

B. Schweizer, Department of Mathematics and Statistics,
University of Massachusetts, Amherst, MA 01003, USA
e-mail: bert@math.umass.edu

A. Sklar, Department of Mathematics,
Illinois Institute of Technology, Chicago, IL 60616, USA

J. Smítal, Mathematical Institute, Silesian University,
746 01 Opava, Czech Republic
e-mail: Jaroslav.Smital@math.slu.cz

DISTRIBUTIONAL (AND OTHER) CHAOS AND ITS MEASUREMENT *

Abstract

After surveying several earlier definitions of “chaos”, this paper is devoted to presenting the recently introduced notion of *distributional chaos* to a non-specialist audience. It is shown that the theory of distributional chaos avoids various shortcomings of the earlier theories and that it allows one not only to distinguish between chaotic and non-chaotic behavior, but also to measure the actual extent of any existing chaotic behavior. The whole is illustrated with numerous examples.

1 Introduction

Since Li and Yorke [27] introduced the term into the vocabulary of mathematics in 1975, several different (rigorous) definitions of “chaos” have been proposed. Each of these definitions yields a “Yes or No” test for the presence

Key Words: Distributional chaos, topological entropy, invariant measures, Li-Yorke chaos

Mathematical Reviews subject classification: Primary: 37-02, 37D45, 26A16 Secondary: 54E70, 54H20

Received by the editors December 13, 1999

*This paper was made possible by visits of B. Schweizer and A. Sklar to the Silesian University in Opava, and A. Sklar and J. Smítal to the University of Massachusetts in Amherst. This research was supported in part by ONR contract N-00014-87-K-0379, and by the Grant Agency of the Czech Republic, Grants No. 201/94/1088 and 201/97/0001. Support of these institutions is gratefully acknowledged

of chaos, but none provides a way of measuring the extent of chaos in a given system, or of comparing the chaotic behavior of two different systems. In addition, as will be seen in the sequel, each of the definitions has certain other deficiencies: in particular, some are such that the very existence of chaos can be wiped out by arbitrarily small perturbations.

In a recent paper [38] Schweizer and Smítal, the first and the third authors of this paper, used ideas from the theory of probabilistic metric spaces to develop a new definition of “chaos” and a corresponding variety of measures of chaos. There, and in a subsequent paper written jointly with F. Balibrea and A. Sklar [5], they elaborated the resulting theory in some detail, showing in particular that, at least for one-dimensional systems, the new definition avoids the deficiencies mentioned above. However, the two papers, addressed as they are to experts in the field, are unavoidably technical and are provided with only a few illustrative examples.

Our aim in the present paper is to introduce this new notion of “distributional chaos” to a wider audience, to illustrate it with an appropriate body of examples, and to outline the salient features of the resulting theory. Among these features is one we may note here: when combined with the results of Schweizer and Sklar [34], the notion of distributional chaos allows us to clearly distinguish between ergodic behavior on the one hand and genuinely chaotic behavior on the other. (See also Sklar and Smítal [44].)

Our paper is organized as follows: After this introduction, Section 2 presents some necessary preliminaries and Section 3 reviews, compares and criticizes some earlier definitions of chaos. In Section 4 we introduce the notion of distributional chaos, discuss some of its properties and define a measure of chaos to which it leads. Section 5 is devoted to the familiar “tent map” τ , defined on the unit interval $[0, 1]$ by $\tau(x) = 1 - |2x - 1|$, and to several closely related functions. Here the fact that we are able to give direct and simple verifications of our main results allows us to lay bare the central ideas without appealing to the heavy machinery of the general theory. (Of course, without the general theory, these examples would not be as meaningful.) Section 6 summarizes the general theory for continuous functions on compact intervals as developed by Schweizer and Smítal in [38] and augmented by Balibrea et al. in [5]. The final Section 7 presents extensions of the general theory, in particular to higher dimensions, and concludes with a discussion of the abovementioned distinction between ergodicity and chaos and a conjecture about their co-occurrence.

2 Preliminaries

The *composite* of two functions f and g is the function $f \circ g$ whose domain is the set of all x in the domain of g for which $g(x)$ is in the domain of f , and whose value for any such x is $f(g(x))$. In this paper we shall generally (but not always) be dealing with functions mapping a given set E into itself. The *iterates* of any such function f are the functions defined recursively by:

f^0 is the identity function on E ;

$$f^{n+1} = f \circ f^n \text{ for all non-negative integers } n.$$

An *orbit* of f is an equivalence class under the equivalence relation \sim_f defined on E by: $x \sim_f y$ if and only if there exist non-negative integers m, n such that $f^m(x) = f^n(y)$. Orbits come in two varieties, cyclic and acyclic. A *cyclic* orbit is one that contains a (necessarily unique) *cycle*, i.e., a set that, for some positive integer n , consists of n distinct points x_1, \dots, x_n such that, if $n = 1$ then $f(x_1) = x_1$, while if $n > 1$ then

$$f(x_m) = x_{m+1} \text{ for } m = 1, 2, \dots, n-1 \text{ and } f(x_n) = x_1.$$

The *order* of a cycle is the number of points in it. Thus a cycle of order 1 is simply a fixed point of f ; and generally, any point in a cycle of order n is a *periodic point* of f , of (exact) *period* n . Clearly, a periodic point of f of period n is a fixed point of f^n . An *acyclic* orbit is one that contains no cycle of any order.

Given x in E , the sequence whose n^{th} term is $f^n(x)$ ($n \geq 0$) is the *trajectory* of x (under f). Note that the sequence, which is infinite, has to be distinguished from the set $\{x, f(x), f^2(x), \dots\}$, which may be finite. Some authors refer to either this set or the sequence as the “orbit” (or “forward orbit”) of x , thus applying the term “orbit” to points rather than to functions; for us, the set of terms of a trajectory is a part, and generally a proper part, of an orbit. The trajectories of any two points in the same orbit ultimately merge; in particular, the trajectories of all points in a cyclic orbit end in the cycle of the orbit.

If f maps a compact space E into itself, then for every x in E , $\omega_f(x)$ is the set of limit points of the trajectory of x , i.e., the set of all y in E such that for each neighborhood V of y there is a positive integer n such that $f^n(x)$ is in V . Any such set is an *ω -limit set* of f . Note that, since any two points in the same orbit of f determine the same ω -limit set, an ω -limit set really belongs to an orbit of f rather than to any individual trajectory. The ω -limit set of a cyclic orbit is the cycle of the orbit; the ω -limit set of an acyclic orbit

may be the entire space; and clearly, distinct orbits can have identical ω -limit sets. Lastly, an ω -limit set is *maximal* if it is not properly contained in any other ω -limit set.

Two functions f and g are *conjugate* if there is an invertible function h whose domain includes the range of f and whose range includes the domain of g , such that

$$h \circ f = g \circ h. \quad (2.1)$$

As is well-known, conjugacy is an equivalence relation.

It follows readily from (2.1) that, for any x and y in the domain of f and any non-negative integers m, n , we have

$$f^m(x) = f^n(y) \text{ if and only if } g^m(h(x)) = g^n(h(y)), \quad (2.2)$$

i.e.,

$$x \sim_f y \text{ if and only if } h(x) \sim_g h(y). \quad (2.3)$$

Thus h is an *orbit-isomorphism*. The converse is also true, i.e., if there is a one-one function h such that (2.3) holds, then f and g are conjugate – simply let $m = 1$ and $n = 0$ in (2.2). This yields the important fact that:

Two functions are conjugate if and only if they are orbit-isomorphic.

If f maps a topological space X into itself, and g maps a topological space Y into itself, then f and g are *topologically conjugate* if the conjugating function h is a homeomorphism from X onto Y . Functions on topological spaces can be conjugate without being topologically conjugate and it will be important for us to observe the distinction between the two notions. Imposing conditions on h yields stronger versions of conjugacy. In particular, if X is a linear topological space and h is a linear function, then we speak of *linear conjugacy* (see Rice et al. [32] and Schweizer and Sklar [37]).

The *topological entropy* of a function f that maps a topological space into itself is a non-negative number $h(f)$ that, roughly speaking, measures the way in which the orbits of f are intertwined. The general definition of the notion, introduced by Adler et al. in [1], is rather complicated, but simple expressions for the value of the topological entropy are known in some special cases. For our purposes, we need only the fact that the topological entropy of a piecewise monotonic map f of an interval into itself is given by

$$h(f) = \lim_{n \rightarrow \infty} \left(\frac{1}{n} \ln c_n \right),$$

where c_n is the number of monotonic pieces (“laps”) of f^n . This was shown by Misiurewicz and Szlenk in [31] extending earlier work of Milnor and Thurston from 1977 (reprinted in 1988 in [28]). See also [2] and [3].

3 A survey of some definitions of “chaos”

On page 17 of his popular book “*Does God Play Dice?*” Ian Stewart, [47], suggests the following dictionary definition of “chaos”:

Stochastic behaviour occurring in a deterministic system.

Mathematically, speaking about chaos means speaking about the iterative behavior of functions that map topological or, more commonly, metric spaces into themselves. In accordance with this, we feel that any reasonable definition of chaos should enjoy the following properties:

1. **Comprehensiveness:** The definition should be meaningful for a large class of functions, including at the very least all functions mapping compact metric spaces into themselves.
2. **Stability:** If a function is chaotic, then any sufficiently small perturbation of the function should also be chaotic.
3. **Invariance:** Any function on a metric space topologically conjugate to a chaotic function on a (possibly different) metric space should also be chaotic.
4. **Familiarity:** Any function universally and unequivocally regarded as either chaotic or non-chaotic should be so classified by the definition.

One of the most familiar chaotic functions is the logistic function L_4 , defined on the unit interval $[0, 1]$ by

$$L_4(x) = 4x(1 - x);$$

L_4 is topologically conjugate to the tent map τ defined on $[0, 1]$ by $\tau(x) = 1 - |2x - 1|$ and is linearly conjugate to the restriction to the interval $[-1, 1]$ of the standard quadratic Čebyšev polynomial $2x^2 - 1$. Indeed, the restrictions to $[-1, 1]$ of the standard Čebyšev polynomials of degree at least 2 are classic examples of chaotic functions.

