

## Noise, chaos, and $(\varepsilon, \tau)$ -entropy per unit time

Pierre Gaspard<sup>a,b</sup> and Xiao-Jing Wang<sup>b</sup>

<sup>a</sup> *Computational and Applied Math Program, The Ryerson Laboratory and The James Franck Institute, The University of Chicago, Chicago, IL 60637, USA*

<sup>b</sup> *Centre for Nonlinear Phenomena and Complex Systems, Université Libre de Bruxelles, Campus Plaine C.P. 231, Blvd du Triomphe, B-1050 Brussels, Belgium*

Received June 1993; editor: I. Procaccia

### Contents:

1. Introduction	293	3.4. Time-discrete, amplitude-continuous random processes	310
2. $(\varepsilon, \tau)$ -entropy per unit time	295	3.5. Time-discrete noisy mappings	313
2.1. Dynamical processes	295	3.6. Time and amplitude continuous random processes	314
2.2. Algorithmic complexity	296	3.7. White noise	322
2.3. Entropy of a process over a time interval $T$ and a partition $\mathcal{A}$	297	3.8. Lévy flights	322
2.4. Entropy per unit time over a partition $\mathcal{A}$	298	4. Application to hydrodynamic turbulence	323
2.5. Partition (P) $(\varepsilon, \tau)$ -entropy per unit time	298	5. Application to statistical mechanics	328
2.6. Shannon–Kolmogorov (SK) $(\varepsilon, \tau)$ -entropy per unit time	299	6. Spacetime processes	331
2.7. Cohen–Procaccia (CP) $(\varepsilon, \tau)$ -entropy per unit time	301	6.1. $(\varepsilon, \tau)$ -entropy per unit time and volume	331
3. Evaluation of the $(\varepsilon, \tau)$ -entropy for random processes	303	6.2. Evaluation of $(\varepsilon, \tau)$ -entropy for spacetime processes	331
3.1. Deterministic processes	303	7. Discussions and conclusions	335
3.2. Bernoulli and Markov chains	305	7.1. Classification of the random processes	335
3.3. Birth-and-death processes	308	7.2. The principle of entropy invariance and some consequences	336
		7.3. Perspectives	341
		References	342

### Abstract:

The degree of dynamical randomness of different time processes is characterized in terms of the  $(\varepsilon, \tau)$ -entropy per unit time. The  $(\varepsilon, \tau)$ -entropy is the amount of information generated per unit time, at different scales  $\tau$  of time and  $\varepsilon$  of the observables. This quantity generalizes the Kolmogorov–Sinai entropy per unit time from deterministic chaotic processes, to stochastic processes such as fluctuations in mesoscopic physico-chemical phenomena or strong turbulence in macroscopic spacetime dynamics.

The random processes that are characterized include chaotic systems, Bernoulli and Markov chains, Poisson and birth-and-death processes, Ornstein–Uhlenbeck and Yaglom noises, fractional Brownian motions, different regimes of hydrodynamical turbulence, and the Lorentz–Boltzmann process of nonequilibrium statistical mechanics. We also extend the  $(\varepsilon, \tau)$ -entropy to spacetime processes like cellular automata, Conway’s game of life, lattice gas automata, coupled maps, spacetime chaos in partial differential equations, as well as the ideal, the Lorentz, and the hard sphere gases. Through these examples it is demonstrated that the  $(\varepsilon, \tau)$ -entropy provides a unified quantitative measure of dynamical randomness to both chaos and noises, and a method to detect transitions between dynamical states of different degrees of randomness as a parameter of the system is varied.

# NOISE, CHAOS, AND $(\varepsilon, \tau)$ -ENTROPY PER UNIT TIME

**Pierre GASPARD<sup>a, b</sup> and Xiao-Jing WANG<sup>b</sup>**

<sup>a</sup> *Computational and Applied Math Program, The Ryerson Laboratory and The James Franck Institute, The University of Chicago, Chicago, IL 60637, USA*

<sup>b</sup> *Centre for Nonlinear Phenomena and Complex Systems, Université Libre de Bruxelles, Campus Plaine C.P. 231, Blvd du Triomphe, B-1050 Brussels, Belgium*



NORTH-HOLLAND

## 1. Introduction

In the last decade, there have been many discussions on the differences and similarities between deterministic chaos and stochastic noises. In order to introduce the problem, let us show typical trajectories of two famous processes, on one hand, a chaotic one, and on the other hand, a stochastic one.

In fig. 1a, we have plotted a solution of the set of ordinary differential equations proposed by Rössler

$$\dot{X}_1 = -X_2 - X_3, \quad \dot{X}_2 = X_1 + aX_2, \quad \dot{X}_3 = bX_1 - cX_3 + X_1X_3, \quad (1.1)$$

for the parameter values  $a = 0.32$ ,  $b = 0.3$ , and  $c = 4.5$  [1]. This trajectory belongs to the chaotic attractor shown in fig. 1b.

In fig. 2, a trajectory is plotted from the Ornstein-Uhlenbeck stochastic process given by the Langevin equation [2]

$$\dot{X} = -aX + c\xi(t), \quad (1.2)$$

for  $a = 2^{-8}$  and  $c = 50$  where  $\xi(t)$  is a  $\delta$ -correlated white noise of zero mean and unit variance. (See below for a more detailed definition.)

Both trajectories appear irregular in time. However, there are important differences between them and, in particular, the second is often recognized as being more irregular than the first. In this regard, the chaotic trajectory is a differentiable curve according to Cauchy's theorem; while the stochastic signal is nowhere differentiable, irregular on arbitrarily small scales if we refer to the strict mathematical construction of the process. Already, the visual comparison reveals that stochastic noises are qualitatively more random than the chaotic signals. From this observation, we may wonder if there exists a quantitative measure of randomness that would capture our visual intuition.

Such a question is of crucial importance in the natural sciences where the time evolution of many physico-chemical and biological phenomena is described by random processes. The Brownian

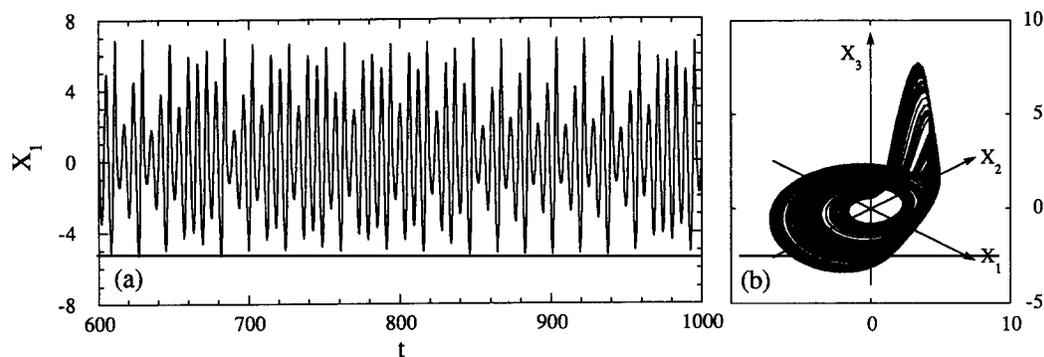


Fig. 1. (a) Trajectory of the Rössler chaotic system (1.1) integrated from the initial conditions  $X_1 = 0$ ,  $X_2 = -4$ , and  $X_3 = 0$ . (b) The corresponding chaotic attractor is depicted in its phase space  $(X_1, X_2, X_3)$  in the inset.

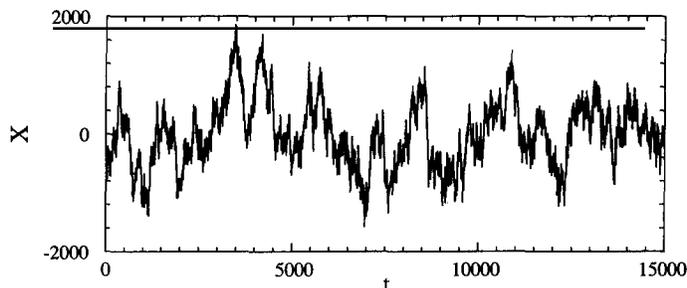


Fig. 2. Typical trajectory of the Ornstein-Uhlenbeck process of stochastic equation (1.2) as an example of a noise signal.

motion has been known for more than a century and has been modelled at the beginning of this century by stochastic differential equations like the Langevin equations [2]. The foundations of stochastic processes were firmly established in the thirties and the forties by Wiener and others who showed the importance of the white noise [2].

It was largely recognized only recently that deterministic nonlinear differential equations like Newton's equations also admit random solutions in the phenomenon called chaos [3]. Here, randomness has its origin in the exponential separation of nearby trajectories. In this context, dynamical randomness is characterized by the Kolmogorov-Sinai (KS) entropy per unit time [4]. This quantity was first introduced by Shannon in what became information theory [5]. In this theory, each natural system is a source of information. A measuring device of finite resolution records the time evolution of the process going on in the system and encodes its observations in binary or decimal form. In general, the amount of data never grows faster than linearly with time. Data accumulate at a certain rate  $C$  which characterizes the measuring device (see fig. 3).

On the other hand, the KS entropy per unit time is the intrinsic rate at which information is produced by the dynamical system and gives the number of bits which is necessary and sufficient to record without ambiguity a trajectory of the chaotic system during a unit time interval. So it sets a lower bound for the required data accumulation rate  $C$ . If a measuring device has a poor resolution and a data accumulation rate which is below the KS entropy we shall not be able to recover the precise trajectory of the system from the recorded data. On the other hand, this recovery is possible if the accumulation rate of the observing device is equal to or greater than the KS entropy. This general reasoning is not limited to the chaotic systems but can be extended to more general stochastic processes.

Our purpose in this paper is to generalize the concept of KS entropy per unit time to a new concept called  $(\varepsilon, \tau)$ -entropy per unit time and to show that it can be applied to a very large class

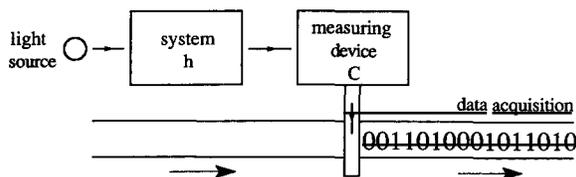


Fig. 3. Schematic representation of the observation of a natural process by a measuring device and of the recording of the data. The observation of the system is optical. The system is a source of information. The intrinsic rate of production of information by the system is the entropy per unit time  $h$ . The data are printed on the advancing band of paper or of magnetic support at the encoding rate  $C$ .

of processes that includes both chaotic and stochastic ones.

The idea is the following. In the definition of the KS entropy per unit time, the phase space is partitioned into small cells. The entropy per unit time is then calculated for the probabilities to visit successively different cells at times separated by a lapse  $\tau$ . Afterwards, the supremum of the entropy is taken for smaller and smaller partitions. This supremum is finite only in deterministic systems, but it is infinite for stochastic processes. It is therefore necessary to suppose that the cells of partition have a diameter  $\varepsilon$ . The entropy per unit time of the processes for such  $\varepsilon$ -partitions is essentially what we call the  $(\varepsilon, \tau)$ -entropy per unit time. For stochastic noises, the  $(\varepsilon, \tau)$ -entropy is thus a function that diverges as  $\varepsilon \rightarrow 0$  or as  $\tau \rightarrow 0$ . The dependence of the divergence on  $\varepsilon$  or  $\tau$  characterizes the stochastic process as will be shown below. In continuous processes, the dependence on the time lapse  $\tau$  disappears so that we recover the concept of  $\varepsilon$ -entropy per unit time which has been intensively studied in the mathematical literature [5–10]. In this paper, one of our purposes will be to place the concept of  $\varepsilon$ -entropy in its physical context.

The paper is organized as follows. In section 2, we shall introduce several concepts from the algorithmic complexity to the  $(\varepsilon, \tau)$ -entropy per unit time. In section 3, we calculate the  $(\varepsilon, \tau)$ -entropy for different chaotic and stochastic processes. In section 4, we briefly summarize the application of the  $\varepsilon$ -entropy in fluid turbulence. In section 5, we discuss the importance of the  $(\varepsilon, \tau)$ -entropy in nonequilibrium statistical mechanics. Section 6 is devoted to spacetime processes. In section 7, we carry out the classification of the different physico-chemical processes according to their degree of randomness. Finally, discussions and conclusions are presented.

## 2. $(\varepsilon, \tau)$ -entropy per unit time

### 2.1. Dynamical processes

A random process is a sequence of random variables defined at successive times. We shall denote the random variables by upper case letters  $X, Y, Z, \dots$ . The random variables take their values – denoted by lower case letters  $x, y, z, \dots$  – in  $\mathbb{R}^d$ .

The process may be discrete or continuous in time. When it is continuous we proceed to a discretization of the signal at small time intervals  $\tau$ . A new multi-time random variable is then defined according to

$$X = [X(t_j), X(t_j + \tau), X(t_j + 2\tau), \dots, X(t_j + N\tau - \tau)], \tag{2.1}$$

which belongs to  $\mathbb{R}^{Nd}$  and which corresponds to the signal during the time period  $T = N\tau$ . We use bold face letters to emphasize their vectorial character into  $N$  components. Each component is itself a  $d$ -dimensional vector corresponding to one of the  $N$  successive times.  $t_j$  is the initial time of the sequence.

From the point of view of probability theory, the process is defined by the  $N$ -time joint probability

$$P(\mathbf{x}; d\mathbf{x}, \tau, N) = \Pr\{\mathbf{x} < X < \mathbf{x} + d\mathbf{x}\} = p(\mathbf{x}) d\mathbf{x}, \tag{2.2}$$

where  $p(\mathbf{x})$  is the probability density for  $X$  to take the value  $\mathbf{x} \in \mathbb{R}^{Nd}$ . If the process is stationary, the joint probability (2.2) does not depend on the initial time  $t_j$ .

In the following section, we shall also use the probability for the random variable  $X$  to be within a distance  $\varepsilon$  of the value  $\mathbf{x}$

$$P(\mathbf{x}; \varepsilon, \tau, N) = \Pr\{\text{dist}[\mathbf{x}, X] < \varepsilon\}. \tag{2.3}$$

This other  $N$ -time probability involves the definition of a distance  $\text{dist}[\cdot, \cdot]$  in  $\mathbb{R}^{Nd}$ , which will be specified later. The joint probabilities (2.2) and (2.3) are closely related and their knowledge provides in principle a full characterization of the random process over the time interval  $T = N\tau$ . In this regard, the integer  $N$  is equivalent to the time  $T$  during which the random process is observed by the measuring device and we shall often replace  $N$  by  $T$  and vice versa.

