# State Predictability and Information Flow in Simple Chaotic Systems

B. Pompe, J. Kruscha, and R. W. Leven
Sektion Physik/Elektronik der Ernst-Moritz-Arndt-Universität, Greifswald, DDR

Z. Naturforsch. **41 a,** 801 – 818 (1986); received November 27, 1985

The possibility of state prediction in deterministic chaotic systems, which are described by 1-D maps, is discussed in the light of information theory. A quantity $h(l)$ is defined which represents the production of uncertainty on a future state by the chaotic dynamics (intrinsic noise) after $l$ time steps have passed. $h(l)$ is related to the Lyapunov characteristic exponent. Moreover, the influence of the measuring process (overlappings of mapped boxes of state space partition) and external noise on the state predictability are investigated quantitatively.

## 1. Introduction

In recent years one of the most intensively studied subjects in physics, chemistry, and other fields have been deterministic, dissipative, dynamical systems displaying a so-called "chaotic" behaviour. Typically, two neighbouring orbits of a chaotic system diverge on an average exponentially as time increases, which is indicated by a positive Lyapunov characteristic exponent (LCE). Thus, if one has no exact knowledge on the initial state of the system (practically this is always the case due to the limited precision of any measurement), a predicted state could differ enormously from the actually observed one, i.e., long range state predictions are impossible.

The time evolution of the states of a deterministic system is usually governed by a set of first-order differential equations or in a discrete time version by a set of difference equations. If they are suitably "well-behaved", various theorems can be applied guaranteeing the uniqueness of any orbit starting from a given area of state space. However, any real system interacts with its environment in a complicated manner. Consequently, deterministic equations have to be modified by stochastic terms in order to describe reality. Chaotic orbits sensitively respond to external fluctuations. Hence, state predictions would be limited, even if an initial state was known exactly.

In this paper we are first of all interested in the effect of a limited knowledge of an initial state on the possibility of state prediction. This problem was investigated already by several authors [1]−[7]. For a detailed investigation Farmer [3] used the quantity $I_t$ representing the (average) information on a future state contained in an initial state. $I_t$ is a well-known quantity in information theory. Some denotations of $I_t$ are "mutual information", "syn-entropy", and "transinformation". (We prefer the latter denotation.) The definition of $I_t$ for a dynamical system requires a partition of state space which is assumed to be induced by a limited precision of measurement and a certain gauging of the (digital) measuring equipment. Thus, $I_t$ characterizes the symbolic system which is induced by a partition of state space (see e.g. [4−6]). Consequently, an investigation of $I_t$ is justified from an experimental point of view. However, in order to characterize the "uncertainty production" of a system itself (and not that of the symbolic system only) one should also try to investigate quantities which are defined independent of a partition of state space. In his paper [1] Shaw has done some stimulating work in this field. He argued that the uncertainty on a future state after one time step is given for simple systems by the LCE. It is the purpose of this paper to make a contribution to this discussion. In particular, we will demonstrate that some of Shaw's concepts do not properly reflect reality as they break down even for simple systems. We argue that uncertainty on a future state is produced in expanding regions of state space only. In contracting regions uncertainty is neither produced (if we neglect "overlappings") nor is uncertainty, which is simultaneously produced in expanding regions,

Reprint requests to Prof. R. W. Leven, Sektion Physik/Elektronik der Ernst-Moritz-Arndt-Universität, Domstraße 10a, DDR-2200 Greifswald.

0340-4811 / 86 / 0600-0801 $ 01.30/0. – Please order a reprint rather than making your own copy.

reduced. From this point of view the LCE is found to be a lower bound of the uncertainty production after one time step.

In this paper we restrict our investigations to 1-D maps. Nevertheless, our ideas can be applied to $s$-D maps ($s = 2, 3, 4, \ldots$) and sets of differential equations as well. The paper is organized as follows: In Sect. 2 we give a review of the definition of the transinformation $I_t$ for dynamical systems. After this we introduce a quantity $h(l)$ which is related to the LCE, and which gives the uncertainty on a future state, if an initial state is known ($l = 0, 1, 2, \ldots$ labels the prediction period). $h(l)$ is defined without partitioning state space, hence $h(l)$ characterizes the system itself. In order to illustrate our ideas, we investigate the so-called "$M$-map" in Section 3. Uncertainty $H_c(l)$ on a future state of a symbolic system may also arise from overlappings of mapped boxes (i.e., divisional points of the partition of state space are not mapped on divisional points). Hence a relation between $h(l)$ defined in Sect. 2 and $H_c(l)$ is of interest. In Sect. 4 we take this fact into account and give lower and upper bounds of $H_c(l)$ which are related to $h(l)$. Moreover, the effect of overlappings is investigated in the limit of an infinitely fine partition. As an illustration some results for the well-known logistic equation ($\alpha = 4$: fully developed chaos) are presented. Section 5 is devoted to an elementary investigation of the effect of external noise on the uncertainty production of the dynamical system. In the appendix we present results for several families of piecewise linear 1-D maps in order to give further illustrations of our main ideas.

## 2. Uncertainty Production of 1-D Maps $z_{n+1} = f(z_n)$

In order to explain our motivation for the definition of a new quantity $h(l)$ describing the uncertainty on a future state, reached $l$ iterations in the future, under the condition that the initial state is known, we start with a review of the definition of the transinformation $I_t$ for dynamical systems.

### *2.1. Transinformation*

As already mentioned above, the definition of $I_t$ requires a partition

$$\mathscr{B} := \{B_\mu\}_{\mu=1}^{M}, \quad B_\nu \cap B_\mu = \emptyset \quad \text{if} \quad \nu \neq \mu,$$

of the domain containing the orbit of the system. This domain is assumed to be bounded. Consequently, $\mathscr{B}$ consists of a finite number, say $M$, of bounded boxes. Throughout this paper we only consider so-called $\varepsilon$-partitions. An $\varepsilon$-partition $\mathscr{B}_\varepsilon$ of the interval $[0, 1[$ we define as follows:

$$\mathscr{B}_\varepsilon := \{B_\mu\}_{\mu=1}^{M}, \quad B_\mu := [(\mu - 1)\,\varepsilon, \mu\,\varepsilon[, \quad \mu = 1, 2, \ldots, M,$$

where $M \equiv 1/\varepsilon$ is a positive integer. Thus, each box of an $\varepsilon$-partition is assumed to be of equal size. (For shortness we sometimes say "partition" instead of "$\varepsilon$-partition".) From a physical point of view the partition may be induced by a (digital) measuring apparatus with a finite precision $1/\varepsilon$ of measurement and a certain gauging corresponding to the position of the partition. The measuring apparatus allows us to determine the box $B_{\mu(n)}$ containing the orbit $\{z_n\}_{n=0}^{+\infty}$ of the system at every time step $n$. Hence the boxes can be characterized as states of the system.

Now, let $p_\mu(n)$ label the probability that the orbit visits the box $B_\mu$ at time $n$. Thus a measurement at time $n$ provides the (average) information

$$H(n) = \sum_{\mu=1}^{M} p_\mu(n) \,\mathrm{ld}\, 1/p_\mu(n), \quad (\mathrm{ld} \equiv \log_2), \tag{1.1}$$

using the well-known formula by Shannon. In the following we distinguish between the expressions "information" and "uncertainty": We say that we have the information $H(n)$ about the state $B_{\mu(n)}$ of the system, if we have determined it as a result of a single measurement at time $n$. On the other hand, if no measurement was carried out, we say that $H(n)$ gives the uncertainty about the state $B_{\mu(n)}$.

Let us assume now that we have measured $B_{\mu(n_0)}$ at the initial time $n_0$. Thus we have obtained the information $H(n_0)$. An interesting question is now whether $H(n_0)$ contains information about the future state $B_{\mu(n_0+l)}$, $l = 0, 1, 2, \ldots$. In order to give an answer to this essential question, we introduce the joint probabilities $p_{\nu\mu}(n_0, l)$ that the orbit is found in $B_\mu$ at time $n_0$ and $l$ time steps later in $B_\nu$, $\nu, \mu = 1, 2, \ldots, M$; $l = 0, 1, 2, \ldots$. Hence the joint uncertainty $H(n_0, l)$ of a certain combination $(B_{\mu(n_0)}, B_{\mu(n_0+l)})$ is given by

$$H(n_0, l) = \sum_{\nu,\mu=1}^{M} p_{\nu\mu}(n_0, l) \,\mathrm{ld}\, 1/p_{\nu\mu}(n_0, l).$$

Of course, $H(n_0, l)$ also gives the information obtainable in two measurements at times $n_0$ and $n_0 + l$. If we have only the information $H(n_0)$ from one measurement at time $n_0$, there will remain the uncertainty

$$H_c(n_0, l) := H(n_0, l) - H(n_0) \quad (2.1)$$

about the future state $B_{\mu(n_0+l)}$. Now we introduce the conditional probabilities

$$p_{\nu/\mu}(n_0, l) := \begin{cases} p_{\nu\mu}(n_0, l)/p_\mu(n_0) & \text{if } p_\mu(n_0) \neq 0 \\ 0 & \text{otherwise,} \end{cases}$$
$$\nu, \mu = 1, 2, \ldots, M$$

which gives the probabilities for the transitions from $B_\mu$ at time $n_0$ to $B_\nu$ at time $n_0 + l$. The uncertainty $H_c(n_0, l)$ defined in (2.1) can now be written as follows:

$$H_c(n_0, l) = \sum_{\mu=1}^{M} p_\mu(n_0)\, H_c(n_0, l, \mu) \quad (2.2)$$

with

$$H_c(n_0, l, \mu) := \sum_{\nu=1}^{M} p_{\nu/\mu}(n_0, l) \operatorname{ld} 1/p_{\nu/\mu}(n_0\ l). \quad (3.1)$$

From well-known relations of information theory (see e.g. [8]) we know that

$$H_c(n_0, l) \leqq H(n_0 + l), \quad (4)$$

i.e., the uncertainty about the future state $B_{\mu(n_0+l)}$ can only be decreased by a measurement at time $n_0$. The equality in (4) holds if the states $B_{\mu(n_0)}$ and $B_{\mu(n_0+1)}$ are statistically independent ($p_{\nu/\mu}(n_0, l) = p_\nu(n_0 + l)$). $H_c(n_0, l)$ is called the conditional uncertainty on a future state $B_{\mu(n_0+l)}$ under the condition that an initial state $B_{\mu(n_0)}$ is known. Shaw [1] refers to this quantity as "information production", i.e. $H_c(n_0, l)$ gives the "new" information (in addition to the "old" information $H(n_0)$) which is obtainable in a second measurement at time $n_0 + l$. From our point of view $H_c(n_0, l)$ should be called "uncertainty production" because it gives the uncertainty on the future state $B_{\mu(n_0+l)}$ which remains, if only the initial state was measured. From (4) follows that

$$I_t(n_0, l) := H(n_0 + l) - H_c(n_0, l) \quad (5.1)$$

is always nonnegative. $I_t(n_0, l)$ is called "transinformation" because it gives the information on the future state $B_{\mu(n_0+l)}$ which is obtained from a measurement of the initial state $B_{\mu(n_0)}$. With other words: $I_t(n_0, l)$ gives the decrease of the uncertainty on the future state by measuring the initial one. $I_t(n_0, l)$ equals zero if these states are statistically independent, i.e. if state predictions are completely impossible. Otherwise it is positive and reaches its maximal value $H(n_0 + l)$ if there remains no uncertainty about $B_{\mu(n_0+l)}$ due to the measurement of $B_{\mu(n_0)}$.