In what follows we shall, for illustrative purposes, be using the functions f_t defined, for $t > 0$ on the interval $[0, t]$ by

$$f_t(x) = x(4xt^{-1} - 3)^2.$$

It follows that for any $t > 0$, $f_t(0) = 0$, $f_t(t) = t$, and $f'_t(0) = f'_t(t) = 9$. Moreover, each f_t is linearly conjugate to the restriction to $[-1, 1]$ of the cubic

Čebyšev polynomial $4x^3 - 3x$. We now proceed to examine some commonly accepted definitions of chaos.

A function f that maps a metric space M with distance function d into itself is *chaotic in the sense of Li and Yorke* if it admits a scrambled set (this is an extension of an inessentially modified version of the original definition by Li and Yorke [27], cf. Kuchta and Smítal [25] or Fedorenko et al. [19]). Here, a *scrambled set* for a function f is a subset S of M containing at least two points, such that, for any two distinct points x, y in S ,

$$\liminf_{n \rightarrow \infty} d(f^n(x), f^n(y)) = 0 \quad \text{and} \quad \limsup_{n \rightarrow \infty} d(f^n(x), f^n(y)) > 0.$$

At first, and even at second glance, this definition seems to capture precisely our intuitive notion of the meaning of chaotic behavior, e.g., sensitivity to initial conditions. However, as shown by Fedorenko et al. [19], the real functions of topological entropy 0 that are chaotic in the sense of Li and Yorke (and there are very many of them, see Smítal [45]) form a set of the first category in the set of all real functions of topological entropy 0. This means that arbitrarily small perturbations of chaotic zero-entropy functions can destroy Li-Yorke chaos. Thus chaos in this sense is not stable.

An *ad hoc* way out of this predicament is to restrict the definition of Li-Yorke chaos to functions of positive topological entropy and, since all the familiar *non*-chaotic functions have topological entropy 0, to simply define a function to be chaotic if and only if it has positive topological entropy. This definition of chaos certainly appears to satisfy all the criteria listed above. Indeed, for continuous functions on intervals, where positive topological entropy is equivalent to the existence of cycles whose order is not a power of 2, see Misiurewicz [29], it is in common use, see Block and Coppel [11].

There is, however, a drawback. Topological entropy, as the name indicates, is a topological, not a metric notion. Consequently, the numerical value of the topological entropy can be a quite misleading indicator of the actual extent of chaotic behavior. For example, using the functions f_t defined above, consider the functions f_m^* defined, for positive integers m , on the unit interval by:

$$f_m^*(x) = \begin{cases} f_{1/m}(x), & x \in [0, \frac{1}{m}], \\ x, & x \in [\frac{1}{m}, 1]. \end{cases}$$

Each f_m^* is continuous and piecewise monotonic, with the number of monotonic pieces of f_m^{*n} being 3^n , independent of m . Hence all the functions f_m^* have the same topological entropy, namely $\ln 3$. Yet the chaotic behavior of f_m^* is confined to the interval $[0, \frac{1}{m}]$, which can be made arbitrarily small by taking m large enough.

By combining m copies of the function $f_{1/m}$, we can construct a continuously differentiable function g_m on $[0, 1]$ via:

$$g_m(x) = f_{1/m}(x - k/m) + k/m, \quad x \text{ in } [k/m, (k+1)/m], \quad k = 0, 1, \dots, m-1.$$

It is not difficult to see that for each non-negative integer n , the n^{th} iterate g_m^n of g_m has $m3^n - m + 1$ monotonic pieces, so that the topological entropy of g_m is again $\ln 3$, and again independent of m . Each g_m is chaotic and the chaotic behavior now extends throughout the unit interval. Yet under g_m , every trajectory is confined to an interval of length $1/m$. Thus for $m \geq 2$, and especially for m large, the function g_m , of topological entropy $\ln 3$, must definitely be considered less chaotic than the function L_4 , of topological entropy $\ln 2$, which has trajectories that are dense in the whole unit interval.

R. L. Devaney [15] and [16] has used a condition equivalent to the existence of trajectories that are dense in the domain of a function as part of his definition of chaos. According to Devaney, a function f that maps a metric space M with distance function d into itself is chaotic if it satisfies the following conditions:

1. The function f has “sensitive dependence on initial conditions”, i.e., there is a $\delta > 0$ such that for any x in M and any neighborhood N of x , there is a y in N and a non-negative integer n such that $d(f^n(x), f^n(y)) > \delta$.
2. The function f is topologically transitive, i.e., for any pair (U, V) of non-empty open subsets of M , there is a positive integer k such that $f^k(U) \cap V$ is non-empty. When M is compact, this condition is equivalent to the existence of trajectories that are dense in M .
3. The set of periodic points is dense in M .

In [7], Banks et al. showed that Conditions 2 and 3 together imply Condition 1. Nevertheless, Devaney’s definition has remained popular in unaltered form, probably precisely because it includes the redundant Condition 1. In particular, the unaltered definition is used in the well-known book by Barnsley [9].

In [10], Berglund and Vellekoop extended the theorem of Banks et al. by showing that, for continuous functions on (not necessarily compact) intervals, Condition 2 by itself implies both Conditions 1 and 3. Since it had earlier been pointed out by Assaf and Gadbois [4] that Conditions 1 and 3 together generally do not imply Condition 2, it can be said that the key element in Devaney’s definition is transitivity or the existence of dense trajectories. But

as the example of g_m for $m \geq 2$ shows, this dense trajectory condition is certainly not necessary for the existence of chaos. Nor is it, by itself, sufficient, as is shown by the example of the rotation of a circle through an irrational multiple of π , where every trajectory is dense in the circle, yet the rotation is clearly non-chaotic. Furthermore, as will be shown in Section 5, the property of topological transitivity can be destroyed by arbitrarily small perturbations, from which it follows that chaos in the sense of Devaney is also not stable.

Another possible definition of chaos is suggested by the fact that many of the familiar chaotic functions are strongly mixing with respect to well-behaved measures. But this possibility is ruled out if, once again, all the functions g_m are to be considered chaotic. This is so because for $m \geq 2$, each of the subintervals $[k/m, (k+1)/m]$, $k = 0, 1, \dots, m-1$, of the unit interval is invariant under g_m , whence g_m cannot even be ergodic, let alone mixing, with respect to any measure that assigns positive measure to each non-empty open subset of the unit interval.

Thus, of the various definitions (and parts of definitions) of chaos surveyed in this section, only positive topological entropy satisfies the criteria given at the beginning of the section. But, as we have seen, while positive topological entropy can be taken as an indicator of the presence of chaos, in no way can its numerical value be considered a measure of the extent of this chaos.

4 Distributional chaos

To introduce the notion of distributional chaos, we begin with a metric space (M, d) and a function f that maps M into itself. For any pair (x, y) of points of M , we define the sequence δ_{xy} by:

$$\delta_{xy}(m) = d(f^m(x), f^m(y)), \quad m = 0, 1, \dots$$

Next, for any pair (x, y) and any positive integer n , we define the functions $F_{xy}^{(n)}$ on the real line by

$$F_{xy}^{(n)}(t) = \frac{1}{n} \#\{m; 0 \leq m \leq n-1 \text{ and } \delta_{xy}(m) < t\}, \quad (4.1)$$

where $\#S$ is the number of elements in the set S . Clearly, each function $F_{xy}^{(n)}$ is non-decreasing, has minimum value 0 (since $F_{xy}^{(n)}(t) = 0$ for all $t \leq 0$), has maximum value 1 (since $F_{xy}^{(n)}(t) = 1$ for all t greater than the maximum of the numbers $\delta_{xy}(m)$, $m = 0, 1, \dots, n-1$), and is left-continuous (by virtue of the strict inequality $\delta_{xy}(m) < t$ in (4.1)). Thus each $F_{xy}^{(n)}$ is a left-continuous *distribution function* whose value at t may be interpreted as the probability

that the distance between the initial segments (of length n) of the trajectories of x and y is less than t .

For any real number a , let ϵ_a be the distribution function given by

$$\epsilon_a(t) = \begin{cases} 0, & t \leq a, \\ 1, & t > a. \end{cases}$$

Then, for any point x we have $F_{xx}^{(n)} = \epsilon_0$ for all n ; for any pair of points x and y forming a 2-cycle of f , we have $F_{xy}^{(n)} = \epsilon_{d(x,y)}$ for all n ; and in general, for any x and y , $F_{xy}^{(n)}$ is a step-function that can be expressed as a convex combination of a finite number of ϵ_a 's.

We are interested in the asymptotic behavior of the functions $F_{xy}^{(n)}$ as n gets large. Accordingly, we consider the functions F_{xy} and F_{xy}^* defined by:

$$F_{xy}(t) = \liminf_{n \rightarrow \infty} F_{xy}^{(n)}(t) \quad \text{and} \quad F_{xy}^*(t) = \limsup_{n \rightarrow \infty} F_{xy}^{(n)}(t).$$

For any pair (x, y) , the functions F_{xy} and F_{xy}^* are distribution functions with $F_{xy}(t) \leq F_{xy}^*(t)$ for all real t . They are not necessarily left-continuous but may, without loss of generality, be normalized to be left-continuous. We can then, without inconsistency, adopt the further convention that $F_{xy} < F_{xy}^*$ means that $F_{xy}(t) < F_{xy}^*(t)$ for some, and hence for all, t in some interval of positive length. We shall refer to F_{xy} as the *lower distribution*, and to F_{xy}^* as the *upper distribution* of x and y . It follows that F_{xy}^* is an asymptotic measure of how close x and y can come together while F_{xy} is an asymptotic measure of their maximum separation. Note also that if $\lim_{n \rightarrow \infty} d(f^n(x), f^n(y)) = a$, then $F_{xy} = F_{xy}^* = \epsilon_a$. We now make the following:

Definition 4.1. *A function f that maps a metric space M into itself is distributionally chaotic (briefly, d -chaotic) if there is a pair of points (x, y) in M such that $F_{xy} < F_{xy}^*$.*

The principal aim of the paper is to illustrate and disseminate the import and some of the consequences of this definition. We begin with a discussion of several basic results due to Schweizer and Smítal [38] (more details will be given in Section 6).

