p_2 p_8$	11011
$p_2 p_4$	1100	p_9	10000
$p_1 p_5$	0101	p_{10}	10001
$p_2 p_5$	1101	p_{11}	10010
p_6	1000	p_{12}	10011
p_7	1001	p_{13}	10110
p_8	1011	p_{14}	10111

$\Lambda_{p,2}, \ldots \Lambda_{p,d}$. A cycle expansion of the Selberg product is obtained by generating all strings up to length n allowed by the symbolic dynamics and constructing the multinomial (13)

$$Z(z) = \sum_{j=0}^{n} \sum_{\epsilon_1 \epsilon_2 \ldots \epsilon_j} \tau_{\epsilon_1 \epsilon_2 \ldots \epsilon_j} + \text{longer cycles}, \quad (17)$$

where ϵ_i range over the alphabet, in the present case $\{0, 1\}$. Factorizing every string $\epsilon_1 \epsilon_2 \ldots \epsilon_j = p_1^{k_1} p_2^{k_2} \ldots p_j^{k_j}$ as in table 1, and substituting $\tau_{p_1^{k_1} p_2^{k_2} \ldots}$ from (14) we obtain a multinomial ap-

proximation to $Z(z)$. For example, $\tau_{001001010101} = \tau_{\underline{001}\,\underline{001}\,\underline{01}\,\underline{01}\,\underline{01}} = \tau_{001^2} \tau_{01^3}$, and $\tau_{01^3}, \tau_{001^2}$ are known functions of the corresponding cycle eigenvalues, given in section 7. The zeros of $Z(z)$ can now be easily determined by standard numerical methods.

The fact that as far as the symbolic dynamics is concerned, the cycle expansion of a Selberg product is simply an average over all symbolic strings makes Selberg products rather pretty. However, compared with the cycle expansions of dynamical ζ functions, they have a small blemish; unlike the ζ functions (see for example (4)) they do not separate into a fundamental part and the (topologically vanishing) curvature corrections. For the dynamical ζ functions the curvature corrections are "shadowing" combinations, consisting of equal numbers of positive and negative terms [7]. The same counting for the Selberg product expansion (13) yields

$$\sum \overline{N_n} z^n = \prod_p \sum_{k=0}^{\infty} (-z^{n_p})^k = \frac{1}{\prod(1 + z^{n_p})}$$
$$= \prod \frac{1 - z^{n_p}}{1 - z^{2n_p}} = \frac{1 - 2z}{1 - 2z^2}$$
$$= 1 - 2z + 2z^2 - 4z^3 + 4z^4 - \ldots . \quad (18)$$

Here $\overline{N_n}$ is the difference in the number of the positive and negative terms in the expansion (13); there is an excess in every order, and the terms cannot be arranged into shadowing combinations. That was the reason why we in ref. [7] concentrated on cycle expansions of the dynamical ζ functions rather than Selberg products. However, as pointed out by Christiansen, Paladin and Rugh [29], the shadowing not only works for the Selberg product expansions, but does so with vengeance – the convergence with cycle length is faster than exponential. The difference is mathematically innocuous (for hyperbolic systems the Selberg products are entire functions; the dynamical ζ functions ratios of entire functions [38, 6]) but in numerical applications the difference is the difference between needing tens or thousands of cycles in order to attain the same accuracy [43].

7. Convergence of $Z(z)$ expansions

A qualitative understanding of the spectrum of a transfer operator can be obtained by considering the simplest non-trivial example, a piecewise linear version of the two-branch repeller (11). In this case the curvatures [7] in the cycle expansion of each ζ_k term in the product (12) vanish,

$$1/\zeta_k(z) = 1 - \left(t_0/\Lambda_0^k + t_1/\Lambda_1^k\right)z,$$

and the eigenvalues (the zeros of the Selberg product) are simply

$$\lambda_k = \frac{1}{t_0/\Lambda_0^k + t_1/\Lambda_1^k}. \tag{19}$$

Asymptotically the spectrum is dominated by the lesser of the two fixed point slopes $\lambda_k \approx \Lambda^k/t$ (here $\Lambda = \Lambda_0$ if $|\Lambda_0| < |\Lambda_1|$, otherwise $\Lambda = \Lambda_1$). In this approximation, the zeros of the Selberg product (12) are given by

