INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE CENTRATOM TRIESTE

SMR.764 - 7

RESEARCH WORKSHOP ON CONDENSED MATTER PHYSICS
13 JUNE - 19 AUGUST 1994

MINIWORKSHOP ON
"NONLINEAR TIME SERIES ANALYSIS"
8 - 12 AUGUST 1994

"Symbolic Dynamics and the Description of Complexity"

HAO Bai-Lin
Institute of Theoretical Physics
Academia Sinica
P.O. Box 2735
Beijing 100080
People's Republic of China

---

and the Description of Complexity

Bai-lin HAO

Symbolic dynamics provides a general framework to describe complexity of dynamical behaviour. After a discussion of the state of the field special emphasis will be given to the role of the transfer matrix (the Stefan matrix) both in deriving the grammar from known symbolic dynamics and in extracting the rules from experimental data. The block structure of the Stefan matrix may serve as another indicator of the complexity of the associated dynamics.

1 Introduction

It is a commonplace nowadays that many physical systems are capable of undergoing sharp transitions to more organized states when control parameters are tuned away from the trivial, near equilibrium, or linear, regimes. Intuitively speaking, these organized states are characterized by lower symmetry, lower entropy, more information, and a higher degree of complexity. By the way, it is the change of symmetry, explicit or hidden, that makes the transition sharp, because symmetry is a property that must either be present or absent; it cannot be accumulated gradually. However, when one comes to the stage of characterizing complexity, we see that this is a far from trivial task.

First of all, complexity is a notion which has been used in so many different contexts and which can hardly be defined in general. On 10th September, 1992, there were 302 books in the U. S. Library of Congress that have the word "complexity" in their title[3]. The most frequent usages include algorithmic complexity (AC), biological complexity, computational complexity (CC), developmental complexity, ecological complexity, economical complexity, evolutional complexity, grammatical complexity (GC), language complexity, etc.

Second, simplicity and reductionism has been a guideline in science. Many scientists believe that the fundamental laws of Nature must be simple. Although they encounter "complex" phenomena everyday and everywhere, they still try hard to reduce them to something simpler. Indeed, science knows a number of ways by which simple things can get more complex. For, example:

1. Projection onto lower-dimensional space may make things look more complex or, put the other way around, adding new dimensions, sometimes in the

[3] We thank Dr. Ming-zhou Ding, Florida Atlantic University, for checking the number.

Springer Series in Synergetics, Vol. 61 **On Self-Organization**
Eds.: R.K. Mishra, D. Maaß, E. Zwierlein

parameter space, sometimes in the configuration space, may simplify the description. Some nonlinear problems may be embedded into higher dimensions as linear ones, some non-Markovian processes may be made Markovian by adding more stochastic variables; even a discretized version of a continuous model may turn out to be more complex.

2. Repeated use of simple rules may lead to more complex behaviour. Everyone knows that iteration of a quadratic polynomial may yield chaos, iteration of "complex" (yet another context of the word) maps may produce the beautiful patterns of the Julia and Mandelbrot sets. Simple nearest neighbour rules of cellular automata may simulate the complexity of universal computers, as was conjectured by S. Wolfram.

Furthermore, complexity appears in our description of Nature simply because we are not clever enough. For instance, the use of a wrong system of reference may bring about unnecessary complications, as was the case with the Ptolemy system compared to Copernicus. The modern notation of the Maxwell equations for the electromagnetic fields, e.g., using the notion of exterior differential d, codifferential δ and differential 2-form Θ:

$$\begin{aligned} d\Theta &= 0, \\ \delta\Theta &= J, \end{aligned} \tag{1}$$

look much "simpler" than the original ones in Maxwell's paper in *Philosophical Transaction* or in his *Treatise*. In order to view things simply, you must stand high.

We see that the problem of complexity and its description are much like the problem of beauty and the appreciation of the beautiful. One needs a definite framework into which to set the problem and an objective way to estimate the complexity. Historically, the quantitative description of information has experienced similar problems. In a sense, it is a correct attitude to start with simple situations, where complex behaviour is generated by a comprehensible mechanism. If we cannot deal with these situations, there is little hope of coping with "real" complexities.

Comparison of various classes of orbits in one-dimensional mappings provides a first challenge toward this goal. Here is the arena where symbolic dynamics and its natural relation to formal language theory may play an instructive role. This development comes concurrently with the need to characterize various chaotic attractors at a more fundamental, "microscopic", level.