Now we reverse the question for the possibility of state prediction and ask for the information about a former state $B_{\mu(n_0)}$ under the condition that a present state $B_{\mu(n_0+l)}$ is known, i.e., we now ask whether a dynamical system is able to remember its previous state. Obviously the uncertainty about a former state $B_{\mu(n_0)}$, under the condition that a present state $B_{\mu(n_0+l)}$ was measured, is given, in analogy to (2.1), by

$$H_c(n_0, -l) := H(n_0, l) - H(n_0 + l).$$

Hence, in analogy to (5.1),

$$I_t(n_0, -l) := H(n_0) - H_c(n_0, -l)$$

gives the transformation on the former state $B_{\mu(n_0)}$, if the present state $B_{\mu(n_0+l)}$ is known. A simple analysis proves that $I_t(n_0, -l)$ equals $I_t(n_0, l)$. Consequently, $I_t(n_0, l)$ can be interpreted formally in the same way as $I_t(n_0, -l)$. More specifically, we say that the system has "forgotten" its initial state $B_{\mu(n_0)}$ after $l$ time steps, if $I_t(n_0, l)$ equals zero.

So far we have considered processes which are not of necessity stationary. Nevertheless, in the following we consider only stationary processes. Hence all quantities defined above are now assumed to be independent of the initial time $n_0$ which is equalized to zero. With other words: Transients are supposed to have died out such that the orbit of the (dissipative) system has reached an attractor $A$, which is assumed to be bounded. We further assume to have an ergodic invariant ("natural") measure $\bar{\mu}$ (which is defined on a suitable $\sigma$-algebra of subsets of $A$) such that

$$\bar{\mu}(B_\mu) = p_\mu, \quad \mu = 1, 2, \ldots, M, \quad (6)$$

gives the probability that an orbit visits the box $B_\mu$ at any time $n$. In the following we say "almost everywhere" and "almost every" with respect to $\bar{\mu}$. Moreover, $\bar{\mu}$ is assumed to be defined by an invariant density $\varrho(z)$ (for more details of the definition of a "natural" or "physical" measure see e.g. [9]). Let $f^{-l}(B_\nu)$ label the set which is mapped on $B_\nu$, if the 1-D map acts $l$ times on it:

$$f^{-l}(B_\nu) := \{z \in A \mid f^l(z) \in B_\nu\}, \quad \nu = 1, 2, \ldots, M,$$
$$(f^l \equiv f \circ f^{l-1}, l = 1, 2, 3, \ldots, f^0 \equiv 1).$$

Hence

$$\frac{\bar{\mu}(f^{-l}(B_\nu) \cap B_\mu)}{\bar{\mu}(B_\mu)} = p_{\nu/\mu}(l), \quad \nu, \mu = 1, 2, \ldots, M \tag{7}$$

is the probability for the transition $B_\mu \to B_\nu$ after $l$ time steps. Because of the stationary situation we have $H(n_0 + l) = H(n_0) \equiv H_\mathrm{i}$ (i refers to "initial"). The transformation defined in (5.1) can now be written as follows:

$$I_\mathrm{t}(l) = H_\mathrm{i} - H_\mathrm{c}(l) \tag{5.2}$$

with

$$H_\mathrm{i} = \sum_{\mu=1}^{M} p_\mu \,\mathrm{ld}\, 1/p_\mu\,, \tag{1.2}$$

$$H_\mathrm{c}(l) = \sum_{\mu=1}^{M} p_\mu H_\mathrm{c}(l,\mu), \quad \text{and} \tag{2.3}$$

$$H_\mathrm{c}(l,\mu) = \sum_{\nu=1}^{M} p_{\nu/\mu}(l) \,\mathrm{ld}\, 1/p_{\nu/\mu}(l)\,. \tag{3.2}$$

Shaw [1] argues that the uncertainty $H_\mathrm{c}(l)$ on a future state is given for simple cases by $l \cdot$ LCE. Especially, he argues that the system has forgotten the initial state (and thus state predictions have become impossible), if $l \cdot$ LCE equals approximately $H_\mathrm{i}$, i.e., if the produced new information (uncertainty) has "pushed away" the initial information. In the following we present a refinement of the argumentation of Shaw for simple 1-D maps.

## *2.2. Uncertainty Production and Lyapunov Characteristic Exponent*

Now we give a heuristic argumentation for a substitution of the uncertainty production $H_\mathrm{c}(l)$ by a new quantity $h(l)$ which will be defined independent of a partition of state space. However, we start with an $\varepsilon$-partition such that the slope $\mathrm{d}f/\mathrm{d}z$ and the invariant density $\varrho$ are nearly constant in most boxes. For a piecewise $C^2$-function $f$ and a piecewise $C^1$-function $\varrho$ this will be attained if $\varepsilon\,|\mathrm{d}^2 f/\mathrm{d}z^2| \ll 1$ and $\varepsilon\,|\mathrm{d}\varrho/\mathrm{d}z| \ll 1$ almost everywhere in [0, 1], respectively, i.e., the $\varepsilon$-partition has to be fine enough. Moreover, if these assumptions do not hold in a few boxes, the error we make is expected to be small, if these "exceptional" boxes have a small measure $\bar{\mu}$ such that they do not essentially contribute to the uncertainty production. In the following $s_\mu := |\mathrm{d}f/\mathrm{d}z|_{z \in B_\mu}$, $\mu = 1, 2, \ldots, M$, labels the absolute value of the slope in the box $B_\mu$. In general, divisional points of the partition are not mapped on divisional points by $f$. Thus, overlappings of mapped boxes have to be taken into account. As will be seen later (Sect. 4), overlappings increase the uncertainty on a future state. In order to obtain a lower bound of $H_\mathrm{c}(l)$, we neglect the overlappings at this stage.

Now consider a box $B_\mu$ with a slope $s_\mu$ greater than one. Thus, $s_\mu$ gives the approximate number of transition probabilities $p_{\nu_i/\mu}(1)$, $i = 1, 2, \ldots, \mathrm{int}(s_\mu)$, which equal approximately $1/s_\mu$, and the rest of them equal zero. Consequently, a lower bound of $H_\mathrm{c}(1,\mu)$ is given by

$$\sum_{i=1}^{\mathrm{int}(s_\mu)} p_{\nu_i/\mu}(1) \,\mathrm{ld}\, 1/p_{\nu_i/\mu}(1) \approx \mathrm{ld}\, s_\mu \quad \text{(cf. (3.2))}.$$

On the other hand, if $f$ contracts the box $B_\mu$ (i.e., $s_\mu \leqq 1$), one of the transition probabilities equals one, while the rest of them equals zero. Consequently, $H_\mathrm{c}(1,\mu)$ equals zero in this case, i.e., there is no uncertainty produced on the future state. A contracted box $B_\mu$ is not resolvable by the measuring apparatus, assuming a fixed finite measuring precision, but the information on the position of $f(B_\mu)$ is still known due to the knowledge of the position of $B_\mu$ and the deterministic action of $f$ on $B_\mu$ (see Figure 1). Thus we are motivated to define the average uncertainty $h(1)$ on a future state after one time step as follows:

$$h(1) := \int_{-\infty}^{+\infty} \Theta(|\mathrm{d}f/\mathrm{d}z| - 1)\,\mathrm{ld}\,|\mathrm{d}f/\mathrm{d}z|\,\varrho(z)\,\mathrm{d}z$$
$$\equiv \int \Theta(|\mathrm{d}f/\mathrm{d}z| - 1)\,\mathrm{ld}\,|\mathrm{d}f/\mathrm{d}z|\,\mathrm{d}\bar{\mu}, \tag{8.1}$$

with

$$\Theta(z) \equiv \begin{cases} 1, & z > 0 \\ 0, & \text{otherwise.} \end{cases}$$

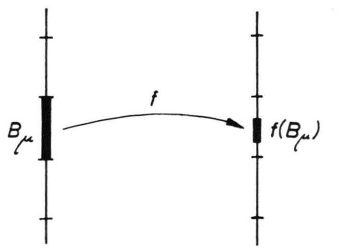


Fig. 1. Schematic representation of a contracting action of the map $f$ without overlappings, i.e. $f(B_\mu)$ contains no divisional point of the $\varepsilon$-partition.

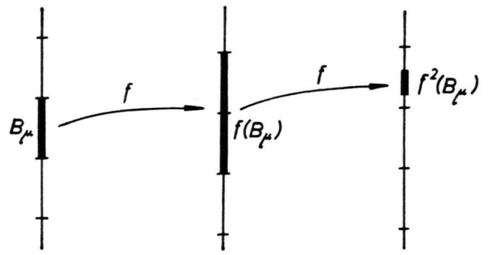


Fig. 2. Schematic illustration of a contracting action of the second iterate of $f$. The uncertainty produced by the first expanding action of $f$ is vanishing in the second step.

Here the sum over $\mu$ (cf. (2.3)) is replaced by an integral in order to obtain an average over all possible initial states. Note that the integral in (8.1) is taken over that part where the absolute value of the slope of $f$ is greater than one, because uncertainty on a future state is produced only in expanded intervals. If $f$ acts expanding almost everywhere on supp $\varrho(z) = A$, (8.1) reduces to the well-known formula for the LCE:

$$\lambda := \int \mathrm{ld}\,|\mathrm{d}f/\mathrm{d}z|\,\mathrm{d}\bar{\mu}. \tag{9.1}$$

Thus, in general we obtain

$$\lambda \leqq h(1) \lesssim H_c(1).$$

$\lambda$ is a lower bound of the uncertainty on a future state $B_{\mu(1)}$.

Now we investigate a lower bound $h(2)$ of the uncertainty $H_c(2)$ on a future state $B_{\mu(2)}$, if an initial state $B_{\mu(0)}$ is known. Let $B_\mu$ be a box which is expanded by the action of $f$ (first iteration: $s_\mu > 1$), as illustrated in Figure 2. Thus the uncertainty ld $s_\mu$ is produced in the first step. Now it may happen that $f$ contracts $f(B_\mu)$ such that in total $f^2$ has a contracting action on $B_\mu$, i.e., $|\mathrm{d}f^2/\mathrm{d}z| \leqq 1$ for $z \in B_\mu$. In this case after the second iteration no uncertainty is produced with reference to the initial state. The uncertainty produced by the first expanding action of $f$ may vanish in the second step. Note the difference to the statement above regarding the uncertainty production after one time step: Contraction cannot effect a reduction of uncertainty which is simultaneously produced in expanded regions, but contraction can reduce uncertainty which arose from the expansion in a former time step. On the other hand, if $f^2$ has an expanding action on $B_\mu$, uncertainty is produced after two iterations. Thus, the expanding or contracting character of the action of $f^2$ determines whether uncertainty on the future state $B_{\mu(2)}$ is produced or not. These considerations can be extended to include prediction periods $l > 2$. Consequently we are inclined to define

$$h(l) := \int \Theta(|\mathrm{d}f^l/\mathrm{d}z| - 1)\,\mathrm{ld}\,|\mathrm{d}f^l/\mathrm{d}z|\,\mathrm{d}\bar{\mu}, \tag{10.1}$$

which is expected to be a lower bound of the uncertainty $H_c(l)$ (see (2.3)) if the assumptions given above hold for $f^l$.

We have assumed ergodicity, hence $\lambda$ can be determined operationally from a very (exactly: infinitely) long time series $\{z_n = f^n(z_0)\}_{n=0}^{m}$:

$$\begin{aligned}\lambda &= \lim_{m\to+\infty} 1/m \sum_{n=0}^{m-1} \mathrm{ld}\,|\mathrm{d}f/\mathrm{d}z|_{z_n}| \\ &= \lim_{m\to+\infty} 1/m\,\mathrm{ld}\,|\mathrm{d}f^m/\mathrm{d}z|_{z_0}|,\end{aligned} \tag{9.2}$$

for almost every $z_0 \in A$.

