

## Discrete chaos

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1999 J. Phys. A: Math. Gen. 32 175

(<http://iopscience.iop.org/0305-4470/32/1/019>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

### Download details:

IP Address: 171.66.16.104

The article was downloaded on 02/06/2010 at 07:24

Please note that [terms and conditions apply](#).

## Discrete chaos

Henri Waelbroeck<sup>†§||</sup> and Federico Zertuche<sup>‡¶¶</sup>

<sup>†</sup> Instituut voor Theoretische Fysica, Universiteit Utrecht, Princetonplein 5, Postbus 80006, 3508 TA Utrecht, The Netherlands

<sup>‡</sup> Instituto de Matemáticas, UNAM, Unidad Cuernavaca, AP 273-3, Admon. 3, 62251 Cuernavaca, Morelos, Mexico

Received 9 March 1998, in final form 3 July 1998

**Abstract.** We propose a theory of deterministic chaos for discrete systems, based on their representations in binary state spaces  $\Omega$ , homeomorphic to the space of symbolic dynamics. This formalism is applied to neural networks and cellular automata; it is found that such systems cannot be viewed as chaotic when one uses the Hamming distance as the metric for the space. On the other hand, neural networks with memory can in principle provide examples of discrete chaos; numerical simulations show that the orbits on the attractor present topological transitivity and a dimensional phase space reduction. We compute this by extending the methodology of Grassberger and Procaccia to  $\Omega$ . As an example, we consider an asymmetric neural network model with memory which has an attractor of dimension  $D_a = 2$  for  $N = 49$ .

### 1. Introduction

Recently, discrete systems with a complex dynamical behaviour have received a great deal of attention, for their relevance in fields ranging from theoretical biology to quantum gravity. For example, asymmetric neural networks [1–4] can have a complicated dynamical behaviour which is reminiscent of ‘chaos’. Also cellular automata display [5] bifurcations between several possible dynamical regimes [6], the most disordered of which has been described as ‘chaotic’ [7]. Yet it is unclear precisely how this type of dynamics in discrete spaces is related to deterministic chaos in a Euclidean phase space.

In this paper we will examine how the definitions of deterministic chaos can be translated to the context of discrete state spaces. This will lead us to a formalism which we call ‘discrete chaos’, that allows one to decide whether or not the complex dynamics of some finite systems can be viewed as chaotic in the limit in which the system grows to infinity.

Unfortunately, for most finite systems there is no convenient quasi-representation in terms of real variables. For example, in neural networks and cellular automata the relevant distance is the Hamming distance; this induces a discrete topology on the space of states that is distinct from the usual topology of  $\mathbb{R}^n$ . There are different points of view on this problem, ranging from the fundamentalist, which concludes that a finite system cannot be viewed as approximately chaotic, to the liberal, which reduces the definition of chaos to the exponential growth of the limit-cycle period with the size of the system.

§ E-mail address: hwael@nuclecu.unam.mx

|| On sabbatical leave from the Institute of Nuclear Sciences, UNAM, Circuito Exterior CU, A. Postal 70-543, DF 04510, Mexico. E-mail address: H.Waelbroeck@fys.ruu.nl

¶¶ E-mail address: zertuche@matcuer.unam.mx

Our feeling is that chaos should not be limited to real variables, as these are idealizations of a reality which could be viewed equally well in terms of finite state spaces. Indeed, the fact that most real numbers have infinite algorithmic information [8] is not really satisfactory from a physicist's point of view. Yet some form of idealization is necessary to define chaos rigorously.

Our purpose in this article is to propose a different idealization, inspired from symbolic dynamics [9–13]. We will assume that one is given a representation of the system through a sequence of  $N$ -bit vectors. For example, one might consider the case where the different binary words carry information about the system at increasing temporal depths, e.g. by giving the  $N$ -bit description of the system at every past tick of a clock. In general, the *state* of the system will be given by

$$S = \{S(0), S(1), \dots, S(n), \dots\}$$

where  $S(n)$  is a vector with components  $S_i(n) = 0, 1, (i = 1, \dots, N)$ . The set of such binary states will be denoted by  $\Omega$ .

The approximation which makes this concept practical, akin to the 128-bit version of floating-point variables, is the truncation of the symbolic states to the first  $n$  words. This is a good approximation if the difference between states which coincide in the first  $n$  words belongs to a small neighbourhood of the origin. We will formalize this demand through the assignment of a base for the topology on  $\Omega$ , related to the cylinders of symbolic dynamics [10, 13].

With this topology, the space  $\Omega$  is homeomorphic to the one-sided shift space of symbolic dynamics. Our main contribution is to provide a definition of chaos for general dynamical maps in  $\Omega$ . In symbolic dynamics one usually considers the shift map  $\sigma$ , which consists of erasing the word  $S(0)$  from the semi-infinite sequence and shifting the other slices by  $S(n) \rightarrow S(n-1)$  [10]. This example satisfies our definition of discrete chaos. But we stress that this is only one of many possible chaotic maps in  $\Omega$ .

We will consider functions which are continuous or discontinuous. Neural networks and cellular automata will turn out to be examples of discontinuous functions. For general discontinuous functions very little is known, basically due to the fact that analytically there is very little that one can prove. However, numerically one can distinguish several types of dynamical behaviour. In  $\mathbb{R}^N$  the Grassberger and Procaccia method is widely used to estimate the fractal dimension of attractors. We will extend its application to the space  $\Omega$  in order to characterize different chaotic behaviours and define an effective attractor dimension.

One important class of maps which we will consider in this paper corresponds to the case when the binary state represents the system at every past tick of a clock, as explained above. To define such a map one must provide a function which allows one to compute the new word  $S(0)$  from the state  $S$ . The left inverse of any such map is the shift map  $\sigma$  of symbolic dynamics. Examples include neural network models and cellular automata. Note that in this case not all points of  $\Omega$  represent possible histories; instead,  $\Omega$  plays the role of an embedding space for the attractor.

The results of this paper can be generalized without difficulty to other alphabets besides the binary one, and also to the case where the space  $\Omega$  is the two-sided shift space [11, 12], where a state is given by a sequence

$$\{\dots S(-1), S(0), S(1), \dots\}.$$

In this case our construction reduces to the invertible shift map when  $S(n)$  is taken to be the binary description of the system but once again we stress that this is only one of several possible dynamical maps  $F : \Omega \rightarrow \Omega$ .