In recent years there has been a new upsurge of interest in quantifying and measuring complexity. Many meetings have been devoted to this problem, see, e.g., [1]-[6]. In particular, the complexity of chaotic attractors and trajectories have been the subject of many discussions. A number of new definitions of complexity have been suggested. We have in mind the use of grammatical complexity to characterize the complexity of cellular automata by Wolfram [7], the extensive study of grammatical complexity of symbolic sequences in one-dimensional maps and in cellular automata by Grassberger and his definition of "set complexity", "true measure complexity", "effective measure complexity" [8], the construction

of logical trees and calculation of "grammatical complexity" in a narrower sense in [9, 10], and the "hierachical approach" in [11], etc. An essential advance is the understanding that some conventional measures of complexity, e.g., entropy, are in fact measures of randomness, not complexity. Although periodicity and pure randomness are two extremes on the scale of randomness, they are close to each other on the scale of complexity. Complexity is always associated with a certain degree of organization and lies somewhere between simple periodicity and pure randomness on the scale of randomness.

Nevertheless, we would like to note that the state of the field is far from satisfactory. For example, some of the definitions mentioned above assign infinite complexity to the orbits at the accumulation point of the period-doubling sequence of one-dimensional maps [8, 12], while we know that the orbits there are only quasiperiodic and the symbolic dynamics is quite simple. The definition in [10] gives finite complexity to the limit of so-called Fibonacci sequence, i.e., the sequence of orbits, whose periods grow as Fibonacci numbers. This limit is also a quasiperiodic orbit and the symbolic dynamics is even more simple. We will return to this issue later.

To provide a remedy for these shortcomings, some authors introduced hierachical definitions which yield non-zero complexity only for systems in higher than one dimensions [12], hence put aside the whole problem of describing complexity of orbits in one-dimensional maps. Anyway, one-dimensional dynamics may be a result of projection from higher dimensions and it may be even more complex.

Furthermore, while there are convincing arguments [8] that there must be infinitely many orbits in one-dimensional maps whose complexity should go beyond "regular language", the lowest step in the Chomsky ladder of language complexity [17], no explicit way of constructing or approaching these orbits has been indicated so far. On the other hand, some seemingly simple languages such as R^nL^n or $R^nM^nL^n$ are more complex than regular (the former being non-regular, and the latter non-context-free), but they are not admissible in any known symbolic dynamics for one-dimensional maps.

We will not go into details of any of the definitions mentioned above, nor analyze their pros and cons. Our main aim is to stress the usefulness of transfer matrices as a bridge between the dynamics and the underlying grammar. Since the most workable definitions of complexity are related to symbolic dynamics and our discussion of grammatical complexity will also be restricted to the context of symbolic dynamics, a few words on symbolic dynamics may be in order.

2 Symbolic Dynamics

Symbolic dynamics is a coarse-grained description of dynamics. Instead of tracing the trajectories in the phase space in full detail, one divides ("coarse-grains") the phase space into a number of regions and labels each region with a letter from some alphabet. The total number may be finite or infinite; only the finite case can be studied in any great detail. The time evolution of a trajectory shows up as a series of letters; a numerical trajectory is replaced by a symbolic sequence. The correspondence may be many-to-one, thus opening the possibility for classification of trajectories. For theorem-proving, the exact way of partitioning does not matter in most cases. However, when one divides the phase space according to the "physics" or "geometry" of the dynamics, many detailed rules may be derived in the case of one- or two-dimensional mappings.

In a sense, symbolic dynamics is nothing but what experimental-physicists do every day. Using an analog-digital converter of, say, 12 bits precision, there are no more than 4096 different readings, i.e., symbols, from the instrument, yet one intends to draw reliable conclusions on invariant, robust properties of the system.

Symbolic dynamics as a chapter in abstract mathematics has a long history from the work of Morse in the 1920s; see [18, 19] for the development until the late 1970s. Applied aspects of symbolic dynamics have been developed mainly since the 1973 paper of Metropolis et al. [20] through the work of many mathematicians and physicists, see, e.g., [13, 15, 23]. Our group at the Institute of Theoretical Physics, Academia Sinica, Beijing, has made some contributions to symbolic dynamics of one-dimensional [21, 22] and two-dimensional [26]-[29] mappings, and in their application to systems, described by ordinary differential equations [21].

The essence of symbolic dynamics is rather simple. I will take a simple example to demonstrate some of the most important ingredients.

In the simplest case of unimodal maps, one divides the phase space (the interval) into two regions: one to the *L*eft of the *C*entral or *C*ritical point, labeled by L, another to the *R*ight of C, labeled by R. These three letters have a natural order

$$L < C < R, \tag{2}$$

which is the basis for ordering all possible symbolic sequences. The ordering rule is simple. Suppose two symbolic sequences

$$\Sigma_1 = \Sigma^* \sigma \cdots$$

and

$$\Sigma_2 = \Sigma^* \tau \cdots$$

have a common leading part Σ^*, and the next letters σ and τ are different. Since they are different, they must have been ordered according to the natural order (2). Then the order of Σ_1 and Σ_2 are the same as σ and τ if the common leading part Σ^* contains an even number of the letter R; or the order be reversed if there is an odd number of R in Σ^*. This rule holds for any one-dimensional map, not only for unimodal ones, provided the R counting in Σ^* extends to counting all letters which represent a decreasing branch of the mapping function.