$$Z(z) \approx \prod_{k=0}^{\infty} \left(1 - \frac{zt}{\Lambda^k}\right) \tag{20}$$

and the eigenvalues λ_k fall off exponentially as $1/\Lambda^k$. The physical implication of the dominance of the smallest of the cycle eigenvalues is that the higher eigenvalues in the spectrum will be more sensitive to presence of almost stable orbits. In contexts such as the Hénon map "strange attractors" and quasi-ergodic Hamiltonian systems such as the x^2y^2 potential, this means that while leading eigenvalues might converge well (for example, in the computation of the Hausdorff dimension of a strange set), the distribution of high eigenvalues is in principle unknowable, as it depends on the (structurally unstable) homoclinic tangencies and presence of arbitrarily miniscule islands of stability. From this point of view it is not at all clear that we should or can worry deeply about the asymptotic eigenvalue distributions, as is the pre-vailing trend in studies of quantum chaos [4]. Indeed, from the vantage point of the Selberg product expansions our main concern is development of accurate cycle expansions for the bottom of quantum spectra, and incorporation of corrections to the semi-classical approximations to the spectrum.

Physically interesting repellers are not piecewise linear, but the above two-slope approximation gives a rough sketch of the eigenvalue spectrum. Refinements are obtained by replacing the second iterate of the map by four linear segments, and so forth; that is precisely the meaning of the finite cycle length truncations of cycle expansions. The eigenvalues are either real or come in complex pairs – the main point is that for hyperbolic systems we expect them to be exponentially spaced [38].

An immediate consequence of the exponential spacing of the eigenvalues is that the convergence of the Selberg product expansion (13) as function of the topological cycle length, $Z(z) = \sum_n C_n z^n$, is faster than exponential. Consider a d-dimensional map for which all Jacobian eigenvalues in (8) are equal: $u_p = \Lambda_{p,1} = \Lambda_{p,2} = \ldots = \Lambda_{p,d}$. Clearly the stability eigenvalues are generally not isotropic; the goal here is to obtain qualitative bounds on the spectrum, by replacing all stability eigenvalues with the least expanding one. In this case the p-cycle contribution to the product (10) reduces to

$$Z_p(z) = \prod_{k_1 \ldots k_d = 0}^{\infty} \left(1 - t_p u_p^{k_1 + k_2 + \ldots + k_d}\right)$$

$$= \prod_{k=0}^{\infty} \left(1 + t_p u_p^k\right)^{m_k},$$

$$m_k = \binom{d-1+k}{d-1} = \frac{(k+d-1)!}{k!(d-1)!}$$

$$= \prod_{k=0}^{\infty} \sum_{l=0}^{m_k} \binom{m_k}{l}\left(-u_p^k t_p\right)^l. \tag{21}$$

In one dimension the expansion can be given in closed form [44]:

$$\prod_{k=0}^{\infty} \left(1 + tu^k\right)$$

$$= 1 + \frac{t}{1-u} + \frac{t^2 u}{(1-u)(1-u^2)}$$

$$+ \frac{t^3 u^3}{(1-u)(1-u^2)(1-u^3)} + \dots$$

$$= \sum_{k=0}^{\infty} t^k \frac{u^{k(k-1)/2}}{(1-u)\dots(1-u^k)}, \quad |u| < 1, \quad (22)$$

and the coefficients C_k in (13) are given by

$$\tau_{p^k} = (-1)^k \frac{u_p^{k(k-1)/2}}{\prod_{j=1}^{k}\left(1 - u_p^j\right)} t_p^k. \quad (23)$$

By this estimate the coefficients in the $Z(z) = \sum_n C_n z^n$ expansion of the Selberg product (12) should fall off faster than exponentially, as $|C_n| \approx u^{n(n-1)/2}$. In contrast, the cycle expansions [8] of dynamical zeta functions fall of "only" exponentially; in numerical applications, the difference is significant [43].

