WORKING PAPER

On observability of chaotic systems: an example

Giovanni B. Di Masi
Andrea Gombani

August 1989
WP-89-63

IIASA
International Institute
for Applied Systems Analysis

On observability of chaotic systems: an example

Giovanni B. Di Masi
Andrea Gombani

August 1989
WP-89-63

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

FOREWORD

The problem of studying the behaviour at infinity of a system has always been central in mathematical modelling, and there is a complete theory for systems converging to a limit point or a limit cycle. The study of systems which still converge to some subsets of points at infinity, but present a more complicated (chaotic) behaviour has only been investigated in the last few decades and still remains to be fully understood. In particular, problems connected to control and observation of chaos are quite new and seem challenging. This paper studies an example of imperfect observation of a chaotic dynamical system.

Alexander B. Kurzhanski
Chairman
System and Decision Sciences Program

On observability of chaotic systems: an example

Giovanni B. Di Masi and Andrea Gombani

On observability of chaotic systems: an example

Giovanni B. Di Masi
Andrea Gombani

Abstract: the concept of observability of a special chaotic system, namely the dyadic map, is studied here in case the observation is not exact. The usual concept of observable subspace does not distinguish among the behaviour of different models. It turns out that a suitable extension of this concept can be obtained using the idea of Hausdorff dimension. It is shown that this dimension increases as the observation error becomes smaller, and is equal to one only if the system is observable.

§1. Introduction

The study of nonlinear dynamical systems has recently attracted the attention of an increasing group of scientists involved in theoretical as well as in applicative fields. In particular there has been a growing interest for chaotic systems since it has been recognized that chaotic and random behaviour of solutions of deterministic systems is an inherent feature of many physical and engineering phenomena.

Since a possible characterization of chaos is that, under a suitable observation mechanism, the output of the system behaves as a purely nondeterministic process, it is of interest to study the observability properties of such systems. Results in this direction can be found in [1],[4],[8].

In this paper we examine the observability properties of a simple chaotic system described by the dyadic map, whose dynamic behaviour can be effectively characterized in terms of symbolic dynamics.

It turns out that a natural extension of the concept of dimension of the observability space for linear systems can be given in terms of Hausdorff dimension of the observable set. The tool of Hausdorff dimension has been used in investigations on chaotic systems in connection with the study of the dimension of strange attractors [6].

§2. The problem

Let I be the unit interval. By *chaotic system* it is usually meant a map f: I $\rightarrow$ I with the following properties [5]:

1. f has sensitive dependence on the initial conditions, i.e. there exists a $\delta > 0$ s.t. for each $x,y \in I$ there exists $n \in \mathbb{N}$, s.t. $|f^n(x) - f^n(y)| > \delta$.

2. periodic points are dense in I.

3. f is topologically transitive, i.e. for any pair of open sets $U,V \subset I$, there exist $k > 0$ s.t. $f^k(U) \cap V \neq \emptyset$.

It is fairly straightforward to check that the dyadic (figure 1)

$$f(x) = \begin{cases} 2x & \text{for } 0 \leq x < 1/2 \\ 2x - 1 & \text{for } 1/2 \leq x \leq 1 \end{cases} \tag{1}$$

satisfies the above requirements.

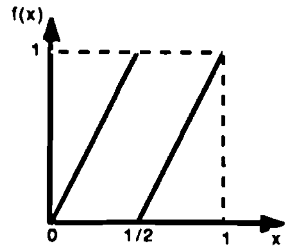


figure 1

There is another way to describe this model which is often used, called *symbolic dynamics* (see [5], [6]). By this term it is meant a representation of f in terms of a shift on a set of binary sequences. Define the set of binary sequences

$$\Sigma_2 := \{ s = (s_0 s_1 s_2 ...) \mid s_j = 0 \text{ or } 1 \}$$

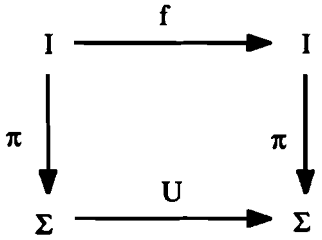
endowed with the following metric:

$$d(s,t) = \sum_{i=0}^{\infty} \frac{| s_i - t_i |}{2^i}$$

On Σ_2 define a shift as follows

$$U(s_0 s_1 s_2 ...) = (s_1 s_2 s_3 ...)$$

Denoting by π the map from I onto Σ_2 which associates to a real number its nonterminating binary representation, it is clear that the following diagram commutes:



