Symbolic Dynamics Approach to Chaos

Bai-lin HAO*
Institute of Theoretical Physics, Academia Sinica
P. O. Box 2735, Beijing 100080, China

Abstract

Symbolic dynamics is a coarse-grained description of the dynamics. By taking into account the "geometry" of the dynamics, it can be cast into a powerful tool for practitioners in nonlinear science. Detailed symbolic dynamics can be developed not only for quadratic, unimodal maps, but also for maps with multiple critical points and even with discontinuities. Their application to differential equations will be mentioned briefly.

1 Introduction

Symbolic dynamics is a *rigorous* way to study complex dynamics with *finite* precision. As an abstract chapter in the mathematical theory of dynamical systems[1, 2], it has a history of more than half of a century. The basic idea is very simple: divide the phase space into a finite number of regions and label each region by a letter from a certain alphabet; instead of following a trajectory one only keeps recording the alternation of letters. One loses a great amount of detailed information on the dynamics, but some essential, robust, features of the motion may be kept, e.g., periodicity or chaoticity of an orbit. This is nothing but what physicists call a coarse-grained description.

The idea of symbolic dynamics applies to dynamics in any finite-dimensional phase space. In many cases, say, for theorem-proving, an arbitrary partition of the phase space would do the job. However, only for one-dimensional mappings symbolic dynamics has been developed more or less completely. This is due to the nice ordering property of real numbers on an interval and due to the possibility of partitioning the "phase space", i.e., the interval, in accordance with the "geometry" of the dynamics. In fact, many useful rules and beautiful results have been derived. Recently, significant progress has been made in symbolic dynamics of two-dimensional maps, but the achievement is still rather limited compared to what has been known in one dimension. Nevertheless, the knowledge of symbolic dynamics in one and two dimensions proves to be quite instructive in understanding the systematics of periodic orbits and chaotic behavior in some higher-dimensional dissipative systems, e.g., the Lorenz model and some periodically forced nonlinear oscillators. The presence of dissipation is essential, since it causes the shrinking of phase space volume, which makes "strange attractors" close to low-dimensional objects at least in some sections.

Chaotic dynamics of dissipative systems provides a rare and lucky case in physics, when one-dimensional systems are not merely toy models, but lead to essential "universal" results which are quite useful in understanding higher dimensional systems. In a sense, everyone who enters the field of chaos should start with the study of one-dimensional maps. We have called the symbolic dynamics approach to chaos, based on low-dimensional maps, *elementary* or *applied symbolic dynamics*[1] [3, 4, 5].

Applied symbolic dynamics originated from a seminal paper by Metropolis, Stein and Stein[6]. The kneading theory of Milnor and Thurston[7], the lecture of Guckenheimer[8], and a paper by Derrida, Gervois, and Pomeau[9], among others, further developed the theory. What had been

*Lectures given at the 1st International Workshop on Chaotic Dynamics, Pohang, Korea, 7 – 9 July 1993.
[1]Ian Percival suggested the name.

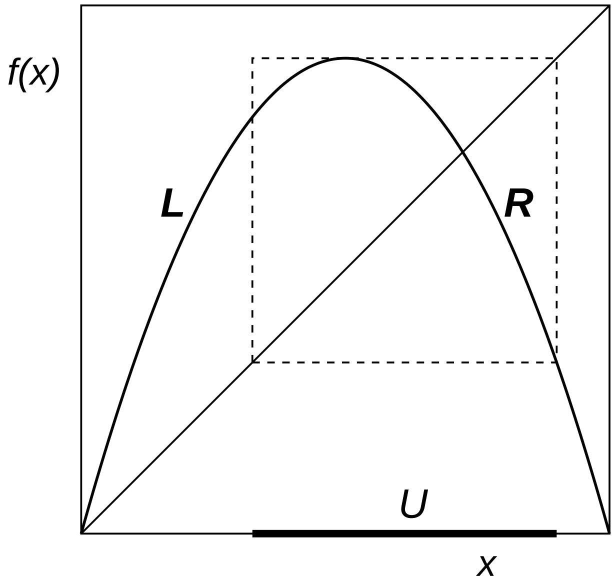

Figure 1: A quadratic map.

known by the end of 1970s was summarized in the book by Collet and Eckmann[10]. There has been significant generalization and simplification of the theory and its application to ordinary differential equations in the 1980s, see, e.g., references in [3, 4, 5].

In these lectures, we will confine ourselves to one-dimensional maps of the general form

$$x_{n+1} = f(\mu, x_n), \tag{1}$$

where $f(\mu, x)$ is a nonlinear "mapping function" of the variable x and μ is a parameter. The function $f(x)$ maps an interval I into itself; it may have several monotone pieces on the interval and may contain discontinuity.

In order to have a better feeling as how symbolic dynamics works, we will explain the essentials of symbolic dynamics on four examples, namely, the quadratic map, the gap map, the antisymmetric cubic map, and the sine-square map. We describe briefly these maps.

1. The quadratic or **logistic map**

$$x_{n+1} = \lambda x_n(1 - x_n), \tag{2}$$

where $\lambda \in (0, 4]$ is a parameter and the variable x_n is confined to the interval $[0, 1]$. We denote the left, monotonically increasing, branch of the mapping function f by a label L, and the right, monotonically decreasing, branch — by R, see Fig. 1. If necessary, we may write explicitly f_L or f_R. Sometimes we use a letter C to denote the central or critical point of the map, where the function f reaches a maximum. The letters R, C, and L, are the symbols we are going to play with in the corresponding symbolic dynamics. They are ordered naturally:

$$L < C < R. \tag{3}$$

Since the mapping function f is nonlinear, its inverse is multi-valued. We can define single-valued pieces of the inverse function by using the labeled branches. In order to simplify the notation, we use the labels themselves to denote these inverse functions:

$$\begin{aligned} L(y) &\equiv f_L^{-1}(y), \\ R(y) &\equiv f_R^{-1}(y). \end{aligned} \tag{4}$$

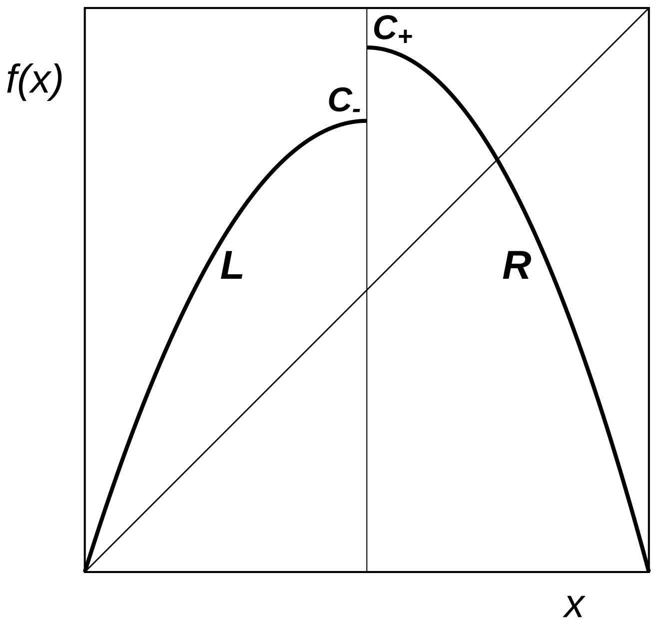

Figure 2: A gap map.

Sometimes it is more convenient to rescale the parameter and variable to write

$$x_{n+1} = \mu - x_n^2. \tag{5}$$

In this case the inverse functions read

$$\begin{array}{l} L(y) = +\sqrt{y-\mu}, \\ R(y) = -\sqrt{y-\mu}. \end{array} \tag{6}$$

We will make use of these functions in the next section.

2. The gap map. By introducing a discontinuity at the central point of the logistic map, we get the gap map, see Fig. 2. Numerical study has revealed some "strange" behavior of this map; even a "new route to chaos" has been proposed[11]. However, all its peculiarities may be fully understood by way of symbolic dynamics[12, 13]. In fact, the gap map happens to be just the next step of modifying the quadratic map, when the number of monotone branches remains the same, only the central point C splits into C_- and C_+. This makes the gap map an instructive example in studying symbolic dynamics. We note that the letters are ordered as follows:

$$L < C_- \leq C_+ < R. \tag{7}$$

3. The antisymmetric cubic map

$$x_{n+1} = Ax_n^3 + (1-A)x_n \tag{8}$$

has three monotone branches R, M, L, and two critical points C and D, see Fig. 3. These letters as ordered in a natural way:

$$L < C < M < D < R. \tag{9}$$

The map (8) is invariant under a discrete transformation $x \to -x$ so the orbits may be symmetric or asymmetric with respect to this transformation. It provides a convenient example of analyzing symmetry breaking and symmetry restoration in dynamical systems. As a matter of fact, this one-dimensional cubic map has a close relation with the Lorenz model of three differential equations in their systematics of periodic solutions.