Although the letter C occurs only at one point, the sequence starting with C, or, to be more precise, starting with the first iterate of C, i.e., $f(C)$, where $f(x)$ is the mapping function, plays an important role. It has acquired a special name "kneading sequence" [13]. A map is best parametrized by its kneading sequence

(or sequences, when there is more than one critical point). Given the kneading sequences, everything about the symbolic dynamics is determined.

Take, for example, a period 5 kneading sequence $(RLLRC)^\infty$ in the unimodal map. This is the symbolic sequence of the superstable period 5 orbit which starts from the critical point C of the map. It is a good convention to denote a symbolic sequence by the first number that starts the iteration. Since iterations of the map correspond to consecutive shifts of the symbolic sequence, we have the following alternation of numbers-sequences:

$$\begin{aligned} x_0 &\equiv C = CRLRRC\cdots, \\ x_1 &= RLRRC\cdots, \\ x_2 &= LRRC\cdots, \\ x_3 &= RRC\cdots, \\ x_4 &= RC\cdots, \end{aligned} \tag{3}$$

then it repeats. Using the ordering rule of symbolic sequences, it is easy to check that these sequences, and consequently the corresponding points, are ordered as follows:

$$x_2 < x_0 < x_3 < x_4 < x_1. \tag{4}$$

Besides the ordering rule, another important issue is the admissibility condition. Obviously, not every arbitrarily chosen symbolic sequence can correspond to a realizable trajectory in a given dynamics. One needs some criterion to check the admissibility. Referring to, e.g., [22]–[23], for a detailed formulation of the admissibility condition, we note only that in case of unimodal maps the condition reduces to shift-maximality of the symbolic sequences.

3 Coarse-Grained Chaos

A good example of how symbolic dynamics embodies the idea that complexity lies somewhere in between simple periodicity and pure randomness is the notion of coarse-grained chaos, based on a generalized composition rule [25].

We start from a superstable fixed point, which corresponds to the kneading sequence C^∞. By disturbing C a little, it goes either to R or to L, according to the natural order (2). Omitting the infinite power in the notations, we have a symbolic fixed point window:

$$(L, C, R). \tag{5}$$

Applying repeatedly the substitutions

$$\begin{aligned} R &\to RL, \\ C &\to RC, \\ L &\to RR, \end{aligned} \tag{6}$$

to the fixed point window (5), we get the symbolic representation of the whole period-doubling cascade.

The substitutions (6) hint at a generalization: find the conditions that the following substitutions

$$R \to \rho,$$
$$L \to \lambda, \tag{7}$$

ρ and λ being strings made of R and L, applied to any admissible symbolic sequence, would yield another admissible sequence. These conditions were found in [25] and they happened to be a natural generalization of the well-known $*$–composition in symbolic dynamics [15], hence the name generalized composition rule.

Now take the much-studied case of chaotic maps, namely, the map

$$x_{n+1} = 1 - 2x_n^2. \tag{8}$$

Its kneading sequence is RL^∞. It has many nice properties, for example:

1. There is a continuous distribution $\rho(x)$ for the orbital points x_i for almost all choices of initial points;
2. Each initial point leads to a different symbolic sequence; there are as many different symbolic sequences as real numbers in the interval $(-1, 1)$ (This is the symbolic statement of sensitive dependence on initial conditions);
3. There exist homoclinic orbits;
4. It is a crisis point;
5. It is a band-ending point, beyond which a chaotic band no longer exists;
6. It is a surjective map;

and so on.

Applying the generalized composition rule (7) to the kneading sequence RL^∞, we get infinitely many kneading sequences of the form $\rho\lambda^\infty$ for infinitely many different choices of ρ and λ. Almost everything said about the map with kneading sequence RL^∞ may be carried over to maps with kneading sequences $\rho\lambda^\infty$. The RL^∞ map is the most random map with maximal topological entropy. The $\rho\lambda^\infty$ maps have lower entropy, but if viewed with lower resolution, i.e., taking each of the strings ρ and λ for a single letter, they are as random as the RL^∞ map. However, if we look at them with higher resolution, we see patterns of organization, embodied in the repeated structures of ρ and λ. This is what we call coarse-grained chaos.

Now we have been prepared enough to introduce the transfer matrices.