In analogy, $h(l)$ can be obtained operationally too:

$$\begin{aligned}h(l) = \lim_{m\to+\infty} \frac{1}{m} \sum_{n=0}^{m-1} \Theta(|\mathrm{d}f^l/\mathrm{d}z|_{z_n}| - 1) \\ \cdot\,\mathrm{ld}\,|\mathrm{d}f^l/\mathrm{d}z|_{z_n}|.\end{aligned} \tag{10.2}$$

it follows from the chain rule applied to $\mathrm{d}f^l/\mathrm{d}z$ that (10.1) can be transformed as follows:

$$h(l) = l \cdot \lambda + \delta(l), \tag{10.3}$$

with

$$\delta(l) := \int \Theta(1 - |\mathrm{d}f^l/\mathrm{d}z|)\,\mathrm{ld}\,1/|\mathrm{d}f^l/\mathrm{d}z|\,\mathrm{d}\bar{\mu}. \tag{11}$$

$\delta(l)$ is a correction term for the formula proposed by Shaw [1] and fulfils the relation $\delta(l) \geqq 0$. The equality holds if $f^l$ has an expanding action ($|\mathrm{d}f^l/\mathrm{d}z| \geqq 1$) almost everywhere in $A$. Thus we have, in general, the relations

$$l \cdot \lambda \leqq h(l) \lesssim H_c(l), \quad l = 1, 2, 3, \ldots. \tag{12}$$

Due to the supposed ergodic chaotic dynamics almost every orbit which starts on the attractor $A$ generates a "typical" orbit such that the limit $\lambda$ of (9.2) is independent of almost every initial value $z_0$. Consequently, the sequence of functions in $z_0$

$$\{1/m\,\mathrm{ld}\,|\mathrm{d}f^m/\mathrm{d}z|_{z_0}|\}_{m=1}^{+\infty} \tag{13}$$

is pointwise convergent for almost every $z_0 \in A$. The pointwise limit function is given by $\lambda = \text{constant} > 0$. If we even assume a uniform convergence of the sequence (13) almost everywhere in $A$, there is a least positive integer $l^*$ which is independent of $z_0$ and which guarantees that for $l > l^*$ the slope $|df^l/dz|$ is greater than one for almost every $z_0 \in A$. Hence, from (11) follows

$$\delta(l) = 0 \quad \text{for} \quad l > l^*,$$

and (10.3) reduces to that one proposed by Shaw [1]:

$$h(l) = l \cdot \lambda \quad \text{for} \quad l > l^*.$$

Nevertheless, in general the correction term $\delta(l)$ is of importance, especially for small values of the prediction period $l$ ($l \leqq H_{\text{initial}}/\lambda < l^*$). In the following (Sect. 3 and the appendix) we present some examples in order to illustrate these statements.

## 3. An Example: The So-Called "*M*-Map"

In order to illustrate our idea of uncertainty production presented above, we define a piecewise linear 1-D map $f_{M,\alpha}$ (called "*M*-map") as follows:

$$f_{M,\alpha}(z) := \begin{cases} z/\alpha, & 0 \leqq z < \alpha \\ -\alpha z/(1/2-\alpha) + 1 + \alpha^2/(1/2-\alpha), & \alpha \leqq z < 1/2 \\ \alpha z/(1/2-\alpha) + 1 - \alpha(1-\alpha)/(1/2-\alpha), & 1/2 \leqq z < 1-\alpha \\ -z/\alpha + 1/\alpha, & 1-\alpha \leqq z \leqq 1 \end{cases}$$

$$\text{for} \quad 0 < \alpha < 1/2.$$

Figure 3 shows the graph of $f_{M,\alpha}$ including that of its twofold iterate $f^2_{M,\alpha}$ (dashed line) for $\alpha = 1/6$. If the control parameter $\alpha$ approaches 1/2, the well-known symmetric tent map appears. The probability density

$$\varrho_{M,\alpha}(z) := \begin{cases} 1/(2-2\alpha), & 0 \leqq z < 1-\alpha, \\ 1/(2\alpha), & 1-\alpha \leqq z \leqq 1, \\ 0, & \text{otherwise}, \end{cases}$$

fulfils the Frobenius-Perron equation

$$\varrho_{M,\alpha}(z) = \sum_i \varrho_{M,\alpha}(z_i)/|df_{M,\alpha}/dz|_{z_i}|,$$

where the sum is taken over all $z_i$ which are mapped on $z$ by $f_{M,\alpha}$. Hence $\varrho_{M,\alpha}$ defines an

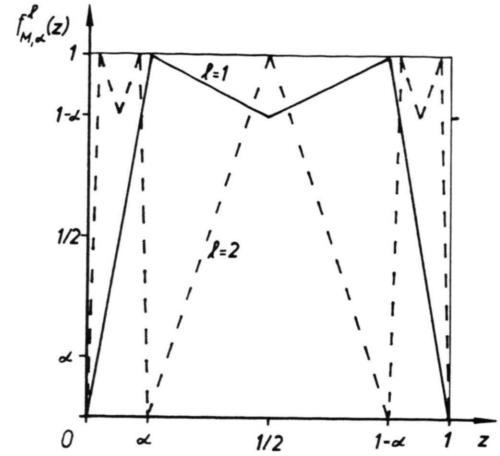


Fig. 3. Graph of the *M*-map $f_{M,1/6}$ including that of its second iterate $f^2_{M,1/6}$.

invariant measure $\bar{\mu}_{M,\alpha}$ which is assumed to be ergodic. Note that the probability to find an orbit in the "small" interval $[1-\alpha, 1]$ is the same (namely 1/2) as to find it in the "great" interval $[0, 1-\alpha]$. This is due to the special action of the map: in the first iteration the interval $[\alpha-\alpha^2, 1-(\alpha-\alpha^2)]$ is condensed to $[1-\alpha, 1]$ and then, in the second iteration, spread over the entire interval [0, 1].

The LCE is easily obtained from (9.1) using the invariant measure $\bar{\mu}_{M,\alpha}$:

$$\lambda_{M,\alpha} = \frac{1}{\alpha - 1} \operatorname{ld} (1/2-\alpha)^{(1/2-\alpha)} \alpha^{\alpha}.$$

For $0 < \alpha < 1/2$ the LCE $\lambda_{M,\alpha}$ is always positive. Hence every system of the family $\{f_{M,\alpha}\}_{\alpha \in ]0,1/2[}$ is a chaotic one. The uncertainty on a future state can be obtained from (8.1):

$$h_{M,\alpha}(1) = \begin{cases} \lambda_{M,\alpha}, & 1/4 < \alpha < 1/2, \\ \dfrac{\operatorname{ld} \alpha}{2\alpha - 2}, & 0 < \alpha \leqq 1/4. \end{cases} \tag{14}$$

If $f_{M,\alpha}$ is expanding in [0, 1] (i.e., $\alpha \in ]1/4, 1/2[$), the uncertainty after one iteration is given by the LCE. On the other hand, for $\alpha \in ]0, 1/4[$ $f_{M,\alpha}$ has a contracting effect in $]\alpha, 1-\alpha[$, and the uncertainty production $h_{M,\alpha}(1)$ is greater than $\lambda_{M,\alpha}$:

$$h_{M,\alpha}(1) - \lambda_{M,\alpha} \equiv \delta_{M,\alpha}(1)$$

$$= \begin{cases} 0 & 1/4 \leqq \alpha < 1/2, \\ \dfrac{1/2-\alpha}{1-\alpha} \operatorname{ld} \dfrac{1/2-\alpha}{\alpha}, & 0 < \alpha < 1/4. \end{cases}$$

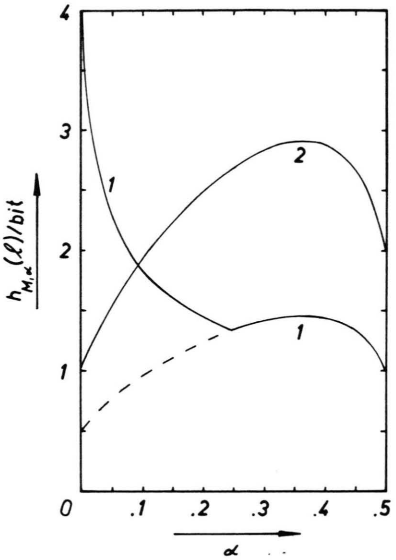


Fig. 4. Uncertainty $h_{M,\alpha}(l)$ on a future state of the $M$-map $f_{M,\alpha}$, if an initial state is known, as a function of the control parameter $\alpha$. Parameter: prediction period $l$ $(= 1, 2)$.

If $\alpha$ approaches zero, $\delta_{M,\alpha}(1)$ goes to infinity, while $\lambda_{M,\alpha}$ remains finite ($\lambda_{M,\alpha} \to 1/2$ if $\alpha \to 0$), as illustrated in Figure 4. Thus the correction term $\delta_{M,\alpha}(1)$ becomes the more important the more $\alpha$ deviates from 1/4 and approaches zero. This is a typical behaviour also for the maps considered in the appendix and will be discussed there in more detail. For $l \geqq 2$ $f^l_{M,\alpha}$ $(0 < \alpha < 1/2)$ is expanding in [0, 1]. Thus we obtain for the uncertainty production

$$h_{M,\alpha}(l) = l \cdot \lambda_{M,\alpha}, \quad l = 2, 3, 4, \ldots .$$

Now we consider special $\varepsilon$-partitions $\mathscr{B}_\varepsilon$, $\varepsilon = \alpha/i$ with $i = 1, 2, 3, \ldots$, of the interval [0, 1], and special maps $f_{M,\alpha}$, $\alpha = 1/(2k+2)$ with $k = 1, 2, 3, \ldots$. Thus, $\mathscr{B}_\varepsilon$ is a Markov partition with respect to $f_{M,\alpha}$, i.e., $f_{M,\alpha}$ is monotonous in every box $B_\mu$ of the partition, and $f_{M,\alpha}$ maps the set $\{\mu \cdot \varepsilon\}_{\mu=0}^{1/\varepsilon}$ of divisional points of $\mathscr{B}_\varepsilon$ onto itself. In this case the initial information $H_{i;M,\alpha}$ is immediately obtained from (6) and (1.2) using the invariant measure $\bar{\mu}_{M,\alpha}$:

$$H_{i;M,\alpha} = \mathrm{ld}\, 1/\varepsilon + 1/2\, \mathrm{ld}\, 4\alpha(1-\alpha).$$

The uncertainty $H_{c;M,\alpha}(1)$ on a future state $B_{\mu(1)}$ is easily obtained from (7) and (2.3):

$$H_{c;M,\alpha}(1) = \frac{\mathrm{ld}\,\alpha}{2\alpha - 2}.$$

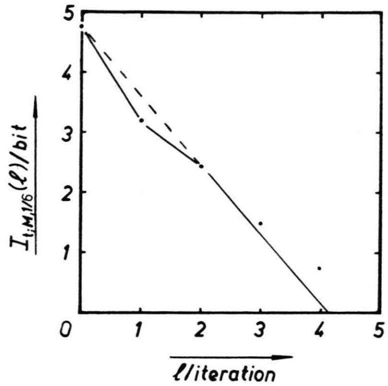


Fig. 5. Transinformation $I_{t;M,1/6}(l) = H_i - H_{c;M,1/6}(l)$ of the $M$-map versus prediction period $l$ using the $\varepsilon$-partition $\mathscr{B}_{1/36}$ (points). The dashed line represents $H_i - l \cdot \lambda_{M,1/6}$, $\lambda_{M,1/6} = 1.15\ldots$, and the solid line $H_i - h_{M,1/6}(l)$. Obviously $h_{M,1/6}(l)$ describes the uncertainty production more properly than $l \cdot \lambda_{M,1/6}$.

$H_{c;M,\alpha}(1)$ is in perfect correspondence with the uncertainty $h_{M,\alpha}(1)$ (see (14)) which was calculated from (8.1). This is due to the special choice of the partition of state space which eleminates the effect of overlappings, and which guarantees that the slope of $f_{M,\alpha}$ is constant in every box of the partition. As mentioned above, we would, in general, expect that $h_{M,\alpha}(1)$ is a lower bound of $H_{c;M,\alpha}(1)$ (see (12)).