The organization of this paper is as follows. 'Discrete chaos' will be defined in section 2 and different types of dynamical maps in  $\Omega$  are discussed. In section 3 we will consider the

correlation function  $C(\rho)$  and the Grassberger–Procaccia method to compute the correlation dimension of the attractor. Numerical examples will be considered in section 4. In section 5 we give the conclusion.

## 2. Chaotic dynamics of binary systems

Binary systems, like cellular automata and neural networks, are described, in general, by a set of  $N$  binary variables  $S_i$ ,  $i = 1, \dots, N$ , or in short  $S$ , that evolve according to dynamical rules. The natural metric for these systems is the Hamming distance

$$d_H(S - S') \equiv \sum_{i=1}^N |S_i - S'_i|.$$

The space  $\{S\}$  has  $2^N$  possible states and so the topology constructed from  $d_H$  is discrete. Generally one is interested in studying these dynamical systems in the limit  $N \rightarrow \infty$  since that is where interesting statistical properties appear, such as phase transitions, and it is possible to use powerful techniques like mean field theory [1–4]. Furthermore numerical simulations which need to be done for finite, but large  $N$ , are understood as approximations of a system with infinite variables, much in the same way as floating point variables in computers are finite approximations of real numbers which generally have an infinite number of digits. Nevertheless for  $N \rightarrow \infty$ ,  $d_H$  is no longer a distance and the topology is ill defined in that limit. That makes our understanding of binary systems quite different from that of dynamical systems in  $\mathbb{R}^d$  or in differentiable manifolds where one works with the usual topology of the real numbers. Here we will overcome this situation by extending the phase space  $\{S\}$  to have an infinite number of states while preserving the equal status that the Hamming distance confers to each of the variables. That is to say, all the variables  $S_i$  give the same contribution to the distance for any  $i$ .

Let us consider the Cartesian product of infinite copies of  $\{S\}$  and call this space  $\Omega$ . We denote the elements of  $\Omega$  by

$$S = (S(0), S(1), S(2), \dots). \quad (1)$$

We make  $\Omega$  a topological space by introducing the following base:

$$\mathcal{N}_n(S) = \{S' \in \Omega | S'(m) = S(m), \forall m < n\} \quad (2)$$

with  $n = 1, 2, \dots$ . These base sets are closely related to the cylinders in one-sided shift spaces and  $\Omega$  is homeomorphic to the space of symbols of the symbolic dynamics with  $2^N$  symbols [10, 11]. It follows that  $\Omega$  is a cantor set. In symbolic dynamics the topology is usually derived from the metric

$$d(S, S') = \sum_{n=0}^{\infty} \frac{1}{2^n} d_n(S - S') \quad (3)$$

where

$$d_n(S - S') \equiv \sum_{i=1}^N |S_i(n) - S'_i(n)| \quad (4)$$

is the Hamming distance of the  $n$ th copy of  $\{S\}$ . One can check that if  $S(m) = S'(m) \forall m < n$  then  $d(S, S') < \frac{N+1}{2^{n-1}}$ , so that (2) and (3) define the same topology.

Here and in the following our purpose is to study dynamical systems in  $\Omega$  generated by a function  $F : \Omega \rightarrow \Omega$ . This function may be continuous or discontinuous, unless explicitly stated below. Allowing discontinuous functions in principle opens the door to a richer variety of systems, which include neural networks and cellular automata.

We begin by generalizing in a natural way the definitions of chaos in subsets of  $\mathbb{R}^N$  (see for example [11]) to  $\Omega$ .

**Definition 1.**  $F$  has sensitive dependence on initial conditions on  $\mathcal{A} \subset \Omega$  if  $\exists n \in \mathbb{N} \ni \forall S \in \mathcal{A}$  and  $\forall \mathcal{N}_m(S) \exists S' \in \mathcal{N}_m(S) \cap \mathcal{A}$  and  $k \in \mathbb{N}$  such that  $F^k(S') \notin \mathcal{N}_n(F^k(S))$ .

**Definition 2.** Let  $\mathcal{A} \subset \Omega$  be a closed invariant set.  $F : \Omega \rightarrow \Omega$  is topologically transitive on  $\mathcal{A} \subset \Omega$  if for any open sets  $U, V \subset \mathcal{A} \exists n \in \mathbb{Z} \ni F^n(U) \cap V \neq \emptyset$ . In the last expression, if  $F$  is non-invertible we understand  $F^{-k}(U)$  with  $k > 0$ , as the set of all points  $S \in \Omega$  such that  $F^k(S) \in U$ .

**Definition 3.** Let  $\mathcal{A} \subset \Omega$  be a compact set.  $F : \mathcal{A} \rightarrow \mathcal{A}$  is chaotic on  $\mathcal{A}$  if  $F$  has sensitive dependence on initial conditions and is topologically transitive on  $\mathcal{A}$ .

**Definition 4.** A closed subset  $\mathcal{M} \subset \Omega$  is called a trapping region if  $F(\mathcal{M}) \subset \mathcal{M}$ .

**Property 1.** If  $F$  is a continuous function in  $\Omega$ ,  $F^n(\mathcal{M})$  is compact and closed  $\forall n \in \mathbb{N}$ .

**Proof.** Since every closed subset of a compact set is compact, it follows that  $\mathcal{M}$  is compact and since  $F$  is continuous  $F^n(\mathcal{M})$  is compact. Since  $\Omega$  is Hausdorff every compact subset of it is closed, so  $F^n(\mathcal{M})$  is closed [14].  $\square$

**Definition 5.** The map  $F : \Omega \rightarrow \Omega$  has an attractor if it admits an asymptotically stable transitive set, i.e. if there exists a trapping region  $\mathcal{M}$  such that

$$\Lambda \equiv \bigcap_{n \geq 0} F^n(\mathcal{M})$$

and  $F$  is topologically transitive on  $\Lambda$ .

Note carefully that the trapping region is defined in the  $\Omega$  space while in the theory of dynamical systems in manifolds, it is defined in the manifold [10–13, 15]. This makes most theorems (as those shown in [15]) concerned with Cantor sets considered as attractors in manifolds to be not applicable.