By *dynamical system* we mean here a mathematical model evolving in time whose trajectories $\{x(t), y(t)\}_{t\in N}$ (behaviours, see [7]) admit a representation through a pair of functions (F,H) on a suitable domain

$$\begin{cases} x(t+1) & = F(x(t)) \\ y(t) & = H(x(t)) \end{cases} \tag{2}$$

The variables x and y are called state and observation of the system. The above system is said to be *observable at time t* if there exists an injection form the range of x(t) into the cartesian product of the observation up to time t $\{y(1), y(2),..., y(t)\}$. The system is *observable* if there exists a t_0 (possibly infinite) such that the system is observable at each $t \geq t_0$. Since F is a deterministic function, in this context, the system is observable as soon as the initial condition can be determined exactly. More generally, even for an unobservable system, we shall say that an initial condition *x is observable* if it is uniquely determined by the observation of the corresponding trajectory.

Consider now a particular example of (2), where F is the dyadic map (1) and H is the following two state observation function

$$h_0(x) = \begin{cases} 0 & \text{for } 0 \leq x < 1/2 \\ 1 & \text{for } 1/2 \leq x \leq 1 \end{cases} \tag{3}$$

We get the following dynamical system

$$(S_0) \quad \begin{cases} x(t+1) = f(x(t)) \\ y(t) = h_0(x(t)) \end{cases} \tag{4}$$

It is easy to see that this model is observable over an infinite interval of time, i.e. the initial condition (and hence the whole trajectory) is determined uniquely by the infinite observation of the system (to this end take the binary representation of an initial state x_0: this will coincide with the history of the observations). One reason why this model is so interesting is that, in spite of its complete observability, any observation over a finite time interval is indistinguishable from the outcome of a coin tossing (see[6]). Another reason is that it introduces the *symbolic dynamics* in very natural way. In fact, the history of the observations of $f^n(x)$ under h is precisely $\pi(x)$. Observability here depends on the fact that the inverse images under h_0 of the states 0 and 1 coincide exactly with the two intervals [0,1/2), [1/2,1) (these are called Markov partitions of f, see [6]). For this reason we call this observation exact. The problem we want to consider now is the following: suppose our observation function does not distinguish exactly between the two intervals, but contains some error $\varepsilon > 0$ as follows:

$$h_\varepsilon(x) = \begin{cases} 0 & \text{for } 0 \le x < 1/2 + \varepsilon \\ 1 & \text{for } 1/2 + \varepsilon \le x \le 1 \end{cases} \tag{5}$$

It can be seen that, for example, the initial conditions x and x + 1/2 are indistinguishable for $x \in [0,\varepsilon)$. We give a precise characterization of this concept in theorem 1 below. However, if $\varepsilon < \eta$, also h_ε yields a better observation than h_η in the sense that more points are distinguishable.The questions we try to answer in this paper are the following:

Does there exist a way to define observability of S_ε such that:

a) this definition generalizes the usual observability concept for dynamical systems.
b) the function which measures observability of S_ε is decreasing in ε.

It should be noted that the first thing one would try, namely the measure of the observable set, fails, as shown in Proposition 1 below.

We will study the special case $\varepsilon = 2^{-n}$, and we write, by abuse of notation, with S_n, h_n, instead of $S_{2^{-n}}$, $h_{2^{-n}}$. We define Ω_n to be the set of observable points of S_n. By the notation $\underbrace{0...0}_{k}$ we mean a sequence of exactly k zeros.

Theorem 1: *Ω_n is the set of points of I whose binary representation has the following properties:*

a) the sequence $01\underbrace{0...0}_{k}1$ never occurs for $k > n-2$
b) the sequence $11\underbrace{0...0}_{k}1$ never occurs for $k > n-1$.

For the proof we need two lemmas. Define by Y_n the set of trajectories of S_n:

$$Y_n := \{s \in \Sigma_2 : s = \{h_n(f^t(x))\}_{t\in N} \text{ for some } x \in [0,1)\}$$

Lemma 1: *the set Y_n consists exactly of the points of Σ_2 in which the sequence $\{1\underbrace{0...0}_{n-1}1\}$ never appears.*

Proof: if $\pi(x)$ never has not more than n-2 zeros in a row, then $h(f^t(x)) = h_n(f^t(x))$, and the number of zeros is preserved. If $\pi(x)$ has more than n-1 zeros, the 1 preceeding the zeros is set to zero by h_n, and so the output sequence will have at least n zeros. Therefore the sequence $1\underbrace{0...0}_{n-1}1$ can never occur as the output of S_n.