As illustrated in Fig. 3, $f^l_{M,\alpha}$ becomes more complicated in shape as $l$ increases. Thus, for a fixed parition $\mathscr{B}_\varepsilon$ there will be a positive integer $l_0$ such that $\mathscr{B}_\varepsilon$ is no Markov partition with respect to $f^l_{M,\alpha}$ if $l \geqq l_0$. Either there occur overlappings, or $f^l_{M,\alpha}$ is not monotonous in each box, or both. In these cases the uncertainty $h_{M,\alpha}(l)$ cannot be related to $H_{c;M,\alpha}(l)$ any longer. This statement is illustrated in Figure 5. $\mathscr{B}_{1/36}$ is a Markov partition with respect to $f^l_{M,1/6}$, $l = 1, 2$, and the slope $\mathrm{d}f^l_{M,1/6}/\mathrm{d}z$ is constant in every box $B_\mu$, $\mu = 1, 2, \ldots, 36$. Hence, the uncertainty $H_{c;M,1/6}(l)$ equals $h_{M,1/6}(l)$. For $l \geqq 3$ overlappings occur and the slope of $f^l_{M,1/6}$ varies within some boxes. As a result we obtain $H_{c;M,1/6}(l) < h_{M,1/6}(l)$, i.e., for $l \geqq 3$ $h_{M,1/6}(l)$ cannot be related to $H_{c;M,1/6}(l)$ any longer.

## 4. Uncertainty Production Due to Overlappings

In Sect. 2 we discussed the uncertainty production of 1-D maps neglecting the fact that uncertainty

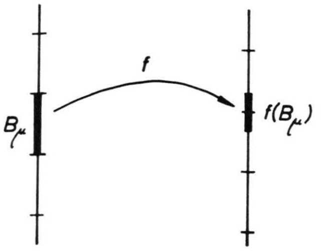


Fig. 6. Schematic representation of a contracting action of $f$. Due to an overlapping 1 bit uncertainty arises in this case.

may also arise, if divisional points of a partition of state space are not mapped on divisional points, i.e., if overlappings of mapped boxes of the partition occur. Figure 6 shows this situation for $s_\mu < 1$, i.e., if $f$ contracts the box $B_\mu$. Due to an overlapping 1 bit uncertainty may arise in this case. (Note the difference to the case of no overlappings where the contracting action of $f$ does not contribute to the uncertainty production.) An investigation of the effect of overlappings appears appropriate because overlappings do occur in experimental measurements, at least in general.

We start our investigations with piecewise linear maps $f$ ($f$: $[0, 1] \to [0, 1]$) assuming that there is a unique invariant density $\varrho$ (which is a piecewise constant function in this case) such that we have a finite set $\{z_i\}_{i=1}^{k^*}$ of so-called "points of discontinuity". We consider $z_i$ to be a point of discontinuity if at least one of the following conditions holds: (i) $f$ is discontinuous at $z_i$, (ii) $\varrho$ is discontinuous at $z_i$, (iii) $\mathrm{d}f/\mathrm{d}z$ does not exist at $z_i$. (E.g., the $M$-map considered in Sect. 3 has the following points of discontinuity: 0, $\alpha$, 1/2, $1-\alpha$, 1, and the map $f_{1,\alpha}$ considered in the appendix has: 0, $\alpha$, $1-\alpha$, 1.) A box $B_\mu$ of the $\varepsilon$-partition $\mathscr{B}_\varepsilon$ of [0, 1[ is called "box of discontinuity", if $B_\mu$ contains at least one point of discontinuity. Otherwise $B_\mu$ is called "box of continuity". Now we introduce a (finite) decomposition $\{I_j\}_{j=1}^{k}$ of [0, 1[ which unities adjoining boxes of equal type such that each interval of the decomposition is either a (finite) union of adjoining boxes of continuity (in this case $I_j$ is called "interval of continuity"), or a (finite) union of adjoining boxes of discontinuity (then $I_j$ is called "interval of discontinuity"). The intervals of continuity are assumed to be as large as possible such that they adjoin only one interval of discontinuity. Moreover, if $I_j$ is an interval of discontinuity, and only the left boundary of $I_j$ is a point of discontinuity, then $I_j$ can be considered in the following as an interval of continuity. More specifically, if all points of discontinuity coincide with divisional points of the $\varepsilon$-partition, we say that we have no interval of discontinuity. If the $\varepsilon$-partition is fine enough, the intervals of continuity typically contain a large number of boxes, and the intervals of discontinuity typically consist of only one box. (Obviously the number of intervals of continuity cannot exceed $k^*$.) In the following, we restrict our attention to intervals of continuity because these intervals typically give the dominant contribution to the uncertainty production, if the $\varepsilon$-partition is fine enough, and if the prediction period $l$ is small enough. We now will derive formulas for an effective calculation of the uncertainty production after one time step ($l = 1$). Thereby we will learn about the influence of overlappings on the uncertainty production.

Let $M_j$ label the number of (adjoining) boxes in the interval $I_j$ of continuity, and $s_j \equiv |\mathrm{d}f/\mathrm{d}z|_{z \in I_j}|$. The situation is shown in Fig. 7a and b for an expanding ($s_j \geqq 1$) and a contracting ($0 < s_j < 1$) action of $f$, respectively. From a somewhat extensive but simple calculation we obtain from (7) and (3.2) the uncertainty production in $I_j$:

*Case a)* ($s_j \geqq 1$):

$$H_{\mathrm{c}}^{I_j,\,\mathrm{exp}}(1) = \mathrm{ld}\, s_j + \frac{1}{s_j M_j} \left\{ \sum_{i=1}^{M_j-1} [q_{ji}^{\mathrm{exp}}\, \mathrm{ld}\, 1/q_{ji}^{\mathrm{exp}} + (1-q_{ji}^{\mathrm{exp}})\, \mathrm{ld}\, 1/(1-q_{ji}^{\mathrm{exp}})] + (\Delta_j/\varepsilon)\, \mathrm{ld}\, \varepsilon/\Delta_j + q_{jM_j}^{\mathrm{exp}}\, \mathrm{ld}\, 1/q_{jM_j}^{\mathrm{exp}} \right\},$$

with

$$q_{ji}^{\mathrm{exp}} := (i \cdot s_j - \Delta_j/\varepsilon) - \mathrm{int}\,(i \cdot s_j - \Delta_j/\varepsilon), \tag{15}$$

*Case b)* $(0 < s_j < 1)$:

$$H_c^{I_j,\,\mathrm{con}}(1) = \frac{1}{M_j} \sum_{i=0}^{N_j} [q_{ji}^{\mathrm{con}} \,\mathrm{ld}\, 1/q_{ji}^{\mathrm{con}} + (1 - q_{ji}^{\mathrm{con}}) \,\mathrm{ld}\, 1/(1 - q_{ji}^{\mathrm{con}})] ,$$

with

$$q_{ji}^{\mathrm{con}} := \frac{\Delta_j + i \cdot \varepsilon}{s_j \cdot \varepsilon} - \mathrm{int}\left(\frac{\Delta_j + i \cdot \varepsilon}{s_j \cdot \varepsilon}\right), \qquad N_j := \mathrm{int}\left(\frac{s_j \cdot M_j \cdot \varepsilon - \Delta_j}{\varepsilon}\right). \tag{16}$$

$\Delta_j$ labels the shift, as illustrated in Figure 7. It fulfils $0 \leqq \Delta_j < \varepsilon$. (If $M_j$ equals 1, the sum over $i$ in (15) is equalized to zero. Moreover, if $s_j\, M_j\, \varepsilon < \Delta_j$ (i.e., $N_j = -1$), we set $H_c^{I_j,\,\mathrm{con}}(1) = 0$.) $H_c^{I_j,\,\mathrm{exp}}(1) - \mathrm{ld}\, s_j$ and $H_c^{I_j,\,\mathrm{con}}(1)$ give the additional uncertainty production due to overlapping in the expanding and contracting case, respectively. The average uncertainty production is now obtained from

$$H_c(1) = \sum_{j=1}^{k} \bar{\mu}(I_j)\, H_c^{I_j}(1) . \tag{17}$$

In (17) the average over the uncertainty production in all intervalls of the decomposition is taken. ($H_c^{I_j}(1) = H_c^{I_j,\,\mathrm{con}}(1)$, if $s_j \geqq 1$, and $H_c^{I_j}(1) = H_c^{I_j,\,\mathrm{con}}(1)$, if $0 < s_j < 1$.) If $I_j$ is an interval of discontinuity, $H_c^{I_j}(1)$ has to be obtained directly from (7) and (3.2). For a sufficiently fine $\varepsilon$-partition the intervals of discontinuity are expected to be of relatively small measure $\bar{\mu}$ such that they do not essentially contribute to the average uncertainty production $H_c(1)$. Hence a "good" estimation of $H_c(1)$ is expected to be obtained, if the average in (17) is taken only over the intervals of continuity.

Now we are looking for some estimations of $H_c^{I_j,\,\mathrm{exp}}(1) - \mathrm{ld}\, s_j$ and $H_c^{I_j,\,\mathrm{con}}(1)$. If we use the relations

$$0 \leqq q \,\mathrm{ld}\, 1/q + (1 - q) \,\mathrm{ld}\, 1/(1 - q) \leqq 1 \quad \text{and}$$

$$0 \leqq q \,\mathrm{ld}\, 1/q \leqq 1/(e \ln 2) = 0.5307 \ldots ,$$

which hold for every $q \in [0, 1]$, we obtain

$$0 \leqq H_c^{I_j,\,\mathrm{exp}}(1) - \mathrm{ld}\, s_j \leqq \frac{1}{s_j}\left[1 + \frac{1}{M_j}\left(\frac{2}{e \ln 2} - 1\right)\right] \leqq \frac{1.0614 \ldots}{s_j} \tag{18}$$

and

$$0 \leqq H_c^{I_j,\,\mathrm{con}}(1) \leqq \frac{N_j + 1}{M_j} \leqq s_j + \frac{1 - \Delta_j/\varepsilon}{M_j} \leqq s_j + 1/M_j . \tag{19}$$

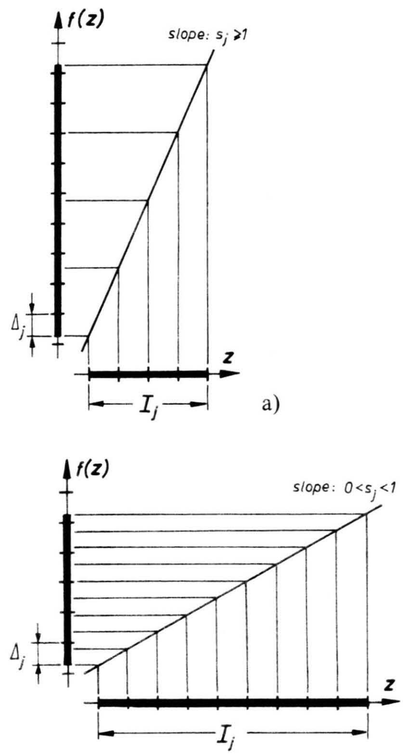


Fig. 7. Schematic representation of an expanding (Fig. 7a) and a contracting (Fig. 7b) action of $f$ on an interval $I_j$ of continuity.

The upper bounds in (18) and (19) often give useful estimations of the maximum additional uncertainty production due to overlappings. Nevertheless, sometimes they exceed 1 bit, though from (15) and (16) follows that the additional uncertainty production due to overlappings cannot exceed 1 bit. The lower bounds in (18) and (19) are reached, if the shift equals zero, and if the slope $s_j$ in case a) ($s_j \geqq 1$) and $1/s_j$ in case b) ($0 < s_j < 1$) are (positive) integers. Under these assumptions no overlappings occur in $I_j$. This situation was already discussed in Section 2.

There we defined $h(1)$ (see (8.1)), which equals exactly $H_c(1)$, if the assumptions given above hold for every interval of continuity, and if there is no internal of discontinuity. From this consideration it becomes evident that there are two different reasons for the occurrence of overlappings which cause an additional uncertainty production: firstly, the shift $\Delta_j$ may differ from zero, and secondly the slope $s_j$ (in expanding) and $1/s_j$ (in contracting regions) may be no integer.