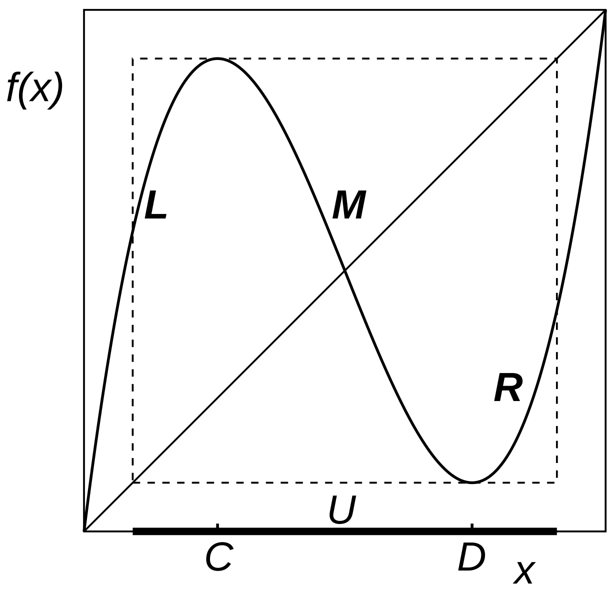

Figure 3: An antisymmetric cubic map.

4. The sine-square map

$$y_{n+1} = A\sin^2(y_n - B), \quad |y_n - B| \le \pi \tag{10}$$

appeared in a model, describing a hybrid optical bistability device, using liquid crystal as the nonlinear medium in the cavity. The structure of the parameter plane $A \sim B$ has been studied several years ago by using symbolic description without constructing the symbolic dynamics[15]. Its symbolic dynamics has been worked out recently[16].

In this case, four letters, R, N, M, and L, are needed to name the inverse of the four monotone branches shown in Fig. 4. Formally, there are five critical points: D_1, D_2, D_3, C_1, and C_2, but all C_i change simultaneously and so do all D_i. The ordering of letters is

$$D_1 < L < C_1 < M < D_2 < N < C_2 < R < D_3. \tag{11}$$

For $K_\pm$ in Fig. 4, see Eq. (40) in Section 3.

2 Symbolic Sequences and Word-Lifting Technique

We consider the general form (1) of one-dimensional map of an interval I into itself. Starting from a given initial point x_0, we get a *numerical sequence* or an *orbit*

$$x_0, x_1 = f(x_0), x_2 = f(x_1) = f^2(x_0), \cdots, x_n = f(x_{n-1}) = f^n(x_0), \cdots \tag{12}$$

where we have dropped the fixed parameter μ and introduced a notation for the iterates of f:

$$f^m(x) = f(f^{m-1}(x)), \;\; m = 1, 2, \cdots; \quad f^0(x) \equiv x.$$

To the numerical sequence (12) we juxtapose a *symbolic sequence* in the following way. If a point x_i falls in a subinterval of I which corresponds to a monotone branch of f, labeled by a letter σ_i, i.e., f_{σ_i}, then we put the letter σ_i in accordance with x_i. If x_i falls exactly at one of the critical points, then σ_i is a letter denoting the corresponding critical point. In short, we have the correspondence

$$\begin{array}{cccccc} x_0, & x_1, & \cdots, & x_{n-1}, & x_n, & \cdots \\ \sigma_0, & \sigma_1, & \cdots, & \sigma_{n-1}, & \sigma_n, & \cdots \end{array} \tag{13}$$

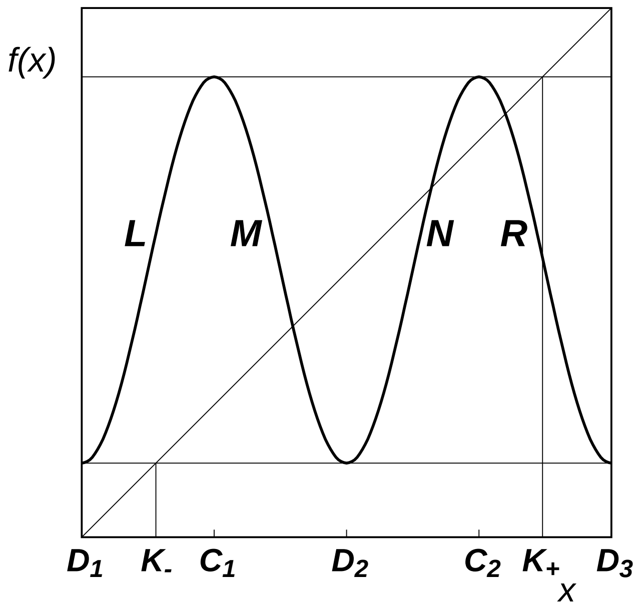

Figure 4: A sine-square map.

The letter σ_i may be one of the symbols, which we wrote in (3), (7), (9), or (11).

Sometimes we denote a finite or infinite symbolic sequence by a single capital Greek or Roman letter, e.g., Σ or X. We will see that it is a good convention to name a symbolic sequence after the number that starts the iteration. For example, we may write

$$x_0 = \Sigma = \sigma_0\sigma_1\cdots\sigma_{n-1}\sigma_n\cdots. \tag{14}$$

Now let us try to reverse the last relation in (12) to get

$$x_0 = f^{-n}(x_n). \tag{15}$$

Of course, this cannot be done so simply, because the inverse map is multi-valued and we do not know which branch to use when taking the inverse. However, when these branches have been labeled by the symbol chosen by its argument at each step of the iteration in (12), we may write

$$\begin{aligned} x_0 &= f_{\sigma_0}^{-1}(x_1) \\ &= f_{\sigma_0}^{-1}\circ f_{\sigma_1}^{-1}(x_2) \\ \cdots \\ &= f_{\sigma_0}^{-1}\circ\cdots\circ f_{\sigma_{n-1}}^{-1}(x_n) \end{aligned} \tag{16}$$

To simplify the notation, we will denote a monotone branch of the inverse map by its label, i.e., letting

$$\sigma(y) \equiv f_\sigma^{-1}(y). \tag{17}$$

Examples of such $\sigma(y)$ are $L(y)$ and $R(y)$ in (6).

Consequently, we may write Eq. (16) as a functional composition

$$x_0 = \sigma_0\circ\sigma_1\circ\cdots\circ\sigma_{n-1}(x_n). \tag{18}$$

Thus we have a correspondence between a numerical orbit (12), a symbolic sequence (14), and a sequence of functional composition (18).

A simple application of this correspondence is the *word-lifting technique*[14, 15] for determining the loci of superstable periodic orbits in the parameter space. A superstable periodic orbit contains at least one of the critical points and represented by, for example, a symbolic sequence $(\Sigma C)^\infty$, where Σ is a string of symbols that do not contain any critical points. According to our convention (14), we may write

$$C = C\Sigma C\cdots, \tag{19}$$

where the first letter C on the left is a number, while the first letter C on the right hand side is a symbol, corresponding to a f^{-1}. Applying the map f to both sides of (19), we get

$$f(C) = \Sigma(C), \tag{20}$$

Now, the C on both sides are numbers, and Σ must be understood as a composite function. For a given map, everything is known in Eq. (20) except for the parameter value where the equality takes place. Therefore, (20) is an equation to determine the parameter of the orbit. When there is only one parameter, it yields a number; when there are more parameters, it defines a curve or surface in the parameter space. For example, the period 5 orbit RL^2RC in the quadratic map (5) lifts into an equation

$$f(C) = R\circ L\circ L\circ R(C), \tag{21}$$

in which $R(y)$ and $L(y)$ have to be understood as the inverse functions (6) and the central point $C=0$. It can be solved by transforming into an iteration scheme

$$\mu_{n+1} = \sqrt{\mu_n + \sqrt{\mu_n + \sqrt{\mu_n - \sqrt{\mu_n}}}},$$

which converges quickly to $\mu = 1.860782522\cdots$, starting from any reasonable initial value, say, $\mu_0 = 2$.