**Property 2.** If  $F$  is a continuous function in  $\Omega$ ,  $\Lambda$  is compact and closed.

**Proof.** From property 1 if  $F$  is continuous,  $\Lambda$  is an intersection of closed sets, so it is closed. Since every closed subset of a compact space  $\Omega$  is compact, it follows that  $\Lambda$  is compact.  $\square$

**Definition 6.**  $\Lambda$  is called a chaotic attractor if  $F$  is chaotic on  $\Lambda$ .

**Lemma.** Let  $F$  be a continuous function in  $\Omega$ , if  $\Lambda$  is a chaotic attractor then it is perfect.

**Proof.** By property 2,  $\Lambda$  is closed. It remains to prove that every point in  $\Lambda$  is an accumulation point of  $\Lambda$ . By contradiction, let  $S_0 \in \Lambda$  be an isolated point, then there exists  $n \in \mathbb{N} \ni \mathcal{N}_n(S_0) \cap \Lambda = \{S_0\}$ . Then, by topological transitivity  $\Lambda$  has an isolated orbit (the orbit of  $S_0$ ) which implies that it is not sensitive to initial conditions on  $\Lambda$ .  $\square$

**Theorem.** If  $F$  is a continuous function in  $\Omega$ , and  $\Lambda$  is a chaotic attractor then it is a Cantor set.

**Proof.** The theorem follows directly from property 2, the lemma and the fact that a subset of a totally disconnected set is also totally disconnected.  $\square$

In the following we will consider some examples of dynamical functions  $f : \Omega \rightarrow \Omega$ . The first one is the one-side shift map  $\sigma$  of symbolic dynamics which we introduce to familiarize the reader with the notation.

(i) The one-sided shift map  $\sigma$ .

The continuous map  $\sigma$  defined by

$$\sigma(S(0), S(1), \dots) = (S(1), S(2), \dots) \tag{5}$$

is chaotic in  $\Omega$  [10]. Note that  $\sigma$  is non-invertible and its action loses the information carried by the binary state  $S(0)$ . The meaning and usefulness of this map is quite clear in the context of symbolic dynamics when the Conley–Moser conditions are satisfied [16]. There one studies, in general, a non-invertible function  $f : \Xi \rightarrow \Xi$  where  $\Xi$  is a Cantor set embedded in  $\mathbb{R}^N$ . The set  $\Xi$  is divided in  $2^N$  sectors  $I_\alpha$   $\alpha = 0, 1, \dots, 2^N$ . Then it is possible to establish a topological conjugation between  $f$  and  $\sigma$  through a homeomorphism  $\psi$ , so that the following diagram commutes [11]

$$\begin{array}{ccc} \Xi & \xrightarrow{f} & \Xi \\ \psi \downarrow & & \downarrow \psi \\ \Omega & \xrightarrow{\sigma} & \Omega \end{array} \tag{6}$$

Moreover, let  $S = \psi(x)$ , then  $S(n)$  is the binary decomposition of the label  $\alpha$ , such that  $f^n(x) \in I_\alpha$ .

(ii) Chaotic maps with non-trivial attractors in  $\Omega$ .

The shift map can be modified to create maps which are homeomorphic to the shift map on an asymptotically stable transitive subset of the space of symbols. We introduce two very simple examples.

Take the space of symbols  $\Omega$  with  $N = 2$ , homeomorphic to  $\Xi \times \Xi$  where  $\Xi$  is the space of symbols with  $N = 1$ , that is the space of semi-infinite sequences  $S = (S_0, S_1, S_2, \dots)$ . Then consider the function  $f_c : \Xi \times \Xi \rightarrow \Xi \times \Xi$  given by  $f_c = \sigma \times \zeta$ . Where  $\sigma$  is the usual shift function and  $\zeta$  is a right inverse of the shift function defined as follows:

$$\zeta(S_0, S_1, S_2, \dots) = (0, S_0, S_1, S_2, \dots).$$

It is easy to check that  $\zeta$  is a continuous function, and of course so is the shift: so  $f_c$  is continuous. The set  $\Xi \times \{0\}$  is an asymptotically stable transitive set, on which the restriction of  $f_c$  is the shift map  $\sigma$ .

As another example, consider the space  $\Omega$  with  $N = 1$ . It can be split into the disjoint union of two Cantor sets  $\Omega = \Lambda_0 \cup \Lambda_1$ . Where  $\Lambda_0$  is the set of sequences such that  $S_0 = 0$  and an analogous fashion for  $\Lambda_1$ . Take the continuous function  $f_\pi = \pi \circ \sigma$ , where  $\sigma$  is the shift map and  $\pi$  projects  $\Omega$  in  $\Lambda_0$  such that:

$$\pi(S_0, S_1, S_2, \dots) = (0, S_1, S_2, \dots).$$

Then the action of  $f_\pi$  is given by,

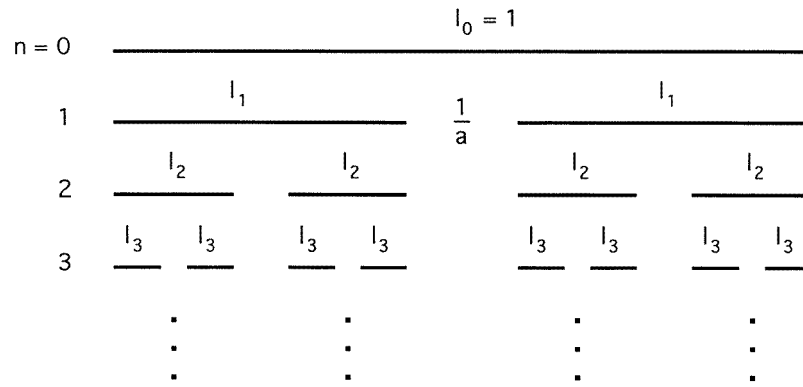
$$f_\pi(S_0, S_1, S_2, \dots) = (0, S_2, S_3, \dots).$$

It is easy to check that  $\Lambda_0$  is a chaotic attractor of  $f_\pi$ .

(iii) Chaotic maps in  $\Omega$  induced through chaotic maps in Cantor subsets of  $\mathbb{R}^N$ .