Lemma 2: *the sequences of Y_n generated by observable points are those which have at most n consecutive zeros.*

Proof: if a sequence $s = h_n(f^t(x))\}_{t\in N}$ has less than n consecutive zeros, then in view of Lemma 1 it has at most n-2 zeros and thus $h_n(f^t(x)) = h(f^t(x))$. If s has exactly n zeros, than it is seen by inspection that it can only be generated by a sequence $\{11\underbrace{0...0}_{n-1}1\}$. If s has r consecutive zeros, r > n, then both the element

$$t_1 = \{11\underbrace{0...0}_{r-1}1\}$$
$$t_2 = \{101\underbrace{0...0}_{r-2}1\}$$

yield the same output and thus the trajectory does not determine uniquely the initial point. ■

Proof of theorem 1: to characterize the observable points of [0,1), observe first that a point x for which $\pi(x)$ has a subsequence of n or more consecutive zeros is not observable. In fact, in this case, $s = h_n(f^t(x))\}_{t\in N}$ will have at least n+1 zeros, and in view of Lemma 2 this trajectory is generated by more than one point. If there are less than n-1 consecutive zeros, then $h_n(f^t(x)) = h(f^t(x))$ for all t, and the point is observable. If a subsequence with exactly n-1 zeros occurs, there are two possibilities:

a) the subsequence is of the form $01\underbrace{0...0}_{n-1}1$ and it is thus indistinguishable either from

$1\underbrace{0...0}_{n}1$ (if there are less than n-1 zeros before the preceeding 1) or from $\underbrace{0...0}_{n+1}1$ (if there are more than n-2 zeros before the preceeding 1).

b) the subsequence is of the form $11\underbrace{0...0}_{n-1}1$. Then the image of the subsequence is $1\underbrace{0...0}_{n}1$, which in view of lemma 2 comes from an observable point. ■

Corollary: $\Omega_n \subset \Omega_{n+1}$. $\quad 11\underbrace{0...0}_{n-1}1$

Proposition 1: *Let Ω_n be the set of observable points for S_n Then,*

$$\mu(\Omega_n) = 0 \tag{7}$$

where μ denotes the Lebesgue measure.

Proof: from theorem 1, Ω_n is the set of points x such that in $\pi(x)$ some sequences never occur. In view of the Borel-Cantelli lemma the measure of this set is zero. ■

This proposition says that the system S_0 is very special with respect to the Lebesgue measure, because it is the only one whose observable set has measure one.

It turns out that a reasonable tool to characterize the magnitude of the observable set is the Hausdorff dimension, as we show below.

§3. Main result

We are going now to define the Hausdorff dimension of a metric space X. The diameter of a set U of X, is defined as

$$\mathrm{diam}(U) = \sup \{d(x,y) : x,y \in U\}$$

Given $\delta > 0$ we denote by $\mathcal{U}_\delta$ a cover of X such that diam (U) < δ for all U in $\mathcal{U}_\delta$.

Definition: *the Hausdorff dimension of a metric space X is*

$$HD(X) = \inf\left\{\alpha: \forall \varepsilon > 0, \exists \text{ a cover } \mathcal{U}_\delta \text{ of } X \text{ s.t. } \sum_{U \in \mathcal{U}_\delta} (\mathrm{diam}\,(U)^\alpha) < \varepsilon\right\} \quad (8)$$

The Hausdorff dimension has several interesting properties (see [3]) :

$$HD\,(X) \leq HD\,(X') \quad \text{if} \quad X \subset X' \qquad (9a)$$

$$HD\,(X) = n \qquad \text{for } X \subset \mathbb{R}^n \text{ if } \mu(X) > 0 \qquad (9b)$$

$$HD\,(\cup_n X_n) = \sup HD\,(X_n) \qquad (9c)$$

The Hausdorff dimension is equal to the usual dimension in the case of a linear space or of a smooth manifold. As a consequence, we have the following:

Proposition 2: *let S be a linear dynamical system with observable space of dimension n. Then also the Hausdorff dimension of this space is n.*

So the Hausdorff dimension is equivalent to the usual one in all classical cases. In general, though, it is a rather difficult object to compute whenever it does not coincide with the usual notion of dimension. Its interest for our application lies in the fact that S_n is not a classical case, but $HD(\Omega_n)$ is still quite easy to compute. Denote, as above, by Ω_n the observable set of S_n

We are now ready to compute the Hausdorff dimension of Ω_n.