In the case of maps with multiple critical points a superstable periodic sequence may contain more than one critical points, e.g., $(\Sigma C\Pi D)^\infty$, where C and D are critical points and Σ, Π are strings without critical points. In this case, word-lifting leads to a pair of equations

$$\begin{aligned} f(D) &= \Sigma(C), \\ f(C) &= \Pi(D). \end{aligned} \tag{22}$$

In particular, when there are only two independent parameters, as is the case with the sine-square map, Eqs. (22) determines an isolated point in the parameter plane. It corresponds to a double-superstable periodic orbit, i.e., it contains two critical points, and is called a *joint*. Double-superstable orbits and joints play an essential role in understanding the dynamics, see Section 4 below.

Superstable periodic symbolic sequences are only a particular class of sequences which starts from a critical point C, or, to be more precise, from the first iterate $f(C)$ of C. If one starts from a critical point C and first gets a nonperiodic string ρ then goes into a periodic repetition of λ^∞, where the strings ρ and λ do not contain critical points, then the sequence $\rho\lambda^\infty$ is called an *eventually periodic* sequence. Generally, any symbolic sequence, which starts from $f(C)$, is called a *kneading sequence*. Kneading sequences play an important role in determining the dynamics. Periodic and eventually periodic kneading sequences are two best-understood classes of kneading sequences.

The word-lifting technique applies not only to periodic kneading sequences, but also to eventually periodic kneading sequences. It is easy to see that a word $\rho\lambda^\infty$ may be lifted into a pair of equations[18]:

$$\begin{aligned} f(C) &= \rho(\nu), \\ \mu &= \lambda(\nu). \end{aligned} \tag{23}$$

As an example, let us determine the parameter where a four-piece chaotic band merges into a two-piece one in the quadratic map (5). The kneading sequence is known to be $K = RLRR(RL)^\infty$, see, e.g., [4]. Consequently, the equations (23) read

$$\begin{aligned} f(C) &= R\circ L\circ R\circ R(\nu), \\ \nu &= R\circ L(\nu). \end{aligned} \tag{24}$$

For the map (5) the inverse functions are given by (6). One has to solve the following pair of iterated equations:

$$
\begin{aligned}
\mu_{n+1} &= \sqrt{\mu_n + \sqrt{\mu_n - \sqrt{\mu_n - \sqrt{\mu_n - \nu_n}}}}, \\
\nu_{n+1} &= \sqrt{\mu_n + \sqrt{\mu_n - \nu_n}}.
\end{aligned} \tag{25}
$$

Using any reasonable initial conditions, say, $\mu_0 = 2.0$, $\nu_0 = 1.95$, a few iterations yield $\mu = 1.4303576\cdots$, $\nu = 1.3248379\cdots$.

Eventually periodic sequences of type $\rho\lambda^\infty$ determine so-called boarder to chaos. However, we will not go into details; for more, see, e.g., [4].

3 Ordering of Symbolic Sequences and The Admissibility Conditions

If a given symbolic sequence, finite or infinite, may be obtained from iterating a map, using suitably chosen parameter and initial value, it is said to be an *admissible sequence* for the map. Clearly, not any arbitrary symbolic sequence is admissible. We have to find the conditions that an admissible symbolic sequence must satisfy. These admissibility conditions are based on the ordering of symbolic sequences, so we first define the ordering rule.

Our convention of denoting a symbolic sequence by the number x_0 which starts the corresponding numerical sequence and understanding a symbol σ as the inverse of the monotone branch f_σ, see Eq. (17), allows us to explain the ordering of symbolic sequences in a rather simple way. Suppose we are required to compare two symbolic sequences

$$x_1 = \Sigma_1 = \Sigma^*\sigma\cdots$$

and

$$x_2 = \Sigma_2 = \Sigma^*\tau\cdots,$$

where Σ^* denotes the common leading string of symbols in the two sequences (Σ^* may be blank). The two subsequent symbols σ and τ must be different. We simply assign the order of the two numbers x_1 and x_2 to the two symbolic sequences: if $x_1 > x_2$ we say $\Sigma_1 > \Sigma_2$ and vice versa. In fact, this ordering does not depend on the numbers x_1 and x_2 and may be read off from the symbols alone.

In order to explain this, we recall a few simple properties of monotone functions:

1. A monotone increasing function preserves the order, i.e., from $x_1 > x_2$ it follows that $f(x_1) > f(x_2)$ and vice versa.
2. A monotone decreasing function reverses the order, i.e., from $x_1 > x_2$ it follows that $f(x_1) < f(x_2)$ and vice versa.
3. A composition of many monotone functions acts as a monotone increasing function if there is an even number of decreasing components, and it acts as a monotone decreasing function if it contains an odd number of decreasing components. Therefore, we may assign an *even parity* (or $+1$) to a monotone increasing function and an *odd parity* (or -1) to a decreasing one. The parity of a functional composition is obtained as the product of the parities of its components.
4. A monotone function and its inverse are both increasing (or decreasing), so they have the same parity. For example, in the sine-square map the letters or functions L and N have even parity, whereas M and R — odd parity.

Since σ and τ are different, they must have been ordered according to the natural order, e.g., one of the realtions (3), (7), (9), or (11). Now, looking at the common leading string Σ^* as a composition of monotone functions, it follows that the order of σ and τ is passed to x_1 and x_2 if Σ^* has even parity, and the order is reversed if the parity is odd.

Hence the **ordering rule**: symbolic sequences are ordered according to the first different symbols after their common leading string Σ^*. If Σ^* has even parity, the order is preserved, if Σ^* has odd parity, the order is reversed. An empty string is considered to be even.

We note that this ordering rule works for any one-dimensional map with a finite number of crtitical points and discontinuities.

Before formulating the admissibility conditions, we introduce the notion of a *dynamically invariant range.* Look at Fig. 1 of the quadratic map, we see that the first and second iterates of C determine a subinterval $[f^2(C), f(C)]$ within the interval I, shown in heavy line and labeled by U in the figure. All initial points in this subinterval lead to orbits, confined to this subinterval. Orbits starting from any initial point outside U will enter it in finite steps, i.e., will show a trivial transient process. Since in chaotic dynamics we are interested in long-time behavior of the dynamics, from now on we will concentrate on the dynamically invariant range U and derive the admissibility conditions for this interval only.

For the antisymmetric cubic map the dynamically invariant range U is shown in Fig. 3. For other maps, the invariant range may be determined from an inspection of the map.

In order to formulate the admissibility condition for the **quadratic map** we define two sets $\mathcal{R}$ and $\mathcal{L}$ for a given, finite or infinite, symbolic sequence Σ, made of R and L. The set R is a collection of all substrings of Σ, which follow any letter R in Σ. The set L is a collection of all substrings of Σ, which follow any letter L in Σ. Take, for example, the sequence

$$\Sigma = RLLRLRL\cdots,$$

the two sets are

$$\begin{aligned} \mathcal{R}(\Sigma) &= \{LLRLRL\cdots, LRL\cdots, L\cdots, \cdots\}, \\ \mathcal{L}(\Sigma) &= \{LRLRL\cdots, RLRL\cdots, RL\cdots, \cdots\}. \end{aligned} \tag{26}$$

In order to remember the origin of these subsets we have put Σ as an argument of R and L. We often drop the argument when no confusion may occur.

From Fig. 1 of the quadratic map we see that the rightmost point of the dynamically invariant range U is the first iterate of the maximum of the map, i.e., $f(C)$. According to our convention, $f(C)$ starts a symbolic sequence which is nothing but the kneading sequence

$$K = f(C) = R\cdots.$$

If a symbolic sequence Σ is admissible, then any of its subsequences must be admissible. Any of these subsequences starts from a point other than C. Consequently, its iterates cannot go beyond the rightmost point of U. Therefore, we have

$$\begin{aligned} \mathcal{R}(\Sigma) \le K, \\ \mathcal{L}(\Sigma) \le K. \end{aligned} \tag{27}$$

Equations (27) give the admissibility conditions we are looking for. The equal sign in (27) is a subtle point. When Σ does not contain the critical point C, the equal sign may be dropped and conditions (27) become inequalities.