Let $C(I, I)$ be the set of continuous functions that map a compact interval I into itself. Then, first of all, the set of d -chaotic functions in $C(I, I)$ is non-empty. Next, f in $C(I, I)$ is d -chaotic if and only if $h(f) > 0$. Specifically, if $h(f) = 0$, which is the case if and only if the order of any cycle of f is a power of 2, see Misiurewicz [29], then $F_{xy} = F_{xy}^*$ for all x, y in I , whence f is not

d -chaotic. On the other hand, if $h(f) > 0$, then there exist points x, y in I and a positive number c such that

$$F_{xy}^* = \epsilon_0 \quad \text{and} \quad F_{xy} \leq \epsilon_c,$$

whence for any t in the open interval $(0, c)$, we have

$$0 = F_{xy}(t) < F_{xy}^*(t) = 1,$$

whence f is d -chaotic. It follows that the pair $\{x, y\}$ is a scrambled set, whence any function f in $C(I, I)$ that is d -chaotic is chaotic in the sense of Li and Yorke. The converse is false since there exist functions that are chaotic in the sense of Li and Yorke but have topological entropy 0, see Smítal [45], so are not d -chaotic.

For functions in $C(I, I)$, the properties “positive topological entropy” and “distributionally chaotic” are equivalent. But since the family of upper and lower distribution functions contains much more information than the single number $h(f)$, we can go further.

If the metric space (M, d) has a finite diameter d_M (in particular, if the space is compact) then, using the notion of distributional chaos, we can introduce a measure of the amount of chaos exhibited by a function f from M into M as follows: We first note that for any pair of points x and y in M we have the inequalities,

$$\epsilon_{d_M} \leq F_{xy} \leq F_{xy}^* \leq \epsilon_0,$$

whence for all real t ,

$$0 \leq F_{xy}^*(t) - F_{xy}(t) \leq \epsilon_0(t) - \epsilon_{d_M}(t).$$

Therefore,

$$0 \leq \frac{1}{d_M} \int_0^\infty (F_{xy}^*(t) - F_{xy}(t)) dt \leq \frac{1}{d_M} \int_0^\infty (\epsilon_0(t) - \epsilon_{d_M}(t)) dt = 1;$$

and this leads us to the following:

Definition 4.2. *Given a function f mapping a metric space (M, d) with finite diameter d_M into itself, the (principal) measure of chaos of f is the number $\mu_p(f)$ in the unit interval $[0, 1]$ given by:*

$$\mu_p(f) = \sup_{x, y \in M} \frac{1}{d_M} \int_0^\infty (F_{xy}^*(t) - F_{xy}(t)) dt.$$

It follows at once that $\mu_p(f) \neq 0$ if and only if f is d -chaotic.

In some instances, e.g., for the tent map (see Section 5), the principal measure of chaos can be determined explicitly. There are also other measures of chaos which we shall define and discuss in Sections 6 and 7.

5 The tent map

Our aim in this section is to illustrate the subject matter of the preceding section with several closely related examples. We begin with the well-known tent map τ which is defined on the unit interval $[0, 1]$ by $\tau(x) = 1 - |2x - 1|$, i.e.,

$$\tau(x) = \begin{cases} 2x, & x \in [0, \frac{1}{2}], \\ 2(1-x), & x \in [\frac{1}{2}, 1]. \end{cases}$$

An easy induction yields that τ^n is given by

$$\tau^n(x) = \begin{cases} 2^n(x - \frac{k}{2^{n-1}}), & x \in [\frac{k}{2^{n-1}}, \frac{2k+1}{2^n}], \\ 2^n(\frac{k+1}{2^{n-1}} - x), & x \in [\frac{2k+1}{2^n}, \frac{k+1}{2^{n-1}}], \end{cases}$$

for $k = 0, 1, \dots, 2^{n-1} - 1$.

The graph of τ^n consists of 2^{n-1} compressed copies of the graph of τ , placed side-by-side over the interval $[0, 1]$. It follows that τ^n has 2^n fixed points. The largest of these is $2^n/(2^n + 1)$ and it belongs to the n -cycle of τ consisting of the points

$$2^k/(2^n + 1), \quad k = 1, 2, \dots, n.$$

For $n > 2$, there are other n -cycles, but it is this one, which we call the *principal n -cycle* of τ , that will concern us here. Since τ has the principal 3-cycle $\{2/9, 4/9, 8/9\}$, it is immediate that τ is chaotic in the sense of Li and Yorke (cf. Section 3). Moreover, since τ^n has 2^n monotonic pieces, the topological entropy of τ is $\ln 2 > 0$, whence τ is d -chaotic.

The function τ has the fixed points 0 and $2/3$; and it is not difficult to show that every binary rational (i.e., number of the form $m/2^n$) in $[0, 1]$, but no other number, is mapped onto 0 by some iterate of τ . In other words, the τ -orbit of 0 is precisely the set of binary rationals in $[0, 1]$. As regards other τ -orbits, it is easy to see that x is in a cyclic τ -orbit if and only if x is rational and that x is actually in the cycle of a cyclic τ -orbit if and only if x is of the form $\frac{\text{even}}{\text{odd}}$.

To proceed further, it is useful to work with the binary representation of numbers in $[0, 1]$. In view of the above, we need only do this for numbers that are not binary rationals. Each such number has a unique representation as a sequence of 0's and 1's that contains an infinite number of 0's and an infinite number of 1's. If s is such a sequence, we let \bar{s} be the sequence obtained from s by interchanging 0 and 1 at each place. From the definition of τ , we see that if x has the binary representation $.0s$ (a sequence beginning with 0 followed by the sequence s), then $\tau(x) = .s$, while if $x = .1s$ then $\tau(x) = \bar{s}$.

Many things follow from this. For one, let s be a sequence that contains all possible finite sequences of 0's and 1's, each such finite sequence being preceded in s by a 0, and let x be the number represented by s . Then since any sequence beginning with 0 and followed by a finite sequence of 0's and 1's is the initial sequence of $\tau^n(x)$ for some n , it follows that the ω -limit set of x is the whole interval $[0, 1]$, i.e., the trajectory of x is dense in $[0,1]$ (whence τ is transitive and $[0,1]$ is the unique maximal ω -limit set of τ). Clearly, there are non-denumerably many such numbers x . Now recall that $.0000\dots$ is a binary representation of the fixed point 0 and $.101010\dots$ the binary representation of the fixed point $2/3$ and consider a sequence s constructed as follows: s begins with a long string of 0's, followed by a much longer, say ten times as long, string of 10's. This is followed by a string of 0's a hundred times as long as the entire preceding initial string of s , followed in turn by a string of 10's a thousand times as long as its preceding initial string of s , and so on, alternating exponentially longer strings of 0's and 10's. It follows that if p_1 is the number represented by s , then as n increases the sequence $\{F_{p_1 0}^{(n)}\}$ alternates between coming closer and closer to ϵ_0 and closer and closer to $\epsilon_{2/3}$. Thus $F_{p_1 0} = \epsilon_{2/3}$ and $F_{p_1 0}^* = \epsilon_0$, which shows directly that τ is d -chaotic and that $\mu_p(\tau) \geq 2/3$.

We can improve this estimate by bringing the principal (and only) 2-cycle $\{2/5, 4/5\}$ of τ into play. We do this by constructing a sequence that alternates long strings of 0's, longer strings of 10's, still longer strings of 1100's ($.110011001100\dots = 4/5$), back to exponentially longer strings of 0's, and so ad infinitum. If p_2 is the number represented by this sequence, then the distribution functions $F_{p_2 0}^{(n)}$ alternately approach the three functions $\epsilon_0, \epsilon_{2/3}$ and $\frac{1}{2}(\epsilon_{2/5} + \epsilon_{4/5})$. It follows that $F_{p_2 0}^*$ is the pointwise maximum of these three functions, which is ϵ_0 , while $F_{p_2 0}$ is the corresponding pointwise minimum, which is $\frac{1}{2}(\epsilon_{2/3} + \epsilon_{4/5})$. Using $F_{p_2 0}$ and $F_{p_2 0}^*$ in place of $F_{p_1 0}$ and $F_{p_1 0}^*$ adds an amount $\frac{1}{2}(\frac{4}{5} - \frac{2}{3}) = \frac{1}{15}$ to the previous lower bound for $\mu_p(\tau)$, yielding $\mu_p(\tau) \geq \frac{11}{15}$.