When the random variables of the process take discrete values like in birth-and-death processes where the random variables are numbers of particles –  $X(t) \in \mathbb{N}^d$  or  $\mathbb{Z}^d$  – the signal is not continuous anymore and there is no need to introduce the infinitesimal quantities  $d\mathbf{x}$  or  $\varepsilon$ . Similarly, when the process is discrete in time the infinitesimal time  $\tau$  is not necessary and can be dropped. In that case the total time of observation  $T$  is related to the number  $N$  of random variables in (2.1) by the average time  $\bar{T}$  between the occurrences of these random variables:  $T = N\bar{T}$ .

In physics, chemistry, or biology, the joint probabilities are not the quantities which are primarily available in the definition of a dynamical process. More often, the dynamical process is defined in terms of ordinary or stochastic differential equations from which we need to calculate the joint probability which is invariant under the time evolution, as shown below.

## 2.2. Algorithmic complexity

It is useful to consider a random process from the constructive point of view where an algorithm or computer program is used to build or draw a trajectory representative of the process. When the system is random the program makes use of a pseudorandom number generator at some stage of its run. The more often the generator is called, the more random the signal is. The call of a pseudorandom number generator may be very subtle. For instance, during the integration of a typical trajectory of a deterministic chaotic systems like (1.1), arbitrary digits are coming out from below the truncation limit of the machine and the program acts itself like as pseudorandom number generator\*). However, the simulation of stochastic systems like (1.2) requires the explicit calls of a pseudorandom number generator.

Another point of view that we take as basically equivalent to the preceding one is provided by algorithmic complexity theory [11, 12]. In this theory, a trajectory over a time  $T$  will be considered as a set of  $N = T/\tau$  integer numbers where  $\tau$  is the interval between the instants when the program issues new coordinates of the trajectory. The algorithmic complexity of a trajectory over a time  $T$  is the length of the smallest program  $\mathcal{P}(T)$  able to reproduce the trajectory on a universal Turing machine [11, 12]

$$K(T) = \min_{\mathcal{P}} |\mathcal{P}(T)|, \quad (2.4)$$

where  $|\cdot|$  denotes the binary length of the code or program.

The algorithmic point of view is also very closely related to the point of view of information theory described in the introduction, where the observation of a natural process during a time  $T$  produces in a first stage  $CT$  bits which are afterwards decoded. In this second stage, the amount of data is compressed to  $I(T)$  bits by taking into account the eventual regularities in the signal. We have already mentioned that the information  $I(T)$  is proportional to the KS entropy per unit

\*) On a digital computer, the digits coming from below the truncation limit are not arbitrary since the computer is purely deterministic and has a finite number of states. However, these digits should be arbitrary in a reliable integration respecting the invariant measure of the dynamical system. The purpose of our remark was to point out the origin of randomness in the integration of a chaotic system.

time in chaotic systems:  $I(T) \simeq h_{KS}T$ . This information can be more rigorously defined with the algorithmic complexity (2.4).

Dynamical randomness and information compression due to the regularities determine how the algorithmic complexity increases with time. If the trajectory is periodic, the program only needs to memorize the pattern of the period and to repeat this pattern over the total time  $T$  so that

$$K(T) \sim \log T \quad (\text{periodicity}). \tag{2.5}$$

Let us remark that, in fact, (2.5) applies also to quasiperiodic and many other aperiodic patterns. As examples, we can mention the binary or decimal expansions of the number  $\pi$  as well as time series generated by the Feigenbaum attractor at the threshold of chaos in the period doubling scenario [13].

On the other hand, if the trajectory is random so that the pseudorandom number generator is called regularly in time, there is no way to reproduce it other than to memorize the whole trajectory. In this case,

$$K(T) \sim T \quad (\text{regular randomness}). \tag{2.6}$$

We then see that the information in the trajectory cannot be compressed.

However, there exist random processes with strong correlation in time where the signal shows long periods of order interspaced by bursts of randomness in such a way that the algorithmic complexity is not extensive with the time  $T$ . We define sporadicity or sporadic randomness by the condition that  $K(T)$  is of the form [14]

$$K(T) \sim T^\alpha \quad (0 < \alpha < 1) \quad \text{or} \quad T/(\log T)^\beta \quad (\beta > 0) \quad (\text{sporadic randomness}). \tag{2.7}$$

Sporadicity manifests itself in the Manneville–Pomeau intermittent maps [15]. In sporadic processes, the information contained in a trajectory of length  $T$  can be compressed to a number  $K(T)$  of bits by taking advantage of the redundancy in the trajectory. However, this compression cannot be as complete as it is the case for periodic processes. Consequently, a certain degree of randomness remains in the sporadic signals [14].

We conclude here by saying that a given process is random if

$$\lim_{T \rightarrow \infty} K(T)/\log T = \infty \quad (\text{randomness}). \tag{2.8}$$

### 2.3. Entropy of a process over a time interval $T$ and a partition $\mathcal{A}$

We now consider a dynamical process defined by a joint probability like (2.2). The time signal is a function  $X(t)$  in the phase space  $\mathbb{R}^d$ .

We partition the phase space into cells,  $\mathcal{A} = \{A_1, \dots, A_M\}$ . Each cell is labelled by an integer  $\{1, 2, \dots, M\}$ . We calculate the probabilities to visit successively the cells  $\omega_0\omega_1 \dots \omega_{N-1}$  at the times  $0, \tau, \dots, (N-1)\tau$ . These probabilities are given by

$$\begin{aligned} P(\omega_0\omega_1 \dots \omega_{N-1}) &= \text{Pr}\{X(0) \in A_{\omega_0}, X(\tau) \in A_{\omega_1}, \dots, X(N\tau - \tau) \in A_{\omega_{N-1}}\} \\ &= \int_{\mathbb{R}^{Nd}} \mathcal{I}_{A_{\omega_0}}(x_0) \dots \mathcal{I}_{A_{\omega_{N-1}}}(x_{N-1}) p(\mathbf{x}) \, d\mathbf{x}, \end{aligned} \tag{2.9}$$

where  $p(\mathbf{x})$  is the joint probability density in (2.2) and  $\mathcal{I}_{A_\omega}(x)$  denotes the characteristic function of the cell  $A_\omega$ .

The entropy of the process with respect to the partition  $\mathcal{A}$  is then defined as usual by [4, 16]

$$H(\mathcal{A}, \tau, T) = - \sum_{\omega_0 \dots \omega_{N-1}} P(\omega_0 \dots \omega_{N-1}) \log P(\omega_0 \dots \omega_{N-1}), \quad (2.10)$$

where we recall that  $T = \tau N$ .

If the partition  $\mathcal{A}$  is used to write the algorithm of reconstruction of the random process, the entropy  $H(T)$  gives an evaluation of the algorithmic complexity

$$K(T) \simeq H(T), \quad (2.11)$$

for almost all trajectories [17].

#### 2.4. Entropy per unit time over a partition $\mathcal{A}$

The entropy  $H(T)$  grows at most linearly with time [16]

$$H(T) \leq T \log M, \quad (2.12)$$

where  $M$  is the number of cells in the partition  $\mathcal{A}$ . The entropy per unit time of the system with respect to the partition  $\mathcal{A}$  is then defined by the limit [16]

$$h(\mathcal{A}, \tau) = \lim_{T \rightarrow \infty} H(\mathcal{A}, T, \tau)/T. \quad (2.13)$$

The entropy per unit time can distinguish between the regularly random processes and the other processes where information is produced more slowly than proportionally with time since we have

$$h(\mathcal{A}, \tau) = 0 \quad (\text{periodicity or sporadic randomness}), \quad (2.14)$$

$$h(\mathcal{A}, \tau) > 0 \quad (\text{regular randomness}). \quad (2.15)$$

In deterministic systems, the KS entropy per unit time is defined according to [4, 16]

$$h_{\text{KS}} = \lim_{\tau \rightarrow 0} \text{Sup}_{\mathcal{A}} h(\mathcal{A}, \tau), \quad (2.16)$$

where the supremum is taken over all countable partitions  $\mathcal{A}$ . A system is then said to be chaotic if  $h_{\text{KS}}$  is positive and finite. In particular, it is known that

$$h_{\text{KS}} = \lim_{T \rightarrow \infty} K(T)/T, \quad (2.17)$$

for almost all trajectories of a chaotic system [17].

Let us make the remark that the basis of the logarithms used in eq. (2.10) fixes the unit of the entropy per unit time according to

$$\begin{aligned} \log_2 : & \quad \text{bits/second,} \\ \log_e = \ln : & \quad \text{nats/second,} \\ \log_{10} : & \quad \text{digits/second.} \end{aligned} \quad (2.18)$$

#### 2.5. Partition $(P)$ $(\varepsilon, \tau)$ -entropy per unit time

However, as we shall see in the next section, no supremum exists for stochastic processes which appear to have a higher degree of randomness than chaotic systems. There is no supremum because

the entropy per unit time  $h(\mathcal{A}, \tau)$  grows indefinitely as the cells of the partition become smaller and smaller. To control this growth, let us suppose that the diameter of the cells  $\{A_i\}$  is smaller than  $\varepsilon$

$$\text{diam}(A_i) \leq \varepsilon, \tag{2.19}$$

and define the  $(\varepsilon, \tau)$ -entropy per unit time according to

$$h_P(\varepsilon, \tau) = \text{Inf}_{\mathcal{A}: \text{diam}(A_i) \leq \varepsilon} h(\mathcal{A}, \tau), \tag{2.20}$$

which is a function of  $\varepsilon$  and  $\tau$  and where the index P refers to the word partition. We remark that such a function depends on the definition of the diameter which is adopted. Accordingly, different criteria may be adopted as will be discussed later.

### 2.6. Shannon–Kolmogorov (SK) $(\varepsilon, \tau)$ -entropy per unit time

#### 2.6.1. General definition

Variants of the previous definition are possible if other criteria are used to introduce the quantity  $\varepsilon$ . A particularly important variant is the Shannon–Kolmogorov  $\varepsilon$ -entropy per unit time which was originally called the rate distortion function by Shannon and was renamed the  $\varepsilon$ -entropy by Kolmogorov. For continuous amplitude stationary sources, its definition is the following [6, 7].

Let us suppose that we want to reproduce the time signal of a process up to a precision  $\varepsilon$ . The exact time signal  $X(t)$  is then approached by an approximate copy  $Y(t)$ . The distortion between them is measured by a cost function

$$\rho_N(X, Y) = \frac{1}{N} \sum_{k=0}^{N-1} \rho[X(k\tau) - Y(k\tau)] \simeq \frac{1}{T} \int_0^T \rho[X(t) - Y(t)] dt, \tag{2.21}$$

where the distance  $\rho(X, Y)$  is chosen to depend only on the difference  $(X - Y)$ . It can be for instance the absolute error function,  $\rho(z) = |z|$ , or the squared-error function,  $\rho(z) = z^2$ . The cost function measures the distance between the copy  $Y(t)$  and the actual process  $X(t)$ . We require that their separation measured with the cost function is not larger than  $\rho(\varepsilon)$ . There remains the important question: what is the general method used to construct the copy  $Y(t)$  from our knowledge of the actual process  $X(t)$ . The following scheme is adopted.

The probability density for the actual signal  $X(t)$  to take successively the values

$$\mathbf{x} = [x(0), x(\tau), \dots, x(N\tau - \tau)], \tag{2.22}$$

is given by  $p(\mathbf{x})$ . The approximate signal  $Y(t)$  is constructed according to a joint probability for  $X(t)$  to take the values  $\mathbf{x}$  while  $Y(t)$  takes the values

$$\mathbf{y} = [y(0), y(\tau), \dots, y(N\tau - \tau)]. \tag{2.23}$$

Its conditional probability density is a general function  $q(\mathbf{y}|\mathbf{x})$ . The average mutual information between  $\mathbf{x}$  and  $\mathbf{y}$  is then

$$J[q] = \int_{\mathbb{R}^{Nd}} \int_{\mathbb{R}^{Nd}} d\mathbf{x} d\mathbf{y} p(\mathbf{x}) q(\mathbf{y}|\mathbf{x}) \log q(\mathbf{y}|\mathbf{x})/q(\mathbf{y}), \tag{2.24}$$

where

$$q(\mathbf{y}) = \int_{\mathbb{R}^{Nd}} d\mathbf{x} p(\mathbf{x}) q(\mathbf{y}|\mathbf{x}). \quad (2.25)$$

The average cost due to the distortion of  $\mathbf{y}$  with respect to  $\mathbf{x}$  is evaluated by

$$\bar{\rho}_N[q] = \int_{\mathbb{R}^{Nd}} \int_{\mathbb{R}^{Nd}} d\mathbf{x} d\mathbf{y} p(\mathbf{x}) q(\mathbf{y}|\mathbf{x}) \rho_N(\mathbf{x}, \mathbf{y}), \quad (2.26)$$

using the definition (2.21). This average cost depends on the way the process  $Y(t)$  is constructed. This dependence appears through the conditional probability density  $q(\mathbf{y}|\mathbf{x})$  which has been arbitrary up till now in a similar way as the partition  $\mathcal{A}$  was arbitrary in (2.10). To deal with this arbitrariness, we consider the set  $\mathcal{Q}(\varepsilon)$  of all conditional probabilities such that the average cost is less than  $\rho(\varepsilon)$ ,

$$\mathcal{Q}(\varepsilon) = \{q(\mathbf{y}|\mathbf{x}): \bar{\rho}_N[q] \leq \rho(\varepsilon)\}. \quad (2.27)$$

The Shannon–Kolmogorov  $(\varepsilon, \tau)$ -entropy is then defined by

$$H_{\text{SK}}(\varepsilon, \tau, T) = \inf_{q \in \mathcal{Q}(\varepsilon)} J[q]. \quad (2.28)$$

Whereupon the SK  $(\varepsilon, \tau)$ -entropy per unit time is

$$h_{\text{SK}}(\varepsilon, \tau) = \lim_{T \rightarrow \infty} H_{\text{SK}}(\varepsilon, \tau, T)/T. \quad (2.29)$$

An important property of the  $(\varepsilon, \tau)$ -entropy per unit time is that it is monotonically non-decreasing in  $\varepsilon$ .

### 2.6.2. The Kolmogorov formula for stationary Gaussian processes

When the process is stationary and Gaussian it is possible to calculate explicitly the  $\varepsilon$ -entropy per unit time thanks to a formula obtained by Kolmogorov.