A kneading sequence K itself must satisfy the admissibility conditions as well, hence

$$\begin{aligned} \mathcal{R}(K) \le K, \\ \mathcal{L}(K) \le K. \end{aligned} \tag{28}$$

The two sets $\mathcal{R}$ and $\mathcal{L}(\mathcal{K})$, taken together, give all *shifts* of K. A shift operator $\mathcal{S}$ is defined by dropping the first letter in a sequence:

$$\mathcal{S}\sigma_1\sigma_2\sigma_3\cdots = \sigma_2\sigma_3\cdots. \tag{29}$$

The operator $\mathcal{S}$ may be applied repeatedly and we have

$$\{\mathcal{R}(K), \mathcal{L}(K)\} = \{\mathcal{S}^n(K) | n = 1, 2, 3, \cdots\}.$$

Therefore, (28) means that all shifts of K must not exceed K. In other words, a kneading sequence K of the quadratic map must be a *shift-maximal* sequence.

We emphasize the twofold meaning of admissibility conditions. One one hand, Eqs. (28) are the conditions for a symbolic sequence K to be a kneading sequence. It deals with the parameter space. If necessary, the corresponding parameter may be calculated by word-lifting for a given map. On the other hand, Eqs. (27) are conditions for a symbolic sequence Σ to be admissible when the kneading sequence K, i.e., the parameter, is given. If Σ satisfies (27), then there must exist an initial point x_0 whose iterations lead to Σ. Here one deals with the phase space. If necessary, the value of x_0 may be determined numerically by a bisection method for a given map. This remark holds for the admissibility conditions of other maps which we are going to study.

The **gap map** has two kneading sequences

$$\begin{aligned} K_+ &= f(C_+), \\ K_- &= f(C_-). \end{aligned} \tag{30}$$

Given the kneading sequences $K_\pm$, the admissibility conditions for a symbolic sequence Σ read

$$\begin{aligned} \mathcal{R}(\Sigma) &\le K_+, \\ \mathcal{L}(\Sigma) &\le K_-. \end{aligned} \tag{31}$$

The conditions for a pair (K_+, K_-) to be kneading sequences consist in:

$$\begin{aligned} \mathcal{R}(K_+, K_-) &\le K_+, \\ \mathcal{L}(K_+, K_-) &\le K_-. \end{aligned} \tag{32}$$

One cannot speak about shift-maximality now.

If we do not impose anti-symmetry on the cubic map, it would have two kneading sequences

$$\begin{aligned} K_C &= f(C), \\ K_D &= f(D), \end{aligned} \tag{33}$$

see Fig. 3. The admissibility conditions for a symbolic sequence Σ, made of the letters R, M, and L, are

$$\begin{aligned} \mathcal{L}(\Sigma), \mathcal{M}(\Sigma) &\le K_C, \\ K_D &\le \mathcal{M}(\Sigma), \mathcal{R}(\Sigma). \end{aligned} \tag{34}$$

By imposing the anti-symmetry requirement ($\overline{C} \equiv D = -C$, in particular) and introducing a symbolic transformation of interchanging R and L but leaving M unchanged:

$$\begin{aligned} R &\Leftrightarrow L, \\ C &\Leftrightarrow \overline{C}, \\ M &\Leftrightarrow M, \end{aligned} \tag{35}$$

we are left with only one kneading sequence

$$K \equiv K_D = f(D), \tag{36}$$

(The other kneading sequence $\overline{K} \equiv K_C$ is obtained by applying the transformation (35) to K.) The admissibility conditions reduce to:

$$\begin{aligned} \mathcal{L}(\Sigma), \mathcal{M}(\Sigma) &\le \overline{K}, \\ K &\le \mathcal{M}(\Sigma), \mathcal{R}(\Sigma). \end{aligned} \tag{37}$$

The admissibility conditions for the kneading sequences become:

$$\begin{aligned} \mathcal{L}(K, \overline{K}), \mathcal{M}(K, \overline{K}) &\le \overline{K}, \\ K &\le \mathcal{M}(K, \overline{K}), \mathcal{R}(K, \overline{K}). \end{aligned} \tag{38}$$

Generally speaking, a map with multiple critical points and discontinuities may contain many parameters, some of which may be abundant. A natural question arises: what is the optimal number of parameters and how to choose them. It turns out that the best way to parameterize a map is to use independent kneading sequences as parameters[4]. When there are discontinuities, one may need additional sequences, which start from the two sides of the discontinuity.

Now it is clear that the **sine-square map** (10), as it is written, is not well suited for deriving symbolic dynamics rules. In order to facilitat the construction of symbolic dynamics, we introduce a new variable

$$x = (y - B)/\pi \tag{39}$$

to rewrite Eq. (10) as

$$x_{n+1} = (K_+ - K_-)\sin^2(\pi x_n) + K_-, \quad |x_n| \le 1, \tag{40}$$

where the two new parameters K_+ and K_- are simply related to the old parameters:

$$K_+ = (A - B)/\pi, \quad K_- = -B/\pi. \tag{41}$$

Now, K_+ and K_- are nothing but the kneading sequences of the map:

$$\begin{aligned} K_+ &= f(C_1) = f(C_2), \\ K_- &= f(D_1) = f(D_2) = f(D_3). \end{aligned} \tag{42}$$

(We will only consider the case $K_+ > K_-$; the opposite case may be treated by a trivial transformation.) The admissibility condition for a symbolic sequence Σ consists in

$$K_- \le \mathcal{S}^k(\Sigma) \le K_+, \quad k = 0, 1, 2, \cdots. \tag{43}$$

This condition works when the parameters of the map are fixed, i.e., when the kneading sequences are given. However, kneading sequences themselves are symbolic sequences, so they must satisfy the admissibility condition as well. This means the sequence K_+ must be shift-maximal:

$$\mathcal{S}^k(K_+) \le K_+, \quad k = 0, 1, 2, \cdots, \tag{44}$$

and K_- be shift-minimal:

$$K_- \le \mathcal{S}^k(K_-), \quad k = 0, 1, 2, \cdots. \tag{45}$$

Of course, we might have written these conditions using the four sets $\mathcal{R}$, $\mathcal{M}$, $\mathcal{N}$, and $\mathcal{L}$.

4 Median Words and Kneading Plane

It follows from the admissibility conditions that once the kneading sequences of a map are known the entire symbolic dynamics is determined. Thus we have to learn how to generate all kneading sequences for a given map. This task can be accomplished if we know how to generate all the admissible sequences, included in between two given kneading sequences, up to a certain length.

In the simplest case of the quadratic map the construction is based on the

Periodic Window Theorem[19]: if a superstable periodic sequence ΣC is admissible, then the letter C may be replaced by R or L and the three kneading sequences $(\Sigma C)^\infty$, $(\Sigma L)^\infty$, and $(\Sigma R)^\infty$, are all admissible.

In fact, these three sequences form a periodic window

$$[(\Sigma C)_-^\infty, \Sigma C, (\Sigma C)_+^\infty], \tag{46}$$

where

$$\begin{aligned} (\Sigma C)_+ &= \max\{\Sigma R, \Sigma L\}, \\ (\Sigma C)_- &= \min\{\Sigma R, \Sigma L\}. \end{aligned} \tag{47}$$

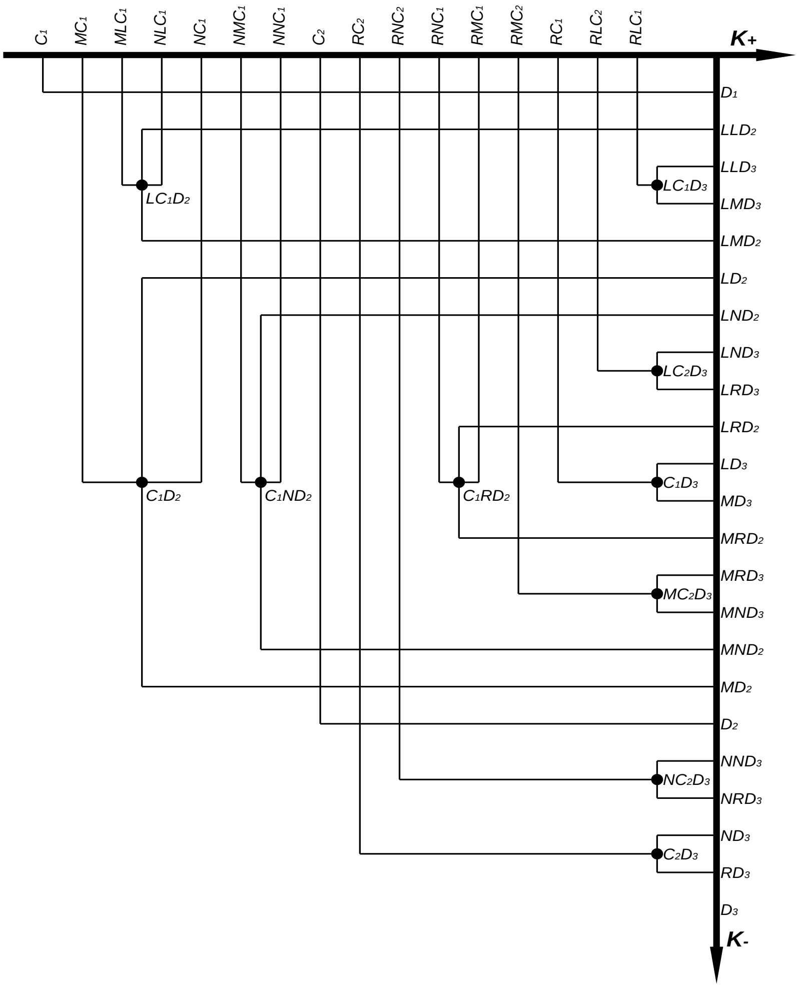

Figure 5: The kneading plane of the sine-square map.