4 Stefan Matrix

The main message I would like to convey [14] is that the transfer matrix, or Stefan matrix, as it is often called in this case [15], may be very useful both in deriving the grammatical rules from a given symbolic dynamics and in inferring the unknown rules from experimental data. Moreover, the Stefan matrix has a direct relation to the transfer functions used in constructing the non-deterministic finite automaton (NDFA), as well as the corresponding deterministic finite automaton (DFA), which accepts the language. The block structure of the matrix

may serve as another indicator of complexity for the automaton and the language. In fact, the topological entropy calculation makes use only of the largest eigenvalue of the matrix, the construction of the automaton utilizes more information in the matrix. I will devote the rest of this paper to the explanation of what has been said.

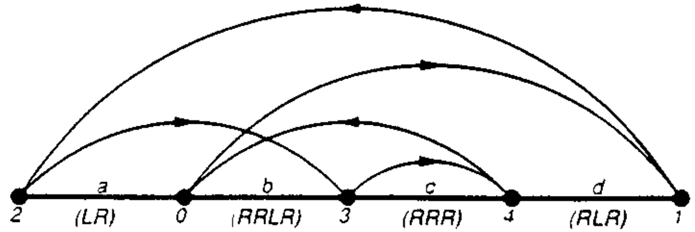

Fig. 1. Construction of the Stefan matrix for the period 5 kneading sequence $(RLRRC)^\infty$.

Although we are more concerned with the characterization of infinite aperiodic sequences, we continue with the example of the period 5 kneading sequence $(RLRRC)^\infty$ for simplicity of presentation. In fact, we will try to construct everything using this example, then the generalizations needed for infinite sequences will become clearer.

First of all, the ordered numbers in (4) are shown in Fig. 1. Also shown in the figure is the way the orbit visits these points. The latter divide the interval into four segments, denoted by a, b, c, and d. If one shifts the initial point of iteration from the superstable orbital points to other points in the interval, then continuity considerations alone tell us that the four segments will transform into one another in the following way:

$$\begin{aligned} a &\to c + d, \\ b &\to d, \\ c &\to b + c, \\ d &\to a. \end{aligned} \tag{9}$$

Written in matrix form, they define the Stefan matrix

$$S = \begin{vmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix}. \tag{10}$$

The trace of S^n in the $n \to \infty$ limit, hence the largest eigenvalue of S, gives the number of different periodic orbits of length n. In fact, the logarithm of the largest eigenvalue, in our example $\lambda = 1.51288$, yields the topological entropy [16], which, as we have said, is rather a measure of randomness, not complexity. However, we will see that other approaches of relating complexity to the underlying grammar of the symbolic sequences just make more use of the same Stefan matrix, going beyond its largest eigenvalue.

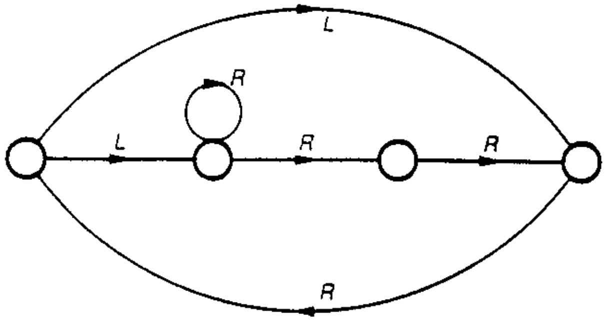

Fig. 2. Non-deterministic finite automaton constructed from the Stefan matrix of the period 5 kneading sequence $(RLRRC)^\infty$.

At each of the numbered points in Fig. 1 we have one of the symbolic sequences given in (3). They contain the letter C. By shifting from C to the left or to the right, we get two symbolic sequences, located on the two sides of the numbered points according to the ordering rule. By comparing the sequences at the two ends of each segment, we get the first few symbols for any sequences starting from that segment. These are the words (LR), $(RRLR)$, (RRR), and (RLR), written under a, b, c, and d, respectively, in Fig. 1. From the meaning of the transfer matrix we deduce that these 4 words appear in the following 6 contexts: $LRRR$, $LRLR$, $RRLR$, $RRRLR$, $RRRR$, and RLR, each corresponding to a '1' in the Stefan matrix.

In the rightmost column of the Stefan matrix S there are always two 1's, one on top of the other. They come from the right and left neighbourhood of the critical point C. Drawing a horizontal line between them, we see that the segments above this line are located to the left of C, and those below the line to the right of C. Taking the letters a through d as denoting different states of an automaton, we can take the Stefan matrix as the definition of transfer function for the automaton, see Table 1.

Table 1. Transfer function for the NDFA associated with $(RLRRC)^\infty$.