For stationary Gaussian processes in one dimension ( $d = 1$ ), the joint probability density is given by

$$p(\mathbf{x}) = \exp(-\frac{1}{2} \mathbf{x}^T \cdot \mathbf{C}_N^{-1} \cdot \mathbf{x}) / (2\pi)^{N/2} (\det \mathbf{C}_N)^{1/2}, \quad (2.30)$$

where  $\mathbf{C}_N$  is the matrix of the correlations

$$[\mathbf{C}_N]_{ij} = \langle X_i X_j \rangle = C(|i - j|) \quad \text{with} \quad \langle X_i \rangle = 0. \quad (2.31)$$

For time continuous processes,  $X_i$  denotes the random variable  $X(t_i)$  at the discretized time  $t_i = i\tau$ . Since the process is stationary, the correlation function depends only on  $|i - j|$  and  $\mathbf{C}_N$  is a symmetric Toeplitz matrix. For Gaussian processes, the amplitude is always a continuous random variable and the dependence on  $\tau$  of the  $(\varepsilon, \tau)$ -entropy per unit time disappears in the limit  $\tau \rightarrow 0$ .

For time-discrete processes, the spectral density is defined by

$$\Phi(\omega) = \sum_{n=-\infty}^{\infty} \exp(-i\omega n) \langle X_n X_0 \rangle, \quad (2.32)$$

which is defined in the frequency interval  $\omega \in (-\pi, +\pi)$ . On the other hand, for time-continuous processes, the spectral density is

$$\Phi(\omega) = \int_{-\infty}^{\infty} \exp(-i\omega t) \langle X(t)X(0) \rangle dt, \tag{2.33}$$

where  $\omega \in \mathbb{R}$ .

The Shannon–Kolmogorov  $\varepsilon$ -entropy can be estimated in the case of the squared-error cost function. This fidelity criterion requires that the approximate signal  $Y(t)$  is close to the exact signal  $X(t)$  according to

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \langle [X(t) - Y(t)]^2 \rangle dt \leq \varepsilon^2. \tag{2.34}$$

The following Kolmogorov formula [6] then gives an exact evaluation of  $h_{SK}(\varepsilon)$  for stationary Gaussian processes

$$\varepsilon^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \min[\theta, \Phi(\omega)] d\omega, \tag{2.35}$$

$$h_{SK}(\varepsilon) = \frac{1}{4\pi} \int_{-\infty}^{\infty} \max[0, \log \Phi(\omega)/\theta] d\omega, \tag{2.36}$$

in terms of the spectral density (2.33). For time-discrete processes the integration should be carried out from  $-\pi$  to  $+\pi$  and the spectral density (2.32) used. The geometry involved in the calculation of the integrals (2.35)–(2.36) is schematically depicted in fig. 4. A derivation of this fascinating formula can be found in the book by Berger [10].

*2.7. Cohen–Procaccia (CP)  $(\varepsilon, \tau)$ -entropy per unit time*

A numerical method to evaluate an  $(\varepsilon, \tau)$ -entropy per unit time is provided by a method described by Cohen and Procaccia [18]. A realization of the process  $X(t)$  over a very long time interval  $L\tau \gg T = N\tau$  is given by the time series  $\{x(k\tau)\}_{k=0}^{L-1}$ . Within this long time series, sequences of length  $N$  are compared with each other. A set of  $R \ll L$  reference sequences, which are also of length  $N$ , is considered

$$\mathbf{x}_i = [x(i\tau), \dots, x(i\tau + N\tau - \tau)], \quad i \in \{i_1, \dots, i_R\}. \tag{2.37}$$

The distance between a reference sequence and another sequence of length  $N$  is defined by

$$\text{dist}_N[\mathbf{x}_i, \mathbf{x}_j] = \max\{|x(i\tau) - x(j\tau)|, \dots, |x(i\tau + N\tau - \tau) - x(j\tau + N\tau - \tau)|\}, \tag{2.38}$$

for  $j = 1, \dots, L' = L - N + 1$ .

The probability (2.3) for this distance to be smaller than  $\varepsilon$  is then evaluated by

$$P(\mathbf{x}_i; \varepsilon, \tau, N) = (1/L') \text{Number}\{\mathbf{x}_j; \text{dist}_N[\mathbf{x}_i, \mathbf{x}_j] \leq \varepsilon\}, \tag{2.39}$$

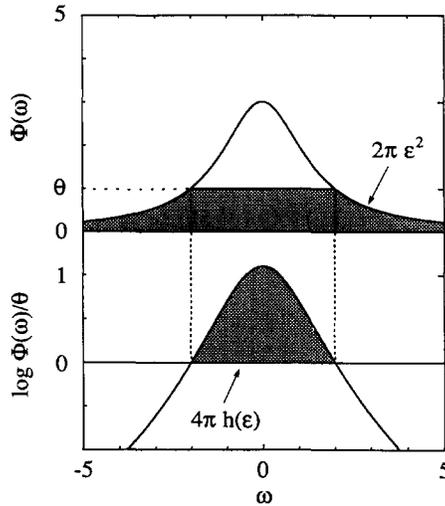


Fig. 4. Geometry of the integrals used to calculate the Shannon-Kolmogorov  $\varepsilon$ -entropy from the spectral density  $\Phi(\omega)$  of a stationary Gaussian process with the Kolmogorov formula (2.35), (2.36). The upper diagram corresponds to the first integral (2.35) which gives  $\varepsilon$  in terms of the intermediate quantity  $\theta$ . The lower diagram corresponds to the second integral (2.36) giving the  $\varepsilon$ -entropy.

where  $L' = \text{Number}\{x_j\}$ . The average of the logarithm of these probabilities over the different reference points  $x_i$  is then calculated

$$\mathcal{H}(\varepsilon, \tau, N) = -\frac{1}{R} \sum_{\{x_i\}} \log P(x_i; \varepsilon, \tau, N), \tag{2.40}$$

where  $R = \text{Number}\{x_i\}$ . The Cohen-Procaccia  $(\varepsilon, \tau)$ -entropy per unit time is then

$$h_{\text{CP}}(\varepsilon, \tau) = \frac{1}{\tau} \lim_{N \rightarrow \infty} \lim_{R, L' \rightarrow \infty} [\mathcal{H}(\varepsilon, \tau, N + 1) - \mathcal{H}(\varepsilon, \tau, N)]. \tag{2.41}$$

Contrary to the Shannon-Kolmogorov  $(\varepsilon, \tau)$ -entropy per unit time, no infimum is here taken but the entropy is defined by the average (2.40).

For chaotic systems, it is known that

$$h_{\text{KS}} = \lim_{\varepsilon, \tau \rightarrow 0} h_{\text{CP}}(\varepsilon, \tau). \tag{2.42}$$

We shall show that the Cohen-Procaccia method can also be applied to noise signals where  $h_{\text{CP}}(\varepsilon, \tau)$  diverges at  $\varepsilon, \tau \rightarrow 0$  in the same way as the Shannon-Kolmogorov  $(\varepsilon, \tau)$ -entropy per unit time but with a different prefactor.

In the following section, we shall evaluate the  $(\varepsilon, \tau)$ -entropy per unit time for different stochastic processes in order to compare their degree of dynamical randomness.

### 3. Evaluation of the $(\epsilon, \tau)$ -entropy for random processes

#### 3.1. Deterministic processes

Let us suppose that the process is governed by the differential equation system

$$dX/dt = F(X), \quad X \in \mathbb{R}^d, \tag{3.1}$$

in the phase space  $\mathbb{R}^d$ . The flow induces a mapping

$$X(t) = \Phi^t(X_0), \tag{3.2}$$

in the phase space according to Cauchy's theorem.

A statistical ensemble of systems can be represented by a nonstationary density which evolves in time according to the Liouville equation

$$\partial_t \mu + \text{div } F \mu = 0. \tag{3.3}$$

The solution of this equation can be written as

$$\mu(x; t) = \int_{\mathbb{R}^d} \delta[x - \Phi^t(x_0)] \mu_0(x_0) dx_0, \tag{3.4}$$

in terms of the mapping (3.2) and of the initial density  $\mu_0$ .  $\delta(z)$  denotes the  $d$ -dimensional Dirac distribution.

In order to obtain the joint probability (2.2) we need an invariant probability, the density of which is solution of  $\partial_t \mu_{\text{st}} = 0$ . This density may be a continuous function as in Hamiltonian systems. In dissipative systems, the density is a distribution since the support of the invariant probability is a chaotic attractor of zero  $d$ -volume in phase space [4].

Among the many ergodic probabilities, we select the natural invariant probability which is defined as the noiseless limit of the unique invariant probability of the Langevin process given by adding a white noise to the right hand side of (3.1) [19]. In Axiom A systems, this natural invariant probability is also obtained by weighting each unstable periodic orbit with a mass of probability inversely proportional to its stability eigenvalues [19].

The joint probability density is then given by a distribution as follows:

$$\begin{aligned} P(\mathbf{x}; d\mathbf{x}, \tau, N) &= \text{Pr}\{x_0 < X(0) < x_0 + d\mathbf{x}_0, \dots, x_{N-1} < X(N\tau - \tau) < x_{N-1} + d\mathbf{x}_{N-1}\} \\ &= \delta[x_{N-1} - \Phi^\tau(x_{N-2})] \dots \delta[x_1 - \Phi^\tau(x_0)] \mu_{\text{st}}(x_0) dx_0 \dots dx_{N-1}. \end{aligned} \tag{3.5}$$

The mapping (3.2) can then be used to calculate the  $(\epsilon, \tau)$ -entropy per unit time. The function  $h(\epsilon) = \lim_{\tau \rightarrow 0} h_P(\epsilon, \tau)$  presents a plateau so that the partition  $(\epsilon, \tau)$ -entropy per unit time converges to the KS entropy per unit time in the limit

$$h_{\text{KS}} = \lim_{\epsilon, \tau \rightarrow 0} h_P(\epsilon, \tau) \tag{3.6}$$

(see fig. 5).

The Pesin theorem [20] states that the KS entropy per unit time of bounded systems is equal to the sum of the positive Lyapunov exponents  $\{\lambda_i\}$ ,

$$h_{\text{KS}} = \sum_{\lambda_i > 0} \lambda_i. \tag{3.7}$$

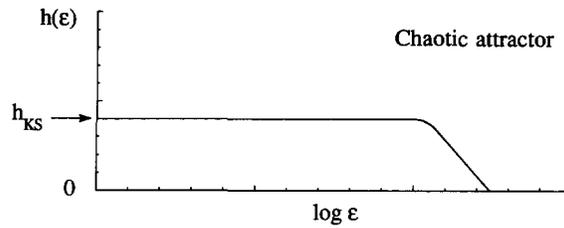


Fig. 5. Schematic behavior of the  $\varepsilon$ -entropy per unit time for deterministic chaos. For small  $\varepsilon$ , the  $\varepsilon$ -entropy reaches a plateau at the value of  $h_{KS}$ .

This theorem establishes a relationship between sensitivity to initial condition and dynamical randomness. When  $h_{KS}$  is positive we talk about chaotic systems.

Figure 6 shows two examples of numerical evaluation of the  $\varepsilon$ -entropy for the 1D logistic and the 2D Hénon map [21] using the the Cohen–Procaccia method.

The  $\varepsilon$ -entropy per unit time is useful to describe systems with very different scales. As an illustration let us consider a mapping like

$$X_{t+1} = 1 - 4(X_t - 1)(X_t - 1 - a), \tag{3.8}$$

We assume that  $a \ll 1$ , so that a chaotic attractor confined in  $1 < X_t < 1 + a$  is very thin (fig. 7). Consequently, very small cells are necessary to resolve the chaos in this system: the diameter of the cells must be  $\varepsilon \ll a \ll 1$ . This leads to the following apparent paradox: as long as our observing device has a resolution  $\varepsilon > 2a$ , the trajectory would always appear to visit the same cell,  $X \in (1 - \varepsilon/2, 1 + \varepsilon/2)$ . Hence the attractor would seem periodic, and the  $\varepsilon$ -entropy would then be nil. Nevertheless, the  $\varepsilon$ -entropy reaches the value  $h_{KS} = \log 2$  when  $\varepsilon \ll a$  (see fig. 7).

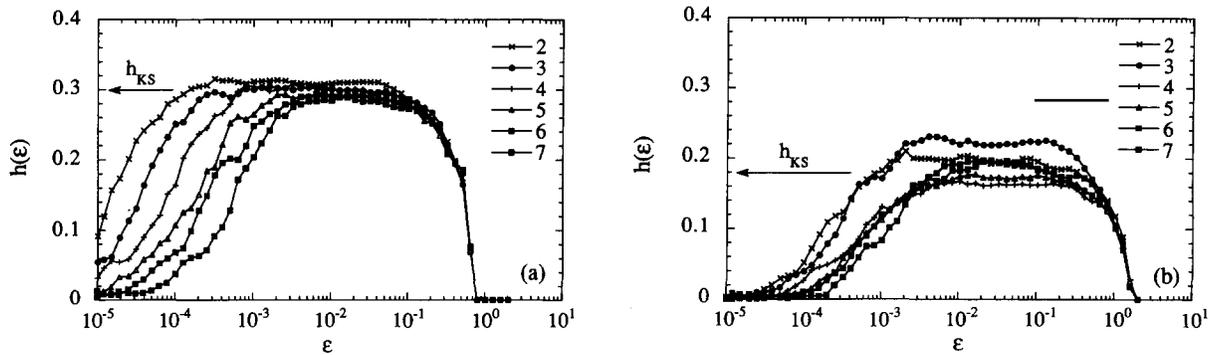


Fig. 6. (a) Numerical evaluation of the KS entropy per unit time for the one-dimensional logistic map  $X_{t+1} = 4X_t(1 - X_t)$  using the Cohen–Procaccia method. The known value of the KS entropy is  $h_{KS} = \lambda = \log 2 = 0.30$  digits/iteration. (b) The same for the  $X$ -component of the two-dimensional Hénon map  $X_{t+1} = 1 + Y_t - 1.4X_t^2$ ,  $Y_{t+1} = 0.3X_t$  where  $h_{KS} = \lambda = 0.18$  digits/iteration. The different curves correspond to the calculation of the entropy for time sequences of lengths  $N = 2 - 7$  [cf. (2.37)–(2.41)]. Since the total length of the time series is limited to  $L = 131072$ , the probability to find a string of length  $N$  decreases as  $N$  increases. As a consequence, the statistics of long strings diminishes at small  $\varepsilon$  and the plateau of the corresponding curve shrinks as seen in the figures.