We will consider a homeomorphism which relates a Cantor set  $\chi \subset \mathbb{R}^N$  to the space  $\Omega$  and allows one to construct chaotic maps in  $\Omega$  from chaotic maps in  $\chi$  through topological conjugation. Let  $\chi \subset \mathbb{R}^N$  be the Cantor set that results from taking the Cartesian product of  $N$  Cantor sets  $\chi_i$ ;

$$\chi = \bigotimes_{i=1}^N \chi_i$$



**Figure 1.** Construction of the Cantor sets  $\Xi_i, i = 1, \dots, N$  by suppressing from  $[0, 1]$  the open middle  $1/a_i$  part,  $1 < a_i < \infty$ . The remaining  $2^n$  intervals at the  $n$ th step of the construction are of length  $l_n = \frac{1}{2^n} (1 - \frac{1}{a_i})^n$ .

where the  $i$ th component  $\chi_i$  is constructed by suppressing from the interval  $[0, 1]$  the open middle  $1/a_i$  part,  $i = 1, \dots, N, a_i > 1$ , and repeating this procedure iteratively with the sub-intervals, see figure 1. Now, we define  $\phi : \Omega \rightarrow \chi$  by:

$$\phi_i(S) = \sum_{n=1}^{\infty} (l_{n-1} - l_n) S_i(n - 1) \tag{7}$$

where

$$l_n = \frac{1}{2^n} \left(1 - \frac{1}{a_i}\right)^n \tag{8}$$

is the length of each of the remaining  $2^n$  intervals at the  $n$ th step of the construction of  $\chi_i$ . If  $\Omega$  is endowed with the metric (3) and  $\chi \subset \mathbb{R}^N$  with the standard Euclidean metric, is easy to show that  $\phi$  is a homeomorphism.

Now, if we have a map  $f : \mathbb{R}^N \rightarrow \mathbb{R}^N$  which is chaotic in  $\chi$  we can construct a map  $F : \Omega \rightarrow \Omega$  which is chaotic in  $\Omega$ , and is defined through the commutation of the diagram

$$\begin{array}{ccc} \chi & \xrightarrow{f} & \chi \\ \phi \uparrow & & \uparrow \phi \\ \Omega & \xrightarrow{F} & \Omega \end{array} \tag{9}$$

This leads to an interesting practical application of the homeomorphism  $\phi$ , to realize computer simulations of chaotic systems on Cantor sets. If, for example, one iterates the logistic map  $f(x) = \mu x(1 - x)$  for  $\mu \geq 4$  with a floating-point variable, the truncation errors nudge the trajectory away from the Cantor set and eventually  $x \rightarrow -\infty$ . The homeomorphism  $\phi$  suggests a natural solution to this, which is to iterate the truncated binary states rather than the floating-point variable. To iterate the dynamics, one computes  $x_i = \phi_i(S) \forall i = 1, \dots, N$  by assuming that the truncated bits are all equal to zero, then applies  $f$  to obtain  $x' = f(x)$ . Since  $x'$  generally does not belong to the Cantor set (because of truncation errors), in the process of constructing  $S' = \phi^{-1}(x')$ , at some  $n$  one will find that this point does not belong to either the interval corresponding to  $S_i(n) = 0$  or to  $S_i(n) = 1$ . This truncation error can be corrected by moving to the extremity of the interval which lies closest to  $x'_i$ . In this way, truncation errors are not allowed to draw the trajectory away from the Cantor set  $\chi \subset \mathbb{R}^N$ .

(iv) Binary systems with memory.

Now we are going to define a map  $\Gamma : \Omega \longrightarrow \Omega$  which is very useful to analyse binary systems with causal deterministic dynamics on  $N$  bits, such as neural networks, cellular automata, and neural networks with memory [1–4, 17]. Let

$$\gamma_i : \Omega \longrightarrow \{0, 1\} \quad (10)$$

$i = 1, \dots, N$ , be a set of continuous or discontinuous functions.  $\Gamma : \Omega \longrightarrow \Omega$  is then defined by:

$$\Gamma_i(S) = (\gamma_i(S), S_i(0), S_i(1), \dots)$$

or in a short hand notation

$$\Gamma(S) = (\gamma(S), S). \quad (11)$$

Such maps have the following properties.

**Property 3.** *The shift map (5) is a left inverse of  $\Gamma$  since from (11)  $\sigma \circ \Gamma(S) = S$ . If  $\Omega$  has an attracting set  $\Lambda \subset \Omega$ , then  $\sigma$  is also a right inverse in the restriction of  $\Gamma$  to  $\Lambda$ , so that,  $\Gamma|_{\Lambda}^{-1} = \sigma$ .*

**Proof.**  $\forall S \in \Lambda \exists S' \in \Lambda$  such that  $\Gamma(S') = S$ . Since

$$\Gamma(S') = (\gamma(S'), S') = S$$

and

$$S = (S(0), S_1)$$

where  $S_1 \equiv (S(1), S(2), \dots)$ , one sees that  $S' = S_1$ . Thus,

$$\Gamma \circ \sigma(S) = \Gamma(S_1) = \Gamma(S') = S.$$

□

**Property 4.**  *$\Gamma$  has an attracting set  $\Lambda$  contained properly in  $\Omega$ .*

**Proof.** Given  $S$  there are  $2^N$  states  $S' = (S'(0), S)$  of which only one,  $\Gamma(S) = (\gamma(S), S)$ , belongs to  $\Gamma(\Omega)$ . Therefore the set

$$\Lambda \equiv \bigcap_{n \geq 0} \Gamma^n(\Omega)$$

is a proper subset of  $\Omega$ .

□

**Property 5.** *If  $\Gamma$  is continuous, then it is not sensitive to initial conditions.*

**Proof.**  $\Gamma$  is a continuous map on a compact set, so it is uniformly continuous. Therefore there exists a  $\delta > 0$  such that for any  $S \in \Omega$ ,  $d(S', S) < \delta \Rightarrow \gamma(S) = \gamma(S')$  and hence  $d(\Gamma(S), \Gamma(S')) < \delta/2$ , where the distance function is given by (3). Applying the same argument to each iterate  $\Gamma^k(S)$  shows that  $d(\Gamma^k(S), \Gamma^k(S')) < \delta/2^k$ , which contradicts sensitivity to initial conditions. □