They are called the *upper* and *lower* sequences of ΣC, respectively. It is easy to see that for the quadratic map the upper sequence always has $-$ parity, while the lower sequence — $+$ parity. If we define a 0 parity for the superstable sequence ΣC itself, then a periodic window has the following *signature*:

$$(+, 0, -). \tag{48}$$

Let us look at two examples. A superstable fixed point must correspond to the symbolic sequence C^∞. It expands into a fixed-point window

$$(L, C, R). \tag{49}$$

After period-doubling, it goes into a period 2 window

$$(RR, RC, RL). \tag{50}$$

They both have the signature (48).

It is interesting to note that at the far end of the chaotic band, at $\mu = 2$ for the quadratic map (5), the kneading sequence is RL^∞. The one-band region splits into a two-band chaotic zone at

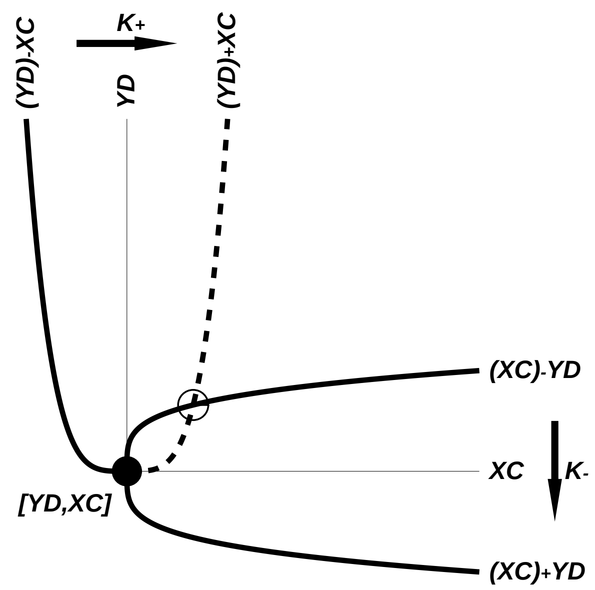

Figure 6: Bones from a joint (schematic).

$K = RL(RR)^\infty$. In general, whenever there is a periodic window

$$(\lambda, \lambda|_C, \rho), \tag{51}$$

where λ and ρ are strings made of R and L, and $\lambda|_C$ denotes a string, obtained by replacing the last symbol in λ by C, there is a chaotic map with kneading sequence $\rho\lambda^\infty$. This observation has led to the establishment of so-called *generalized composition rule*[20], which is a far-reaching generalization of the $*-$composition rule, introduced in [9]. The generalized composition rule opens a new way to derive many more admissible words from known ones and provides a link to renormalization group equations. We will not touch on this development in these lectures.

The periodic window theorem allows us to find easily the shortest admissible sequence between two given admissible sequences $\Sigma C < \Pi C$. We compare the upper sequence $(\Sigma C)_+^\infty$ of the smaller sequence ΣC with the lower sequence $(\Pi C)_-^\infty$ of the larger ΠC. If they are identical, then they are adjacent words and there are no median word in between. In fact, Π is the period-doubled regime of ΣC. We note that only an upper sequence, whose parity is always $-$, is capable to undergo period-doubling. If these two sequences are not the same, then their common leading string appended with a C is the median sequence we are looking for. This method of construction of median words deals only with admissible sequences at its intermediate steps and is easier to be generalized to maps with multiple critical points. The old method of MSS[6], using so-called harmonic and antiharmonic sequences, has become obsolete[19].

How to generalize the notion of upper and lower sequences to maps with multiple critical points? It goes straightforward, using the ordering rule and continuity consideration. Take, for example, a kneading sequence K_+ of the **sine-square map** (40). It starts with a C_i, $i = 1, 2$ and we check the subsequent letters. When another C_i is encountered, K_+ must be a periodic sequence. The periodic window theorem still works and we have the upper and lower sequences:

$$\begin{aligned}
(\Sigma C_1)_+ &= \max\{\Sigma L, \Sigma M\}, \\
(\Sigma C_1)_- &= \min\{\Sigma L, \Sigma M\}, \\
(\Sigma C_2)_+ &= \max\{\Sigma N, \Sigma R\}, \\
(\Sigma C_2)_- &= \min\{\Sigma N, \Sigma R\}.
\end{aligned} \tag{52}$$

As in the case of the quadratic map, the upper sequences are always of $-$ parity.

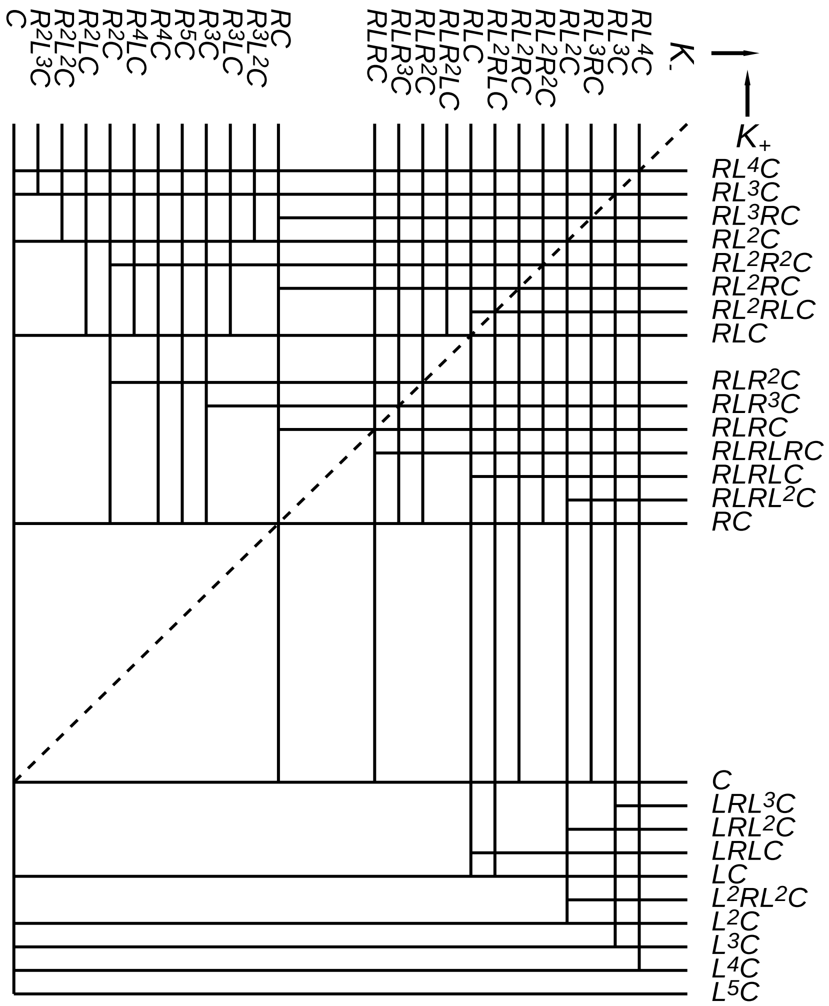

Figure 7: The kneading plane of the gap map. Periods up to 6 are shown.

What happens when

$$K_+ = \Pi D_j \cdots$$

contains one of the critical points D_j, $j = 1, 2, 3$? Since the letter after D_j must be a symbol corresponding to $f(D_j)$, it follows from (42) that the kneading sequence is

$$K_+ = \Pi D_j K_-. \tag{53}$$

If K_- ends with a C_i, then this defines a double superstable periodic sequence, i.e., a "joint" which we have mentioned in section 2.