Now we assume that $s_j$ (in case a)) and $1/s_j$ (in case b)) are integers, but the shift $\Delta_j$ does not of necessity equal zero ($0 \leqq \Delta_j < \varepsilon$). In this case (15) and (16) can be written as follows:

$$H_c^{I_j,\exp}(1) - \text{ld}\, s_j$$
$$= \frac{1}{s_j}\left[(\Delta_j/\varepsilon)\,\text{ld}\,\varepsilon/\Delta_j + (1-\Delta_j/\varepsilon)\,\text{ld}\,1/(1-\Delta_j/\varepsilon)\right] \leqq \frac{1}{s_j},$$

and

$$H_c^{I_j,\text{con}}(1) = \frac{N_j+1}{M_j}\left[q_j\,\text{ld}\,1/q_j + (1-q_j)\,\text{ld}\,1/(1-q_j)\right]$$

with

$$q_j := \frac{\Delta_j}{s_j \cdot \varepsilon} - \text{int}\left(\frac{\Delta_j}{s_j \cdot \varepsilon}\right).$$

In the following we assume that the shift $\Delta_j$ is vanishing, and $s_j \cdot M_j$ $(= N_j)$ is assumed to be an (positive) integer. Hence just $N_j$ boxes of the $\varepsilon$-partition are covered by $f(I_j)$. At first we consider the case of an expanding action of $f$ ($s_j \geqq 1$). The slope $s_j$ can now be written as follows:

$$s_j = \text{int}(N_j/M_j) + n_j/m_j,$$

$n_j$ and $m_j$ ($n_j < m_j$) being assumed to be prime to each other. A simple calculation reveals that $H_c^{I_j,\exp}(1)$ in (15) can now be written as follows:

$$H_c^{I_j,\exp}(1) - \text{ld}\, s_j = \frac{2}{s_j\, m_j}\sum_{i=1}^{m_j}\frac{i}{m_j}\,\text{ld}\,\frac{m_j}{i}. \quad (20)$$

The sum in (20) can be estimated as follows:

$$\frac{2}{m_j}\sum_{i=1}^{m_j}\frac{i}{m_j}\,\text{ld}\,\frac{m_j}{i} < 2\int_0^1 q\,\text{ld}\,1/q\,\text{d}q = \frac{1}{2\ln 2}.$$

Hence we have a new upper bound for $H_c^{I_j,\exp}(1) - \text{ld}\, s_j$:

$$H_c^{I_j,\exp}(1) - \text{ld}\, s_j < \frac{1}{s_j}\cdot\frac{1}{2\ln 2} = \frac{0.7213\ldots}{s_j}, \quad (21)$$

i.e., in this case due to overlappings less than $1/(s_j\, 2\ln 2)$ bit uncertainty is produced in addition to ld $s_j$. In the same way an analogous result is obtained in the case of a contracting action of $f$. The inverse of the slope can then be written as follows:

$$1/s_j = \text{int}(M_j/N_j) + m_j/n_j,$$

$m_j$ and $n_j$ ($m_j < n_j$) being relatively prime. For $H_c^{I_j,\text{con}}(1)$ in (16) we now obtain

$$H_c^{I_j,\text{con}}(1) = \frac{2s_j}{n_j}\sum_{i=0}^{n_j}\frac{i}{n_j}\,\text{ld}\,\frac{n_j}{i} < \frac{s_j}{2\ln 2}. \quad (22)$$

The greater $m_j$ in case a) and $n_j$ in case b), the more $H_c^{I_j,\exp}(1) - \text{ld}\, s_j$ and $H_c^{I_j,\text{con}}(1)$ approach their respective upper bounds.

If the shift $\Delta_j$ is not vanishing, the additional uncertainty production due to overlappings can be greater than the upper bounds in (21) and (22). However, if $m_j$ and $n_j$ are large compared to 1, $1/(s_j\, 2\ln 2)$ and $s_j/(2\ln 2)$ give good approximations for $H_c^{I_j,\exp}(1) - \text{ld}\, s_j$ and $H_c^{I_j,\text{con}}(1)$, respectively, also for $\Delta_j \neq 0$.

So far we have only considered the prediction period $l = 1$. It should be noted that similar considerations can be made also for $f^l$, $l = 2, 3, 4, \ldots$, in order to obtain the additional uncertainty production due to overlappings after $l$ time steps. The decomposition of [0, 1[ into intervals of continuity and discontinuity varies, in general, for different values of $l$ – typically the number of points of discontinuity grows, if the prediction period $l$ increases till there are intervals of discontinuity only. However, if $l$ is sufficiently small such that the uncertainty production in intervals of continuity predominates, we have the following estimation:

$$h(l) \leqq H_c(l) \leqq h(l) + 1 \text{ bit}. \quad (23)$$

Up to now piecewise linear maps have been considered. In the following we use the results derived above for a somewhat heuristic investigation of the additional uncertainty production due to overlappings in 1-D maps which are not piecewise linear. Let $f^l$ be "essentially nonlinear" which means that $f^l$ has no linear piece (i.e., $\text{d}^2 f^l/\text{d}z^2 \neq 0$ almost everywhere in [0, 1]). We assume to have a fine enough $\varepsilon$-parition such that the contributions of intervals of discontinuity to the uncertainty production can be neglected, and [0, 1[ can be nearly completely composed of a very (infinitely) large number of intervals of continuity where $f^l$ is nearly linear. If the slope $s^l \equiv |\mathrm{d}f^l/\mathrm{d}z|$ varies continuously from one interval to the other, the majority of these intervals will be characterized by large values of $m_j$ and $n_j$ (cf. (20) and (22)). Consequently, $1/(s^l \, 2 \ln 2)$ and $s^l/(2 \ln 2)$ give good approximations of the additional uncertainty production due to overlappings in expanding and contracting regions, respectively. Hence we expect in the limit $\varepsilon \to 0$:

$$H_c(l) \underset{\varepsilon \to 0}{\longrightarrow} \tilde{H}_c(l), \quad \text{with}$$

$$\tilde{H}_c(l) := h(l) + \frac{1}{2 \ln 2} \int \Theta\left(\left|\frac{\mathrm{d}f^l}{\mathrm{d}z}\right| - 1\right) \cdot \left|\frac{\mathrm{d}f^l}{\mathrm{d}z}\right|^{-1} + \Theta\left(\left|\frac{\mathrm{d}f^l}{\mathrm{d}z}\right|^{-1} - 1\right) \cdot \left|\frac{\mathrm{d}f^l}{\mathrm{d}z}\right| \, \mathrm{d}\bar{\mu}. \quad (24)$$

Obviously $\tilde{H}_c(l)$ fulfils the inequality:

$$h(l) \leqq \tilde{H}_c(l) \leqq h(l) + \frac{1}{2 \ln 2}.$$

To demonstrate the validity of our estimation, we now present some results for the well-known logistic map (see e.g. [10])

$$f_{L,\alpha}(z) = \alpha z (1 - z), \quad 0 \leqq z \leqq 1, \quad 0 \leqq \alpha \leqq 4.$$

For $\alpha = 4$ an ergodic invariant measure $\bar{\mu}_{L,4}$ is known which is derived from the invariant density

$$\varrho_{L,4}(z) = \begin{cases} \dfrac{1}{\pi \sqrt{z(1-z)}}, & 0 < z < 1 \\ 0, & \text{otherwise.} \end{cases}$$

Consequently, the LCE $\lambda_{L,4}$ is readily obtained from (9.1): $\lambda_{L,4} = +1$ bit. The uncertainty production $h_{L,4}(1)$ is greater than $\lambda_{L,4}$ because $f_{L,4}$ has a contracting action for $z \in ]3/8, 5/8[$. From (8.1) we obtain $h_{L,4}(1) = (1.23042 \pm 0.00002)$ bit. The uncertainty production including overlappings in the limit $\varepsilon \to 0$ is estimated using (24): $\tilde{H}_{c;L,4}(1) = (1.52564 \pm 0.00002)$ bit. In order to test these results, we have calculated $H_{c;L,4}(1)$ via (6), (7), (3.2), and (2.3), using the invariant ergodic measure $\bar{\mu}_{L,4}$, for different $\varepsilon$-partitions ($1/\varepsilon = 5 \cdot 2^1, 5 \cdot 2^2, \ldots, 5 \cdot 2^{15}$). An illustration is given in Figure 8. Obviously we obtain the expected behaviour: the uncertainty $H_{c;L,4}(1)$ approaches $\tilde{H}_{c;L,4}(1)$ if the precision $1/\varepsilon$ of measurement is increased. $\bar{h}_{L,4}(1)$ and $\underline{h}_{L,4}(1)$ are defined in [11]. They give an estimated upper and lower bound of $H_{c;L,4}(1)$ which are expected to be valid even for relatively small precisions $1/\varepsilon$ of measurement. $\bar{h}_{L,4}(1)$ and $\underline{h}_{L,4}(1)$ were derived from the assumption that for every box of the partition the most unfavourable and favourable case of overlapping occurs, respectively, whereas, in contrast to the explanation above, no use was made of the fact that possibly not for all boxes the most unfavourable respectively favourable case is realized at the same time.

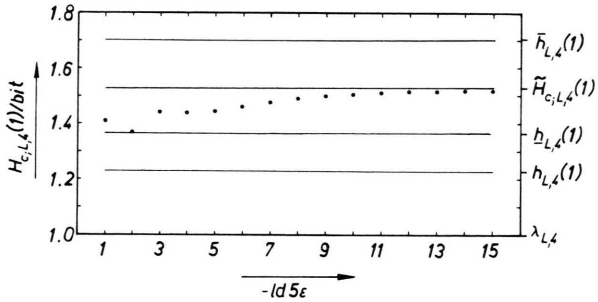


Fig. 8. Logistic equation $f_{L,4}$: Uncertainty $H_{c;L,4}(1)$ about a future state $B_{\mu(1)}$, if an initial state $B_{\mu(0)}$ is known, as a function of the precision $1/\varepsilon$ of measurement (points). Obviously, $h_{L,4}(1)$ describes the uncertainty production more properly than $\lambda_{L,4}$. $\underline{h}_{L,4}(1)$ and $\bar{h}_{L,4}(1)$ are estimated lower and upper bounds of $H_{c;L,4}(1)$ (see [11]) which take into account the effect of overlappings. $\tilde{H}_{c;L,4}(1)$ is the estimated uncertainty production which should approach $H_{c;L,4}(1)$ as $1/\varepsilon \to +\infty$.

## 5. The Influence of External Noise on the Possibility of State Prediction

In this section we present some elementary investigations of the effect of additive external noise on the uncertainty production after one time step ($l = 1$). We assume that the noise is given by a time discrete stationary stochastic process $\{\xi_n\}_{n=0}^{+\infty}$ such that every random variable $\xi_n$, $n = 1, 2, 3, \ldots$, is independent of the former $\xi_m$, $m = 0, 1, 2, \ldots, n-1$, and all random variables $\xi_n$ are assumed to be uniformly distributed over the interval $[-1, 1]$. In the following the noise power is given by $\sigma$. A realization of the stochastic process is multiplied by $\sigma$ and then added to the right side of the 1-D map $f$ which maps the interval [0, 1] onto itself:

$$z_{n+1} = f(z_n) + \sigma \cdot \xi_n \pmod 1 \equiv f_\sigma(z_n). \quad (25)$$

The (mod 1)-transformation guarantees that the orbit remains in [0, 1[.

In general the stationary probability density $\varrho_\sigma$, defining the stationary measure $\bar{\mu}_\sigma$, differs from the invariant density $\varrho$ belonging to the deterministic map $f$. $\varrho_\sigma$ is expected to be typically "more smooth" than $\varrho$ (see e.g. [12]). In the following we presume $\varepsilon$-partitions which are fine enough such that the slope $\mathrm{d}f/\mathrm{d}z$ and the invariant density $\varrho$ are nearly constant in most boxes $B_\mu$, $\mu = 1, 2, \ldots, M$, of the partition.