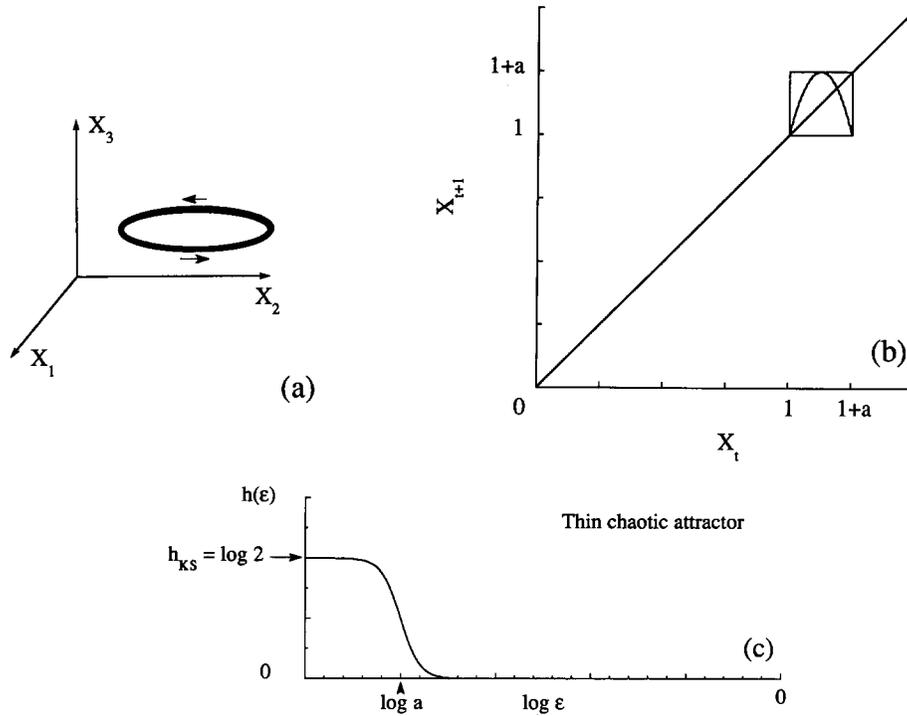


Fig. 7. Schematic behavior of the  $\varepsilon$ -entropy in the case of an extremely thin chaotic attractor. (a) Schematic thin attractor in phase space. (b) One-dimensional first return map (3.8) in a Poincaré section transverse to the thin attractor (a). (c) Behavior of the  $\varepsilon$ -entropy which becomes positive only below  $\varepsilon = a$  where it takes the value  $h_{KS} = \log 2$ .

### 3.2. Bernoulli and Markov chains

These random processes are defined by a matrix of transition probabilities  $P_{\alpha\beta}$  satisfying

$$\sum_{\beta} P_{\alpha\beta} = 1, \quad \sum_{\alpha} p_{\alpha} P_{\alpha\beta} = p_{\beta}, \quad (3.9)$$

where  $\{p_{\alpha}\}$  is the invariant probability vector giving the probabilities to find the system in the states  $\alpha \in \{1, \dots, M\}$  [22].

Markov chains are time and amplitude discrete random processes. The random events occur at time intervals separated by the non-infinitesimal average time constant  $\bar{T}$ .

Bernoulli chains are particular Markov chains for which

$$P_{\alpha\beta} = p_{\beta}. \quad (3.10)$$

The joint probability to visit successively different states is

$$P(\omega_0 \cdots \omega_{N-1}) = p_{\omega_0} P_{\omega_0 \omega_1} \cdots P_{\omega_{N-2} \omega_{N-1}}. \quad (3.11)$$

Accordingly, the  $(\varepsilon, \tau)$ -entropy per unit time is a constant giving the KS entropy per unit time [16, 22]

$$h_{KS} = -\bar{T}^{-1} \sum_{\alpha} p_{\alpha} \log p_{\alpha} \quad (\text{Bernoulli chains}), \quad (3.12)$$

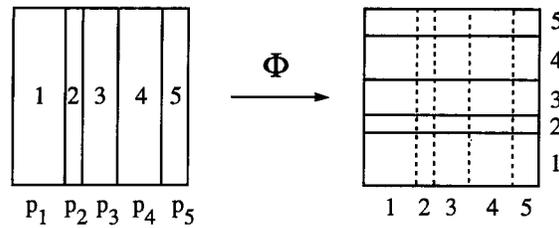


Fig. 8. Two-dimensional map of the unit square which is isomorphic to the Bernoulli process of probabilities  $\{p_\alpha\}$ .

$$h_{KS} = -\bar{T}^{-1} \sum_{\alpha} p_{\alpha} P_{\alpha\beta} \log P_{\alpha\beta} \quad (\text{Markov chains}). \tag{3.13}$$

The finiteness of the entropy per unit time has its origin in the discreteness of these random processes in time: the random choices between the different states generate a finite amount of information per time step  $\bar{T}$ .

The  $(\varepsilon, \tau)$ -entropy per unit time is bounded because Markov chains belong to the same class as the previous deterministic systems with a continuous phase space. Indeed, Markov chains are equivalent to area-preserving mappings of the unit square onto itself.

Let us first prove the simpler statement that a Bernoulli chain is equivalent to an area-preserving map of the unit square. We suppose that the unit square is divided into  $M$  vertical rectangles  $\mathcal{R}_\alpha$  of area  $p_\alpha$  and corresponding to the  $M$  states of the chain (see fig. 8). The map acts on these rectangles by horizontal stretching into horizontal rectangles  $\Phi(\mathcal{R}_\alpha)$  of unit width in such a way that the area of each of them is preserved. The height of the horizontal rectangle  $\Phi(\mathcal{R}_\alpha)$  is therefore  $p_\alpha$ . The first horizontal rectangle is placed at the bottom of the square. The next one on top of it and so on. Let us denote by  $\mathcal{R}_\beta$  the vertical rectangles on the right-hand square which correspond to the states  $\beta$  after one iteration  $\Phi$ . The joint probability to visit successively the states  $\alpha$  and  $\beta$  is the area  $A_{\alpha\beta}$  of the rectangle  $(\alpha, \beta) = \Phi(\mathcal{R}_\alpha) \cap \mathcal{R}_\beta$ . This rectangle  $(\alpha, \beta)$  has a height  $p_\alpha$  and a width  $p_\beta$ . Whereupon, the transition probability from the state  $\alpha$  to the state  $\beta$  is

$$P_{\alpha\beta} = A_{\alpha\beta}/p_\alpha = p_\alpha p_\beta / p_\alpha = p_\beta, \tag{3.14}$$

which agrees with the definition (3.10) of the Bernoulli chain.

The construction is similar for the Markov chains (see fig. 9). Vertical rectangles  $\mathcal{R}_\alpha$  of unit height are associated with the different states  $\alpha$ . The areas of these rectangles are given by the invariant probabilities  $\{p_\alpha\}$ , which are thus equal to their widths. The vertical division lines of the

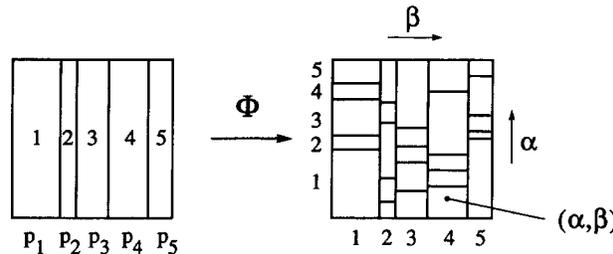


Fig. 9. Two-dimensional map of the unit square which is isomorphic to the Markov chain (3.9) of probabilities  $\{p_\alpha\}$  and of transition matrix  $P_{\alpha\beta}$ .

rectangles are reported on the right-hand square where they now represent the states  $\beta$  after the transition. Let us consider the transition from the state  $\alpha = 1$  to the state  $\beta$ . Since the transition probabilities  $P_{\alpha\beta}$  are now depending on both  $\alpha$  and  $\beta$ , the stretching can no longer be uniform within the rectangle  $\mathcal{R}_1$ . This vertical rectangle  $\mathcal{R}_1$  is therefore mapped on a series of  $M$  small rectangles, each one of them occupying the bottom of each vertical rectangles  $\mathcal{R}_\beta$  but with different heights that we must now determine. The area of the rectangle  $(\alpha, \beta) = \Phi(\mathcal{R}_\alpha) \cap \mathcal{R}_\beta$  divided by the area of the rectangle  $\mathcal{R}_\alpha$  should be equal to the transition probability  $P_{\alpha\beta}$  as for the Bernoulli chains

$$P_{\alpha\beta} = A_{\alpha\beta}/p_\alpha. \quad (3.15)$$

Since the rectangle  $(\alpha, \beta)$  has a width equal to  $p_\beta$  by our previous construction, its height should be equal to

$$A_{\alpha\beta}/p_\beta = (p_\alpha/p_\beta)P_{\alpha\beta}, \quad (3.16)$$

which ends the construction of the equivalent area-preserving map. Summing over the heights of all the rectangles  $(\alpha, \beta)$  in the same vertical rectangle  $\beta$ , we must recover the unity

$$\sum_{\alpha} (p_\alpha/p_\beta)P_{\alpha\beta} = 1, \quad (3.17)$$

which is the case according to the definition (3.9) of the invariant probability vector.

The preceding construction shows that every Markov chain on a countable number of states is isomorphic to a deterministic area-preserving map of the unit square. This result is the converse of a series of results obtained by Sinai [16], Ornstein [23], and others who showed that certain uniformly hyperbolic mappings like the baker transformation or the Arnold cat are isomorphic to a Markov chain [24].

These results are fundamental in showing that Markov chains and deterministic systems are equivalent sources of information. This statement was reinforced by Kolmogorov who proved that the KS entropy is an invariant quantity for isomorphisms between two dynamical processes, i.e. the KS entropy of the Markov chain is equal to the KS entropy of the isomorphic deterministic system [16]. Both dynamical processes, the random Markov chains and the deterministic chaotic systems, are therefore sources of information of the same degree of dynamical randomness.

For long, a fundamental difference was made between discrete Markov chains and deterministic systems. In particular, it was shown that Markov chains obey a H-theorem. Nowadays, it is clear that this H-theorem is based on the coarse graining of the chain into its states  $\alpha$  at a single time [25]. If we refine the partition into multiplet states  $(\omega_0, \omega_1, \dots, \omega_{N-1})$  the H-theorem loses its validity if it is applied after the point limit  $N \rightarrow \infty$  is taken and the well-known paradox of the time-invariance of the standard entropy is recovered for the Markov chains.

The reciprocal of the above statement is true. If a deterministic mapping is shown to be equivalent to a Markov chain for an appropriate partition then a H-theorem applies to this coarse-grained partition [26].

From the preceding discussion, we see that a totally new point of view must be adopted. Both Markov chains and deterministic systems must be considered as random processes, if both have a positive KS entropy per unit time. We shall be able to say that one is more random than the other if there is a difference between the KS entropies per unit time of two random processes. Whether the dynamical process is defined by a deterministic map or by a Markov chain becomes irrelevant

to the question of randomness under the condition that the KS entropies of the map and of the Markov chain take the same value.

### 3.3. Birth-and-death processes

These processes are often encountered in kinetic theories [25, 27]. Certainly, the simplest and most well-known among them is the Poisson random process which is defined as follows [28].

The time axis is divided into small time intervals  $\tau$ . During each small interval, several events may occur. There is a random variable  $N_k$  which is the number of events in a given time interval  $k$  and which is defined by the probabilities  $p_n = \Pr\{N_k = n\}$ ,

$$p_0 = 1 - w\tau + O(\tau^2), \quad p_1 = w\tau + O(\tau^2), \quad p_2 = p_3 = \dots = O(\tau^2), \quad (3.18)$$

where  $w$  is the rate at which the events occur. Finally, it is assumed that the random variables  $N_k$  are independent so that the discretized process forms a Bernoulli chain. Since there is a random variable for each infinitesimal time interval  $\tau$ , the process has a  $\tau$ -entropy per unit time which will diverge as  $\tau \rightarrow 0$ . We can calculate it using (3.12) with  $\bar{T}$  here replaced by  $\tau$

$$h(\tau) = -\tau^{-1} \sum_{n=0}^{\infty} p_n \ln p_n, \quad (3.19)$$

and we obtain

$$h(\tau) = w \ln(e/w\tau) + O(\tau), \quad (3.20)$$

where  $e$  is the basis of natural logarithms.

The same result can be obtained in another way using the property that the random variable  $T_k$ , which is the time between two successive events, has an exponential distribution of parameter  $w$

$$\Pr\{T_k > t\} = \exp(-wt). \quad (3.21)$$

The different random times  $T_k$  are independent. The expectation value of these times is  $\langle T \rangle = 1/w$ . The probability density is  $p(t) = w \exp(-wt)$ . A  $\tau$ -entropy per unit time can be evaluated using a discretization of this continuous variable distribution along the  $t$ -axis. Each cell centered on  $t_i = i\tau$  has a probability mass equal to  $P_i = p(t_i)\tau$  and we can use the following definition for the  $\tau$ -entropy per unit time:

$$\tilde{h}(\tau) = -\frac{1}{\langle T \rangle} \sum_{i=0}^{\infty} P_i \ln P_i, \quad (3.22)$$

where the tilde is there to recall that this definition is a priori different from the preceding one (3.19). If we introduce the definitions we made into eq. (3.22) it is easy to see that (3.22) gives the same leading term as (3.20) with the logarithmic divergence as  $\tau \rightarrow 0$ .