Theorem 2: $HD\,(\Omega_n) = \dfrac{2^{n+1} - 6}{2^{n+1} - 3}$

Proof: we need first the following result (see [3], Theorem 14.1). Let $u_k(x)$ be the subinterval of I of the form $\left[\frac{j}{2^k}, \frac{j+1}{2^k}\right)$ containing x. Let ν be a probability measures on I such that $\nu(X) = 1$, and let μ denote the Lebesgue measure. If

$$X \subset \left\{x : \lim_{k \to \infty} \frac{\log \nu(u_k(x))}{\log \mu(u_k(x))} = \delta\right\} \qquad (10)$$

then HD (X) = δ. The problem thus becomes to construct the measure ν on the set Ω_n. A standard procedure (see [3], p.143), is to use the measure induced by a Markov chain whose trajectories belong almost surely to Ω_n. Denoting by p_{ij} the transition probabilities and by p_i the invariant measure, it is easily seen that

$$\lim_{k \to \infty} \frac{\log \nu(u_k(x))}{\log \mu(u_k(x))} = -\frac{1}{\log 2} \sum_{i,j=1}^{n} p_i p_{ij} \log p_{ij} \tag{11}$$

In our case, Ω_n is the set of all numbers whose binary expansion never has the sequences $01\underbrace{0...0}_{k}1$ for k >n-2 and the sequence $11\underbrace{0...0}_{k}1$ for k >n-1. We now construct the Markov chain z(t) as follows: if the first two digits of x(t) are 01 followed by i-1 zeros, set z(t) = i for i = {1,..., n-1}. If x(t) terminates with exactly i zeros preceeded by 11 then set z(t) = n+i. it is easy to see that z(t) has transition probabilities

$$[p_{ij}] = \begin{cases} 1/2 & \text{for } j=i+1, i\neq n-1,\ 2n-1 \\ & \text{for } i=j=n \text{ and } i=1,\ j=n \\ & \text{for } i=2,\dots,n-2, n+1,\dots,2n-2 \text{ and } j=1 \\ 1 & \text{for } i=n-1,\ i=2n-1, j=1 \text{ and } j=1 \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

The invariant measure for $[p_{ij}]$ is seen to be

$$p_1 = p_n = \frac{2^{n-1}}{2^{n+1}-3}$$

$$p_i = \frac{2^{n-i}}{2^{n+1}-3} \qquad i = 2,\dots, n-1$$

$$p_i = \frac{2^{2n-i-1}}{2^{n+1}-3} \qquad i = n+1,\dots, 2n-1$$

a simple substitution in (11) yields the result.

We still need to justify the choice of (12). It is easy to see that, when we condition on $\{x \in \Omega_n\}$, the probability measure induced by the Lebesgue measure is exactly the measure induced by the Markov chain (12). To see that this conditional probability is indeed the one with support Ω_n, we refer the reader to the original paper of Billingsley [2]. ■

Another and perheaps more natural way to look at the observability problem is the one concerned with the set of possible output trajectories, Y_n. This set of binary strings can be inbedded in [0,1) in an obvious way, and we can thus define, with abuse of notation, the Hausdorff dimension of Y_n. In a fashion completely similar to that of theorem 2 we can prove the following

Theorem 3: the Hausdorff dimension of Y_n *is* $\frac{2^{n+1}}{2^{n+1}+1}$

We would like to remark that the dimensions computed in theorems 2 and 3 converge to 1 as n goes to infinity, yielding thus that consistency which was seeked in the beginning of the paper.

§4. Conclusions.

We have presented an example where the definition of dimension of the observable subspace of a dynamical system is extended to the case of noninteger numbers. We conjecture that this procedure can be generalized to a system of the form (2) whenever the function F admits a Markov partition on its domain and H takes only finite values. This problem is currently being investigated by the authors.

Acknowledgments: it is a pleasure to thank Professor Tom Taylor for drawing our attention on this subject, and for the useful discussions we had together.

This research took parly place while the second author was visiting the Division of Systems and Decision Sciences, IIASA, Laxenburg, Austria.

References

[1] Aeyels D., Generic Observability of Differentiable Systems, *SIAM J. Control and Optimization,* Vol 19, 1981, 595-603.

[2] Billingsley P., Hausdorff dimension in probability theory II, *Illinois Journal of Mathematics,* **5**,1961, 291-298.

[3] Billingsley P., *Ergodic theory and information,* Wiley, 1965.

[4] Drager L., Martin C., Global observability of a class of nonlinear discrete time systems, *Systems and Control Letters,* **6**, 1985,65-68.

[5] Devaney R.L., *Chaotic Dynamical systems,* Benjamin/Cummings 1986.

[6] Guckenheimer J., Holmes P., *Nonlinear Oscillations, Dynamical Systems, and Bifurcation of Vector Fields,* Springer-Verlag, 1983.

[7] Willems J.C., From time series to linear systems, Part I. Approximate modelling. *Automatica,* Vol 22, N.5, 1986, 561-580.

[8] Taylor T.J.S., On observation of chaotic systems and randomness, Proc. MNTS 1987.