However, in order to write down the upper and lower sequences of K_+ it is not enough to replace D_j by its neighboring letters. For instance, look at the case of ΠD_2. The two neighbors of D_2 are M and N, but ΠM and ΠN are not the closest sequences to ΠD_2, as it is possible to insert other sequences in between. Since for the sine-square map the largest sequence is RL^∞ and the smallest — L^∞, the smallest sequence, which is larger than ΠD_2, is $(\Pi D_2)_+ L^\infty$. Similarly, the largest sequence, which is smaller than ΠD_2, is $(\Pi D_2)_- L^\infty$. We list these sequences in descending order:

$$
\begin{array}{l}
(\Pi D_2)_+ K_- \\
\cdots \\
(\Pi D_2)_+ L^\infty \\
\Pi D_2 K_- \\
(\Pi D_2)_- L^\infty \\
\cdots \\
(\Pi D_2)_- K_-
\end{array}
$$

Therefore, we have found the upper and lower sequences of ΠD_2:

$$
\begin{array}{l}
(\Pi D_2)_+ L^\infty, \\
(\Pi D_2)_- L^\infty,
\end{array}
\tag{54}
$$

where, as usual,

$$
\begin{array}{l}
(\Pi D_2)_+ = \max\{\Pi M, \Pi N\}, \\
(\Pi D_2)_- = \min\{\Pi M, \Pi N\}.
\end{array}
$$

Now, the lower sequence always has a $-$ parity.

Since the letter D_3 has only a smaller neighbor R, a kneading sequence $K+ = \Pi D_3$ has a lower sequence

$$(\Pi D_3)_- = \Pi R L^\infty,$$

only when ΠR has $-$ parity.

Analogous discussion may be carried out for K_-. We list only the results:

1. The upper and lower sequences of $K_- = \Sigma D_2$ are $(\Sigma D_2)_+^\infty$ and $(\Sigma D_2)_-^\infty$, respectively, with the usual definition for $(\ldots)_\pm$.

2. The lower sequence of ΣD_3, if exists, is $(\Sigma R)^\infty$.

3. The sequences, closest to $K_- = \Pi C_i K_+$, are (in descending order)

$$
\begin{array}{l}
\cdots \\
(\Pi C_i)_+ R L^\infty \\
(\Pi C_i)_+ D_3 \\
\Pi C_i K_+ \\
(\Pi C_i)_- D_3 \\
(\Pi C_i)_- R L^\infty \\
\cdots
\end{array}
$$

Knowing how to construct all upper and lower sequences of the superstable sequences, we can generate all the kneading sequences K_+ and K_- up to a given length. Take K_+ for example. From the ordering rule it follows that among all K_+ the maximal one is $K_+ = D_3$ and the minimal one — $K_+ = L^\infty$. Using the lower sequence of D_3, i.e., R^∞, and inserting all possible critical points in the first column according to the natural order (11), we have (in descending order):

$$
\begin{array}{rcl}
 & & D_3 \\
R^\infty & = & RRR\cdots \\
 & & \cdots \\
 & & C_2 \\
 & & \cdots \\
 & & D_2 \\
 & & \cdots \\
 & & C_1 \\
 & & \cdots \\
L^\infty & = & LLL\cdots
\end{array}
$$

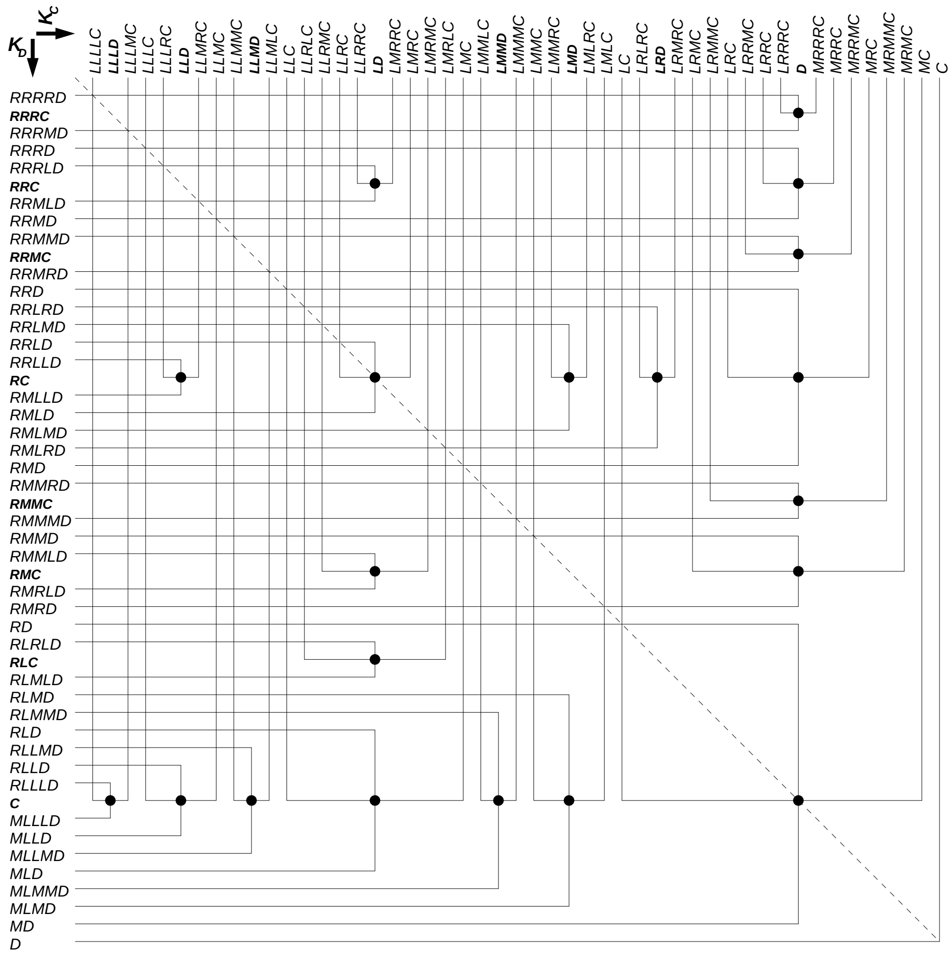

Figure 8: The kneading plane of a general cubic map. Joints and bones up to period 5 are shown.

Now, putting in the upper and lower sequences of C_2, D_2, and C_1, and inserting critical points in the second column according to the natural order (11), we get all superstable kneading sequences up to period 2, and so on, and so forth.

In this way we can get all K_+ and K_- up to a certain period. Without checking their compatibility, these K_+ and K_- give the axes which span the kneading plane. In order to continue these kneading sequences into the plane, we have to check whether a pair (K_+, K_-) satisfies the condition (43) or not. All kneading sequences up to period 3 are shown in Fig. 5. Solid circles in the figure are joints; the corresponding double superstable sequences are labeled explicitly.

Which bones grow from a joint may be determined from continuity consideration, i.e., by changing one of the critical points into its upper and lower sequences. This is shown in Fig. 6, where a general joint $[YD, XC]$ is drawn. The dash line vanishes for (YD_3), since no upper sequence $(YD_3)_+$ exists. For YD_2 the dash line should be understood as a solid one.

Not going into further details of this type of construction, we show the kneading planes of a few maps.

Fig. 7 shows the kneading plane of the gap map up to period 6. All the unusual behavior observed in [11] may be understood by inspecting this figure.