Now we decompose the stochastic action of $f_\sigma$ on the box $B_\mu$ in two steps: In the first step $B_\mu$ is mapped deterministically by $f$ on an interval $[a_\mu, b_\mu[ = f(B_\mu)$. Under the assumptions given above this first step is equivalent to mapping an ensemble of points, which are approximately uniformly distributed over $B_\mu$, on $[a_\mu, b_\mu[$. Because the slope $\mathrm{d}f/\mathrm{d}z$ is assumed to be nearly constant in $B_\mu$, the points of the ensemble will be nearly uniformly distributed over $[a_\mu, b_\mu[$ after the action of $f$ as well. In the second step we add the random variable $\sigma \cdot \xi_1$ which is uniformly distributed over $[-\sigma, \sigma]$. In order to obtain the resulting transition probability density $d_\mu(z)$, we now have to distinguish between two cases:

(i) $(2\sigma < b_\mu - a_\mu)$

$$d_{\mu,\,(\mathrm{i})}(z) = \begin{cases} \dfrac{z - a_\mu + \sigma}{(b_\mu - a_\mu)\, 2\sigma}, & a_\mu - \sigma \leqq z < a_\mu + \sigma, \\[2ex] \dfrac{1}{b_\mu - a_\mu}, & a_\mu + \sigma \leqq z < b_\mu - \sigma, \\[2ex] \dfrac{b_\mu + \sigma - z}{(b_\mu - a_\mu)\, 2\sigma}, & b_\mu - \sigma \leqq z \leqq b_\mu + \sigma, \\[2ex] 0, & \text{otherwise}; \end{cases}$$

(ii) $(2\sigma \geqq b_\mu - a_\mu)$

$$d_{\mu,\,(\mathrm{ii})}(z) = \begin{cases} \dfrac{z - a_\mu + \sigma}{(b_\mu - a_\mu)\, 2\sigma}, & a_\mu - \sigma \leqq z < b_\mu - \sigma, \\[2ex] \dfrac{1}{2\sigma}, & b_\mu - \sigma \leqq z < a_\mu + \sigma, \\[2ex] \dfrac{b_\mu + \sigma - z}{(b_\mu - a_\mu)\, 2\sigma}, & a_\mu + \sigma \leqq z \leqq b_\mu + \sigma, \\[2ex] 0, & \text{otherwise}. \end{cases}$$

If the support of $d_\mu$ (case (i) or (ii)) exceeds the interval $[0, 1]$, a modified density

$$d_\mu^*(z) := \begin{cases} \sum\limits_{i=-\infty}^{+\infty} d_\mu(z + i), & 0 \leqq z < 1, \\ 0, & \text{otherwise}, \end{cases}$$

has to taken. This modification of $d_\mu$ corresponds to the (mod 1)-transformation in (25). Now we obtain the transition probabilities $p_{\nu/\mu}(1)$, $\nu = 1, 2, \ldots, M$, as follows (using the abbreviation $s_\mu \equiv |\mathrm{d}f/\mathrm{d}z|_{z \in B_\mu}|$, and hence $b_\mu - a_\mu = s_\mu \cdot \varepsilon$):

$$p_{\nu/\mu}(1) = \begin{cases} \int\limits_{B_\nu} d^*_{\mu,\,(\mathrm{i})}(z)\, \mathrm{d}z, & 2\sigma < s_\mu \cdot \varepsilon, \\ \int\limits_{B_\nu} d^*_{\mu,\,(\mathrm{ii})}(z)\, \mathrm{d}z, & 2\sigma \geqq s_\mu \cdot \varepsilon. \end{cases} \tag{26}$$

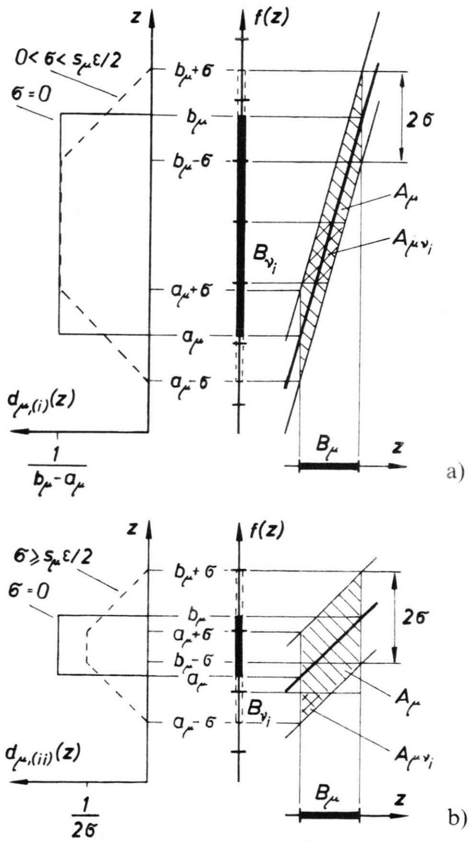


Fig. 9. Schematic illustration of the effect of additive external noise (equally distributed in $[-\sigma, \sigma]$) on the transition probability density $d_\mu(z)$. For sufficiently small noise power (Fig. 9a) external noise effects a decrease of the steepness of the flanks of $d_\mu(z)$ only. For sufficiently strong noise (Fig. 9b) also the plateau of $d_\mu(z)$ decreases – in this case the uncertainty production is determined mainly by the external noise.

A survey on the considered situation is given in Fig. 9, assuming that the support of $d_\mu$ does not exceed the interval [0, 1[. In the noisy case the graph of $f$ can be considered to be a "noisy band" such that a vertical line cuts the noisy band at a length $2\sigma$. Two vertical lines which are passing the left and right boundary of the box $B_\mu$ cut an area $A_\mu$ out of the noisy band which is marked by a lining. On the other hand, two horizontal lines which pass the upper and lower boundary of a "target box" $B_{v_i}$ cut an area $A_{\mu v_i}$ out of $A_\mu$ which is marked by a double lining. Hence, the relative content of $A_{\mu v_i}$ with respect to that of $A_\mu$ gives the transition probability $p_{v_i/\mu}(1)$.

There is a variety of situations which include overlappings. Nevertheless, in order to obtain "convenient" results we neglect overlappings now. At first we consider the case (i). Thus, in the central interval $[a_\mu + \sigma, b_\mu - \sigma]$ approximately $\operatorname{int}(s_\mu - 2\sigma/\varepsilon)$ boxes $B_{v_i}$, $i = 1, 2, \ldots, \operatorname{int}(s_\mu - 2\sigma/\varepsilon)$, can be placed, and the corresponding transition probabilities $p_{v_i/\mu}(1)$ amount to $1/s_\mu$. (Note that this is the same transition probability as in the noiseless case which was considered in Section 2.) One "flank" of $d_{\mu,\,(\mathrm{i})}$ contains approximately $\operatorname{int}(2\sigma/\varepsilon)$ boxes $B_{v_j}$, $j = 1, 2, \ldots, \operatorname{int}(2\sigma/\varepsilon)$, and the corresponding transition probabilities equal $[1 - \varepsilon(j - 1/2)/(2\sigma)]/s_\mu$. For the remaining flank we find the transition probability $[2\sigma - \varepsilon \cdot \operatorname{int}(2\sigma/\varepsilon)]^2/(4 s_\mu\, \varepsilon\, \sigma)$. Via (3.2) we now obtain for the uncertainty production

$$H_c(1,\mu) \approx \left(1 - \frac{2\sigma}{s_\mu\,\varepsilon}\right) \operatorname{ld} s_\mu - 2 \cdot \sum_{j=1}^{\operatorname{int}(2\sigma/\varepsilon)} \left\{ \frac{1}{s_\mu}\left[1 - \frac{\varepsilon}{2\sigma}(j - 1/2)\right] \ldots \ldots \operatorname{ld} \frac{1}{s_\mu}\left[1 - \frac{\varepsilon}{2\sigma}(j - 1/2)\right]\right\} - \frac{(2\sigma - \varepsilon \operatorname{int}(2\sigma/\varepsilon))^2}{2 s_\mu\, \varepsilon\, \sigma} \operatorname{ld} \frac{(2\sigma - \varepsilon \operatorname{int}(2\sigma/\varepsilon))^2}{4 s_\mu\, \varepsilon\, \sigma}. \quad (27)$$

For a small noise power $(2\sigma/\varepsilon < 1)$ we obtain from (27):

$$H_c(1,\mu) \approx \operatorname{ld} s_\mu + \frac{2\sigma}{s_\mu\,\varepsilon} \operatorname{ld} \frac{\varepsilon}{\sigma}. \quad (28)$$

The Equations (27) and (28) provide exact results if $s_\mu - 2\sigma/\varepsilon$ is an (positive) integer, and if $a_\mu + \sigma$ coincides with a divisional point of the $\varepsilon$-partition, which guarantees that overlappings will not occur. From (28) follows that $H_c(1,\mu)$ approaches the noiseless uncertainty production $\operatorname{ld} s_\mu$, if $\sigma \ll s_\mu\, \varepsilon$. I.e., if the noise power $\sigma$ is small compared to the length of $f(B_\mu)$, the noise induced uncertainty production is expected to be negligible as compared to the uncertainty production by the deterministic chaotic dynamics.

Let $s_{\min}$ label the minimum value of the slope of $f$: $s_{\min} := \inf_{z \in [0,1]} |\mathrm{d}f(z)/\mathrm{d}z|$. If $s_{\min}$ is greater than zero the precision $1/\varepsilon$ of measurement has to fulfil $1/\varepsilon \ll s_{\min}/\sigma \equiv 1/\varepsilon_{\min}$ in order to have an uncertainty production which is mainly caused by the deterministic action of $f$. Otherwise the external noise may considerably affect the uncertainty production. (It should be mentioned that the condition $\varepsilon \gg \varepsilon_{\min}$ can be mitigated, if the absolute value of the slope of $f$ is small (especially zero) in an area $W$ which has only a small stationary measure $(\bar{\mu}_\sigma(W) \ll 1)$.)

A similar investigation can be made for the case (ii) (strong noise). From a simple analysis we obtain in this case

$$H_c(1,\mu) \approx \left(1 - \frac{s_\mu\,\varepsilon}{2\sigma}\right) \operatorname{ld} \frac{2\sigma}{\varepsilon} - 2 \cdot \sum_{j=1}^{\operatorname{int} s_\mu} \left\{ \frac{\varepsilon}{2\sigma}\left[1 - \frac{j - 1/2}{s_\mu}\right] \operatorname{ld} \frac{\varepsilon}{2\sigma}\left[1 - \frac{j - 1/2}{s_\mu}\right]\right\} - \frac{\varepsilon(s_\mu - \operatorname{int} s_\mu)^2}{2 s_\mu\, \sigma} \operatorname{ld} \frac{\varepsilon(s_\mu - \operatorname{int} s_\mu)^2}{4 s_\mu\, \sigma}. \quad (29)$$

If the slope is small $(s_\mu < 1)$, we obtain from (29)

$$H_c(1,\mu) \approx \operatorname{ld} 2\sigma/\varepsilon + \varepsilon\, s_\mu/(2\sigma) \operatorname{ld} 2/s_\mu. \quad (30)$$

The Equations (29) and (30) provide exact results, if $2\sigma/\varepsilon - s_\mu$ is a positive integer and if $a_\mu + \sigma$ is a divisional point of the partition. If we have even $s_\mu\, \varepsilon/2 \ll \sigma$ (and $s_\mu < 1$), the uncertainty production is mainly noise determined. Under this assumption we obtain

$$H_c(1,\mu) \approx \operatorname{ld} \frac{2\sigma}{\varepsilon}. \quad (31)$$

In order to illustrate the influence of external noise on the uncertainty production, we have calculated the transinformation $I_{\mathrm{t};\, M_\sigma,\, 1/36}$ for the noisy $M$-map. In Fig. 10 the results are plotted for different values of the noise power $\sigma$. For $\sigma =$

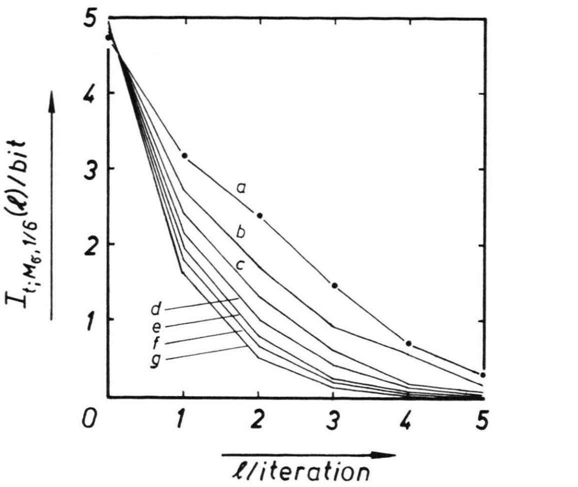


Fig. 10. Transinformation $I_{t;M_\sigma,1/6}(l)$ of the noisy $M$-map versus prediction period $l$ using the $\varepsilon$-partition $\mathscr{B}_{1/36}$. Parameter: noise power $\sigma/\sqrt{3} = 0$ (a), 0.01 (b), 0.02 (c), 0.03 (d), 0.04 (e), 0.05 (f), 0.06 (g). (Curve (a) was obtained using the invariant measure $\bar{\mu}_{M,1/6}$, while all other curves were obtained operationally from time series of length $m = 20\,000$.)