**Property 6.** *If  $\Gamma$  is continuous, then the attractor  $\Lambda$  is finite.*



**Proof.** From property 4 we know that  $\Lambda$  exists. The property then follows from property 5 above. Indeed, if  $\Gamma$  is not sensitive to initial conditions, then there is a  $n > 0$  such that  $\forall S \in \Omega$

$$\lim_{k \rightarrow \infty} d(\Gamma^k(S) - \Gamma^k(S')) = 0$$

$\forall S' \in \mathcal{N}_n(S)$ . The set  $A \subset \Omega$  defined by  $S \in A$  iff  $\forall m > n, S(m) = 0$ , has a finite number of elements, namely  $2^{N \times n}$ . The whole space  $\Omega$  is the union of the  $n$ -neighbourhoods of each element of  $A$ , and as we just showed the map  $\Gamma$  is contracting in each such neighbourhood, so the number of points in the attractor cannot be greater than the number of elements of  $A$ , namely  $2^{N \times n}$ .  $\square$

Neural networks and cellular automata are binary dynamical systems in which the values of the state variables  $S_i, i = 1, \dots, N$ , at time  $t$  depend on the state variables at time  $t - 1$ . These systems are described by a function  $\Gamma$  such that the functions  $\gamma_i$  depend only on the components  $S(0)$ . Therefore, all points  $S' \in \mathcal{N}_n(S)$  for  $n > 0$  have the same evolution so that these systems are not sensitive to initial conditions. One can recover a very rough approximation of sensitive dependence on initial conditions by considering the growth of *Hamming distance* with time, rather than the metric (3) of symbolic dynamics. However, one cannot describe the behaviour of these systems to be approximately chaotic: they are well known to have attractors that consist of a collection of periodic limit-cycles, and as we will see in section 4, the points of these limit-cycles are scattered over configuration space without any effective lower-dimensional structure. In particular, given any one point on the attractor there is usually no other point ‘nearby’, even in the weak sense of the Hamming distance, that also belongs to the attractor. This fact makes most practical uses of chaos theory in prediction and control inapplicable.

(v) A compact topology for neural networks and cellular automata.

Since neural networks and cellular automata in general are systems in which all the variables have the same type of interactions, it is natural to consider the Hamming distance as the metric (it is in fact the most widely used metric in the literature, see for instance [1–4] and the references therein). We have already seen that the topological structure which the Hamming distance confers to the phase space does not conduce to chaotic behaviour in the sense that we understand it even if we extend the phase space to  $\Omega$ . However, not all the neural network and cellular automata models confer the same type of interactions to neurons, so the use of the Hamming distance for the metric is not so compelling. The use of a different metric can lead to a completely different topology. The resulting system will in general display a very different dynamical behaviour. For example the map  $x_{n+1} = \alpha x_n$  produces quite different dynamical behaviours for  $x_n \in \mathbb{R}$  and  $x_n \in S^1$ .

So, let us consider systems which evolve according to the rule

$$R_i(t + 1) = f_i(R(t)) \quad (12)$$

$R_i = 0, 1; i = 1, \dots, M$  and take for the metric

$$d(S, S') = \sum_{n=0}^M \frac{1}{2^n} d_n(S - S'). \quad (13)$$

These systems include neural networks and cellular automata as particular examples, but where the weight of the different neurons drops off as  $2^{-n}$ . The metric (13) remains well defined in the limit  $M \rightarrow \infty$  and once again we obtain the space  $\Omega$ . In fact (12) and (13) with  $M \rightarrow \infty$  are equivalent to (3) and (4) with  $N = 1$  and  $S_1(n) = R_n$ . As we will see in the next section these systems can have a correlation dimension which is less than or equal to one.

### 3. Correlation function

In the theory of dynamical systems in  $\mathbb{R}^N$  one is interested in calculating the fractal dimension of the attractor in which the system evolves. To do so, following the method of Grassberger and Procaccia [18] one defines the correlation function  $\mathcal{C}(\rho)$  as the average of the number of neighbours  $S_t, S_{t'}$ , with  $S_t = F^t(S)$ , which have a distance smaller than  $\rho$ . Since in  $\mathbb{R}^N$  the volume of a sphere of radius  $\rho$  grows like  $\rho^N$ , one identifies the correlation dimension  $D_a$  of the attractor with the growth rate in  $\mathcal{C}(\rho) \sim \rho^{D_a}$ . This leads to the definition of the correlation dimension as

$$D_a = \lim_{\rho, \rho' \rightarrow 0} \left( \frac{\log(\mathcal{C}(\rho)) - \log(\mathcal{C}(\rho'))}{\log(\rho) - \log(\rho')} \right). \tag{14}$$

In order to have an analogous methodology to compute correlation dimensions in  $\Omega$ , it is necessary to know how many states  $S'$  are within a distance less than  $\rho$  from a given point  $S$ . Since  $\Omega$  is homogeneous we can take  $S = 0$ . To do the calculation we make  $\Omega$  into a finite space by truncating the semi-infinite sequence to only  $T$  slices, and take the limit  $T \rightarrow \infty$  in the end, that is:

$$\mathcal{C}(\rho) = \lim_{T \rightarrow \infty} \frac{1}{2^{NT}} \sum_{\{S\}} \Theta(\rho - d(S, 0)) \tag{15}$$

where the distance is given by (3). Expressing  $\Theta(x)$  in terms of its Fourier transform  $\omega(k) = \pi \delta(k) - \frac{1}{k}$  we have

$$\mathcal{C}(\rho) = \lim_{T \rightarrow \infty} \frac{1}{2^{NT}} \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk \omega(k) e^{ik\rho} \sum_{\{S\}} e^{-ikd(S,0)}.$$

The sum over  $\{S\}$  can be evaluated easily obtaining

$$\sum_{\{S\}} e^{-ikd(S,0)} = 2^{NT} e^{-ikN} \left( \prod_{n=0}^T \cos \frac{k}{2^{n+1}} \right)^N.$$

Using the identity  $\sin k/k = \prod_{n=0}^{\infty} \cos \frac{k}{2^{n+1}}$  we obtain the integral

$$\mathcal{C}(\rho) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk \omega(k) \left( \frac{\sin k}{k} \right)^N e^{ik(\rho-N)}$$

which may be evaluated by standard complex variable methods, to obtain the final result for the correlation function in  $\Omega$ ,

$$\mathcal{C}(\rho) = \frac{1}{2^N N!} \sum_{k=0}^{[\rho/2]} (-1)^k \binom{N}{k} (\rho - 2k)^N. \tag{16}$$

So we see that the scaling in  $\Omega$  is not a power law as in  $\mathbb{R}^N$ . However, in the definition of the attractor dimension one is interested in calculating  $\mathcal{C}(\rho)$  for  $\rho \rightarrow 0$ . For  $\rho \leq 2$  equation (16) has the form

$$\mathcal{C}(\rho) = \frac{1}{2^N N!} \rho^N. \tag{17}$$

Therefore, the same techniques applied in  $\mathbb{R}^N$  can be used in  $\Omega$ . In particular an effective ‘attractor dimension’ will be given by (14).