We can continue the process as follows: For any positive integer r , let 0^r denote a string of r zeros; and for any positive integer n , let n^r denote the string of r repetitions of the string $1\dots 10\dots 0$ consisting of n ones followed by n zeros (note that the infinite string $n^1 n^1 n^1 \dots$ is the binary representation of the number $2^n/(2^n + 1)$); next let $\{k_m\}$ be a strictly increasing sequence of positive integers that grows so fast that the sequence whose m 'th term is $(k_1 + k_2 + \dots + k_m)/k_{m+1}$ decreases monotonically to 0 (e.g., take $k_{m+1} = (m + 1)(k_1 + k_2 + \dots + k_m)$); and lastly form the infinite binary sequence

$$0^{k_1} 0^{k_2} 1^{k_3} 0^{k_4} 1^{k_5} 2^{k_6} 0^{k_7} 1^{k_8} 2^{k_9} 3^{k_{10}} 0^{k_{11}} 1^{k_{12}} 2^{k_{13}} 3^{k_{14}} 4^{k_{15}} 0^{k_{16}} \dots$$

where the rule of formation should be clear (in particular, a string of type n^{km} makes its first appearance in the sequence when $m = (n + 1)(n + 2)/2$). This binary sequence represents a number p in $(0, 1)$, and it readily follows that $F_{p0}^* = \epsilon_0$ while, for any t in $[0, 1]$,

$$F_{p0}(t) = \inf\{F_{c(n)0}(t); n = 1, 2, \dots\}$$

$$= \begin{cases} 0, & 0 \leq t \leq c(1), \\ 1 - \frac{1}{n}, & c(n-1) < t \leq c(n), \text{ for } n \geq 2, \end{cases}$$

where $c(n) = 2^n/(2^n + 1)$. Consequently,

$$\begin{aligned} \mu_p(\tau) &\geq \int_0^\infty (F_{p0}^*(t) - F_{p0}(t))dt \\ &= \frac{2}{3} + \sum_{n=2}^\infty \frac{1}{n} \frac{2^{n-1}}{(2^n + 1)(2^{n-1} + 1)} \approx 0.7861937\dots \end{aligned}$$

We now proceed to show that the distribution function F_{p0} is minimal in the set of all distribution functions F_{uv} for u, v in $[0, 1]$, whence the sum of the above infinite series is in fact equal to $\mu_p(\tau)$. For $n \geq 2$, $t > c(n - 1)$, and arbitrary u, v in $[0, 1]$, consider those non-negative integers m (if any) for which $\delta_{uv}(m) \geq t$. This can happen only if one of $f^m(u), f^m(v)$ is in the interval $[0, 1/(2^{n-1} + 1))$ and the other is in the interval $(c(n - 1), 1]$. It follows from the definition of τ that each of the numbers

$$f^{m+1}(u), f^{m+2}(u), \dots, f^{m+n-1}(u), f^{m+1}(v), f^{m+2}(v), \dots, f^{m+n-1}(v)$$

is in the interval $[0, c(n - 1))$, whence each of the $n - 1$ numbers $\delta_{uv}(m + 1), \delta_{uv}(m + 2), \dots, \delta_{uv}(m + n - 1)$ is less than $c(n - 1)$. This means that any 2 distinct m 's for which $\delta_{uv}(m) \geq t$ must differ by at least n , which in turn means that for any positive integer l , the number of such m 's that are less than l cannot exceed $[\frac{l}{n}] + 1$, where for any real number s , $[s]$ is the greatest integer not exceeding s . Therefore, the number of non-negative integers m that are less than l and are such that $\delta_{uv}(m) < t$ is at least

$$l - 1 - \left[\frac{l}{n}\right] - 1,$$

whence

$$F_{uv}^l(t) \geq 1 - \frac{1}{l} \left[\frac{l}{n}\right] - \frac{2}{l} \geq 1 - \frac{1}{n} - \frac{2}{l}.$$

Letting $l \rightarrow \infty$, we obtain

$$F_{uv}(t) \geq 1 - \frac{1}{n}, \text{ for } n \geq 2 \text{ and } t > c(n - 1).$$

Thus $F_{uv} \geq F_{p0}$ for all u, v in $[0, 1]$, whence F_{p0} is the unique minimal lower distribution function for the tent map τ .

Note that we can insert any finite sequence whatever of 0's and 1's at the beginning of the binary expansion of p without affecting the upper and lower distributions F_{p0}^* and F_{p0} . This means that the set of numbers q such that $F_{q0} = F_{p0}$ is dense in $[0, 1]$. Similarly, given any such q , we can insert any finite sequence whatever of 0's and 1's at the beginning of the binary expansion of 0 without affecting the distributions F_{q0}^* and F_{q0} . Consequently, the set of pairs (x, y) such that $F_{xy} = F_{p0}$ and $F_{xy}^* = F_{p0}^*$, and *a fortiori*, the set of all d -chaotic pairs is dense in $[0, 1]^2$.

Next, we briefly consider the logistic function L_4 . Recall that for x in $[0, 1]$, $L_4(x) = 4x(1 - x)$. As mentioned before, L_4 is topologically conjugate to the tent map τ ; specifically, $L_4 = \lambda^{-1} \circ \tau \circ \lambda$ where, for x in $[0, 1]$,

$$\lambda(x) = \frac{1}{\pi} \arccos(1 - 2x) = \frac{2}{\pi} \arcsin(\sqrt{x}),$$

and

$$\lambda^{-1}(x) = \frac{1}{2}(1 - \cos \pi x) = (\sin(\frac{\pi}{2}x))^2.$$

Hence the qualitative behavior of τ (transitivity, existence of cycles of all orders, existence of d -chaos, etc.) carries over to L_4 . In particular, since λ is strictly increasing and an orbit isomorphism, it follows that for any positive integer n , the largest fixed point of L_4^n lies to the right of the corresponding fixed point of τ^n , so it is immediate that $\mu_p(L_4) > \mu_p(\tau)$.

We now turn to the trapezoid functions. For r in $(0, 1)$, the (symmetric) trapezoid function τ_r (so-called from the shape of its graph) is defined on $[0, 1]$ by:

$$\tau_r(x) = \begin{cases} \frac{2}{1-r}x, & x \text{ in } [0, \frac{1-r}{2}], \\ 1, & x \text{ in } [\frac{1-r}{2}, \frac{1+r}{2}], \\ \frac{2}{1-r}(1-x), & x \text{ in } [\frac{1+r}{2}, 1]. \end{cases}$$

A striking difference between the trapezoid functions and the tent map ($r = 0$) is the fact that, for any $r > 0$, the τ_r -orbit of the fixed point 0 consists of the points 0, 1, and a countable collection of disjoint closed intervals whose interiors form the complement of a Cantor set C_r of Lebesgue measure 0. From this observation it follows at once (see Section 2) that no τ_r is conjugate – and, *a fortiori*, topologically conjugate – to the tent map τ . However, any two trapezoid functions are topologically conjugate and the conjugating function can be taken to be piecewise linear and strictly increasing on the complement of the appropriate Cantor set, see Schweizer and Sklar [36]. In particular, for any r , τ_r is conjugate to $\tau_{1/3}$.

Now for $\tau_{1/3}$, the set $C_{1/3}$ is the classical middle-third Cantor set, the endpoints of whose complementary intervals are those ternary rationals (i.e., those numbers of the form $m/3^n$) whose trajectories under $\tau_{1/3}$ include either $\frac{1}{3}$ or $\frac{2}{3}$. Let $C'_{1/3}$ denote the subset of $C_{1/3}$ consisting of all points in $C_{1/3}$ except the ternary rationals whose trajectories include $\frac{1}{3}$. Thus $C'_{1/3}$ contains the points $\frac{2}{3}, \frac{2}{9}, \frac{7}{9}, \frac{2}{27}, \frac{25}{27}, \frac{7}{27}, \frac{20}{27}, \dots$, but does not contain the points $\frac{1}{3}, \frac{1}{9}, \frac{8}{27}, \frac{1}{27}, \frac{26}{27}, \frac{8}{27}, \frac{19}{27}, \dots$. Now define a function σ from $[0, 1]$ onto $C'_{1/3}$ as follows: If x in $[0, 1]$ is not a binary rational, replace all the 1's in the binary expansion of x by 2's and take the result as the ternary expansion of $\sigma(x)$. If x is a binary rational in $(0, 1)$ of the form $(4m + 1)/2^n$, again replace all the (finitely many) 1's in the binary expansion of x by 2's to get the ternary expansion of $\sigma(x)$. If x in $(0, 1)$ has the form $(4m + 3)/2^n$, replace all except the last 1 in the binary expansion of x by 2's and keep the last 1 as it is to get the ternary expansion of $\sigma(x)$. Finally, set $\sigma(0) = 0$ and $\sigma(1) = 1$. It can be shown that σ conjugates the tent map with $\tau_{1/3}$ restricted to the set $C'_{1/3}$, and so maps the orbit structure of τ one-to-one onto this restriction of $\tau_{1/3}$. A number of consequences flow from these observations.

First of all, no τ_r is transitive on $[0, 1]$. Thus, no τ_r is chaotic in the sense of Devaney; and since the trapezoids τ_r , for $r \rightarrow 0$, converge uniformly to the tent map τ (which is transitive on $[0, 1]$), it follows that chaos in the sense of Devaney is not stable. Note also that for τ the maximal ω -limit set is the unit interval $[0, 1]$ whereas for τ_r it is C_r , a nowhere-dense perfect set of measure 0.

Next, for any r , the restriction of τ_r to C_r has the same structure of periodic points as τ . Thus each τ_r has cycles of all orders, hence has positive topological entropy (viz., $\ln 2$) and is therefore d -chaotic. This latter fact can also be established directly by counting the number of monotonic pieces of τ_r .

Lastly, the largest point in any n -cycle of τ_r is

$$2^n / [2^n + (1 - r)^n].$$

This point varies continuously with r and converges to $c(n)$, the largest point in the n -cycle of τ , as $r \rightarrow 0$. Furthermore, using the abovementioned binary to ternary expansion replacement, the topological conjugacy of any two trapezoids, and the argument leading to the series representing $\mu_p(\tau)$, it is readily seen that

$$\mu_p(\tau_r) = \frac{2}{3 - r} + \sum_{n=2}^{\infty} \frac{1}{n} \frac{2^{n-1}(1 - r)^{n-1}(1 + r)}{(2^n + (1 - r)^n)(2^{n-1} + (1 - r)^{n-1})}.$$

Thus $\mu_p(\tau_r)$ is a strictly increasing, continuous function of r with

$$\lim_{r \rightarrow 1^-} \mu_p(\tau_r) = 1 \text{ and } \lim_{r \rightarrow 0^+} \mu_p(\tau_r) = \mu_p(\tau).$$

Other functions related to the tent map are the truncated tents $\bar{\tau}_\lambda$ defined, for λ in $(0, 1)$ by

$$\bar{\tau}_\lambda(x) = \min(\tau(x), \lambda) \text{ for all } x \text{ in } [0, 1].$$

The first thing to notice is that, since no trajectory can be dense in $[\lambda, 1]$, no $\bar{\tau}_\lambda$ is transitive; and since the transitive tent map can be transformed into a non-transitive truncated tent by an arbitrarily small perturbation on an arbitrarily small interval, this observation provides another illustration of the fact that transitivity is unstable.