	L	R
a	$c+d$	
b		d
c		$b+c$
d		a

The table reads, for example, state a on accepting a letter L goes into either state c or state d, etc. Taking the states as nodes, we draw the graph in Fig. 2 to visualize the table. This is an NDFA, because, first, there is no distinguished node to start with, one can start traveling along the graph from any node; second, at some nodes there is more than one choice of where to go next on seeing the same input letter.

In formal language theory, see, e.g., [17], there is a standard method to derive DFA from NDFA — the subset construction. It is simpler to continue with our example than formulating the rules. Instead of treating single states such as a or b, we take a certain combination (subset) of states as a new state and see what happens according to the transfer function. Let us start with $\{abcd\}$, i.e., the set of all single states. On seeing the letter L it goes into the set $\{cd\}$, while on seeing the letter R it remains unchanged. We put these observations into Table 2 for a new transfer function:

Table 2. Transfer function for the DFA associated with $(RLRRC)^\infty$.

	L	R
$\{abcd\}$	$\{cd\}$	$\{abcd\}$
$\{cd\}$		$\{abc\}$
$\{abc\}$	$\{cd\}$	$\{bcd\}$
$\{bcd\}$		$\{abcd\}$

The automaton, drawn in accordance with this table, is shown in Fig. 3. Now there is a starting node, representing the state $\{abcd\}$. Beginning with the starting node, encircled twice in the figure, there is a unique choice as to where to go next on seeing an R or an L in the input. Amongst all the DFA accepting the same language there is one with a minimal number of nodes. Wolfram [7] took the logarithm of the minimal number of nodes as a measure of the complexity.

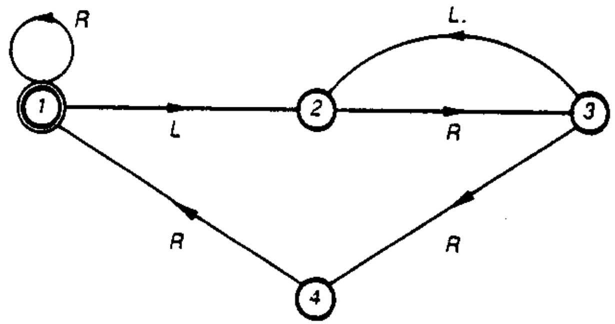

Fig. 3. Deterministic finite automaton corresponding to the period 5 kneading sequence $(RLRRC)^\infty$.

If we derive the Stefan matrix for a map with a $\rho\lambda^\infty$ type of kneading sequence, there will be a transient part in the final DFA. In order to deal with the stationary grammar, Grassberger [8] suggested to drop the transients. However, when we turn to infinite Stefan matrices and the infinite limits of the corresponding automata, sometimes the limit comes from the transient part of the finite automata. So some caution is required.

For any periodic kneading sequence $(\Sigma C)^\infty$ or eventually periodic kneading sequence $\rho\lambda^\infty$, at any admissible choice of the finite strings Σ, ρ, and λ, one always gets a finite automaton. Therefore, they belong to the lowest level of grammatical complexity — regular languages. In order to go beyond the regular level, one must turn to other types of kneading sequences, of which our knowledge is rather limited at present. A convenient way to look for a breakthrough is to construct a series of finite automata, then study the infinite limit.

Having mastered the construction of Stefan matrices and their relation to automata, both NDFA and DFA, we can work only at the transfer matrix level in our search for more complex limits. We continue with examples.

We first look at the $k \to \infty$ limit of $RL^k(RL^{k-1}R)^\infty$ type kneading sequences, which includes the period 3 band-merging point $RLL(RLR)^\infty$, studied in [10]. The limit is clear: RL^∞. The $(2k+2) \times (2k+2)$ Stefan matrix has a fixed structure with two simply growing parts. We show the matrix in Fig. 4. Consecutive '1's are drawn as a thick solid line and a single '1' is represented by a filled circle; all blanks are zero.

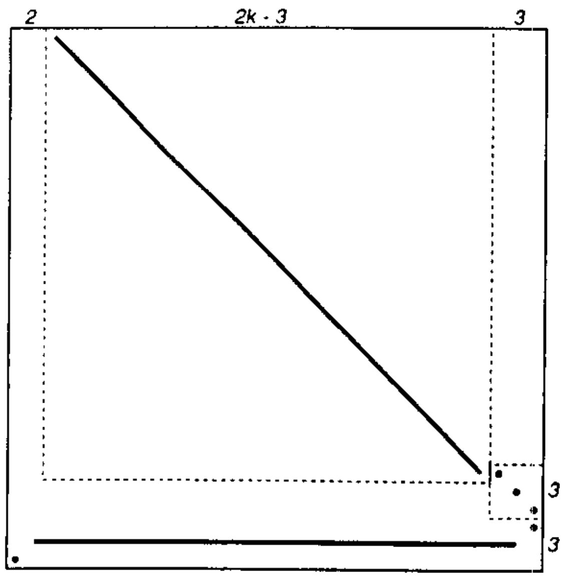

Fig. 4. The Stefan matrix for $RL^k(RL^{k-1}R)^\infty$

We see that the structure of the matrix remains the same, only the size of one upper block and one horizontal line grows with k. This tells us that in the resulting automaton many similar nodes may be combined into one, and the grammar does not get more complex with growing k.