The previous analysis can be extended to general birth-and-death processes governed by the Pauli master equation [27],

$$dp_\beta/dt = \sum_{\alpha} p_{\alpha} W_{\alpha\beta}, \quad (3.23)$$

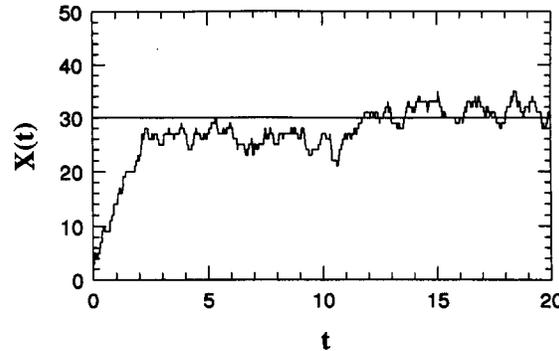


Fig. 10. Typical trajectory of the birth-and-death process  $X \leftrightarrow Y$  with reaction rates  $k_{\pm}$ . The total number of particles is assumed to be constant  $X + Y = N$ . The master equation is  $\dot{p}_X = k_+(X + 1)p_{X+1} + k_-(N - X + 1)p_{X-1} - [k_+X + k_-(N - X)]p_X$ . The constants were chosen to be  $k_+ = 0.2$ ,  $k_- = 0.3$ , and  $N = 50$ . The horizontal line indicates the mean value  $\langle X \rangle = Nk_-/(k_- + k_+) = 30$ .

where  $p_{\alpha}(t)$  is the continuous time probability to be in the state  $\alpha$  while  $W_{\alpha\beta}$  is the rate for a transition between the states  $\alpha$  and  $\beta$ . These rates are here assumed to be constant in time. The conservation of probability implies

$$W_{\alpha\alpha} = - \sum_{\beta(\neq\alpha)} W_{\alpha\beta}. \tag{3.24}$$

Accordingly, the solution of the master equation is

$$\sum_{\alpha} p_{\alpha}(0) P_{\alpha\beta}(t) = p_{\beta}(t), \tag{3.25}$$

with

$$P_{\alpha\beta}(t) = [\exp Wt]_{\alpha\beta}, \tag{3.26}$$

where  $W$  is the matrix composed of the elements  $W_{\alpha\beta}$ . Figure 10 shows a typical trajectory of a birth-and-death process.

For the purpose of calculating the  $\tau$ -entropy per unit time, we shall follow the first method. If we discretize the time axis the probability to be in the state  $\alpha$  at the time  $t_k = k\tau$  obeys a Markov chain according to (3.25). The transition probabilities (3.26) are expanded in Taylor series

$$P_{\alpha\beta}(\tau) = \delta_{\alpha\beta} + W_{\alpha\beta}\tau + O(\tau^2). \tag{3.27}$$

As for the Poisson process, the  $\tau$ -entropy per unit time is defined using (3.13) with  $\bar{T}$  replaced by  $\tau$ ,

$$h(\tau) = -\tau^{-1} \sum_{\alpha\beta} p_{\alpha}^0 P_{\alpha\beta}(\tau) \ln P_{\alpha\beta}(\tau), \tag{3.28}$$

where  $p_{\alpha}^0$  are the stationary probabilities, solutions of

$$\sum_{\alpha} p_{\alpha}^0 W_{\alpha\beta} = 0. \tag{3.29}$$

In the limit of small  $\tau$ , we get

$$h(\tau) = \left( \sum_{\alpha \neq \beta} p_\alpha^0 W_{\alpha\beta} \right) \ln(e/\tau) - \sum_{\alpha \neq \beta} p_\alpha^0 W_{\alpha\beta} \ln W_{\alpha\beta} + O(\tau), \quad (3.30)$$

so that the divergence is logarithmic in  $\tau$

$$h(\tau) \sim \ln(1/\tau), \quad (3.31)$$

like for the Poisson process.

Hence, we are in the presence of a  $\tau$ -entropy per unit time. The discreteness of the random variables suppresses the need of any dependence on  $\varepsilon$ . However, the process being continuous in time we need to keep the dependence on the small time interval  $\tau$ . As  $\tau \rightarrow 0$ , the entropy per unit time slowly increases toward infinity so that the birth-and-death processes are more random than the Markov chains and the chaotic systems.

### 3.4. Time-discrete, amplitude-continuous random processes

Let us now suppose that the random process is discrete in time but continuous in space. The random variables  $\{\dots, X_{-1}, X_0, X_1, X_2, \dots\}$  take their values in  $\mathbb{R}^d$  and are defined by the joint probability

$$\begin{aligned} \Pr\{x_0 < X_0 < x_0 + dx_0, \dots, x_{N-1} < X_{N-1} < x_{N-1} + dx_{N-1}\} \\ = p(x_0, \dots, x_{N-1}) dx_0 \dots dx_{N-1}. \end{aligned} \quad (3.32)$$

#### 3.4.1. Independent variables

Let us first suppose that the variables are independent in time so that

$$p(x_0, \dots, x_{N-1}) = p(x_0) \dots p(x_{N-1}). \quad (3.33)$$

We calculate the  $\varepsilon$ -entropy per unit time (2.13) for a partition of  $\mathbb{R}^d$  into small hypercubes  $C_j$  of side  $\varepsilon$ ,

$$h_{\mathbb{P}}(\varepsilon) = - \sum_j \left[ \int_{C_j} p(x) dx \right] \log \left[ \int_{C_j} p(x) dx \right], \quad (3.34)$$

where  $\mathbb{P}$  refers to the particular partition into hypercubes. We see that the  $(\varepsilon, \tau)$ -entropy per unit time is here independent of  $\tau$ .

Assuming that the density  $p(x)$  admits second derivatives, we have

$$\int_{C_j} p(x) dx = \varepsilon^d p(x_j) + O(\varepsilon^{d+2}), \quad (3.35)$$

where  $x_j$  is the center of the hypercube  $C_j$ . The partition  $\varepsilon$ -entropy per unit time is therefore

$$h_{\mathbb{P}}(\varepsilon) = d \log(1/\varepsilon) - \int_{\mathbb{R}^d} dx p(x) \log p(x) + O(\varepsilon). \quad (3.36)$$

We recall that  $d$  is the dimensionality of the space where the random variables  $X_k$  take their values. Compared with the chaotic processes, we see that the time-discrete amplitude-continuous processes are more random since their  $\varepsilon$ -entropy per unit time diverges as  $\varepsilon \rightarrow 0$ . This divergence originates from the fact that the probability distribution is continuous. Indeed, the outcome of each new random event can arise in any small interval of  $\mathbb{R}^d$  irrespective of the previous outcome. This is in contrast with chaotic systems which are constrained by the deterministic differential equation so that if the trajectory is passing in an infinitesimal ball of size  $\varepsilon$  of a Poincaré surface of section, the next passage is constrained to occur within a ball which is determined by the flow and which is still of infinitesimal size  $\varepsilon \exp(\lambda \bar{T})$ . The Lyapunov exponent  $\lambda$  limits the  $\varepsilon$ -entropy to a finite value, independent of  $\varepsilon$ .

The result (3.36) can be generalized to processes composed of independent random variables  $X_k$  whose probability measure has a fractal set for support. If we suppose that  $d_I$  is the information dimension of this measure the partition  $\varepsilon$ -entropy per unit time (3.34) is then given by

$$h_P(\varepsilon) \simeq d_I \log(1/\varepsilon). \tag{3.37}$$

The result (3.37) can also be interpreted as the limit of the  $(\varepsilon, \tau)$ -entropy per sampling time  $\tau$  of a chaotic attractor when the time interval  $\tau$  between the successive observations is taken to be larger than the relaxation time of the chaotic attractor and when  $\varepsilon$  is not too small. The time series is then a sequence of quasi-independent random variables taking their values on a fractal set.

When the process is moreover assumed to be Gaussian, the Shannon-Kolmogorov  $\varepsilon$ -entropy (2.29) can be calculated for a squared-error cost function  $\rho(z) = z^2$ . We have [9]

$$h_{SK}(\varepsilon) = d \log(1/\varepsilon) - \int_{\mathbb{R}^d} dx p(x) \log p(x) - d \log \sqrt{2\pi\varepsilon} + O(\varepsilon), \tag{3.38}$$

where  $p(x)$  is the  $d$ -dimensional Gaussian density.

When  $d = 1$ , this result can also be obtained by virtue of (2.35), (2.36). Since the random variables are independent we have  $\langle X_k X_0 \rangle = \langle X^2 \rangle \delta_{n0}$  so that  $\Phi(\omega) = \langle X^2 \rangle$  for all frequencies in  $(-\pi, +\pi)$  according to (2.32). We get

$$h_{SK}(\varepsilon) = \log(\langle X^2 \rangle^{1/2}/\varepsilon). \tag{3.39}$$

For time-discrete processes with independent variables, we can also evaluate the Cohen-Procaccia  $\varepsilon$ -entropy per unit time. When  $d = 1$ , the joint probability (2.3) appearing in (2.39) is given by

$$\begin{aligned} P(\mathbf{x}; \varepsilon, N) &= \Pr\{\text{dist}[x_0, X_0] \leq \varepsilon \dots \text{dist}[x_{N-1}, X_{N-1}] \leq \varepsilon\}, \\ &= \prod_{k=0}^{N-1} \int_{C(x_k, \varepsilon)} p(y) dy \simeq (2\varepsilon)^{Nd} \prod_{k=0}^{N-1} p(x_k), \end{aligned} \tag{3.40}$$

where  $\text{dist}[\cdot, \cdot]$  is the maximum distance (2.38) but here in  $\mathbb{R}^d$ .  $C(x_k, \varepsilon)$  denotes the hypercubic cell  $\text{dist}[x_k, X_k] \leq \varepsilon$  of side  $2\varepsilon$ . The  $\varepsilon$ -entropy is now

$$h_{CP}(\varepsilon) \simeq d \log(1/2\varepsilon) - \int_{\mathbb{R}^d} dx p(x) \log p(x), \tag{3.41}$$

which appears to be approximately equal to  $h_P(2\varepsilon)$  up to  $O(\varepsilon)$ . For one-dimensional Gaussian random variables, we have  $h_{CP}(\varepsilon) \simeq \log[(2\pi e \langle X^2 \rangle)^{1/2}/(2\varepsilon)]$ . (See fig. 11.)

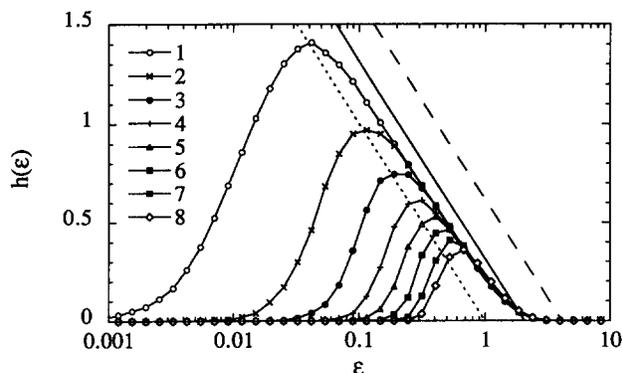


Fig. 11. Numerical evaluation of the CP  $\varepsilon$ -entropy per unit time for the random process of independent one-dimensional Gaussian variables of variance  $\langle X^2 \rangle = 1$ . The numerical CP  $\varepsilon$ -entropy is compared with its theoretical value (3.41) (solid line) as well as with the P (long dashed line) and the SK (short dashed line)  $\varepsilon$ -entropies respectively given by (3.36) and (3.39) ( $d = 1$ ). The different curves correspond to the calculation of the entropy for time sequences of lengths  $N = 1-8$  [cf. (2.37)–(2.41)].

Comparing eqs. (3.36), (3.38) and (3.41) for Gaussian processes, we have therefore the inequalities

$$h_{SK}(\varepsilon) \leq h_{CP}(\varepsilon) \leq h_P(\varepsilon), \tag{3.42}$$

for small  $\varepsilon$ . In spite of their differences, these  $\varepsilon$ -entropies have the same dominant behavior in  $d \log(1/\varepsilon)$ . In this regard, the various definitions of  $\varepsilon$ -entropy are equivalent since

$$\lim_{\varepsilon \rightarrow 0} h_P(\varepsilon)/h_{CP}(\varepsilon) = \lim_{\varepsilon \rightarrow 0} h_{CP}(\varepsilon)/h_{SK}(\varepsilon) = \lim_{\varepsilon \rightarrow 0} h_{SK}(\varepsilon)/h_P(\varepsilon) = 1. \tag{3.43}$$

As a consequence, the divergence

$$h_P(\varepsilon) \simeq h_{CP}(\varepsilon) \simeq h_{SK}(\varepsilon) \simeq d \log(1/\varepsilon), \tag{3.44}$$

is intrinsic to the random process and it characterizes the time-discrete amplitude-continuous random processes.

### 3.4.2. Correlated Gaussian variables

The Shannon–Kolmogorov  $\varepsilon$ -entropy can also be calculated for time-discrete stationary processes with correlated Gaussian random variables for which the joint probability density is given by (2.30). We assume that the random variables are one-dimensional,  $d = 1$ . The Shannon–Kolmogorov  $\varepsilon$ -entropy per unit time is here [9]

$$h_{SK}(\varepsilon) = \log(1/\varepsilon) + \frac{1}{4\pi} \int_{-\pi}^{\pi} \log \Phi(\omega) \, d\omega, \tag{3.45}$$

in terms of the spectral density (2.32). The divergence is again in  $\log(1/\varepsilon)$ .

### 3.5. Time-discrete noisy mappings

We suppose that a deterministic mapping is perturbed by some noise described by a time-discrete amplitude-continuous Gaussian random process. The following example illustrates the complexity of this case:

$$Z_t = X_t + aY_t, \quad \text{where} \quad X_{t+1} = 4X_t(1 - X_t), \tag{3.46}$$

while the random variables  $Y_t$  are independent Gaussians of zero mean and unit variance.

When  $a = 0$ , the random process is purely deterministic so that the  $\epsilon$ -entropy per unit time presents a plateau at the constant value of the KS entropy  $h_{KS} = \log 2$  (see fig. 12). However, when  $a$  is not vanishing but small the signal becomes noisy at small scales  $\epsilon$  where the  $\epsilon$ -entropy per unit time now increases like  $\log(a/\epsilon)$  according to (3.39). Nevertheless, at values of  $\epsilon$  which are larger than the size  $a$  of the fluctuations the noise cannot be resolved so that there remains the plateau at  $h(\epsilon) \simeq h_{KS}$ . For larger noise amplitudes  $a$ , the rise in the  $\epsilon$ -entropy starts at larger values of  $\epsilon$  so that the plateau shrinks.