It is easy to verify that for λ in the interval $(0, \frac{2}{3}]$, λ is a fixed point of $\bar{\tau}_\lambda$; that for λ in $(\frac{2}{3}, \frac{4}{5}]$, λ is a periodic point of period 2; and that for λ in $(\frac{4}{5}, \frac{14}{17}]$, λ is a periodic point of period 4. This pattern extends. There is an increasing sequence $\{\lambda_n\}$ of rational numbers such that if λ is in the interval $(\lambda_{m-1}, \lambda_m]$, then:

- (i) λ is a periodic point of period 2^m ;
- (ii) $\bar{\tau}_\lambda$ has cycles of order $1, 2, 2^2, \dots, 2^m$ but no cycles of any other order;
- (iii) $\bar{\tau}_\lambda$ is neither Li-Yorke chaotic nor d -chaotic.

On the other hand, for $\lambda \geq \frac{52}{63}$, the 6-tuple $(\frac{52}{63}, \frac{22}{63}, \frac{44}{63}, \frac{38}{63}, \frac{50}{63}, \frac{26}{63})$ is a cycle of order 6 of $\bar{\tau}_\lambda$, so that each such $\bar{\tau}_\lambda$ is both Li-Yorke chaotic and d -chaotic. It follows that the sequence $\{\lambda_n\}$ has a limit λ_∞ which is between $\frac{14}{17} = .8235294\dots$ and $\frac{52}{63} = .8253968\dots$. Indeed, Misiurewicz and Smítal [30] have shown that $\lambda_\infty = .8249080\dots$, that $h(\bar{\tau}_\lambda) > 0$ for $\lambda > \lambda_\infty$, and that $h(\bar{\tau}_{\lambda_\infty}) = 0$. Therefore $\bar{\tau}_\lambda$ is d -chaotic for $\lambda > \lambda_\infty$, but $\bar{\tau}_{\lambda_\infty}$ is not d -chaotic. On the other hand, as Misiurewicz and Smítal [30] have also shown, $\bar{\tau}_{\lambda_\infty}$ is chaotic in the sense of Li and Yorke; and since an arbitrarily small perturbation can transform $\bar{\tau}_{\lambda_\infty}$ into a $\bar{\tau}_\lambda$ with $\lambda < \lambda_\infty$, this example shows that chaos in the sense of Li and Yorke is unstable.

6 Functions in $C(I, I)$ in general; proximal trajectories

In this section we present, in outline and without proof, but with further illustrative examples, the principal results of Schweizer and Smítal [38] and Balibrea et al. [5], concerning distributional chaos for arbitrary functions in $C(I, I)$. These results rely heavily on the characterization of various types of maximal ω -limit sets that goes back to the 1966 work of A. N. Sharkovsky, [39, 40, 41], where such sets were partitioned into

- (a) those sets that are finite;

- (b) those sets that are infinite but contain no cycles; and
- (c) those sets that are infinite and contain a cycle.

He referred to the sets of type (a) and (b) as “maximal ω -limit sets of the first kind” and to those of type (c) as “maximal ω -limit sets of the second kind”. The latter are now generally called *basic sets*, see [12].

By a theorem of B. Barna [8], greatly extended and with a simple proof by M. Kuczma [26], a finite ω -limit set must be a cycle. Thus, if for some f in $C(I, I)$ and x, y in I , $\omega_f(x)$ and $\omega_f(y)$ are both finite, then (trivially) $F_{xy} = F_{xy}^*$, and this common distribution function is a step function.

Definition 6.1. *Let f be in $C(I, I)$. Then the points x, y in I have proximal f -trajectories if $\liminf_{n \rightarrow \infty} d(f^n(x), f^n(y)) = 0$. The set of all such pairs of points is denoted by $\text{Prox}(f)$.*

Note that if $(x, y) \in \text{Prox}(f)$, then x and y belong either to the same maximal ω -limit set or to adjacent maximal ω -limit sets, where two maximal ω -limit sets are *adjacent* if they have a nonempty intersection. (It is well-known that the intersection of any two different maximal ω -limit sets is always finite.) Note also that if (x, y) is not in $\text{Prox}(f)$, then there is a $t > 0$ such that $d(f^n(x), f^n(y)) \geq t$ for all sufficiently large n , whence $\lim_{n \rightarrow \infty} F_{xy}^{(n)}(t) = 0$. Thus, if $F_{xy}^* = \epsilon_0$ then (x, y) is in $\text{Prox}(f)$. The converse is false, as the following example shows:

Example 6.2. *Let y be a point whose trajectory with respect to the tent map τ is close to the fixed point $2/3$ most of the time and such that 0 belongs to $\omega_\tau(y)$. Then the pair $(0, y)$ is in $\text{Prox}(\tau)$ while $F_{0y} = F_{0y}^* = \epsilon_{2/3}$.*

The next result is a restatement of Schweizer and Smítal’s Theorem 2.1 of [38].

Theorem 6.3. *Let f be in $C(I, I)$ and suppose that x, y in I are such that each of the sets $\omega_f(x)$, $\omega_f(y)$ is contained in a maximal ω -limit set of type (a) or type (b). Then*

- (i) $F_{xy} = F_{xy}^*$.
- (ii) *If, in addition, (x, y) is in $\text{Prox}(f)$, then $F_{xy} = F_{xy}^* = \epsilon_0$.*

It follows that if f in $C(I, I)$ is such that none of its maximal ω -limit sets is a basic set, then f is not d -chaotic. Regarding the converse, Sharkovsky, in [40], showed that the presence of a basic set is equivalent to the existence of an f -cycle whose order is not a power of 2 (see Janková and Smítal [22]

for a shorter proof); later Misiurewicz, in [29], showed that either of these conditions is equivalent to $h(f) > 0$; and Schweizer and Smítal used these results in [38] to establish the fact, already mentioned in Section 4, that f is d -chaotic if and only if $h(f) > 0$. Thus the study of distributional chaos becomes interesting only when basic sets are in the picture. To present the central results in this case, we begin with:

Definition 6.4. *Let f be in $C(I, I)$ and x, y in I . Then the pair (x, y) is isotectic (with respect to f) if, for every positive integer n , the ω -limit sets $\omega_{f^n}(x)$ and $\omega_{f^n}(y)$ are subsets of the same maximal ω -limit set of f^n .*

Note that (x, x) is isotectic for every x in I ; note further that if x and y are distinct points of a maximal ω -limit set of type (a), then there is an integer k such that $\omega_{f^k}(x) = \{x\}$ and $\omega_{f^k}(y) = \{y\}$ and, since each of these singleton sets is a maximal ω -limit set of f^k , x and y are not isotectic. Finally note that if f is transitive, then I is the unique maximal ω -limit set, whence all pairs (x, y) are isotectic with respect to f .

Equivalently, the pair (x, y) is isotectic if both $\omega_f(x)$ and $\omega_f(y)$ are contained in the same maximal ω -limit set and if, for any periodic interval J such that ω is a subset of the set $J \cup f(J) \cup f^2(J) \cup \dots \cup f^{n-1}(J)$, where n is the period of J , there is a $j \geq 0$ such that both $f^j(x)$ and $f^j(y)$ belong to J .

In the earlier papers of Schweizer and Smítal [38], and Balibrea et al. [5], the word “synchronous” was used to describe the property specified in Definition 6.4. However, we now feel that this term has connotations that are somewhat misleading, since what we wish to describe is more a property of being “under the same roof” than a property of “keeping the same time”.

Now let

- (i) $Iso(f) = \{(x, y) \text{ in } I \times I; (x, y) \text{ is isotectic}\}$, and
- (ii) $D(f) = \{F_{xy}; (x, y) \text{ in } Iso(f)\}$.

Then the *spectrum* of f , denoted by $\Sigma(f)$, is the set of minimal elements of $D(f)$.

To illustrate, for the tent map τ , the unit interval $[0, 1]$ is the unique maximal ω -limit set, whence all pairs (x, y) in $[0, 1] \times [0, 1]$ are isotectic and, as shown in Section 5, $\Sigma(\tau) = \{F_{p0}\}$.

We can now state the central result of Schweizer and Smítal from [38].

Theorem 6.5. *For any f in $C(I, I)$, the spectrum $\Sigma(f)$ is non-empty and finite, and:*

- (i) *If ϵ_0 belongs to $\Sigma(f)$ - in which case $\Sigma(f) = \{\epsilon_0\}$ and $F_{xy} = F_{xy}^* = \epsilon_0$ for all (x, y) in $Iso(f)$ - then f is not d -chaotic; and conversely, if f is not d -chaotic, then $\Sigma(f) = \{\epsilon_0\}$.*

(ii) If ϵ_0 does not belong to $\Sigma(f)$, then f is d -chaotic. In this case, for each distribution function F in $\Sigma(f)$ there is a non-empty perfect set P_F , and a positive number $t(F)$ such that any pair (x, y) of distinct points of P_F belongs to $\text{Prox}(f)$ and

$$F_{xy} = F \leq \epsilon_{t(F)} < \epsilon_0 = F_{xy}^*.$$

Example 6.6. For $0 < a < b < 1$, let f_{ab} be a piecewise linear map in $C(I, I)$ determined via $f_{ab}(0) = f_{ab}(b) = 0$, $f_{ab}(a/2) = b$, $f_{ab}(a) = a$ and $f_{ab}(1) = 1$. The restriction of f_{ab} to the interval $[0, b]$ is topologically conjugate to the tent map τ . Exploiting this, it follows that f_{ab} has two maximal ω -limit sets - the singleton set $\{1\}$ and the interval $[0, b]$ - and that $\Sigma(f_{ab}) = \{\Phi_{ab}\}$, where $\Phi_{ab}(x) = 0$ if $x \in [0, a]$, $\Phi_{ab}(x) = 1$ if $x \in [b, 1]$, and $0 < \Phi_{ab}(x) < 1$ if $x \in (a, b)$.