In higher dimensions the expansions are not quite as compact. The leading power of u and its coefficient are easily evaluated by use of binomial expansions (21) of the $(1 + tu^k)^{m_k}$ factors. More precisely, the leading u^n terms in t^k coefficients are of the form [46]

$$\prod_{k=0}^{\infty} \left(1 + tu^k\right)^{m_k}$$

$$= \dots + u^{m_1 + 2m_2 + \dots + jm_j} t^{1+m_1+m_2+\dots+m_j} + \dots$$

$$= \dots + \left(u^{m_d/(d+1)} t\right)^{\binom{d+m}{m}} + \dots$$

$$\approx \cdots + u^{[\sqrt{d!}/(d-1)!]n^{(d+1)/d}} t^n + \dots.$$

Hence the coefficients in the $Z(z)$ expansion fall off faster than exponentially, as $u^{n^{1+1/d}}$, in agreement with the estimates of Fried [39] for the Fredholm determinants of d-dimensional expanding flows. The Selberg products are entire func-

tions in any dimension, provided that the symbolic dynamics is a finite subshift, and all cycle eigenvalues are bounded sufficiently away from 1.

The case of particular interest are the 2D Hamiltonian mappings; their symplectic structure implies that $u_p = \Lambda_{p,1} = 1/\Lambda_{p,2}$, and the Selberg product (10) is the two-dimensional case of (21),

$$Z(z) = \prod_p \prod_{k=0}^{\infty} \left(1 - t_p u_p^k\right)^{k+1}. \quad (24)$$

In this case the expansion (16) is given by [41]

$$\prod_{k=0}^{\infty} \left(1 + tu^k\right)^{k+1}$$

$$= \sum_{k=0}^{\infty} \frac{F_k(u)}{(1-u)^2(1-u^2)^2\dots(1-u^k)^2} t^k$$

$$= 1 + \frac{1}{(1-u)^2} t + \frac{2u}{(1-u)^2(1-u^2)^2} t^2$$

$$+ \frac{u^2(1 + 4u + u^2)}{(1-u)^2(1-u^2)^2(1-u^3)^2} t^3 + \dots.$$

$$(25)$$

$F_k(u)$ is a polynomial in u, and the coefficients fall off asymptotically as $C_n \approx u^{n^{3/2}}$.

The technology developed above for the classical Selberg products can be taken pretty much in toto over to the quantum Selberg products [32, 31]. For the Hamiltonian flows, the quantum square root $1/\sqrt{\det(\mathbf{1} - \mathbf{J}_p)}$ weight for orbits leads to somewhat prettier Selberg products; for example, (24) is in the quantum case replaced by

$$Z(z) = \prod_p \prod_{k=0}^{\infty} \left(1 - \frac{e^{-iS_p/\hbar + \nu_p}}{\sqrt{|\Lambda_p|}} u_p^k\right). \quad (26)$$

For the low eigenvalues we expect the convergence of cycle expansions to be as good as in the

classical case. However, unlike the classical product (24), which is exact, the quantum Selberg product (26) is the leading term of a semiclassical approximation, and the size of corrections to it is not known.

8. Homework assignment

The above coarse approximations suffice for establishing the main result of this section, that the coefficients in the cycle expansions of Selberg products fall off faster than exponentially. The spectrum of $\mathscr{L}$ for the piecewise-linear approximate maps is only indicative of the spectrum for the exact non-linear map; the details are subtle and the reader is referred to refs. [38, 39] for more careful convergence estimates. In particular, our estimates depended on the assumption that the symbolic dynamics is a subshift of finite type, that the cycle weight is multiplicative along the flow, and that the flow is smooth, so that nearby trajectories have nearby weights.