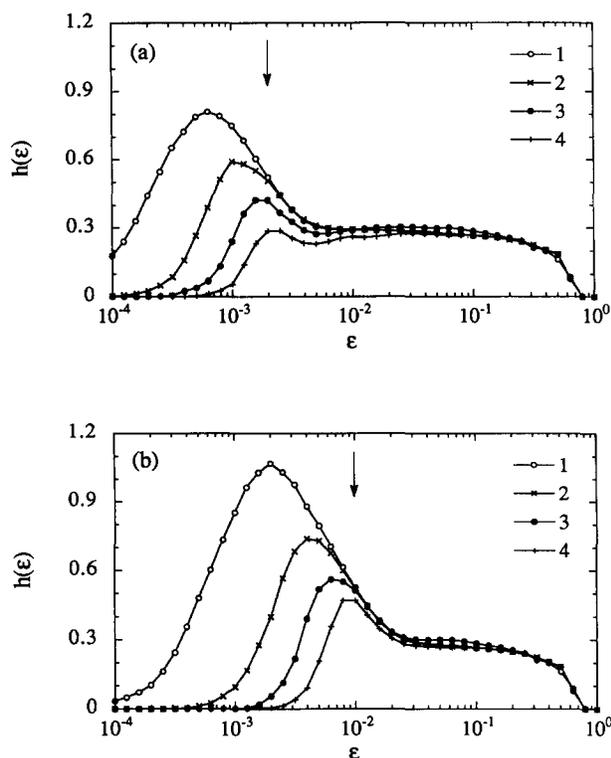


Fig. 12. Numerical CP  $\epsilon$ -entropy for the noisy maps (3.46) with (a)  $a = 0.002$ ; (b)  $a = 0.01$ . The arrows show the corresponding values of the noise amplitude  $a$ . Note the plateau at the value of the KS entropy of the logistic map ( $h_{KS} = \log 2 = 0.30$  digits/iteration) (cf. fig. 6a) followed by the rise as  $\log(1/\epsilon)$  for smaller  $\epsilon$ . The different curves correspond to the calculation of the entropy for time sequences of lengths  $N = 1-4$  [cf. (2.37)–(2.41)].

### 3.6. Time and amplitude continuous random processes

#### 3.6.1. Langevin processes

These processes are defined by stochastic differential equations like [4]

$$dX/dt = F(X) + \xi(t), \quad (3.47)$$

where  $X \in \mathbb{R}^d$  and the components  $\xi_i(t)$  of  $\xi(t)$  are white noises with

$$\langle \xi_i(t) \rangle = 0, \quad \langle \xi_i(t) \xi_j(t') \rangle = 2D_{ij} \delta(t - t'). \quad (3.48)$$

When these processes are stationary there exists an invariant probability measure, the density of which is the unique stationary solution of the Fokker–Planck equation

$$\partial_t \mu + \operatorname{div} F \mu = \sum_{i,j=1}^d D_{ij} \partial_i \partial_j \mu. \quad (3.49)$$

The general solution of (3.49) is given by

$$\mu(x; t) = \int_{\mathbb{R}^d} G(x, x_0; t) \mu_0(x_0) dx_0, \quad (3.50)$$

in terms of the Green function  $G(x, x_0; t)$  and of the initial density  $\mu_0$ . The joint probability (2.3) is then

$$\begin{aligned} P(x; dx, \tau, N) &= \Pr\{x_0 < X(0) < x_0 + dx_0, \dots, x_{N-1} < X(N\tau - \tau) < x_{N-1} + dx_{N-1}\} \\ &= G(x_{N-1}, x_{N-2}; \tau) \cdots G(x_1, x_0; \tau) \mu_{\text{st}}(x_0) dx_0 \cdots dx_{N-1}. \end{aligned} \quad (3.51)$$

The factorization of the joint probability shows that the stationary Langevin processes are Markovian.

#### 3.6.2. Ornstein–Uhlenbeck process

In order to fix the general ideas, we consider the Ornstein–Uhlenbeck process which describes the velocity of a Brownian particle. It is a stationary Gaussian process defined by the Langevin equation [2]

$$dX/dt + aX = c\xi(t), \quad (3.52)$$

where  $\xi(t)$  is a white noise of zero mean and of correlation  $\langle \xi(0) \xi(t) \rangle = \delta(t)$ . The corresponding Fokker–Planck equation ruling the time evolution of a density of particles is [2]

$$\partial \mu / \partial t - a \partial (x \mu) / \partial x = \frac{1}{2} c^2 \partial^2 \mu / \partial x^2, \quad (3.53)$$

with the diffusion coefficient  $D = c^2/2$ . The Green function of (3.53) is the probability density for the particle to travel from  $x_0$  to  $x$  during the time  $t$ ,

$$G(x, x_0; t) = \frac{1}{c\sqrt{2\pi b}} \exp\{-[x - x_0 \exp(-at)]^2 / 2c^2 b\}, \quad (3.54)$$

with

$$b = (1/2a)[1 - \exp(-at)]. \tag{3.55}$$

Consequently, the variance of  $X$  is

$$\langle X(t)^2 \rangle = c^2/2a, \tag{3.56}$$

while the correlation function is

$$\langle X(t)X(0) \rangle = (c^2/2a) \exp(-a|t|), \tag{3.57}$$

so that the spectral density (2.33) is

$$\Phi(\omega) = c^2/(\omega^2 + a^2). \tag{3.58}$$

The process may also be decomposed into its frequencies according to [29]

$$X(t) = \int_{-\infty}^{\infty} \exp(i\omega t) g(\omega) d\zeta(\omega), \tag{3.59}$$

where  $d\zeta(\omega)$  is a complex random variable such that  $d\zeta^*(\omega) = d\zeta(-\omega)$  and

$$\langle d\zeta/d\omega \rangle = 0, \quad \langle [d\zeta(\omega)/d\omega] d\zeta^*(\omega')/d\omega' \rangle = \delta(\omega - \omega'). \tag{3.60}$$

The complex function  $g(\omega)$  satisfies  $g(\omega) = g^*(-\omega)$  and is given by

$$g(\omega) = c/\sqrt{2\pi}(a + i\omega). \tag{3.61}$$

It is related to the spectral density according to  $\Phi(\omega) = 2\pi|g(\omega)|^2$ .

The process can be described equivalently in terms of the response function which is the Fourier transform of  $g(\omega)$  [29]

$$h(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega t} g(\omega) d\omega = \begin{cases} 0 & t < 0, \\ c \exp(-at) & t > 0, \end{cases} \tag{3.62}$$

so that

$$X(t) = \int_{-\infty}^{\infty} h(t-s)\zeta(s) ds. \tag{3.63}$$

The  $\varepsilon$ -entropy can be evaluated by different methods. First, we derive its general behavior by a qualitative argument. The  $\varepsilon$ -entropy will probe the fine scales of the noise corresponding to the high frequencies where the spectral density decreases like  $\Phi(\omega) \simeq c^2/\omega^2$ . Thus, for small  $\varepsilon$ , the  $\varepsilon$ -entropy does not depend on the relaxation rate  $a$ . We base our first qualitative argument on Einstein's relation  $\langle (\Delta X_t)^2 \rangle \simeq 2Dt$ , which holds on arbitrarily small scales. Cells of size  $\varepsilon$  are crossed on average in a time  $t_\varepsilon \sim \varepsilon^2/2D = \varepsilon^2/c^2$ . Since the  $\varepsilon$ -entropy per unit time has the physical unit of

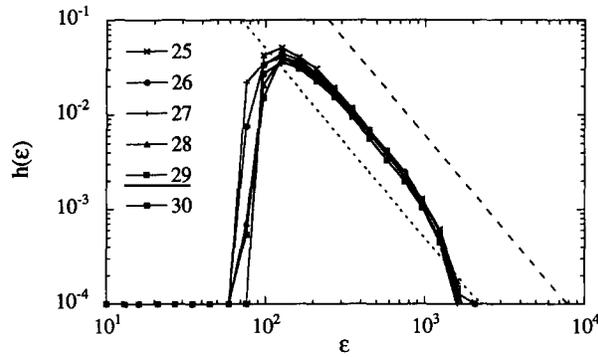


Fig. 13. Numerical CP  $(\varepsilon, \tau)$ -entropy per unit time for the Ornstein-Uhlenbeck process with  $a = 2^{-8}$ ,  $c = 50$ , and  $\tau = 1$ . The numerical values are compared with the SK (short dashed line) and the FP (long dashed line)  $\varepsilon$ -entropies respectively given by (3.65) and (3.76) for the Ornstein-Uhlenbeck process. The different curves correspond to the calculation of the entropy for time sequences of lengths  $N = 25-30$  [cf. (2.37)–(2.41)].

the inverse of a time and measures the number of transitions from cell to cell during one unit of time, we have

$$h(\varepsilon) \sim 1/t_\varepsilon \sim c^2/\varepsilon^2. \quad (3.64)$$

We see that the  $\varepsilon$ -entropy diverges like  $\varepsilon^{-2}$ , much more rapidly than the previous random processes.

Since the Ornstein-Uhlenbeck process is Gaussian, the Shannon-Kolmogorov  $\varepsilon$ -entropy can be estimated by virtue of the Kolmogorov formula (2.35), (2.36). We find

$$h_{\text{SK}}(\varepsilon) \simeq 2c^2/\pi^2\varepsilon^2, \quad (3.65)$$

in nats per unit time, which confirms the  $\varepsilon^{-2}$  dependence.

We have carried out the numerical evaluation of the  $\varepsilon$ -entropy with the Cohen-Procaccia method (see fig. 13). The scaling is confirmed there, but the prefactor turns out to be different because the CP  $\varepsilon$ -entropy is not identical to the SK  $\varepsilon$ -entropy. Indeed, the CP  $\varepsilon$ -entropy is based on the consideration of the probabilities

$$\begin{aligned} & \Pr\{|X_0 - x_0| \leq \varepsilon \dots |X_{N-1} - x_{N-1}| \leq \varepsilon\} \\ & \simeq \frac{(2\varepsilon)^N}{(2\pi)^{N/2}(\det C_N)^{1/2}} \exp\left(-\frac{1}{2}\mathbf{x}^T \cdot C_N^{-1} \cdot \mathbf{x}\right), \end{aligned} \quad (3.66)$$

where the variables  $x_i = x(i\tau)$  represent a given realization of the process, while  $X_i = X(i\tau)$  are the random variables. The correlation matrix is of the form [2]

$$C_N = \frac{c^2}{2a} \begin{pmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \dots & \alpha^{N-1} \\ \alpha & 1 & \alpha & \alpha^2 & \dots & \alpha^{N-2} \\ \alpha^2 & \alpha & 1 & \alpha & \dots & \alpha^{N-3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha^N & \alpha^{N-1} & \alpha^{N-2} & \alpha^{N-3} & \dots & 1 \end{pmatrix}, \quad (3.67)$$

where  $\alpha = \exp(-a\tau)$ . The inverse matrix is

$$C_N^{-1} = \frac{2a}{c^2(1-\alpha^2)} \begin{pmatrix} 1 & -\alpha & 0 & 0 & \dots & 0 & 0 \\ -\alpha & 1+\alpha^2 & -\alpha & 0 & \dots & 0 & 0 \\ 0 & -\alpha & 1+\alpha^2 & -\alpha & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1+\alpha^2 & -\alpha \\ 0 & 0 & 0 & 0 & \dots & -\alpha & 1 \end{pmatrix}, \tag{3.68}$$

in agreement with the Markovian property (3.51). Its determinant is therefore

$$\det C_N = (c^2/2a)^N (1-\alpha^2)^{N-1}, \tag{3.69}$$

so that, for small  $\tau$ ,

$$\det C_N \simeq (c^2\tau)^N / 2a\tau. \tag{3.70}$$

We then obtain the following estimate for the CP  $(\varepsilon, \tau)$ -entropy:

$$h_{CP}(\varepsilon, \tau) \simeq \tau^{-1} \ln(c\sqrt{2\pi\varepsilon\tau}/2\varepsilon) \quad \text{for } \tau \gg \varepsilon^2/c^2. \tag{3.71}$$

However, this result cannot be extended to  $\tau \rightarrow 0$  because the assumption (3.66) is too crude. The approximate equality in (3.66) must be taken cautiously because it ignores correlation effects between the random variables at successive times. If the time  $\tau$  is smaller than the crossing time  $\varepsilon^2/c^2$  of a cell of size  $\varepsilon$ , the particle remains in the cell  $\varepsilon$  because of the correlations  $C_N$ . The ellipsoid of correlations may be elongated so that the actual probabilities may then be larger than in the right-hand member of (3.66). The assumption (3.66) will thus overestimate the actual  $\varepsilon$ -entropy when  $\tau$  is smaller than the crossing time  $\varepsilon^2/c^2$ . Consequently, for  $\tau$  smaller than  $\varepsilon^2/c^2$ , (3.71) cannot be valid otherwise the  $(\varepsilon, \tau)$ -entropy would grow indefinitely as  $\tau \rightarrow 0$ . Nevertheless, the preceding calculation provides a crude estimate of the  $\varepsilon$ -entropy if  $\tau$  is replaced by its value  $\varepsilon^2/c^2$  at the limit of validity of (3.71). We get

$$h_{CP}(\varepsilon) \sim c^2/\varepsilon^2, \tag{3.72}$$

as expected.

Numerically, we found that

$$h_{CP}(\varepsilon) = (1.15 \pm 0.10)c^2/\varepsilon^2, \tag{3.73}$$

in nats per unit time (fig. 13).

Still another evaluation of the  $\varepsilon$ -entropy is based on the first-passage method which is concerned with the random events caused by the first passage of a trajectory of the process at a particular boundary of the system [28, 31]. Indeed, the Cohen–Procaccia method has some similarity with the problem of escape of a pair of Ornstein–Uhlenbeck particles from the interval  $|X(t) - Y(t)| \leq \varepsilon$ . Since the two particles are independent this is equivalent to the escape of an Ornstein–Uhlenbeck particle from the interval  $|X(t)| \leq \varepsilon/\sqrt{2}$ . Moreover, we have [30]

$$\Pr\{|X(t)| \leq \delta; 0 < t < T\} \sim \exp[-\lambda_0(\delta)T], \tag{3.74}$$

where  $\lambda_0(\delta)$  is the smallest eigenvalue of the Ornstein–Uhlenbeck stationary operator (3.53) with absorbing boundaries at  $x = \pm\delta$  where the density of particles is required to vanish. We have

$$\lambda_0(\delta) \simeq \pi^2 c^2 / 8\delta^2, \quad (3.75)$$

so that this  $\varepsilon$ -entropy calculated from a first-passage argument is

$$h_{\text{FP}}(\varepsilon) \simeq \pi^2 c^2 / 4\varepsilon^2, \quad (3.76)$$

in nats per unit time.