#### 4. Numerical examples

We have examined numerically several of the binary systems which have been considered in the literature, including the random  $k = 4$  cellular automata [7], and some neural network models such as those studied by Crisanti *et al* in the context of Shannon's entropy [2].

Random cellular automata of rank  $k$  consist of  $N$  binary variables  $S_i = 0, 1$  with the following dynamical rule. For each binary variable one chooses at random a Boolean function  $f_i : \mathbb{Z}_2^k \rightarrow \mathbb{Z}_2$  from  $k$  binary variables into one, among the  $2^{2^k}$  possible functions. Then, for each  $i = 1, \dots, N$ , a set of  $k$  numbers  $\{i_1, i_2, \dots, i_k\}$  is selected at random from  $\{1, \dots, N\}$ . These numbers are interpreted as labels for the  $k$  inputs of the Boolean function at that  $i$ . The evolution of the system in time is given by applying the Boolean rules synchronously at the  $N$  variables:

$$S_i(t+1) = f_i(S_{i_1}(t), \dots, S_{i_k}(t)). \quad (18)$$

For  $k \geq 4$  the system has very long cycles with periods of order  $e^N$  and present the phenomenon of damage spreading which is the standard way in which a first 'Lyapunov exponent' is assigned to such systems.

In [2] Crisanti *et al* studied a binary neural network described by variables  $S_i = \pm 1$ . The variables evolve in parallel according to the rule

$$S_i(t+1) = \text{sgn}\left(\sum_{j=1}^N J_{ij} S_j(t)\right) \quad (19)$$

where

$$J_{ij} = J_{ij}^S + k J_{ij}^A$$

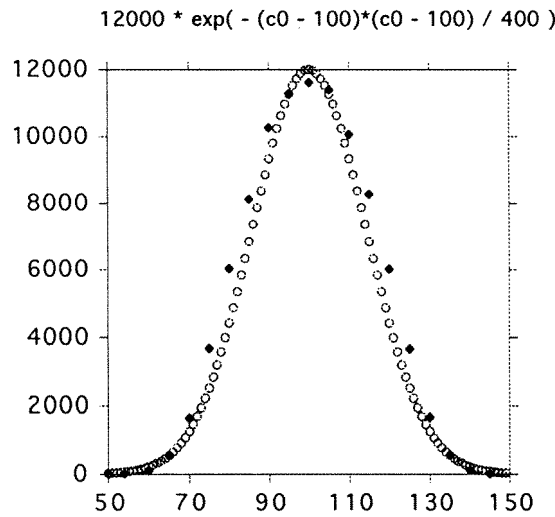
is the synaptic matrix, with  $J_{ji}^S = J_{ij}^S$  and  $J_{ji}^A = -J_{ij}^A$  being random independent Gaussian variables with mean zero and variance  $\sigma^2 = 1/(N-1)(1+k)$ . The parameter  $k$  measures the amount of asymmetry of the synapses. For  $k = 0$  the matrix is symmetric and the network has fixed points as attractors. For  $k > 0$  it is asymmetric and the network can have limit cycles as attractors. When  $k > k_c = 0.5$ , long limit cycles are obtained, with period of order  $e^N$  but with fluctuations of the same order. In [2] the Shannon entropy has been calculated numerically in the range  $k_c < k < 0.9$  and the scaling was found to be given by

$$h \sim (k - k_c)^{1/2}.$$

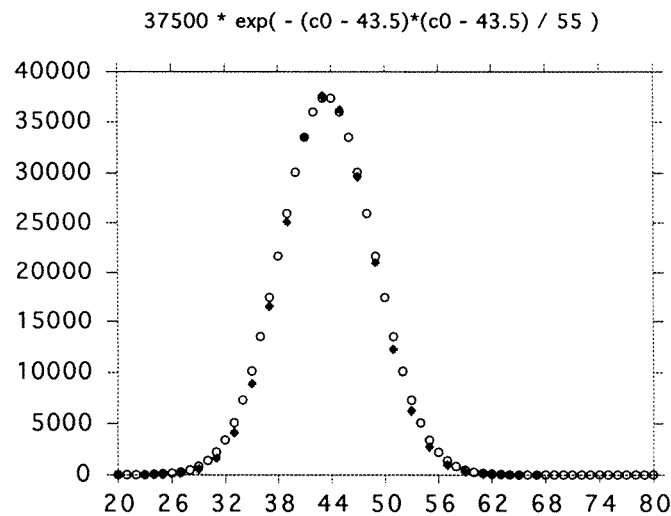
In the limit  $k \rightarrow 1$ ,  $h$  attains a value which is very close to the maximum value  $\log(2)$ , characteristic of a random process. All of this indicates a high degree of complexity.

However, for both of the dynamical systems (18) and (19), the dynamics is not topologically transitive on the limit set of all periodic orbits: as we will see below, points on this set are isolated in Hamming distance, so most points do not even have any 'near-neighbours' that might attempt to satisfy the conditions in the definition of topological transitivity.

The number of returns to within a Hamming distance  $d_H$  of an initial point on one of the long periodic orbits is given in figures 2 and 3 for the cellular automata (18) with  $k = 4$  and the neural network (19) with  $k = 1.2$ . We have also graphed the best-fitting Gaussian curve for comparison. The first obvious result is that there are no returns with  $d_H < 50$  in an automata with  $N = 200$ , in  $5 \times 10^6$  iterations of the dynamical map. The neural network was run with  $N = 100$  neurons for  $5 \times 10^5$  iterations, and again no returns were found with  $d_H < \frac{N}{4}$ . Both the value of the nearest return and the fit to a Gaussian are consistent with a random process which produces patterns all over the configuration space without any restriction to a possible 'attracting subspace'. This indicates a very high degree of algorithmic complexity [8] in the



**Figure 2.** The number of times a  $k = 4$  random cellular automata with  $N = 200$  returns to a Hamming distance  $d_H$  of a point half-way along the trajectory is represented as a function of  $d_H$  (full circles). The best-fitting Gaussian is also given for comparison (open circles).