In the above we have reviewed the general properties of the cycle expansions; those have been developed at some length in ref. [7] and applied in ref. [8] to a series of examples of low-dimensional chaos: 1D strange attractors, the period-doubling repeller [43], the Hénon-type maps and the mode locking intervals for circle maps. The Selberg product expansion method [29] used here is superior to the method of locating zeros of finite products of dynamical ζ functions [6] of refs. [7, 8], and in that sense the present note supercedes the above references. The cycle expansions have also been applied to the irrational windings set of critical circle maps [45], to the Hamiltonian period-doubling repeller [47], to a Hamiltonian three-disk pinball [35], to the three-disk quantum scattering resonances [34, 36, 41] and to the extraction of correlation exponents [29]. Feasibility of analysis of experimental strange sets in terms of cycles is discussed in ref. [48].

Elsewhere in this conference, Smale has given a list of ten outstanding problems of the dynami-

cal systems theory. As the ostensible topic of this conference is the non-linear science of the next decade (maybe wonderful, but unlikely to resemble much the conjurings of today), I conclude cautiously with a homework assignment (due date May 22, 2000):

1. Topology. Develop optimal sequences ("continued fraction approximants") of finite subshift approximations to generic dynamical systems. Apply to (a) the Hénon map, (b) the Lorentz flow and (c) the Hamiltonian standard map.

2. Non-hyperbolicity. Incorporate power-law (marginal stability orbits, "intermittency") corrections into cycle expansions. Apply to long-time tails in the Hamiltonian diffusion problem.

3. Phenomenology. Carry through a convincing analysis of a genuine experimentally extracted data set in terms of periodic orbits.

4. Invariants. Prove that the scaling functions, or the cycles, or the spectrum of a transfer operator are the maximal set of invariants of an (physically interesting) dynamically generated strange set.

5. Field theory. Develop a periodic orbit theory of systems with many unstable degrees of freedom. Apply to (a) coupled lattices, (b) cellular automata, (c) neural networks.

6. Tunneling. Add complex time orbits to quantum mechanical cycle expansions (WKB theory for chaotic systems).

7. Unitarity. Evaluate corrections to the Gutzwiller semiclassical periodic orbit sums. (a) Show that the zeros (energy eigenvalues) of the appropriate Selberg products are real. (b) Find physically realistic systems for which the "semiclassical" period orbit expansions yield the exact quantization.

8. Atomic spectra. Compute the helium spectrum from periodic orbit expansions.

9. Symmetries. Include fermions, gauge fields into the periodic orbit theory.

10. Quantum field theory. Develop quantum theory of systems with infinitely many classically unstable degrees of freedom. Apply to (a) quark confinement, (b) the early universe, (c) the brain.