Now, if $a < a' < b' < b$, then F_{ab} and $F_{a'b'}$ are incomparable. Using this fact, for any positive integer n , it is easy to construct a function f such that $\Sigma(f)$ has exactly n elements. Indeed, for $1 \leq i \leq n$, let I_i be the interval $[\frac{i-1}{n}, \frac{i}{n}]$, let $0 < a_1 < a_2 < \dots < a_n < b_n < b_{n-1} < \dots < b_1 < 1$ be given, let $f_i \in C(I_i, I_i)$ be the affine copy of $f_{a_i b_i}$ and define f on $[0, 1]$ via $f(x) = f_i(x)$ for $x \in I_i$. Then $\Sigma(f) = \{F_1, \dots, F_n\}$ where $F_i(x) = \Phi_{a_i b_i}(nx)$.

On the other hand, if f has exactly one basic set, in particular if f is transitive on I , then $\Sigma(f)$ is a singleton set; and if the endpoints of this basic set, say c and d , are fixed points of f , then $\Sigma(f) = \{\epsilon_{d-c}\}$.

Definition 6.7. For any f in $C(I, I)$, the spectral measure of chaos of f is the number $\mu_s(f)$ in the interval $[0, 1]$ given by

$$\mu_s(f) = \max \left\{ \frac{1}{|I|} \int_0^{|I|} (1 - F(t)) dt; F \in \Sigma(f) \right\},$$

where $|I|$ denotes the length of the interval I .

Theorem 6.8. Let f be in $C(I, I)$ and let F be a distribution function in $\Sigma(f)$. Then there are sequences $\{p_n\}_{n=1}^\infty$ and $\{q_n\}_{n=1}^\infty$ of periodic points such that

$$F = \liminf_{n \rightarrow \infty} F_{p_n q_n}.$$

In other words, any distribution function in the spectrum of f can be approximated by distribution functions generated by periodic points. This means, in effect, that the spectrum of f is determined by the cycle structure of f .

The spectrum exhibits the following stability property found by Balibrea et al. in [5].

Theorem 6.9. *Let f in $C(I, I)$ and $\epsilon > 0$ be given. Then there is a $\delta > 0$ such that for every $g \in C(I, I)$ satisfying the condition $\sup_{x \in I} |f(x) - g(x)| < \delta$, the following property holds:*

For any F in $\Sigma(f)$ there is a G in $\Sigma(g)$ such that

$$\int_0^{|I|} \max\{0, G(t) - F(t)\} dt < \epsilon.$$

Corollary 6.10. *The spectral measure of chaos is lower semi-continuous, i.e., for any f in $C(I, I)$ and any $\epsilon > 0$, there is a $\delta > 0$ such that*

$$\mu_s(f) - \epsilon < \mu_s(g),$$

for any g in $C(I, I)$ satisfying $\sup_{x \in I} |f(x) - g(x)| < \delta$.

As in the case of length in the Calculus of Variations, the stability described above is definitely one-sided. Chaos cannot disappear when the original function is replaced by a slightly perturbed function; however, small perturbations can significantly increase, even “explode”, the spectral measure of chaos, as the following example shows:

Example 6.11. *Let f be the identity function on $[0, 1]$, i.e., $f(x) = x$ for any x in $[0, 1]$, and let g be a topologically transitive map that is arbitrarily close to f and has 0 and 1 as fixed points. (Such functions always exist; a piecewise linear example is given by Smítal and Smítalová in Theorem 1 of [46].) Then $\Sigma(g) = \{\epsilon_1\}$, whence $\mu_s(g) = 1$ while $\mu_s(f) = 0$.*

Example 6.12. *The conclusion of Theorem 6.9 cannot be strengthened to obtain $G(t) < F(t) + \epsilon$ for all t in I . To see this, consider the tent map τ and, for δ in $(0, 1)$, let g be its dilation to $[0, 1 - \delta]$, extended to $[0, 1]$ by letting $g(x) = 0$ on $[1 - \delta, 1]$. Then $\Sigma(g) = \{G\}$, where G is the dilation of F_τ to $[0, 1 - \delta]$. We then have $G(t) - F_\tau(t) \geq \frac{1}{2}$ for $\frac{2}{3}(1 - \delta) < t \leq \frac{2}{3}$.*

The two measures of chaos, μ_s and μ_p are related as follows:

Theorem 6.13. *For any f in $C(I, I)$,*

$$\mu_s(f) \leq \mu_p(f) \leq 2\mu_s(f).$$

The inequality $\mu_s(f) \leq \mu_p(f)$ is immediate, as is the fact that $\mu_s(f) = \mu_p(f)$ whenever all pairs (x, y) in $I \times I$ are isotectic with respect to f . The following example shows that we may also have $\mu_p(f) = 2\mu_s(f) > 0$.

Example 6.14. Let T be the piecewise-linear function defined on $[0, 1]$ by

$$T(x) = \begin{cases} 2x + \frac{1}{2}, & x \in [0, \frac{1}{4}], \\ \frac{3}{2} - 2x, & x \in [\frac{1}{4}, \frac{3}{4}], \\ 2x - \frac{3}{2}, & x \in [\frac{3}{4}, 1]. \end{cases}$$

It is worth noting that T has no cycles of odd order greater than 1, but has cycles of all other orders, including 20 cycles of order 6.

The unit interval $[0, 1]$ is the single maximal ω -limit set for T , but each even-numbered iterate of T , i.e., T^2, T^4, \dots , has two maximal ω -limit sets, namely $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$. It follows that (x, y) is an isotectic pair for T if and only if either $(x, y) \in [0, \frac{1}{2}] \times [0, \frac{1}{2}]$ or $(x, y) \in [\frac{1}{2}, 1] \times [\frac{1}{2}, 1]$. Next, we have that $\sum(T) = \{F_T\}$ where F_T is related to F_τ via $F_T(t) = F_\tau(2t)$ for all t . Thus F_T is the minimal distribution derived from isotectic pairs, and $\mu_s(T) = \frac{1}{2}\mu_p(\tau)$. On the other hand, it can be shown that there are non-isotectic pairs (x, y) such that

$$F_{xy} = F_\tau < F_T < \epsilon_0 = F_{xy}^*$$

and that F_τ is the unique minimal lower distribution of T , whence $\mu_p(T) = \mu_p(\tau)$ and this yields $\mu_p(T) = 2\mu_s(T)$

This and other examples, in particular those involving adjacent basic sets with a common fixed point, indicate that, when dealing with arbitrary functions in $C(I, I)$, it is convenient to supplement the notion of spectrum. In [38] Schweizer and Smítal do this as follows: For f in $C(I, I)$, let

$$D_w(f) = \{F_{uv}; (u, v) \text{ in } \text{Prox}(f)\}.$$

Then the weak spectrum of f , denoted by $\sum_w(f)$ is the set of all minimal elements of $D_w(f)$. Thus, for instance, in Example 6.14 we have $\sum_w(T) = \{F_\tau\}$ (since $F_{xy}^* = \epsilon_0$ implies that (x, y) is in $\text{Prox}(T)$).

Theorem 6.15. For any f in $C(I, I)$, the weak spectrum $\sum_w(f)$ is non-empty and finite. Moreover, for any F in $\sum_w(f)$,

(i) there is at least one pair (u, v) in $\text{Prox}(f)$ such that

$$F = F_{uv} < F_{uv}^* = \epsilon_0, \text{ and}$$

(ii) there are sequences $\{p_n\}_{n=1}^\infty$ and $\{q_n\}_{n=1}^\infty$ of periodic points such that $F = \liminf_{n \rightarrow \infty} F_{p_n q_n}$, whence $\sum_w(f)$ is also determined by the cycle structure of f .

A distribution function F in $\sum_w(f)$ need not be in $\sum(f)$ and, if it is not, then in contrast to the case of functions in $\sum(f)$, not only is there no perfect set associated with F , there is not even a triple of distinct points such that any two of them determine F .

In general, there is no simple relation between the number of functions in $\sum(f)$ and $\sum_w(f)$. In fact, one can modify Example 6.6 to show that each of the sets

$$\sum(f) \setminus \sum_w(f), \quad \sum_w(f) \setminus \sum(f), \quad \sum(f) \cap \sum_w(f)$$

can simultaneously have any finite number, 0 included, of elements, subject only to the following restrictions: (1) at least one of the three sets is non-empty; (2) if $\sum(f) \setminus \sum_w(f)$ is non-empty, then so is $\sum_w(f) \setminus \sum(f)$, i.e., $\sum_w(f)$ cannot be a proper subset of $\sum(f)$.

If the set $\sum_w(f) \setminus \sum(f)$ is non-empty, then it does not enjoy the stability properties of $\sum(f)$ that are stated in Theorem 6.9 and Corollary 6.10; for there will be functions g in $C(I, I)$, arbitrarily close to f such that $\sum_w(g) \setminus \sum(g)$ is empty (see Remark 2 in Section 4 of [5]). On the other hand, if in Theorem 6.9 we replace the phrase “for every g in $C(I, I)$ ” by “there exists a g in $C(I, I)$ ”, then replacing “ $\sum(f)$ ” and “ $\sum_w(f)$ ” by “ $\sum_w(f) \setminus \sum(f)$ ” and “ $\sum(g) \setminus \sum_w(g)$ ”, respectively, does yield a valid statement (see Remark 3 in Section 4 of [5]).

As we have seen, for some functions in $C(I, I)$, e.g., the tent map τ , the principal measure of chaos is determined by the spectrum; for others, e.g., the function in Example 6.14, it is determined by the weak spectrum. However, as the next example shows, there are functions in $C(I, I)$ for which the principal measure of chaos is determined by neither the spectrum nor the weak spectrum.