$\sqrt{3} \cdot 0.03 \approx 0.052$ (case d in Fig. 10) the uncertainty production after one iteration ($l = 1$) can be readily estimated: In the intervals $I_1 \equiv [0, 1/6[$ and $I_3 \equiv [5/6, 1[$ the absolute value of the slope equals 6. Hence we obtain $s_\mu\, \varepsilon/2 = 0.08\bar{3} > \sigma$, i.e. we have to deal with case (i) for the boxes $B_\mu$, $\mu = 1, 2, \ldots, 6, 31, 32, \ldots, 36$. From (27) we obtain $H_c^{I_1, I_3}(1, \mu) \approx 3.06$ bit. In the interval $I_2 \equiv [1/6, 5/6[$ $s_\mu$ equals 1/2, and we obtain $s_\mu\, \varepsilon/2 \approx 0.007 \ll \sigma$, i.e. we can use (31) which leads to $H_c^{I_2}(1, \mu) \approx 1.90$ bit. Numerical investigations show that the stationary density $\varrho_{\sigma;M_\sigma,1/36}$ does not essentially differ from the invariant density $\varrho_{M,1/36}$. Hence $H_c^{I_1, I_3}(1, \mu)$ and $H_c^{I_2}(1, \mu)$ have a statistical weight of about 3/5 and 2/5, respectively. Thus we obtain the average uncertainty production $H_{c;M_\sigma,1/36}(1) \approx 2.60$ bit. This roughly estimated value is in good correspondence with the operationally obtained value $H_{c;M_\sigma,1/36}(1) = (2.66 \pm 0.01)$ bit. Of course, we do not expect that we always obtain such a good correspondence between the estimated and the actually observed values because in (27)−(31) overlappings were neglected. Exact values of $H_c(1)$ can be obtained using (26) and (3.2). However, arguments similar to those used in Section 4 suggest that the additional uncertainty production by overlappings is at most of the order of 1 bit also in the noisy case.

This consideration makes evident that external noise typically increase the (average) uncertainty production. Nevertheless, also the reverse may be true. This is the case in maps showing a so-called "noise-induced order" (see [13]), where the stationary density $\varrho_\sigma$ may differ considerably from the invariant density $\varrho \equiv \varrho_{\sigma=0}$ such that some regions which are characterized by a large value of the slope of $f$ are less frequently visited by the noisy orbit. Hence external noise may also improve the state predictability.

## 6. Conclusion

An essential question in many sciences is that for the prediction of future states of a dynamical system. To give an answer to this question scientists usually proceed as follows: First they construct a state space, then they derive (deterministic) equations of motion governing the time evolution of the system, and finally they measure the initial state in order to specify just one orbit of an ensemble. A precise implementation of this algorithm would provide a satisfactory answer to the question for future state, but this may involve immense problems. Especially it may be very difficult to find a complete state space and a manageable set of equations of motion. However, in this paper we assume that these problems have been satisfactorily solved and that we have a low-(one-)dimensional map which describes the time evolution at every discrete time step $n = 0, 1, 2, \ldots$. (In Sect. 5 the effect of external noise is investigated, which may represent the influence of "disregarded degrees of freedom".) Moreover, scientists have to take into consideration that all states of a dynamical system can be measured only within a finite precision. If the system is a stable one, a small change of the initial state would not essentially alter its time evolution. A future state would be called "predictable" in this case because it can be predicted within the precision of measurement. On the other hand, already Henri Poincaré [14] pointed out that there may be quite different situations. It may happen that we have to deal with an unstable system such that any error in the measurement of the initial state causes a great error in the prediction of the future state, i.e., state predictions are limited or even impossible. This is the situation in chaotic systems.

In this paper we have investigated the consequences of a limited precision of measurement for the possibility of state prediction in the light of information theory. The term "state" as specified in this paper means that states which are not resolvable by the measuring apparatus can be considered to be equivalent to each other. Then we have asked for the information on a future state under the condition that an initial state is known. This information is given by $I_t(l) = H_i - H_c(l)$, where $H_i$ corresponds to the maximum information which is attainable on a future state, and $H_c(l)$ gives the uncertainty on a future state under the condition that an initial state is known ($l = 0, 1, 2, \ldots$ labels the prediction period). For a stable periodic motion of period $p$ the maximum information $H_i$ cannot exceed $\mathrm{ld}\, p$ bit. This limit is attained if the precision $1/\varepsilon$ of measurement is large enough such that the periodic attractor is resolved. A further increase of $1/\varepsilon$ would not provide more information. The transformation $I_t(l)$ of a stable periodic motion is periodic with a period $p^* \leqq p$. If the periodic attractor is resolved, we have $I_t(l) = H_i$, i.e., from the measurement of the initial state, all future states are known.

The situation is quite different if we have a motion on a chaotic attractor. In this case the information $H_i$ goes to infinity as the precision of measurement goes to infinity ($H_i \to -D_1\, \mathrm{ld}\, \varepsilon$ if $\varepsilon \to 0$, where $D_1$ is the information dimension of order 1 [8]). On the other hand, the uncertainty production $H_c(l)$ will remain finite but positive at every level of measuring precision. There are three sources that contribute to the production of uncertainty on a future state (under the condition that an initial state is known within a finite precision): Firstly, an expanding action of phase flow ("intrinsic noise") contributes to the uncertainty production. For chaotic systems the mean rate of exponential expansion is given by the Lyapunov characteristic exponent (LCE). In Sect. 2 of this paper we have investigated relations between the uncertainty production and the LCE. We argue that uncertainty is produced only in regions of state space which have been expanded by the flow of the dynamical system, whereas in regions which have been contracted no uncertainty on the future state is produced. Moreover, contracted regions cannot effect a reduction of the uncertainty which is simultaneously produced in expanded regions. From this point of view we define a quantity $h(l)$ (see (10.1)) which is related to the LCE (see (10.3), (11), and (12)), and which gives the uncertainty production by the intrinsic noise. Secondly, uncertainty on a future state can be produced as a result of the measuring process (overlappings). From Sect. 4 we see that at most 1 bit uncertainty is produced in addition to $h(l)$ by overlappings. For a sufficiently fine $\varepsilon$-partition a formula is derived (see (24)) which estimates the uncertainty production including overlappings for 1-D maps characterized by continuous variations of the slope. Thirdly, uncertainty can be produced by external noise (see Sect. 5). However, if the noise power is small compared to $|df/dz|\, \varepsilon$, and if the stationary density of the noisy map does not essentially differ from the invariant density of the noiseless map, then the noise induced uncertainty production is expected to the negligible in comparison to that caused by intrinsic noise.

## Appendix: Some More Examples

For a further illustration of the main results of this paper we give some more examples: In Fig. A 1 we present three one-parameter families $f_{j,\alpha}$, $j = 1, 2, 3$, of 1-D maps including their invariant densities $\varrho_{j,\alpha}$. (The time-correlation function of $f_{2,\alpha}$ was already considered in [15].) From (9.1) we obtain the LCEs:

$$\lambda_{1,\alpha} = \frac{1}{\alpha - 1}\, \mathrm{ld}\, \alpha^{\alpha} (1 - 2\alpha)^{\frac{1-2\alpha}{2}} > 0 \quad \text{for } 0 < \alpha < 1/2,$$

$$\lambda_{2,\alpha} = \frac{1}{\alpha - 1}\, \mathrm{ld}\, \alpha^{\alpha} (1 - \alpha)^{(1-\alpha)} > 0 \quad \text{for } 0 < \alpha < 1,$$

$$\lambda_{3,\alpha} = -\tfrac{1}{2}\, \mathrm{ld}\, \alpha^{2\alpha^2} (1 - \alpha^2)^{(1-\alpha^2)} > 0 \quad \text{for } 0 < \alpha < 1.$$

The uncertainty production follows from (10.3) and (11):

$$h_{j,\alpha}(l) = l\, \lambda_{j,\alpha} + \delta_{j,\alpha}(l)$$

with

$$\delta_{1,\alpha}(l) = \begin{cases} \dfrac{(1-2\alpha)^{(l+1)/2}}{2(1-\alpha)}\, \mathrm{ld}\, \dfrac{(1-2\alpha)^{(l+1)/2}}{\alpha} & \text{if } l = 1, 3, 5, \ldots \text{ and } \dfrac{(1-2\alpha)^{(l+1)/2}}{\alpha} > 1, \\ 0 & \text{otherwise,} \end{cases}$$