Acknowledgements

References

[1] H. Poincaré, Les Méthodes Nouvelles de la Méchanique Céleste (Guthier-Villars, Paris, 1892–99).
[2] M.C. Gutzwiller, J. Math. Phys. 8 (1967) 1979; 10 (1969) 1004; 11 (1970) 1791; 12 (1971) 343.
[3] R. Balian and C. Bloch, Ann. Phys. (NY) 85 (1974) 514.
[4] M.V. Berry and K.E. Mount, Rep. Prog. Phys. 35 (1972) 315.
[5] R.S. MacKay and J.D. Miess, Hamiltonian Dynamical Systems (Adam Hilger, Bristol, 1987).
[6] D. Ruelle, Statistical Mechanics, Thermodynamic Formalism (Addison-Wesley, Reading, MA, 1978).
[7] R. Artuso, E. Aurell and P. Cvitanović, Nonlinearity 3 (1990) 325.
[8] R. Artuso, E. Aurell and P. Cvitanović, Nonlinearity 3 (1990) 361.
[9] S. Smale, Bull. Am. Math. Soc. 73 (1967) 747
[10] J. Guckenheimer and P. Holmes, Non-linear Oscillations, Dynamical Systems and Bifurcations of Vector Fields (Springer, New York, 1986).
[11] C. Mira, Chaotic Dynamics (World Scientific, Singapore, 1987).
[12] P. Cvitanović, G.H. Gunaratne and I. Procaccia, Phys. Rev. A 38 (1988) 1503.
[13] P.J. Myrberg, Ann. Acad. Sc. Fenn., Ser. A 259 (1958) 1.
[14] A. Sharkovskii, Ukr. Mat. Z. 16 (1964).
[15] N. Metropolis, M.L. Stein and P.R. Stein, J. Comb. Theory A 15 (1973) 25.
[16] J. Milnor and W. Thurston, On iterated maps of the interval, in: Dynamical Systems, Proceedings, Univ. of Maryland 1986–87, eds. A. Dold and B. Eckmann, Springer Lecture Notes in Mathematics 1342 (Springer, Berlin, 1988) p. 465.
[17] P. Collet and J.-P. Eckmann, Iterated Maps on the Interval as Dynamical Systems (Birkhauser, Boston, 1980).
[18] P. Grassberger, Z. Naturforsch. 43a (1988) 671.
[19] P. Cvitanović, in preparation.
[20] P. Grassberger, H. Kantz and U. Moening, J. Phys. A 22 (1989) 5217.
[21] M. Nordahl, unpublished.
[22] J.E. Hopcroft and J.D. Ullman, Introduction to Automata Theory, Languages, and Computation (Addison-Wesley, Reading, MA, 1979).
[23] K.T. Hansen, unpublished.
[24] D.M. Cvetković, M. Doob and H. Sachs, Spectra of Graphs (Academic Press, New York, 1980).
[25] P. Cvitanović, in: Nonlinear Physical Phenomena, Brasilia 1989 Winter School, eds. À. Ferraz, F. Oliveira and R. Osorio (World Scientific, Singapore, 1990).
[26] L. Kadanoff and C. Tang, Proc. Natl. Acad. Sci. USA 81 (1984) 1276.
[27] D. Ruelle, J. Stat. Phys. 44 (1986) 281.
[28] W. Parry and M. Pollicott, Ann. Math. 118 (1983) 573.
[29] F. Christiansen, G. Paladin and H.H. Rugh, Phys. Rev. Lett. 65 (1990) 2087.
[30] V. Baladi, J.-P. Eckmann and D. Ruelle, Nonlinearity 2 (1989) 119.
[31] M.C. Gutzwiller, J. Phys. Chem. 92 (1984) 3154.
[32] A. Voros, J. Phys. A 21 (1988) 685.
[33] P. Gaspard and S.A. Rice, J. Chem. Phys. 90 (1989) 2225; 90 (1989) 2242; 90 (1989) 2255.
[34] P. Cvitanović and B. Eckhardt, Phys. Rev. Lett. 63 (1989) 823.
[35] P. Cvitanović and B. Eckhardt, J. Phys. A 24 (1991) L237.
[36] P. Cvitanović, B. Eckhardt and P. Scherer, in preparation.
[37] R. Bowen, Equilibrium States and the Ergodic Theory of Anosov-Diffeomorphisms, Springer Lecture Notes in Mathematics Vol. 470 (1975).
[38] D. Ruelle, Invent. Math. 34 (1976) 231.
[39] D. Fried, Ann. Scient. Ec. Norm. Sup. 19 (1986) 491.
[40] A. Grothendieck, La théorie de Fredholm, Bull. Soc. Math. France 84 (1956) 319.
[41] G. Russberg, in preparation.
[42] E. Artin and B. Mazur, Ann. Math. 81 (1965) 82.
[43] F. Christiansen, P. Cvitanović and H.H. Rugh, J. Phys. A 23 (1990) L713.
[44] L. Euler, Opera Omnia (Teubner, 1922).
[45] P. Cvitanović, G.H. Gunaratne and M.J. Vinson, Nonlinearity 3 (1990) 873.
[46] H.H. Rugh, unpublished.
[47] E. Aurell, Göteborg preprint 89-10, submitted to Phys. Rev. A.
[48] P. Cvitanović, Phys. Rev. Lett. 61 (1988) 2729.