Example 6.16. *Using the functions f_t defined in Section 2, let φ be the function defined on $[0, 1]$ by:*

$$\varphi(x) = \begin{cases} f_{2/5}(x), & x \in [0, \frac{2}{5}], \\ f_{1/5}(x - \frac{2}{5}) + \frac{2}{5}, & x \in [\frac{2}{5}, \frac{3}{5}], \\ f_{2/5}(x - \frac{3}{5}) + \frac{3}{5}, & x \in [\frac{3}{5}, 1]. \end{cases}$$

It is easy to verify that φ is continuously differentiable, has topological entropy $\log 3$, and satisfies the functional equation $\varphi(x) + \varphi(1 - x) = 1$. There are 3 maximal ω -limit sets: $[0, \frac{2}{5}]$, $[\frac{2}{5}, \frac{3}{5}]$, $[\frac{3}{5}, 1]$, and it can be shown that $\sum(\varphi) = \{\epsilon_{2/5}\}$, that $\sum_w(\varphi) = \{\epsilon_{3/5}\}$, and that there are pairs (x, y) with x in $[0, \frac{2}{5}]$ and y in $[\frac{3}{5}, 1]$ such that $F_{xy} = \epsilon_1$ and $F_{xy}^ = \epsilon_{1/5}$. Hence $\mu_p(\varphi) \geq \frac{4}{5}$, which is larger than the number obtained from the spectrum ($\mu_s(\varphi) = \frac{2}{5}$), and the number obtained from the weak spectrum ($\frac{3}{5}$).*

Examples such as this indicate that the theory developed by Schweizer and Smítal in [38] should be extended as follows:

Let T_c be the translation through c defined by $T_c(t) = t + c$ for all t . For any function f mapping a metric space into itself, define the set $D_p(f)$ of distribution functions by

$$D_p(f) = \{F_{xy} \circ T_c : F_{xy}^* = \epsilon_c \text{ for some } c \geq 0\},$$

Note that in view of Theorems 6.5 and 6.15, for any f in $C(I, I)$, both $D(f)$ and $D_w(f)$ are subsets, generally proper, of $D_p(f)$. Now define $\sum_p(f)$, the *principal spectrum* of f , to be the set of minimal elements of $D_p(f)$. Thus, for the function φ in Example 6.16, we have $\sum_p(\varphi) = \{\epsilon_{4/5}\}$ – which shows, incidentally, that no one of the three sets $\sum(f)$, $\sum_w(f)$ and $\sum_p(f)$ need be a subset of either of the other two.

Lastly, there is the following result, whose proof will appear elsewhere:

Theorem 6.17. *For any f in $C(I, I)$, the principal spectrum $\sum_p(f)$ is non-empty and finite. Moreover:*

- (i) f is d -chaotic if and only if $\sum_p(f)$ does not contain ϵ_0 .
- (ii) $\sum_p(f)$ (like $\sum_w(f)$) does not share the stability properties of $\sum(f)$.
- (iii) For any F in $\sum_p(f)$, there is at least one pair of points (u, v) such that, for some $c > 0$,

$$F = F_{uv} \circ T_c < F_{uv}^* = \epsilon_c.$$

- (iv) For any F in $\sum_p(f)$, there are sequences $\{p_n\}_{n=1}^\infty$ and $\{q_n\}_{n=1}^\infty$ of periodic points such that

$$F = \liminf_{n \rightarrow \infty} F_{p_n q_n}.$$

- (v) The principal measure of chaos, $\mu_p(f)$, is given by

$$\mu_p(f) = \max \left\{ \frac{1}{|I|} \int_0^{|I|} (1 - F(t)) dt; F \in \sum_p(f) \right\}.$$

PROOF. (i) is clear, (ii) is proved in [5], and (iii) - (v) have recently been proved by K. Janková.

7 Concluding remarks

1. In the preceding sections we have concentrated our attention on the numerical measures μ_p and μ_s . Clearly, however, no single number can adequately describe a phenomenon as complex as chaos. For distributional chaos, a detailed description is given by the finite collection of pairs of lower and upper distributions that are determined by the various spectra. The information supplied by these pairs can be condensed by means of the associated numerical measures given by the (suitably normalized) integrals $\int (F^*(t) - F(t))dt$; and then, as we have done, further condensed to the measures μ_p , μ_s , etc.. In addition, for pairs of the form (F, ϵ_c) one could – and should – also consider the means, medians, etc., of the translated lower distributions.

2. Recently people have begun to investigate d -chaos for continuous functions defined on spaces other than intervals. One significant result of these investigations is the fact that the equivalence between d -chaos and positive topological entropy does not extend to higher dimensions. Specifically, consider the family of functions G defined on the closed unit square by

$$G(x, y) = (\lambda x, g(x, y)),$$

where λ is in $(0, 1)$ and g is a continuous function with $g(0, y) = y$ for all y in $[0, 1]$. Each such function is a special triangular map of the square (see Kolyada [24]). Because of the contraction in the first coordinate, all the functions in this family have topological entropy 0. Furthermore, as shown by Forti et al. [21], there is a particular function in this family that has topological sequence entropy 0 for all sequences; and, as shown by Forti and Paganoni in [20], this function is d -chaotic.

3. The second and third authors have recently shown in [44] that any continuous function on a compact metric space that has either the specification property (see Sigmund [43], or Denker et al. [14, §21]) or the generalized specification property (see Balibrea et al. [5]) is d -chaotic. In particular, for any $n \geq 2$, every hyperbolic automorphism of the n -torus is d -chaotic. Incidentally, functions with the specification property are known to have positive topological entropy, Denker et al. [14, §21] and it is reasonable to conjecture that the same holds for functions with the generalized specification property.

4. To date, hardly a beginning has been made in the investigation of d -chaos for discontinuous functions or for functions on non-compact metric spaces. There are, of course, examples. Perhaps the simplest is the discontinuous analog of the tent map, namely the sawtooth function β defined on the

half-open interval $[0, 1)$ by

$$\beta(t) = \begin{cases} 2t, & \text{for } 0 \leq t < \frac{1}{2}, \\ 2t - 1, & \text{for } \frac{1}{2} \leq t < 1. \end{cases}$$

Using the binary representation of real numbers, one readily constructs a number x in $(0, 1)$ such that $F_{0x} = \epsilon_1$ and $F_{0x}^* = \epsilon_0$, whence β is d -chaotic with $\mu_p(\beta) = 1$. However, as a somewhat isolated example, β offers few clues as to further directions of inquiry.

It should be noted that β must be carefully distinguished from the corresponding function β^0 on the topological circle obtained by identifying the endpoints of the closed interval $[0, 1]$. The functions β and β^0 are conjugate but not topologically conjugate, since their respective domains are not homeomorphic; and while β^0 is also d -chaotic, we have $\mu_p(\beta^0) < 1 = \mu_p(\beta)$.

5. The tent map τ and the function β defined above are both strongly mixing with respect to Lebesgue measure. Similarly, the logistic map L_4 and the hyperbolic automorphisms of the torus are strongly mixing with respect to well-behaved measures on their domains. These facts lead to a seemingly paradoxical state of affairs. To describe it, we need a few results from the theory of probabilistic metric spaces.

Let (M, d) be a metric space and f a function that maps M into itself. Let \mathfrak{S} be the mapping from $M \times M$ into the space of probability distribution functions defined via

$$\mathfrak{S}(x, y) = F_{xy}, \text{ for any } x, y \in M,$$

where F_{xy} is the lower distribution of x and y . Then the pair (S, \mathfrak{S}) is a *transformation-generated probabilistic pseudo-metric* space (see Schweizer and Sklar [35, §11.1]). Such spaces were introduced by Erber et al. [18] and by Schweizer and Sklar in [34] and studied from the point of view of ergodic theory. (For a discussion of how these studies later led to the notion of distributional chaos, see Schweizer [33].) In these papers it was shown that if M is a separable metric space endowed with a probability measure P defined on a σ -algebra that contains the Borel sets of M , and if the function f from M into M is measure-preserving with respect to P , then

$$F_{xy} = F_{xy}^*$$

for almost all pairs (x, y) in $M \times M$ (endowed with the product measure P^2). Thus if f is measure-preserving, then almost all pairs of points in $M \times M$ are non-chaotic.

If, in addition to being measure-preserving, f is strongly mixing with respect to P , then there is a unique distribution function G_f such that

$$F_{xy} = F_{xy}^* = G_f,$$

for almost all pairs (x, y) . Furthermore, the Birkhoff Ergodic Theorem yields that, for any $t \geq 0$,

$$G_f(t) = P^2\{(x, y) \in M \times M; d(x, y) < t\},$$

so that almost all pairs are not only non-chaotic, but are non-chaotic in exactly the same way. In a statistical sense, this is extremely orderly behavior. Note further that if f is mixing and if every open set has positive measure, then f is transitive.

Now if such a mixing transformation f is d -chaotic, then the chaos, which is a highly significant feature of its overall behavior, is generated by a set of measure zero – a set that is deliberately disregarded in ergodic theory! If, in particular, f is in $C(I, I)$, then both the d -chaotic behavior and the topological entropy of f are determined by the points in the cycles (see Theorems 6.8 and 6.15, and Block and Coppell [11, p. 217]). In many cases of interest these points form a countable set which is dense in I , so that both the chaotic and the non-chaotic pairs are dense in $I \times I$.

To illustrate, the tent map τ is mixing with respect to Lebesgue measure on $[0, 1]$. Thus, for any t in $[0, 1]$, it follows that $G_\tau(t)$ is the Lebesgue measure of the set $\{(x, y) \in [0, 1]^2 : |x - y| < t\}$. Specifically,

$$G_\tau(t) = \begin{cases} 0, & t \leq 0, \\ 2t - t^2, & 0 \leq t \leq 1, \\ 1, & 1 \leq t. \end{cases}$$

Comparing G_τ with the lower and upper distributions F_{p0} and F_{p0}^* determined in Section 5, we see that, for all t in $(0, 1)$,

$$F_{p0} < G_\tau(t) < F_{p0}^*.$$

Furthermore, as shown at the end of Section 5, the set of pairs of points (x, y) for which $F_{xy} = F_{p0}$ and $F_{xy}^* = F_{p0}^*$ is dense in $[0, 1]^2$.