The next example is the limit of the Fibonacci sequence we mentioned before. Taking the letter R for *R*abbit, and L for *L*ittle rabbit, and imposing a rule that after a time step a *R*abbit gives birth to a *L*ittle rabbit while an existing *L*ittle rabbit grows up into a *R*abbit, we get from a single R the following sequence:

$$\begin{array}{l} R \\ RL \\ RLR \\ RLRRL \\ RLRRLRLR \\ RLRRLRLRRLRRL \\ \cdots \end{array} \qquad (11)$$

Note that the substitutions

$$\begin{aligned} R &\to RL, \\ L &\to R, \end{aligned} \qquad (12)$$

which generate the symbolic sequences in (11), do not satisfy the conditions of the generalized composition rule. Consequently, not all sequences in (11) are admissible. However, since we are interested in periodic orbits, it is always possible to make them shift-maximal by cyclic permutation. In this way we get the Fibonacci sequence. For a member of period F_n (the n-th Fibonacci number), the Stefan matrix is a $(F_n - 1) \times (F_n - 1)$ table. The case for $F_8 = 21$ is given in Fig. 5.

No matter how large the chosen Fibonacci number, this matrix always has the fixed structure shown above. Horizontally, from left to right, the block sizes are F_{n-2}, F_{n-3}, and $F_{n-2} - 1$; vertically, from top to bottom, the sizes are F_{n-3}, $F_{n-2} - 1$, and F_{n-2}. This implies that in the resulting automaton many nodes may be combined and the final effective automaton and its $F_n \to \infty$ limit cannot be more complex, a conclusion drawn in [10] with some effort by explicitly constructing the automata.

Our last example illustrates a case to which many authors assigned infinite complexity [8, 12], namely, the accumulation point of the period-doubling cascade. It is the result of repeatedly applying the substitutions (6) to (5). The symbolic dynamics notation of this limit, using the $*$-composition of [15], is simply $R^{*\infty}$. It can be reached either from the period 2^n orbits, or from the $2^n \to 2^{n-1}$ band-merging points. In the latter case all the symbolic sequences are of $\rho\lambda^\infty$ type and may be obtained by applying the substitutions (6) to the kneading sequence of the surjective map (8), i.e., to RL^∞. The construction of Stefan matrices is straightforward in both cases. We give the period 16 Stefan matrix in Fig. 6.

This matrix has a more complicated, yet quite regular, block structure. The

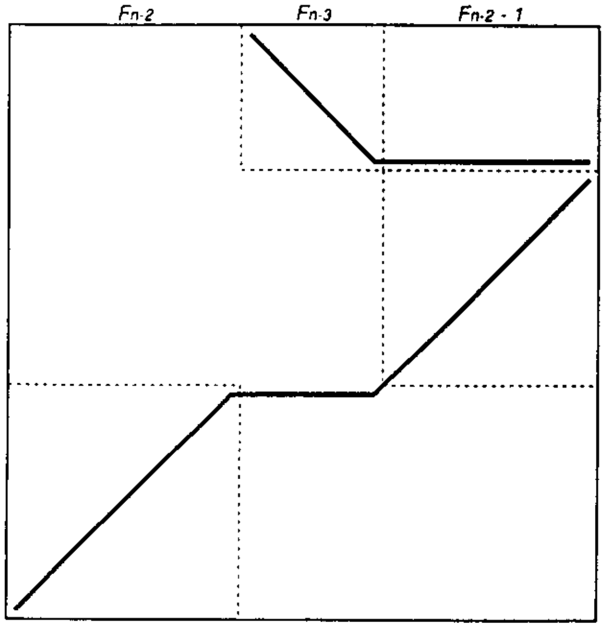

Fig. 5. The Stefan matrix for period F_n in the Fibonacci sequence.