The  $\varepsilon$ -entropies per unit time we calculated differ by their prefactor so that we have the inequalities

$$h_{\text{SK}}(\varepsilon) \leq h_{\text{CP}}(\varepsilon) \leq h_{\text{FP}}(\varepsilon), \quad (3.77)$$

for small  $\varepsilon$ . However, they are equivalent in the sense that they all have the same divergence in  $\varepsilon^{-2}$ ,

$$h_{\text{SK}}(\varepsilon) \sim h_{\text{CP}}(\varepsilon) \sim h_{\text{FP}}(\varepsilon) \sim 1/\varepsilon^2, \quad (3.78)$$

which characterizes the dynamical randomness of the Ornstein–Uhlenbeck process.

### 3.6.3. Yaglom noises

These random processes form a whole family of time and amplitude continuous stationary Gaussian processes which embed and generalize the Ornstein–Uhlenbeck process. They are the stationary analogues of Mandelbrot’s fractional Brownian motions [31]. Their correlation function is taken as [29]

$$\langle X(t)X(0) \rangle = (c^2/a\sqrt{2\pi})|at|^H K_H(|at|), \quad (3.79)$$

where  $K_\nu(z)$  is the modified Bessel function of the third kind and of order  $\nu$ . The exponent  $H$  satisfies  $0 < H < 1$ ,  $a$  and  $c$  are positive constants. Examples of Yaglom noises for different values of the exponent  $H$  are given in fig. 14.

The spectral density is

$$\Phi(\omega) = c^2 2^{H-1/2} a^{2H-1} \Gamma(H + \frac{1}{2}) / (a^2 + \omega^2)^{H+1/2}, \quad (3.80)$$

so that the spectral density (3.58) of the Ornstein–Uhlenbeck process is recovered in the limit  $H = 1/2$ .

The function  $g(\omega)$  entering into the formula (3.59) is given by

$$g(\omega) = A / (a + i\omega)^{H+1/2}, \quad (3.81)$$

where the constant is

$$A = c\pi^{-1/2} 2^{H/2-3/4} a^{H-1/2} \sqrt{\Gamma(H + \frac{1}{2})}. \quad (3.82)$$

The response function for (3.63) is then

$$h(t) = \begin{cases} 0, & t < 0, \\ c \frac{2^{H/2-1/4}}{\sqrt{\Gamma(H + \frac{1}{2})}} (at)^{H-1/2} \exp(-at) & t > 0, \end{cases} \quad (3.83)$$

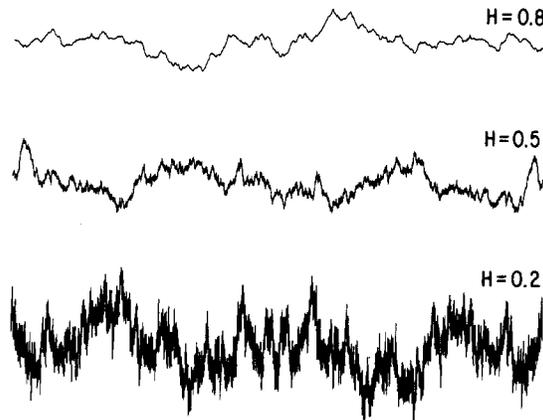


Fig. 14. Typical time signals of Yaglom noises for different values of the exponent  $H$ .

which reduces to the response function (3.62) of the Ornstein–Uhlenbeck process in the limit  $H = 1/2$ .

Let us also mention that the Yaglom noises are continuous in probability with a Hölder exponent  $H$ ,

$$|X(t + \tau) - X(t)| \leq \text{constant} \times \tau^H, \tag{3.84}$$

for almost all realizations  $X(t)$  and for small  $\tau$ .

Applying the Kolmogorov formula (2.35), (2.36) to this stationary Gaussian process we find its  $\varepsilon$ -entropy per unit time to be

$$h_{\text{SK}}(\varepsilon) \simeq a \frac{2H+1}{2\pi} \left( \frac{2^{H-1/2} \Gamma(H+3/2)}{H\pi} \right)^{1/2H} \left( \frac{c}{\varepsilon\sqrt{a}} \right)^{1/H}, \tag{3.85}$$

in nats per unit time, so that (3.65) is recovered when  $H = 1/2$ .

We observe that the exponent of  $\varepsilon^{-1}$  increases indefinitely as  $H \rightarrow 0$ . In that limit, the power spectrum given by (3.80) approaches the  $1/f$ -noise limit. In this sense, the  $1/f$  noise may be said to have the highest degree of dynamical randomness, or to be of maximal complexity.

We have carried out numerical evaluations of the Cohen–Procaccia  $\varepsilon$ -entropy for  $H = 0.4, 0.5$ , and  $0.8$  (see figs. 15–18). The expected exponent  $1/H$  is fairly well reproduced when  $1/2 < H < 1$ . However, the fluctuations become very large when  $0 < H \leq 1/2$  and the exponent is more difficult to obtain. Nevertheless, we observe that, in the case  $H = 0.4$ , the exponent is significantly larger than 2, the value for the Ornstein–Uhlenbeck process (fig. 18).

The Yaglom noises also exist for the Hölder exponent  $H$  larger than 1. The trajectories of these processes become smoother as the exponent  $H$  increases.

### 3.6.4. Fractional Brownian motions

The  $\varepsilon$ -entropy per unit time can also be estimated for these nonstationary time and amplitude continuous Gaussian processes [31]. The Brownian motion in one dimension is obtained from the Ornstein–Uhlenbeck processes in the limit  $a = 0$

$$dX/dt = c\xi(t), \tag{3.86}$$

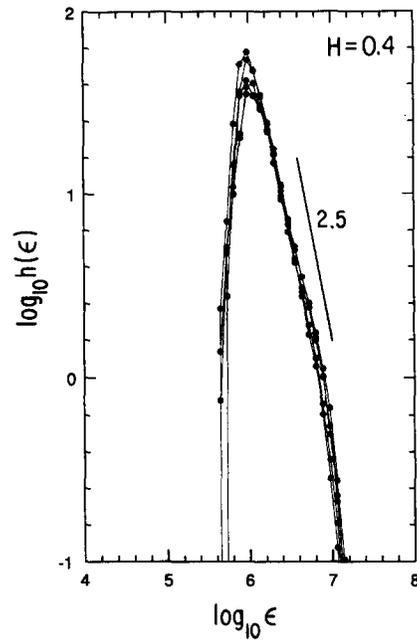


Fig. 15. Numerical CP  $(\varepsilon, \tau)$ -entropy per unit time for the Yaglom noise with  $H = 0.4$ . The different curves correspond to the calculation of the entropy for time sequences of lengths  $N = 31, 33, 35, 37, 39$  [cf. (2.37)–(2.41)].

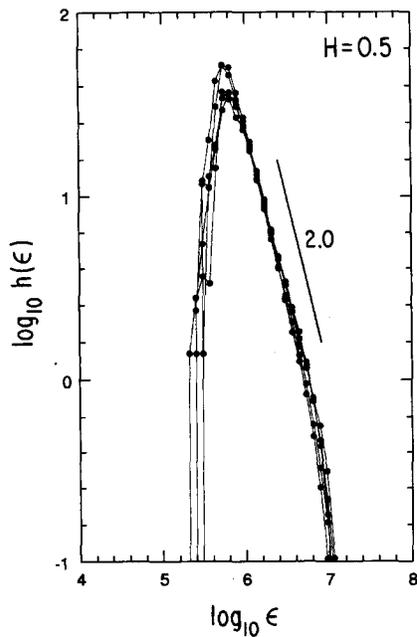


Fig. 16. Same as fig. 15 when  $H = 0.5$ .

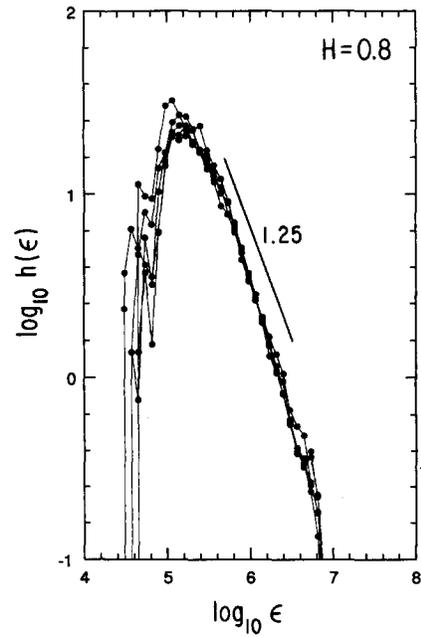


Fig. 17. Same as Fig. 15 when  $H = 0.8$ .

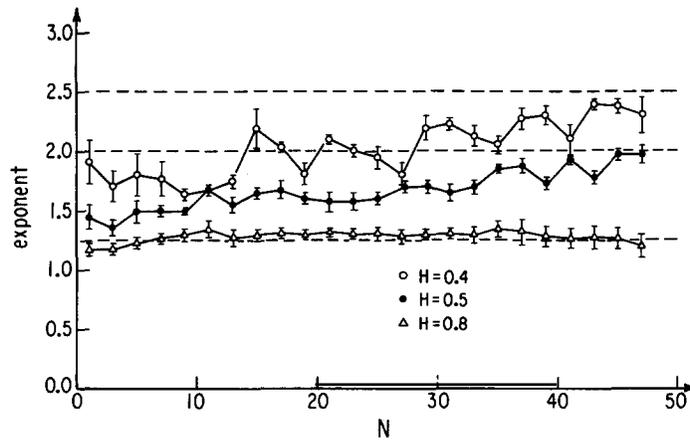


Fig. 18. Scaling exponent of the  $\epsilon$ -entropy versus the length  $N$  of the sequences used in (2.41) for the three different Yaglom noises with  $H = 0.4, 0.5, 0.8$  of figs. 15–17.

for which the Fokker–Planck equation is simply the diffusion equation

$$\partial \mu / \partial t = \frac{1}{2} c^2 \partial^2 \mu / \partial x^2. \tag{3.87}$$

The nonstationarity manifests itself in the fact that the particle wanders in space at arbitrarily large distances so that there is no normalizable invariant probability density in space.

The average square of the signal is

$$\langle [X(t) - X(0)]^2 \rangle = c^2 |t|, \tag{3.88}$$

for the Brownian motion with  $H = 1/2$ , but

$$\langle [X(t) - X(0)]^2 \rangle = \frac{c^2 \sqrt{\pi}}{a \sqrt{2} \sin(\pi H) \Gamma(H + 1)} \left( \frac{|at|}{\sqrt{2}} \right)^{2H}, \tag{3.89}$$

in general for  $0 < H < 1$ , which is obtained in the limit  $|at| \ll 1$  of the correlation function (3.79) of the Yaglom noises.

Mandelbrot and Van Ness gave the following definition for the fractional Brownian motion [31]

$$\begin{aligned} \tilde{X}_H(t) = \frac{B}{\Gamma(H + 1/2)} & \left( \int_{-\infty}^0 (|t - s|^{H-1/2} - |s|^{H-1/2}) \xi(s) ds \right. \\ & \left. + \int_0^t |t - s|^{H-1/2} \xi(s) ds \right), \end{aligned} \tag{3.90}$$

in terms of the white noise  $\xi(t)$  and where  $\beta$  is an appropriate constant. The correlation function is then

$$\langle \tilde{X}_H(t) \tilde{X}_H(s) \rangle = C (|t|^{2H} + |s|^{2H} - |t - s|^{2H}), \tag{3.91}$$

where  $C = B^2 \Gamma(1 - 2H) \cos \pi H / (2\pi H)$  [32].

The fractional Brownian motions have power spectra in  $\omega^{-(2H+1)}$ . Accordingly, their  $\varepsilon$ -entropy per unit time satisfy

$$h(\varepsilon) \sim (1/\varepsilon)^{1/H}. \quad (3.92)$$

Berger has calculated  $h_{\text{SK}}(\varepsilon)$  for the Brownian motion  $H = 1/2$  and found the same result (3.65) as for the stationary Ornstein–Uhlenbeck process [33]. We explain this feature by the fact that the small scales of the Ornstein–Uhlenbeck process are identical with those of the Brownian motion. We conclude this section with the comment that the stationarity or nonstationarity of these time and amplitude continuous Gaussian processes does not modify the  $\varepsilon$ -entropy per unit time.

### 3.7. White noise

The white noise is a stationary Gaussian random process which is a distribution, rather than a function, with respect to time. The random variable  $\xi(t)$  of the white noise is of zero mean and its correlation is  $\langle \xi(0)\xi(t) \rangle = \delta(t)$ .

We have already seen that processes whose realizations are continuous functions have an  $(\varepsilon, \tau)$ -entropy which is independent of  $\tau$ . On the other hand, for the birth-and-death processes whose realizations are discontinuous functions of time, the  $(\varepsilon, \tau)$ -entropy depends on  $\tau$ , in that case smoothly like  $\log(1/\tau)$ . We may thus expect that the white noise will have a strong dependence on  $\tau$ .

A random process approximating the white noise can be constructed by discretization of the time axis into small intervals  $\tau$  and by considering that  $X_k = \xi(k\tau)$  are independent Gaussian variables of zero mean and unit variance. Therefore, using the result (3.44) of section 3.4 which gives the entropy per time interval  $\tau$ , we find the entropy per unit time to be

$$h(\varepsilon, \tau) \simeq \tau^{-1} \log(1/\varepsilon) \quad (\text{white noise}), \quad (3.93)$$

which increases indefinitely as  $\tau \rightarrow 0$ .

Another method to generate random functions which approximate the white noise is to consider a Gaussian process with a constant spectral density in a large but finite bandwidth  $|\omega| < \omega_c$  up to an ultraviolet cutoff beyond which the spectral density is zero ( $\omega_c = \pi/\tau$ ). The SK  $\varepsilon$ -entropy of this process leads to the same result as (3.93) [10].

### 3.8. Lévy flights

We have recently studied a model of anomalous diffusion by Lévy flight, due to Klafter, Blumen and Shlesinger [34]. In the model, a particle undergoes random walks which consist of straight steps interrupted by jumps. The probability density for a single step  $r$  in time  $t$  is given,

$$\Psi(r, t) = \psi(r)\delta(r - t^\nu), \quad (3.94)$$

and the steps are independent of each other. The function  $\psi$  presents a power law decay  $\psi(r) \sim r^{-\mu}$  as  $r \rightarrow \infty$  and the process is assumed to occur in a  $d$ -dimensional space. The realization of these processes is a trajectory in space which is continuous and piecewise linear. The velocity is then piecewise constant with discontinuities at the jumps. Different regimes exist according to the exponent  $\alpha = \nu(\mu - d + 1) - 1$ .