The kneading plane of a general cubic map is shown in Fig. 8. (It first appeared in [17] as

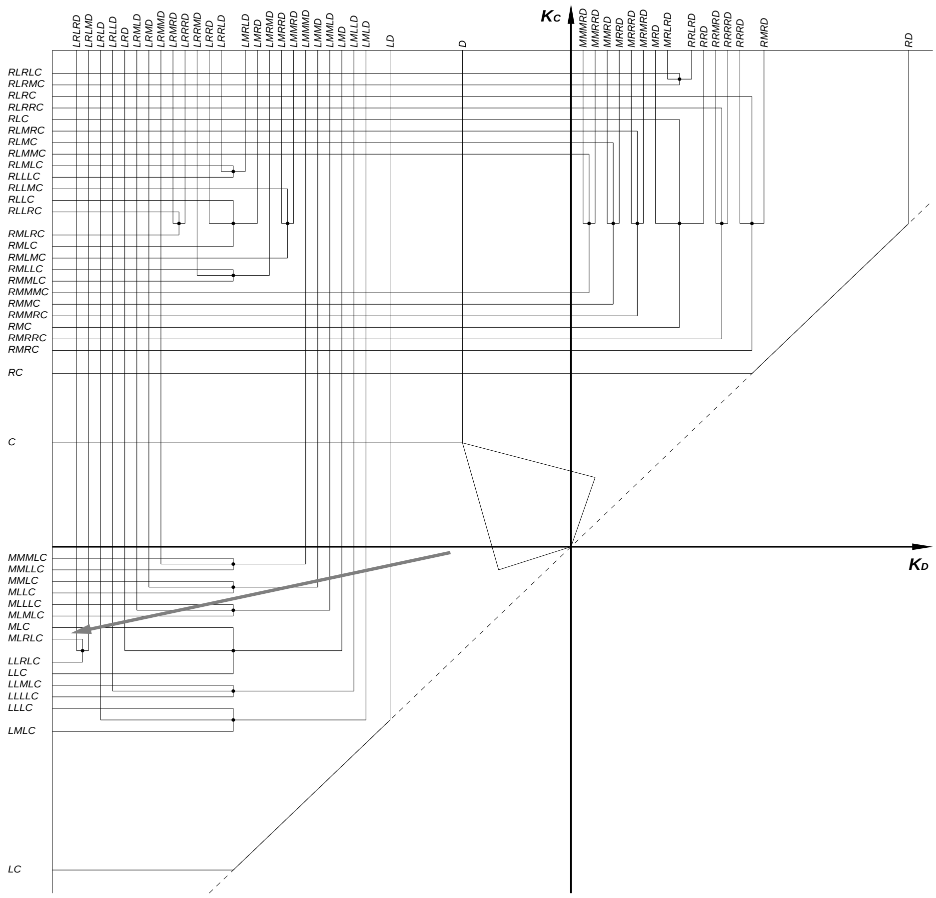

Figure 9: Kneading plane of a $(-,+,-)$ type cubic map. For the thick arrow see text.

Fig. 8 without explicitly labeling the bones. A pair of missing bones has been added.) Bones up to period 5 are drawn and joints are shown as solid circles. In a sense, a kneading plane tells almost everything about the global structure of the parameter space. It applies to all maps of the same topological type. In order to draw it, there is no need to know explicitly the mapping function. All we need is the ordering rule and admissiblity conditions. To transform a kneading plane to the parameter plane of a given map, it is enough to invoke the word-lifting technique, which in the case of two independent kneading sequences yields the location of all bones and joints.

Figure 8 corresponds to a cubic map with two increasing and one decreasing branches, so-called map of $(+,-,+)$ type. One can as well consider a cubic map of $(-,+,-)$ type. The ordering rule and admissibility conditions lead to the kneading plane, shown in Fig. 9. It first appeared in [17] as Fig. 12 without explicitly labeling the bones. We use this figure to comment on modeling antimonotonicity, i.e., reversal of period-doubling cascades in nonlinear systems, by cubic maps[21]. While in higher dimensional systems antimonotonicity is an inevitable and essential phenomenon, in one-dimensional maps with multiple critical points, it depends on how a parameter is varied in the parameter space. In Fig. 9 a thick arrow is superimposed on the kneading plane, which shows the range of parameter, used in one of the bifurcation diagrams in [21] to show the appearance of antimonotonicity. We note that there are infinitely many ways to exhibit as well as to avoid antimonotonicity. When antomonotonicity arises, we can read off the name of periodic orbits invloved from the kneading plane.

5 Symmetry Breaking and Symmetry Restoration

Symmetry breaking and symmetry restoration are common phenomena in physical systems with a certain kind of symmetry. An equation or a thermodynamical potential may possess a higher symmetry, but a particular solution or an equilibrium state may exhibit only a lower symmetry. All these asymmetric solutions (or states), taken together, restore the original symmetry. It is interesting to note that a simple form of symmetry breaking and restoration appears in the bifurcation structure of many dynamical systems. In particular, the existence of symmetric orbits, which first undergo a symmetry-breaking bifurcation into asymmetric orbits, and then enjoy period-doubling, has been observed in many ordinary and partial differential equations as well as in laboratory experiments. The fact that symmetry breaking always precedes period-doubling has been called "precursor"[22] to period-doubling or "suppression"[23] of period-doubling. Moreover, it is easy to see that orbits of even periods are capable of symmetry breaking, but not all even periods can do so. All this can be explained by symbolic dynamics on the example of the antisymmetric cubic map (8)[24].

The bifurcation diagram of the antisymmetric cubic map is shown in Fig. 10, where the symmetry breaking of a period 2 orbit is most clearly seen. The asymmetric period 2 orbit develops a period-doubling cascade with its own reversed period-halving sequence of chaotic bands. In the chaotic regime the symmetry suddenly restores.

The two kneading sequences K and $\overline{K}$ are related by the symmetry transformation (35), so it is enough to look at one of them, say, $\overline{K}$. If a periodic $\overline{K}$ contains only one critical point $\overline{C}$:

$$\overline{K} = \Sigma\overline{C}, \tag{55}$$

then it must correspond to an asymmetric orbit for the simple reason that it does not contain the letter C. Applying the transformation (35) to $\overline{K}$, we get

$$K = \overline{\Sigma}C. \tag{56}$$

This is the other asymmetric orbit, located symmetrically to (55).

A superstable symmetric kneading sequence must contain both $\overline{C}$ and C:

$$\overline{K} = \Sigma C\overline{\Sigma}\overline{C}. \tag{57}$$

First, this is an orbit of even period; only those even periods which can be decomposed like (57) can undergo symmetry breaking. Second, it is invariant under the transformation (35), since it can be brought back by cyclic permutation.

Take, for example, the simplest case when Σ is blank. We have a doubly superstable orbit $C\overline{C}$. We can shift C a little into one of its neighbors, R or M. To keep the symmetry, we must change $\overline{C}$ into L or M, respectively. Similar to the periodic window theorem, now we have a window of symmetric period 2 orbit

$$(MM, C\overline{C}, RL). \tag{58}$$

Applying the transformation (35) to (58), we get

$$(MM, \overline{C}C, LR),$$

which goes back to (58) under cyclic permutation. Therefore, it is indeed a symmetric orbit. However, the signature of this window is (note that in the cubic map only the letter M has $-$ parity)

$$(+, 0, +),$$

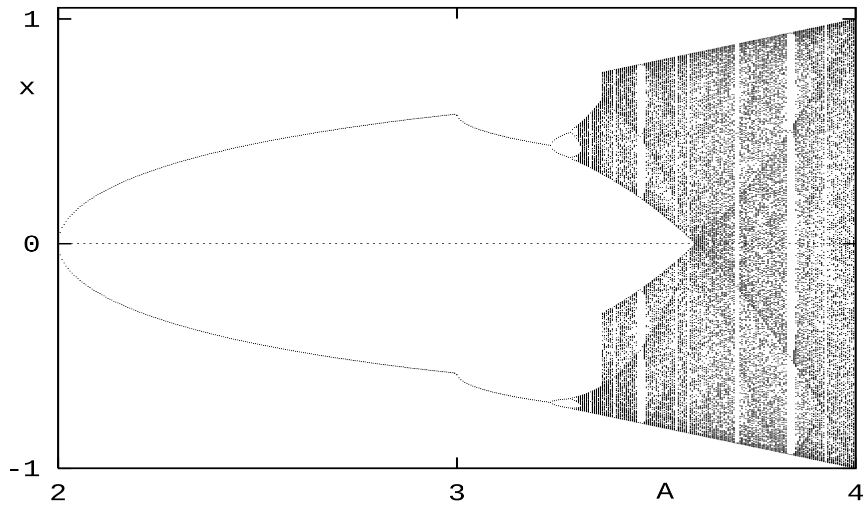

Figure 10: Bifurcation diagram of the antisymmetric cubic map.

which can never undergo a period-doubling bifurcation. A symmetry breaking takes place when an "inverse point"

$$x^* = -f(\mu, x^*), \tag{59}$$

which is analogous to a fixed point $x^* = f(\mu, x^*)$, loses stability. Symbolic dynamics only deals with monotonicity and continuity of the mapping function, it does not know where a symmetry breaking really happens. Consequently, the symbolic sequence RL in the symmetric window (58) continues to the other side of the symmetry breaking point, giving birth to

$$(MM, C\overline{C}, RL)\ (RL, R\overline{C}, RM). \tag{60}$$

We have used continuity argument again, so the letter L goes first into $\overline{C}$, then into M. Applying now the symmetry transformation (35) to (60), we get

$$(MM, \overline{C}C, LR)\ (LR, LC, LM). \tag{61}$$

The second triple (LR, LC, LM) cannot be brought back to $(RL, R\overline{C}, RM)$ by cyclic permutation. They correspond to the asymmetric period 2 orbits, which exist in its own basin. The asymmetric window has a signature

$$(+, 0, -),$$

therefore, it is capable to undergo further period-doubling. Take RL and RM as λ and ρ in (51), the corresponding chaotic map is given by $\rho\lambda^\infty = RM(RL)^\infty$. The precise parameter may be calculated by using the word-lifting technique. It is nothing but the point where symmetry restoration takes place. The mechanism of symmetry restoration consists in the collision of the asymmetric chaotic attractor with the unstable symmetric orbit and the attractor accquires the symmetry of the latter.