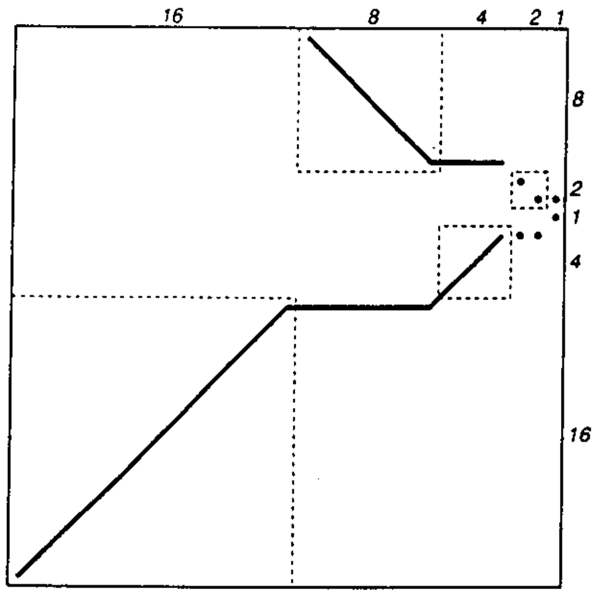

Fig. 6. The Stefan matrix for period 16 in the period-doubling cascade.

$2^{n-2}, \cdots, 8, 4, 2, 1$ along the horizontal, and the same numbers alternating along the vertical from top and bottom toward the two '1's in the rightmost column. Even when each block is represented by a single effective node in the resulting automaton, one still needs an infinite number of effective nodes to realize the automaton. No wonder some authors get infinite complexity for this case. It certainly goes beyond finite automaton, but the structure is too regular to be called "infinitely" complex. Is it reasonable that a quasiperiodic orbit at the accumulation of period-doublings is more complex than any chaotic orbit? In order to discover more complex orbits, one must look at other types of infinite Stefan matrices.

However, in doing so one should note that the overall structure of the Stefan matrices is subject to strong restrictions, originating from the underlying map. If one rotates the last three matrices anticlockwise by 90 degrees, there appears a kind of "monotonicity". To the left of the two central '1's, the location of '1's is "non-decreasing", while to the right it is "non-increasing". Of course, this observation holds only for unimodal maps.

Before concluding the paper, we would like to say a few words on the application of transfer matrices to experimental data. Based on the assumption that the system under investigation has been in a stationary state and the sampled data reflect a typical trajectory, one applies the standard phase space reconstruction technique and draws experimental Poincaré sections. The first return maps for one of the reconstructed coordinates may reveal the outline of an underlying map. If it is close to one-dimensional, as is usually the case with dissipative systems, then there is a good hope to introduce a partition, using a few letters. Then one can try to extract the grammatical rules from the symbolic sequence, obtained from the original time series. The simplest cases deal with two or three letters (in the presence of a discrete symmetry, as our experience with the Lorenz model shows [30]).

Suppose we are lucky enough to start with a symbolic dynamics of two letters, say, A and B, for *A*bove and *B*elow the average. The first thing to do is to count the occurrence of A and B, then the occurrence of the pairs AA, AB, BA, and BB, etc., and so forth. The absence of a pair, e.g., BB would imply a grammatical rule "two consecutive B's are not allowed in the language". This rule alone would cut the number of longer strings. The number of 3-letter strings would be reduced from 8 to 5; that of 4-letter strings from 16 to 8; that of n-letter strings from 2^n to F_{n+2} – the Fibonacci number again. This rule alone determines a smallest transfer matrix:

$$\begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix}, \tag{13}$$

which, in turn, would give an upper bound for the topological entropy: logarithm of the golden mean $1.618\cdots$. If new rules were to be discovered during the counting of longer strings, they would change the number of allowed combinations of even longer strings, and yield larger approximants for the transfer matrices and better estimates for the entropy.

This way of thinking essentially follows the "pre-entropy" experiment of Kahlert and Rössler [31], and the best results so far, to our knowledge, have been obtained with the Belousov-Zhabotinskii reaction data by Lathrop and Kostelich [32], see also the recent work [33] along the line of Badii [9]. Nevertheless, the whole business of extracting grammatical rules and comparing the complexity of experimental chaotic attractors is still in its infancy.

As regards the application of symbolic dynamics to description of complexity in two and more dimensions, much more effort has to be made, since the symbolic dynamics itself has not been well-developed.

Acknowledgements

This work is a part of a project supported by the Chinese Natural Science Foundation and the Open Laboratories Project of Academia Sinica. The work has been continued and the paper written during a visit to the International Centre for Theoretical Physics, Trieste. A kind arrangement by the Kaiserlautern University encouraged me to summarize the work at its present stage. The author thanks all the above-mentioned institutions.