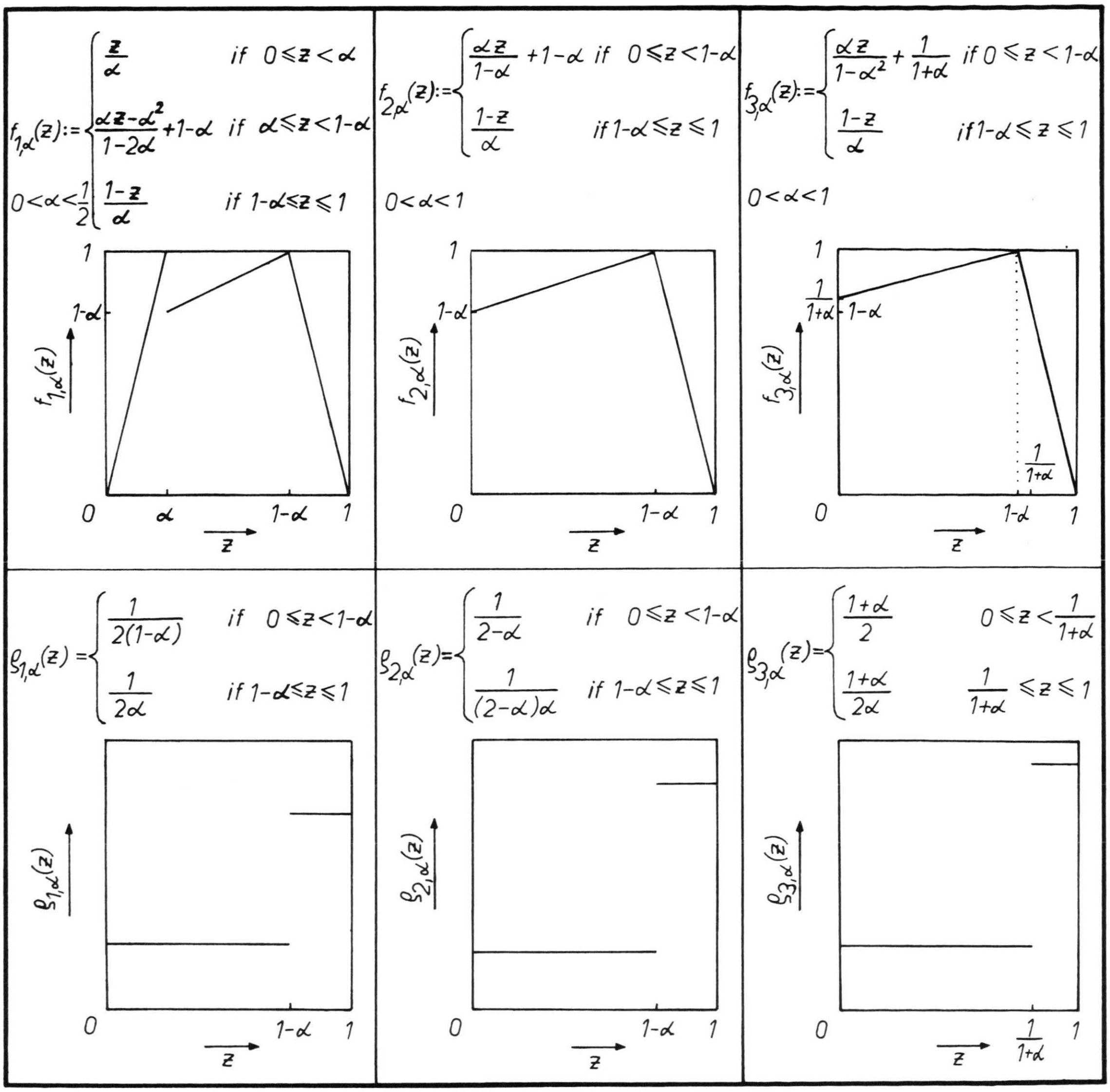


Fig. A1. Three one-parameter families of 1-D maps including their invariant densities. The maps act essentially alternately expanding and contracting.

$$\delta_{2,\alpha}(l) = \begin{cases} \dfrac{(1-\alpha)^{(l+1)/2}}{2-\alpha} \operatorname{ld} \dfrac{(1-\alpha)^{(l+1)/2}}{\alpha} & \text{if} \quad l = 1, 3, 5, \ldots \quad \text{and} \quad \dfrac{(1-\alpha)^{(l+1)/2}}{\alpha} > 1, \\ 0 & \text{otherwise}, \end{cases}$$

$$\delta_{3,\alpha}(l) = \begin{cases} \dfrac{(1-\alpha^2)^{(l+1)/2}}{2} \operatorname{ld} \dfrac{(1-\alpha^2)^{(l+1)/2}}{\alpha} & \text{if} \quad l = 1, 3, 5, \ldots \quad \text{and} \quad \dfrac{(1-\alpha^2)^{(l+1)/2}}{\alpha} > 1, \\ 0 & \text{otherwise}. \end{cases}$$

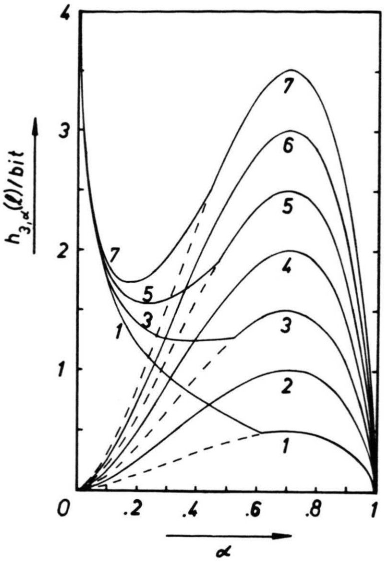

Fig. A2. Uncertainty $h_{3,\alpha}(l)$ on a future state, if an initial state is known, as a function of the control parameter $\alpha$ for the map $f_{3,\alpha}$. Parameter: prediction period $l$ (= 1, 2, ..., 7). The dashed lines represent $l \cdot \lambda_{3,\alpha}$. For sufficiently large values of $\alpha$, $l \cdot \lambda_{3,\alpha}$ coincides with $h_{3,\alpha}(l)$. However, for small values of $\alpha$, $h_{3,\alpha}(l)$ differs considerably from $l \cdot \lambda_{3,\alpha}$ because $f_{3,\alpha}$ has a contracting action in certain regions.

An illustration of $h_{3,\alpha}(l)$ is given in Figure A2. (We will consider only $f_{3,\alpha}$ because $f_{1,\alpha}$ and $f_{2,\alpha}$ are characterized by properties similar to those of $f_{3,\alpha}$.) If the prediction period $l$ is an even (positive) integer, the uncertainty production $h_{3,\alpha}(l)$ is given by $l \cdot \lambda_{3,\alpha}$ because in this case $f^l_{3,\alpha}$ shows an expanding effect almost everywhere in [0, 1]. The same applies, if $l$ is an odd integer and the value of the control parameter $\alpha$ is large enough such that $\alpha$ fulfils the inequality $(1-\alpha^2)^{(l+1)/2} > \alpha$. On the other hand, if $l$ is odd and small enough, $f^l_{3,\alpha}$ has a contracting effect in certain regions of the interval [0, 1]. Consequently, the uncertainty production is greater than $l \cdot \lambda_{3,\alpha}$ in this case. Moreover, for any fixed value of the control parameter $\alpha \in \,]0, 1[$ we find a smallest integer $l^*_{3,\alpha}$ such that $h_{3,\alpha}(l) = l \cdot \lambda_{3,\alpha}$ for $l > l^*_{3,\alpha}$, i.e., $f^l_{3,\alpha}$ has an expanding action almost everywhere in [0, 1] for $l > l^*_{3,\alpha}$. If $\alpha$ approaches zero, $l^*_{3,\alpha}$ goes to infinity and the correction term $\delta_{3,\alpha}(l)$ ($l$ odd and fixed) goes to infinity as well. On the other hand, $l \cdot \lambda_{3,\alpha}$ approaches zero as $\alpha$ goes to zero. Thus the correction term $\delta_{3,\alpha}(l)$ becomes the more important the more $\alpha$ approaches zero.

In order to explain the great uncertainty production after one time step for small values of $\alpha$ we

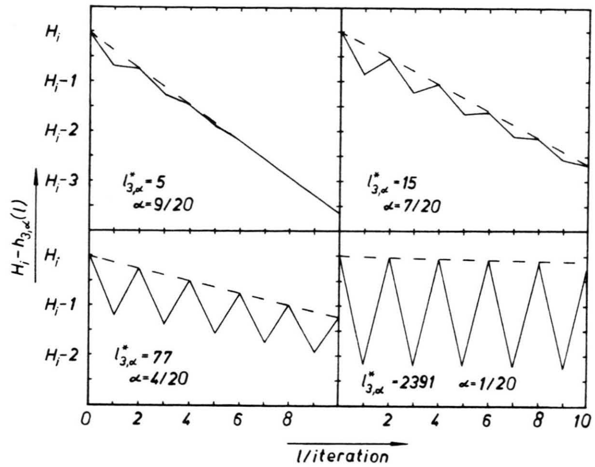

Fig. A3. Estimation of the transinformation $I_{t;3,\alpha}(l) \approx H_i - h_{3,\alpha}(l)$ as a function of the prediction period $l$ for the map $f_{3,\alpha}$. The dashed line gives the plot $H_i - l \cdot \lambda_{3,\alpha}$.

assume an ("fine") $\varepsilon$-partition such that $1/M \equiv \varepsilon = 1 - \dfrac{1}{1+\alpha}$ ($M$ integer). Hence just one box, say $B_M$, covers the small intervall $I_2 \equiv [1/(1+\alpha), 1[$. The initial information $H_i$ is easily obtained from (1.2): $H_i = 1 + \frac{1}{2}\,\mathrm{ld}\,(M-1)$. One bit corresponds to the information whether the orbit is in $B_M$ or not. $\frac{1}{2}\,\mathrm{ld}\,(M-1)$ is the information on that box out of $\{B_\mu\}_{\mu=1}^{M-1}$ which contains the orbit. The factor 1/2 refers to the statistical weight of these boxes. The uncertainty production after one time step can be found as follows: If the orbit starts in $I_1 \equiv [0, 1/(1+\alpha)[$, no uncertainty is produced after one time step because the orbit can reach only one box, namely $B_M$. On the other hand, if the orbit starts in $B_M$, it may reach any point in $I_1$ after one time step. The corresponding uncertainty production is given by $\mathrm{ld}\,(M-1)$. The statistical weight of this event is 1/2. Consequently, the average uncertainty $H_{c;3,\alpha}(1)$ equals $\frac{1}{2}\,\mathrm{ld}\,(M-1)$. (Of course, $H_{c;3,\alpha}(1)$ could be obtained immediately from (2.3) as well.) Hence, $H_i - H_{c;3,\alpha}(1) = 1$ bit gives the average information on a future state after one time step. This is the information whether the orbit is in $B_M$ or not. After a further time step ($l = 2$) most uncertainty, produced by the drastical expansion of $B_M$ after the first iteration, is destroyed due to the contracting action of $f_{3,\alpha}$ on $f_{3,\alpha}(B_M)$. Hence the initial state contains much more information on the position of the orbit after two iterations than on that after one time step. This consideration can be repeated for $l > 2$.

An alternating expansion and contraction and the corresponding information (resp. uncertainty) flow described above is a typical property of the maps investigated in this paper ($M$-map and $f_{j,\alpha}$, $j = 1, 2, 3$). In a plot of the transinformation $I_{t;3,\alpha}(l) \approx H_i - h_{3,\alpha}(l)$ as a function of the prediction period $l$ (see Fig. A3), this behaviour is reflected by the "zig-zag" graph. In [6] we present results for the chaotic motion of a parametrically excited pendulum on an attractor consisting of two bands of distinct largeness which are alternately visited. In this case the flow has an alternating expanding and contracting action. Consequently, $I_t(l)$ was found to be "zig-zag" as well.

[1] R. Shaw, Z. Naturforsch. **36a,** 80 (1981).
[2] D. Farmer, J. Crutchfield, H. Froehling, N. Packard, and R. Shaw, New York Acad. Sci. Annals **357,** 453 (1980).
[3] J. D. Farmer, Z. Naturforsch. **37a,** 1304 (1982).
[4] J. P. Crutchfield and N. H. Packard, Physica **7D,** 201 (1983).
[5] R. W. Leven und B. Pompe, Über die Möglichkeit der Zustandsvoraussage chaotischer Systeme. Ann. Phys. Leipzig **43** (1986) (to be published).
[6] B. Pompe and R. L. Leven, Transinformation of Chaotic Systems, Physica Scripta **33** (1986) (to be published).
[7] H.-P. Herzel and W. Ebeling, Phys. Lett. **111A,** 1 (1985).
[8] A. Rényi, Wahrscheinlichkeitsrechnung – mit einem Anhang über Informationstheorie, 5. Auflage, Deutscher Verlag der Wissenschaften, Berlin 1977.
[9] J.-P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57,** 617 (1985).
[10] S. Grossmann and S. Thomae, Z. Naturforsch. **32a,** 1353 (1977).
[11] B. Pompe, Chaotische Bewegungen – Experimentelle und numerische Untersuchungen an einem parametrisch erregten Pendel, Zustandsvoraussagbarkeit einfacher chaotischer Systeme. Dissertation, Greifswald 1986.
[12] J. P. Crutchfield, J. D. Farmer, and B. A. Huberman, Phys. Rep. **92,** 45 (1982).
[13] K. Matsumoto and I. Tsuda, J. Stat. Phys. **31,** 87 (1983). – J. Stat. Phys. **34,** 111 (1984).
[14] H. Poincaré, Les Methodes Nouvelles de la Mechanique, Céléste, Gauthier-Villars, Paris 1892.
[15] H. Mori, B.-C. So, and T. Ose, Progr. Theor. Phys. **66,** 1266 (1981).