6 Two-Dimensional Maps and Differential Equations

In order to cut a long story short we will only mention a few new developments of symbolic dynamics in higher dimensions by giving some references.

In two- and higher-dimensional systems the nice ordering property of real numbers and the simple partition of an interval, which have played crucial role in symbolic dynamics of one-dimensional maps, no longer exist. The partition for the Hńon map[25], using tangencies of the invariant manifolds, was first discussed by Grassberger and Kantz[26]. Then Cvitanović, Gunaratne, and Procaccia[27] discussed its symbolic dynamics. Later the role of forward and backward foliations of the map in determining the partition lines has been recognized by Zheng and collaborators. The simplest case turns out to be the two-dimensional version of the sawtooth map, introduced by Tél[28]. Its symbolic dynamics was constructed in [29]. The piecewise linear counterpart of the Hénon map, so-called Lozi map[30], may be treated in a similar way[31]. The two piecewise linear maps helped to reach a deeper understanding of the symbolic dynamics of the Hénon map[32, 33].

Some years ago we have applied symbolic dynamics of one-dimensional maps to the systematics of periodic orbits in differential equations. In particular, the ordering of periodic orbits of the periodically forced Brusselator has been compared to that of the quadratic map, using symbolic dynamics of two letters. See Chapter 5 of [3] for details and references. The systematics of periodic orbits in the autonomous Lorenz model has been juxtaposed with the ordering of kneading sequences in the antisymmetric cubic map (8), i.e., symbolic dynamics of three letters[34, 35]. Our main argument for so doing lies in the shrinking of phase space volume due to dissipation. However, the Poincaré maps of both systems are essentially two-dimensional and there is no *a priori* reason that the two-dimensional nature will not show off. Having reached a better understanding of symbolic dynamics of two-dimensional maps, we have undertaken the job of justifying the previous one-dimensional approach and revealing the cases where a two-dimensional study leads to essentially new insight. Our first results seem to be promising and will be published elsewhere.

Acknowledgements

We thank W.-M. Zheng and F.-G. Xie for collaboration and discussion on many problems touched in these lectures. This work has been supported by the Open Laboratories Project of Academia Sinica. We also express our gratitude to the Pohang Institute of Science and Technology for hospitality and support, which has made the participation in the workshop on "Nonlinear Dyanmics and Chaos" a pleasant experience.

References

[1] M. Morse, and G. A. Hedlund, *Am. J. Math.* **60**, 815(1938); reprinted in *Collected Papers of M. Morse*, vol. 2, World Scientific, 1986.

[2] A. M. Alekseev, and M. V. Yakobson, *Phys. Repts.* **75**, 287(1981).

[3] Hao Bai-lin, *Elementary Symbolic Dynamics and Chaos in Dissipative Systems*, World Scientific, Singapore(1989).

[4] W.-M. Zheng, and B.-L. Hao, in *Experimental Study and Characterization of Chaos*, vol. 3 in *Directions in Chaos*, ed. by Hao Bai-lin, World Scientific, 363-459(1990).

[5] W.-M. Zheng, and B.-L. Hao, *Applied Symbolic Dynamics*, a monograph in Chinese, Shanghai Science and Technology Education Press, 1993, in press.

[6] N. Metropolis, M. L. Stein, and P. R. Stein, *J. Combinat. Theory* **A15**, 25(1973).

[7] J. Milnor, and W. Thurston, in *Dynamical Systems*, ed. by J. C. Alexander, *Lect. Notes in Math.*, No. 1342, Springer-Verlag(1988).

[8] Guckenheimer J, Moser J, and Newhouse S E (1980), *Dynamical Systems*, C. I. M. E. Lectures, Progress in Mathematics **8**, Birkhäuser.

[9] B. Derrida, A. Gervois, and Y. Pomeau, *Ann. Inst. H. Poincaré* **29A**, 305(1978).

[10] P. Collet, and J.-P. Eckmann, *Iterated Systems on the Interval as Dynamical Systems*, Birkhäuser, Boston(1980).

[11] M. C. de Sousa Vicira, E. Laso, and C. Tsallis, *Phys. Rev.* **A35**, 945.

[12] W.-M. Zheng, *Phys. Rev.* **A39** (1989), 6608.

[13] W.-M. Zheng, and L.-S. Lu, *Commun. Theor. Phys.* **15** (1991), 161.

[14] W.-Z. Zeng, B.-L. Hao, G.-R. Wang, and S.-G. Chen, *Commun. Theor. Phys.* **3** (1984) 283.

[15] H.-J. Zhang, P.-Y. Wang, J.-H. Dai, C.-D. Jin, and B.-L. Hao, *Chinese Phys. Lett.* **2**, 5(1985); *Commun. Theor. Phys.* **8**, 281(1987).

[16] F.-G. Xie, and B.-L. Hao, *Chaos, Solitons and Fractals*, **3** (1993), 47.

[17] R. S. MacKay, and C. Tresser, *Physica* **27D** (1987), 412.

[18] B.-L. Hao, and W.-M. Zheng, *Int. J. Mod. Phys.* **B3**, (1989), 235.

[19] W.-M. Zheng, *Int. J. Mod. Phys.* **B3** (1989), 1703.

[20] W.-M. Zheng, *J. Phys.* **A22** (1989), 3307.

[21] S. P. Dawson, and C. Grebogi, *Chaos, Solitons and Fractals* **1** (1991), 137; S. P. Dawson, C. Grebogi, J. A. Yorke, and H. Kocak, *Phys. Lett.* **A162** (1992), 249.

[22] D. D'Humieres, M. R. Beasley, B. A. Huberman, and A. Libchaber, *Phys. Rev.* **A26** (1982), 3483; K. Kumar, A. K. Agarwal, J. K. Bhattacharjee, and K. Banerjee, *Phys. Rev.* **A35** (1987), 2334.

[23] K. A. Wiesenfeld, E. Knobloch, R. P. Miraoky, and J. Clark, *Phys. Rev.* **A29** (1984), 2102; J. S. Swift, and K. A. Wiesenfeld, *Phys. Rev. Lett.* **52** (1984), 705.

[24] W.-M. Zheng, and B.-L. Hao, *Int. J. Mod. Phys.* **B3**, (1989), 1183.

[25] M. Hénon, *Commun. Math. Phys.* **50** (1976), 69.

[26] P. Grassberger, and H. Kantz, *Phys. Lett.* **113A** (1985), 235.

[27] P. Cvitanović, G. H. Gunaratne, and I. Procaccia, *Phys. Rev.* **A38** (1988), 1503.

[28] T. Tél, *Phys. Lett.* **97A** (1983), 219.

[29] W.-M. Zheng, *Commun. Theor. Phys.* **17** (1992), 167.

[30] R. Lozi, *J. de Physique* **39C** (1978), 9.

[31] W.-M. Zheng, *Chaos, Solitons and Fractals* **1** (1991), 243; **2** (1992), 461.

[32] H. Zhao, W.-M. Zheng, ang Y. Gu, *Commun. Theor. Phys.* **17** (1992), 263.

[33] H. Zhao, and W.-M. Zheng, *Commun. Theor. Phys.* **19** (1993), 21.

[34] M.-Z. Ding, and B.-L. Hao, *Commun. Theor. Phys.* **9** (1988), 375.

[35] H.-P. Fang, and B.-L. Hao, in preparation.