References

1. Livi R, Ruffo S, Ciliberto S, and Buiatti M (1988), eds., *Chaos and Complexity*, World Scientific.
2. Peliti L, and Vulpiani A (1988), eds., *Measures of Complexity*, Lecture Notes in Phys. **314**, Springer-Verlag.
3. J. Bernasconi, M. Droz, A. Malaspinas, and M. Ryter, eds., Proceedings of Conference on "Complex Systems", *Helv. Phys. Acta* **62** (1989), 461-630.
4. N. B. Abraham, A. M. Albano, A. Passamante, and P. E. Rapp, eds., *Measures of Complexity and Chaos*, Plenum, 1989.
5. Bohr H (1990), ed., *Characterizing Complex Systems – Interdisciplinary Workshop on Complexity and Chaos*, World Scientific.
6. Zurek W (1990), ed., *Complexity, Entropy, and Physics of Information*, Addison-Wesley.
7. S. Wolfram, *Commun. Math. Phys.* **96** (1984) 15.
8. P. Grassberger, *Int. J. Theor. Phys.* **25** (1986) 907; *Physica* **A140** (1986) 319; *Z. Naturforsch.* **43a** (1988) 671; *Helv. Acta Phys.* **62** (1989) 489; and in [2].
9. R. Badii, Europhys. Lett. **13** (1990) 599; and in [4], 313.
10. D. Auerbach, and I. Procaccia, *Phys. Rev.* **A41** (1990) 6602.
11. G. D'Alessandro, and A. Politi, *Phys. Rev. Lett.* **64** (1990) 1609.
12. J. P. Crutchfield, and K. Young, *Phys. Rev. Lett.* **63** (1989) 105.
13. J. Milnor, and W. Thurston, in *Dynamical Systems*, ed. by J. C. Alexander, *Lect. Notes in Math.*, No. 1342, Springer-Verlag(1988).
14. B.-L. Hao, *Physica* **D51** (1991) 161.
15. B. Derrida, A. Gervois, and Y. Pomeau, *Ann. Inst. H. Poincaré* **29A** (1978) 305.
16. J. P. Crutchfield, and N. H. Packard, *Physica* **D7**, (1983) 201.
17. J. E. Hopcroft, and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesley, 1979

18. M. Morse, and G. A. Hedlund, *Am. J. Math.* **60**, 815(1938); reprinted in *Collected Papers of M. Morse*, vol. 2, World Scientific, 1986.
19. A. M. Alekseev, and M. V. Yakobson, *Phys. Repts.* **75**, (1981) 287.
20. N. Metropolis, M. L. Stein, and P. R. Stein, *J. Combinat. Theory* **A15** (1973) 25.
21. Hao Bai-lin, *Elementary Symbolic Dynamics and Chaos in Dissipative Systems*, World Scientific, 1989.
22. W.-M. Zheng, and B.-L. Hao, "Applied Symbolic Dynamics", in *Experimental Study and Characterization of Chaos*, vol. 3 of *Directions in Chaos*, ed. by Hao Bai-lin, World Scientific, 1990, 363-459.
23. P. Collet, and J.-P. Eckmann, *Iterated Systems on the Interval as Dynamical Systems*, Birkhäuser, Boston, 1980.
24. B.-L. Hao, and W.-M. Zheng, *Int. J. Mod. Phys.* **B3**, (1989) 235.
25. Wei-mou Zheng, J. Phys. **A22** (1989), 3307.
26. Wei-mou Zheng, *Chaos, Solitons and Fractals*, **1** (1991) 243.
27. Wei-mou Zheng, "Symbolic dynamics for the Tél map", ASITP Preprint 92-001.
28. Hong Zhao, Wei-mou Zheng, and Yan Gu, "Determination of partition lines from dynamical foliations for the Hénon map", ASITP Preprint 92-002; Hong Zhao, and Wei-mou Zheng, "Symbolic analysis of the Hénon map at $a = 1.4$ and $b = 0.3$", ASITP Preprint 92-040.
29. Wei-mou Zheng, *Chaos, Solitions and Fractals*, 2(1992)461
30. M.-Z. Ding, and B.-L. Hao, *Commun. Theor. Phys.*, **9** (1988) 375.
31. C. Kahlert, and O. E. Rössler, *Z. Naturforsch.* **39a** (1984) 1200.
32. D. P. Lathrop, and E. J. Kostelich, *Phys. Rev.* **A40** (1989) 4028; and in [4], 147.
33. M. Finardi, L. Flepp, J. Parisi, R. Holzner, R. Badii, and E. Brun, *Phys. Rev. Lett.* **68** (1992) 2898.

27. Appeared in Commun. Theor. Phys. 17 (1992) 167-174.

28. ASITP 92-002 appeared in Commun. Theor. Phys. 17 (1992) 263-268.

ASITP 92-040 appeared in Commun. Theor. Phys. 19 (1993) 21-26.