**Figure 3.** The number of returns to Hamming distance  $d_H$  is shown, for an asymmetric neural network with  $N = 100$  (full circles). The best-fitting Gaussian is also given for comparison (open circles).

time-series, which reflects a lack of predictability, like Shannon entropy which is a statistical measure of disorder.

Without anything analogous to a transitive ‘attractor’, none of the practical applications of chaos theory can carry through for large values of  $N$ . The phase space reconstruction methods and other versions of the ‘method of analogues’ [19] fail because *one finds no good analogue* in any finite data set, for large  $N$ . The lack of close returns in binary systems can often be related to the failure to find an attractor on which the dynamics is topologically transitive.

Another example, more in the spirit of the maps  $\Gamma$  is an asymmetric neural network with

state-dependent synapses originally designed to recognize sequences of patterns, and described in [4]. As shown there, this system has a transition from a stable sequence reproduction to a disordered behaviour. We shall modify the dynamical rule by introducing a memory in an analogous way as has been done in [17], as follows:

$$S_i(t+1) = \text{sgn} \left( \sum_{n=0}^{T-1} \frac{1}{2^n} \sum_{j=1}^N J_{ij}^{(n)} S_j(t-n) \right) \quad (20)$$

where the synapses is given by

$$J_{ij}^{(n)} = \frac{1}{N} \sum_{\mu=1}^p \xi_i^{\mu+n+1} \xi_j^{\mu} \quad (21)$$

and

$$S_{\mu} = \frac{1}{N} \sum_{i=1}^N S_i \xi_i^{\mu}$$

is the correlation of the state of the network with the pattern  $\xi^{\mu}$ . The patterns  $\xi^{\mu} = \pm 1$  with  $\mu = 1, \dots, p+T$  are random independent, equiprobable variables, and  $p$  is a parameter of the model. The reader not familiar with the notation of Hopfield-type neural networks may refer to [4].

It is easy to show that the argument of the sign function referred to above as function  $\gamma_i$ , does in fact vanish in  $\Omega$  when  $T \rightarrow \infty$ . So the map  $\Gamma : \Omega \rightarrow \Omega$  is discontinuous, and one cannot immediately rule out the possibility of non-periodic orbits [20]. In practice, however, one always uses a finite memory in computer simulations, so the continuity is recovered and it is the sensitivity to initial conditions which is not valid.

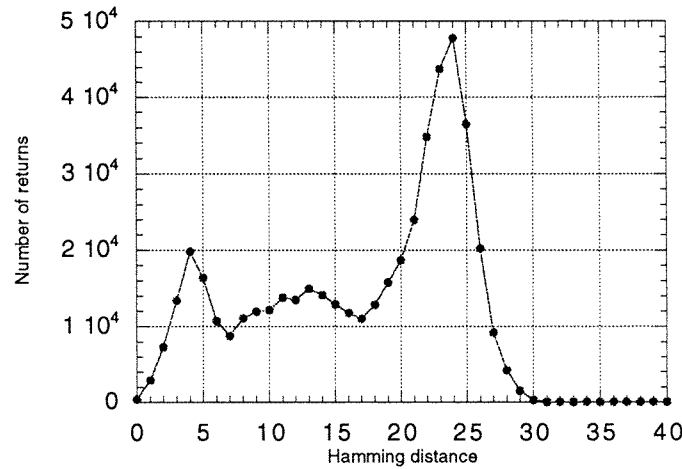
In applications it is often the case that two points can be considered to be distinguishable only if their mutual distance is greater than a small ‘cut-off’ value  $\lambda$ . If  $\lambda > 1/2^T$  one can then claim that the map is ‘effectively’ sensitive to initial conditions, which suggests that the ensuing dynamics may be correctly described as being ‘approximately chaotic’. This is best analysed by computer simulation.

We have run this system with  $T = 30$ ,  $N = 49$  neurons and  $p = 19$  and find evidence for a non-trivial attractor with a low effective dimension. Unlike in the examples above there are substantially more near-returns than for a random sequence, as shown in figure 4, and the correlation function  $\mathcal{C}(\rho)$  produces an effective dimension  $D_a \approx 2$  in the range  $0.002 < \rho < 0.02$  (see figure 5).

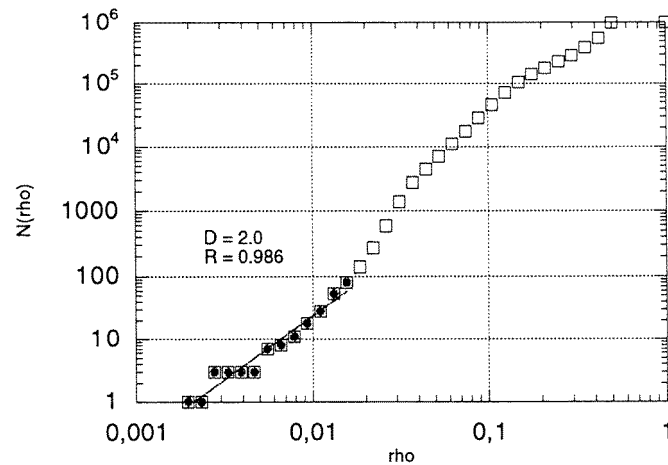
For both the neural network (19) and the  $k = 4$  cellular automata (18), we observe a very different behaviour, as expected. The correlation graph  $\mathcal{C}(\rho)$  coincides for all  $\rho$ , with (16), within an accuracy of the 98% which means that  $D_a \approx N$ . This means that the orbits are scattered in the whole space  $\Omega$  as one would expect from the fact that the distribution of returns to Hamming distance  $d_H$  is approximately Gaussian.

## 5. Conclusion

From an initial ansatz, to replace the usual idealization of physical states as ‘points’ on a differentiable manifold by another idealization as infinite ‘binary states’, we proceeded to define a topology which makes the truncation to finite states a valid approximation, in the same sense that the usual topology on  $\mathbb{R}^N$  allows one to approximate a real coordinate by a finite string of digits or bits. This lead us to a space  $\Omega$  which is homeomorphic to the space of symbolic dynamics.