Similarly, the logistic map L_4 is mixing with respect to an absolutely continuous probability measure. The interval $[0, 1]$ is the unique maximal ω -limit set for L_4 and the unique minimal distribution in $\sum(L_4)$ is smaller than the corresponding distribution in $\sum(\tau)$ (see Section 5). The “ergodic” limiting distribution G_{L_4} is considerably more complicated. It was first determined

explicitly by Johnson and Sklar in [23] for the conjugate quadratic Čebyšev polynomial $x^2 - 2$ on the interval $[-2, 2]$; and an appropriate Möbius transformation of the integrand by Schweizer and Sklar in [35, 11.3.8] yields:

$$G_L(t) = \begin{cases} 0, & t < 0, \\ \frac{8}{\pi^2} \int_{\frac{1-t}{1+t}}^1 \frac{K(v)}{1+v} dv, & 0 \leq t \leq 1, \\ 1, & t > 1, \end{cases}$$

where K is the complete elliptic integral of the first kind.

The function β defined in Remark 4 above is, like τ , mixing with respect to Lebesgue measure (on $[0, 1)$), whence $G_\beta = G_\tau$. Furthermore, $\epsilon_1 < G_\beta < \epsilon_0$. Note that neither the topological entropy nor the limiting distribution determined by the mixing behavior distinguish between τ and β , whereas $\mu_p(\tau) < \mu_p(\beta)$; note also that both “ergodic” limiting distributions differ markedly from the corresponding “chaotic” lower and upper distributions.

The existence of examples such as the above (which we conjecture to be the rule rather than the exception) invites speculation: Distribution functions that arise in “real world” situations are in principle observable. Suppose now that an actual physical system is modeled by a function that is both strongly mixing and d -chaotic. Will the observed distributions more often approximate the “ergodic” ones, the “ d -chaotic” ones, or neither? This question is distinct from questions relating to computer simulations. Actual computers are finite objects, whence any simulation must terminate in a cycle after a finite number of steps, and this finite number can often be surprisingly small (see, e.g., Darsow et al. [13] and Dyson and Falk [17]).

Returning to the more solid ground of mathematics, the existence of functions with, simultaneously, dense sets of pairs of periodic points, dense sets of “ergodic” pairs, and dense sets of d -chaotic pairs means that, although one must carefully distinguish between “order” (in both the usual and statistical sense) and “chaos”, one often cannot have one without the other.

References

- [1] R. L. Adler, A. G. Konheim and M. H. McAndrew *Topological entropy*, Trans. Amer. Math. Soc. **114** (1965), 309 – 319.
- [2] R. L. Adler and B. Marcus, *Topological entropy and equivalence of dynamical systems*, AMS Memoirs **20**, 1979.
- [3] L. Alsedà, J. Llibre and M. Misiurewicz *Combinatorial Dynamics and Entropy in Dimension One*, World Scientific Press, 1993.

- [4] D. Assaf and S. Gadbois, *Definition of chaos*, Amer. Math. Monthly **99** (1992), 865.
- [5] F. Balibrea, B. Schweizer, A. Sklar and J. Smítal, *Generalized specification property and distributional chaos*, Int. J. Bifur. Chaos (to appear).
- [6] F. Balibrea and J. Smítal, *A chaotic continuous map generates all probability distributions*, J. Math. Anal. Appl. **180** (1993), 587 – 598.
- [7] J. Banks, J. Brooks, G. Cairns, G. Davis and P. Stacey, *On Devaney’s definition of chaos*, Amer. Math. Monthly **99** (1992), 332 – 334.
- [8] B. Barna, *Über die Iteration reeller Funktionen I*, Publ. Math. Debrecen **7** (1960), 16 – 40.
- [9] M. F. Barnsley, *Fractals everywhere.*, 2nd ed., Academic Press, 1993.
- [10] R. Berglund and M. Vellekoop, *On intervals, transitivity = chaos*, Amer. Math. Monthly **101** (1994), 353–355.
- [11] L. Block and W. A. Coppel, *Dynamics in One Dimension*, Lecture Notes in Mathematics, **1513**, 1992.
- [12] A. M. Blokh, *The “spectral” decomposition for one-dimensional maps*, Dynam. Reported **4** (1995), 1 – 59.
- [13] W. F. Darsow, T. Erber, M. J. Frank and T. M. Rynne, *The simulation of random processes on digital computers: unavoidable order*, J. Comput. Phys. **49** (1983), 394 – 419.
- [14] M. Denker, C. Grillenberger and K. Sigmund, *Ergodic Theory on Compact Spaces*, Lecture Notes in Mathematics, **527**, 1976.
- [15] R. L. Devaney, *An Introduction to Chaotic Dynamical Systems*, 2nd edn., Addison-Wesley, 1989.
- [16] R. L. Devaney, *Dynamics of simple maps*, in *Chaos and Fractals*, ed. by R. L. Devaney and L. Keen, AMS Proc. of Symposia in Applied Math. **39** (1989), 1 – 24.
- [17] F. J. Dyson and H. Falk, *Period of a discrete cat mapping*, Amer. Math. Monthly **99** (1992), 603 – 614.
- [18] T. Erber, B. Schweizer and A. Sklar, *Mixing transformations on metric spaces*, Comm. Math. Phys. **29** (1973), 311 – 317.

- [19] V. V. Fedorenko, A. N. Sharkovsky and J. Smítal, *Characterizations of weakly chaotic maps of the interval*, Proc. Amer. Math. Soc. **110** (1990), 141 – 148.
- [20] G. L. Forti and L. Paganoni, *A distributionally chaotic triangular map with zero sequence topological entropy*, Math. Pannon. **9** (1998), 147–152.
- [21] G. L. Forti, L. Paganoni and J. Smítal, *Strange triangular maps of the interval*, Bull. Austral. Math. Soc. **51** (1995), 395 – 415.
- [22] K. Janková and J. Smítal, *A theorem of Sharkovsky characterizing continuous maps of the interval with zero topological entropy*, Math. Slovaca **39** (1989), 261 - 265.
- [23] P. Johnson and A. Sklar, *Recurrence and dispersion under iteration of Čebyšev polynomials*, J. Math. Anal. Appl. **54** (1976), 752 – 771.
- [24] S. F. Kolyada, *On dynamics of triangular maps of the square*, Ergodic Theory Dynam. Syst. **12** (1992), 749 – 768.
- [25] M. Kuchta and J. Smítal, *Two point scrambled set implies chaos*, in Proc. Europ. Conf. on Iteration Theory, Spain 1987, ed. by C. Alsina, J. Llibre, Ch. Mira, C. Simó, G. Targonski and R. Thibault, World Scientific Press, 1989, 427 – 430.
- [26] M. Kuczma, *On a theorem of Barna*, Aequationes Math. **21** (1980), 173 – 178.
- [27] T. Y. Li and J. Yorke, *Period three implies chaos*, Amer. Math. Monthly **82** (1975), 985–992.
- [28] J. Milnor and W. Thurston, *On iterated maps of the interval*, in Dynamical Systems, ed. by J. C. Alexander, Lecture Notes in Mathematics **1342** (1988), 465 – 563.
- [29] M. Misiurewicz, *Horseshoes for mappings of the interval*, Bull. Acad. Polon. Sci. Sér. Math. **27** (1979), 167 – 169.
- [30] M. Misiurewicz and J. Smítal, *Smooth chaotic maps with zero topological entropy*, Ergodic Theory Dynam. Systems **8** (1988), 421 – 424.
- [31] M. Misiurewicz and W. Szlenk, *Entropy of piecewise monotone mappings*, Studia Math. **67** (1980), 45 – 63.
- [32] R. E. Rice, B. Schweizer and A. Sklar, *When is $f(f(z)) = az^2 + bz + c$?*, Amer. Math. Monthly **87** (1980), 252 – 263.

- [33] B. Schweizer, *The genesis of the notion of distributional chaos*, Rend. Seminario Mat. e Fis. di Milano **66** (1996), 159 - 167.
- [34] B. Schweizer and A. Sklar, *Probabilistic metric spaces determined by measure-preserving transformations*, Z. Wahrsch. Verw. Geb. **26** (1973), 235 - 239.
- [35] B. Schweizer and A. Sklar, *Probabilistic Metric Spaces*, North-Holland, 1983.
- [36] B. Schweizer and A. Sklar, *Continuous functions that conjugate trapezoid functions*, Aequationes Math. **28** (1985), 300 - 304.
- [37] B. Schweizer and A. Sklar, *Invariants and equivalence classes of polynomials under linear conjugacy*, in Contributions to General Algebra **6** (1988), 253-257.
- [38] B. Schweizer and J. Smítal, *Measures of chaos and a spectral decomposition of dynamical systems on the interval*, Trans. Amer. Math. Soc. **344** (1994), 737 - 754.
- [39] A. N. Sharkovsky, *The partially ordered system of attracting sets*, Soviet Math. Dokl. **7** (1966), 1384 - 1386.
- [40] A. N. Sharkovsky, *The behavior of a map in a neighborhood of an attracting set (Russian)*, Ukrain. Mat. Zh. **18** (1966), 60 - 83.
- [41] A. N. Sharkovsky, *Continuous mapping on the set of ω -limit sets of iterated sequences (Russian)*, Ukrain. Mat. Zh. **18** (1966), 127 - 130.
- [42] K. Sigmund, *Generic properties of invariant measures for axiom A diffeomorphisms*, Inv. Math. **11** (1970), 99 - 109.
- [43] K. Sigmund, *On dynamical systems with the specification property*, Trans. Amer. Math. Soc. **190** (1974), 285 - 299.
- [44] A. Sklar and J. Smítal, *Distributional chaos on compact metric spaces via specification properties*, J. Math. Anal. Appl. **241** (2000), 181 - 188.
- [45] J. Smítal, *Chaotic functions with zero topological entropy*, Trans. Amer. Math. Soc. **297** (1986), 269 - 282.
- [46] J. Smítal and K. Smítalová, *Structural stability of nonchaotic difference equations*, J. Math. Anal. Appl. **90** (1982), 1 - 11.
- [47] I. Stewart, *Does God play dice?*, Basil Blackwell, 1989.