**Figure 4.** The number of returns to Hamming distance  $d_H$  is given for a neural network model with memory, with  $N = 49$ . Unlike the previous examples, which correspond to dynamical systems without memory, we find many analogues. There is one return with  $d_H = 0$ ; the system did not fall on a limit-cycle at that point because the dynamics also considers binary words further back in time.



**Figure 5.** The correlation graph  $N(\rho)$  gives the effective attractor dimension for the neural network with memory,  $D_a \approx 2$  in the range  $0.002 < \rho < 0.02$ . The distance  $\rho$  is given by equation (3).

Continuous or discontinuous dynamical maps on the space of symbolic states can lead to attracting sets within  $\Omega$ , in which case an attractor is defined in the usual way. The dynamical map is said to be chaotic on the attractor if it is sensitive to initial conditions and topologically transitive.

Finite systems such as neural networks and cellular automata without memory that depend only on the previous time-step and for which the different bits have comparable importance do *not* provide good approximations of chaos when the Hamming distance is used as the metric. The absence of even an approximate manifestation of chaos has important practical consequences—for example we found that prediction models based on a search for analogous examples in a data set are not applicable because no good analogues are found in any reasonable

amount of data.

The practical value of the analysis of dynamical maps on the space  $\Omega$  is probably limited to special complex systems problems where an extended binary description is more natural than a continuum description. Discrete spacetime formulations of quantum gravity offer another potentially rewarding area of applicability. There, the evidence for discrete small-scale structure, combined with the perceived need of a spacetime ‘sum over histories’ interpretation, leads to a formalism where one defines ‘states’ to be truncated discrete spacetime histories. An interesting example is the causal set formalism, where a partially ordered set or Poset is conjectured to constitute the minimal required structure to formulate a theory of quantum gravity.

A priority in the continuation of this work is to further elucidate the chaotic properties of neural networks and cellular automata when a compact metric compatible with the topology of  $\Omega$  is given.

### Acknowledgments

The authors would like to thank C Stephens and M Socolovsky for helpful comments and discussions; V Domínguez, E Sacristán, and F Toledo for computational advice. HW would also like to thank Th W Ruijgrok for helpful conversations at Utrecht, and acknowledge the financial support of the Spinoza grant from NWO. FZ thanks S Cosentino for their clarifying comments in relation to symbolic dynamics. This work is supported in part by the DGAPA-UNAM grant IN105197.

### References

- [1] Coolen A C C and Sherrington D 1992 Competition between pattern reconstruction and sequence processing in non-symmetric neural networks *J. Phys. A: Math. Gen.* **25** 5493–526
- [2] Crisanti A, Falcioni M and Vulpiani A 1993 Transition from regular to complex behaviour in a discrete deterministic asymmetric neural network model *J. Phys. A: Math. Gen.* **26** 3441
- [3] Coolen A C C and Ruijgrok Th W 1988 Image evolution in Hopfield networks *Phys. Rev. A* **38** 4253–5
- [4] Zertuche F, López-Peña R and Waelbroeck H 1994 Recognition of temporal sequences of patterns with state-dependent synapses *J. Phys. A: Math. Gen.* **27** 5879–87
- [5] Kauffman S 1993 *The Origins of Order: Self-Organization and Selection in Evolution* (New York: Oxford University Press)
- [6] Derrida B 1987 Dynamical phase transitions in random networks of automata *Chance and Matter* ed J Souletie *et al* (Amsterdam: North-Holland)
- [7] Weisbuch G 1991 *Complex Systems Dynamics* (Redwood City, CA: Addison-Wesley)
- [8] Chaitin G J 1994 Randomness and complexity in pure mathematics *Int. J. Bifurcation Chaos Appl. Sci. Eng.* **4** 3–15  
Chaitin G J 1995 *Guanajuato Lectures on Complex Systems and Binary Networks (Springer Lecture Notes Series)* ed R López Peña *et al* (Berlin: Springer)
- [9] Hao B-L 1989 *Elementary Symbolic Dynamics and Chaos in Dissipative Systems* (Singapore: World Scientific)
- Bruin H 1995 Combinatorics of the Kneading map *Int. J. Bifurcation Chaos Appl. Sci. Eng.* **5** 1339–49
- [10] Devaney R 1989 *An Introduction to Chaotic Dynamical Systems* (Reading, MA: Addison-Wesley)
- Katok A and Hasselblatt B 1995 *Introduction to the Modern Theory of Dynamical Systems* (Cambridge: Cambridge University Press)
- Robinson C 1995 *Dynamical Systems* (Boca Raton, FL: Chemical Rubber Company)
- [11] Wiggins S 1990 *Dynamical Systems and Chaos* (New York: Springer)
- [12] Wiggins S 1988 *Global Bifurcations and Chaos* (New York: Springer)
- [13] Lind D and Marcus B 1995 *An Introduction to Symbolic Dynamics and Coding* (Cambridge: Cambridge University Press)
- [14] Munkres J R 1975 *Topology a First Course* (Englewood Cliffs, NJ: Prentice-Hall)
- [15] Buescu J and Stewart I 1995 Liapunov stability and adding machines *Ergod. Theor. Dynam. Syst.* **15** 271–90

- Buescu J 1997 *Exotic Attractors, Liapunov Stability and Riddled Basins (Progress in Mathematics 153)* (Basel: Birkhäuser)
- [16] Moser J 1973 *Stable and Random Motions in Dynamical Systems* (Princeton, NJ: Princeton University Press)  
Alekseev V M 1968 Quasirandom dynamical systems I–III *Math. USSR-Sb.* **5** 73–128  
Alekseev V M 1968 Quasirandom dynamical systems I–III *Math. USSR-Sb.* **6** 505–60  
Alekseev V M 1969 Quasirandom dynamical systems I–III *Math. USSR-Sb.* **7** 1–43
- [17] Sompolinsky H and Kanter I 1986 Temporal association in asymmetric neural networks *Phys. Rev. Lett.* **57** 2861  
Riedel U, Kühn R and van Hemmen J L 1988 Temporal sequences and chaos in neural nets *Phys. Rev. A* **38** 1105
- [18] Grassberger P and Procaccia I 1983 *Physica D* **9** 189–208  
Grassberger P and Procaccia I 1983 *Phys. Rev. Lett.* **50** 346–9
- [19] Lorenz E N 1991 Dimension of weather and climate attractors *Nature* **353** 241 and references therein
- [20] Ruijgrok Th W Private communication