The  $\varepsilon$ -entropy  $H(\varepsilon, T)$  for the velocity process during a time interval  $T$  was found to be

$$H(\varepsilon, T) = \begin{cases} Td \log(1/\varepsilon) & \text{for } \alpha > 1, \\ \frac{T}{\log T} d \log(1/\varepsilon) & \text{for } \alpha = 1, \\ T^\alpha d \log(1/\varepsilon) & \text{for } \alpha < 1. \end{cases} \quad (3.95)$$

Accordingly, the process is regularly random for  $\alpha > 1$  with a positive  $\varepsilon$ -entropy per unit time, but it is sporadically random for  $\alpha \leq 1$  where the  $\varepsilon$ -entropy per unit time vanishes

$$h(\varepsilon) = \begin{cases} d \log(1/\varepsilon) & \text{for } \alpha > 1, \\ 0 & \text{for } \alpha \leq 1. \end{cases} \quad (3.96)$$

We see that when  $\alpha > 1$  the process has the same degree of randomness in  $\log(1/\varepsilon)$  as the time discrete process considered in section 3.4. We remark also that, according to (3.94), the choice of a random vector  $r$  in  $\mathbb{R}^d$  is equivalent to a random choice of a time interval  $t$  and to a random direction on the  $(d-1)$ -sphere. Therefore  $\varepsilon^d \sim \tau \delta^{d-1}$  where  $\tau$  is an infinitesimal time interval while  $\delta$  is the diameter of an infinitesimal cell of the  $(d-1)$ -sphere. When  $d = 1$  and  $\alpha > 1$ , we have equivalence with the type of processes of section 3.3 like the Poisson process.

#### 4. Application to hydrodynamic turbulence

Fluid turbulence is an important random process which has recently attracted considerable interest. It is usually assumed that the fluid obeys the Navier–Stokes (NS) equations and that the thermodynamic fluctuations do not play a role in the random character of fluid turbulence in incompressible fluids. The fluid randomness is thought to have its origin in the nonlinearity of the NS equations themselves.

The NS equations are supposed to have an attractor whose dimension increases with the Reynolds number. While the attractor may be of low dimensionality near the threshold of turbulence, in the regime of chaos, the dimension soon becomes very large in the regimes of developed turbulences. If an attractor exists in the state space of the fluid, it would be characterized by a finite KS entropy per unit time so that an  $\varepsilon$ -entropy per unit time should ultimately saturate at the value of the KS entropy [35].

However, if this value is very large and if many degrees of freedom are active in the fluid, the actual measure of the KS entropy or of the Hausdorff dimension of the attractor turns out to be practically not possible, because that measure would require to probe a time signal with high resolution on very small temperature or velocity variations and on an extremely long time. On the contrary, with the resolutions and the time series that are currently available experimentally, the  $\varepsilon$ -entropy can be evaluated from the measurements. It turns out to present scaling behaviors showing that the fluid can be effectively described as a stochastic process where the deterministic character has disappeared.

When this noisy character of developed turbulence becomes dominant, the  $\varepsilon$ -entropy is very useful. More important perhaps is the fact that the crossover from the chaotic regime toward various turbulent regimes can be followed with the  $\varepsilon$ -entropy, revealing the way in which the deterministic character disappears. We have applied this idea to the recent experiments on Rayleigh–Bénard convection; where chaos, and two turbulent regimes (soft and hard) have been observed as the Rayleigh number was increased [36]. Our main results reported in ref. [37] may be summarized

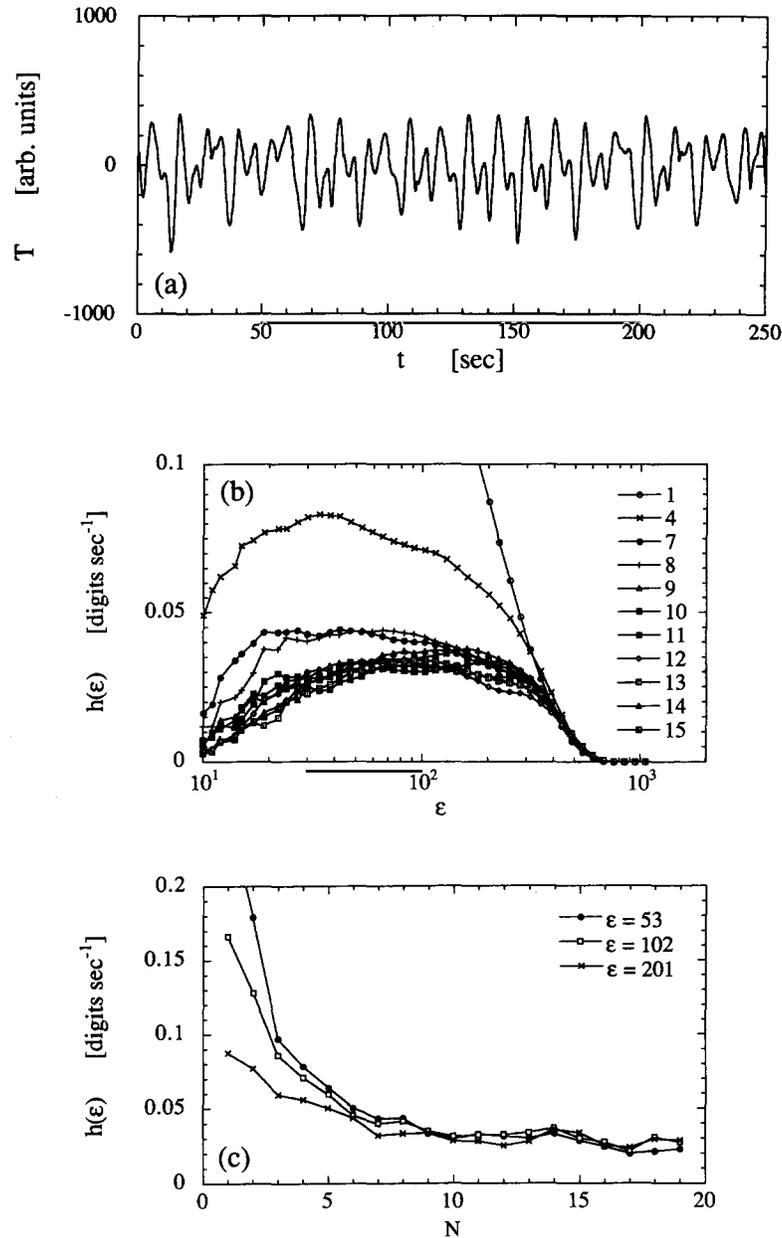


Fig. 19. (a) Temperature versus time in Rayleigh convection in He in a chaotic regime at  $Ra = 1.78 \cdot 10^5$  and inverse diffusivity time  $\kappa/L^2 = 9.4 \cdot 10^{-3}$  Hz. The total length of the time series is of 327 680 points separated by  $\Delta t = 1/10$  s. The experimental data have been kindly provided us by Prof. A. Libchaber (ref. [36]). (b) Numerical CP  $(\epsilon, \tau)$ -entropy per unit time of the signal (a) calculated over 256 reference points with  $\tau = 25\Delta t$ . The different curves correspond to the calculation of the entropy for time sequences of lengths  $N = 1, 4, 7-15$  [cf. (2.37)–(2.41)]. A plateau appears at large  $N$  after a decrease due to aliasing effects. (c) CP  $(\epsilon, \tau)$ -entropy per unit time versus the length  $N$  of the time sequences at different values of  $\epsilon$  in the  $\epsilon$ -interval where the plateau appears.

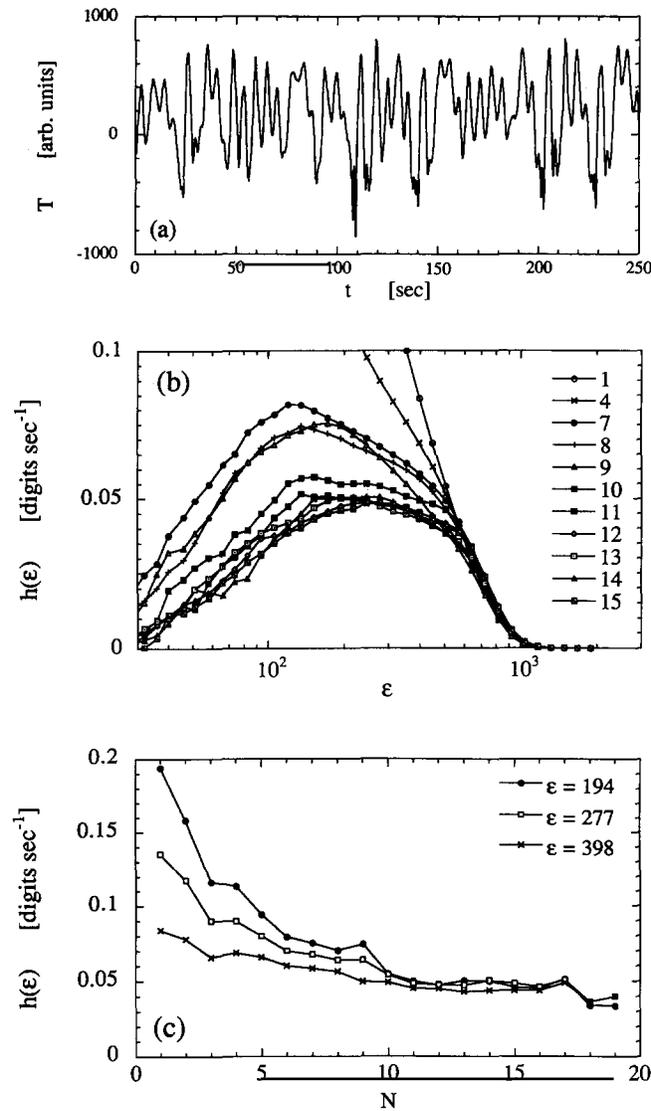


Fig. 20. (a) Temperature versus time in Rayleigh convection in He in a chaotic regime at  $Ra = 1.97 \times 10^5$  and inverse diffusivity time  $\kappa/L^2 = 9.3 \times 10^{-3}$  Hz. The total length of the time series is of 327 680 points separated by  $\Delta t = 1/10$  s (data of Prof. A. Libchaber [36]). (b) Numerical CP  $(\epsilon, \tau)$ -entropy per unit time of the signal (a) calculated over 256 reference points with  $\tau = 20\Delta t$ . Note that the plateau appears at a larger value of the entropy so that this regime is more chaotic than the one of fig. 19. (c) CP  $(\epsilon, \tau)$ -entropy per unit time versus the length  $N$  of the time sequences at different values of  $\epsilon$  in the  $\epsilon$ -interval where the plateau appears.

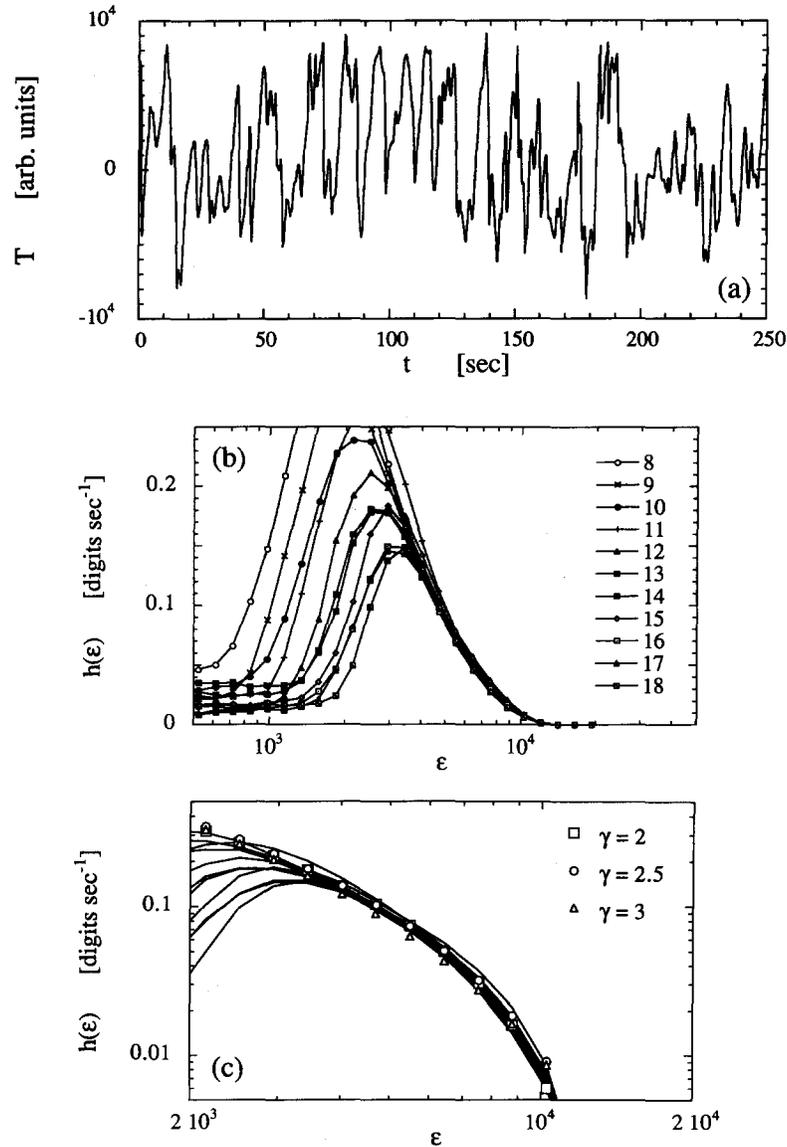


Fig. 21. (a) Temperature versus time in Rayleigh convection in He in a regime of soft turbulence at  $Ra = 9.9 \times 10^5$  and inverse diffusivity time  $\kappa/L^2 = 3.13 \cdot 10^{-3}$  Hz. The total length of the time series is of 204 800 points separated by  $\Delta t = 1/25.6$  s (data of Prof. A. Libchaber [36]). (b) Numerical CP  $(\epsilon, \tau)$ -entropy per unit time of the signal (a) calculated over 256 reference points with  $\tau = 20\Delta t$ . The different curves correspond to the calculation of the entropy for time sequences of lengths  $N = 8-18$  [cf. (2.37)–(2.41)]. Note that the plateau has disappeared whereas the curves are increasing where they coincide, like in fig. 11. (c) Same as (b) but with a logarithmic scale for the entropy axis. Three different fits of the form  $h(\epsilon) = a[\log(\epsilon_0/\epsilon)]^\gamma$  with the exponents  $\gamma = 2, 2.5,$  and  $3$  are superimposed.

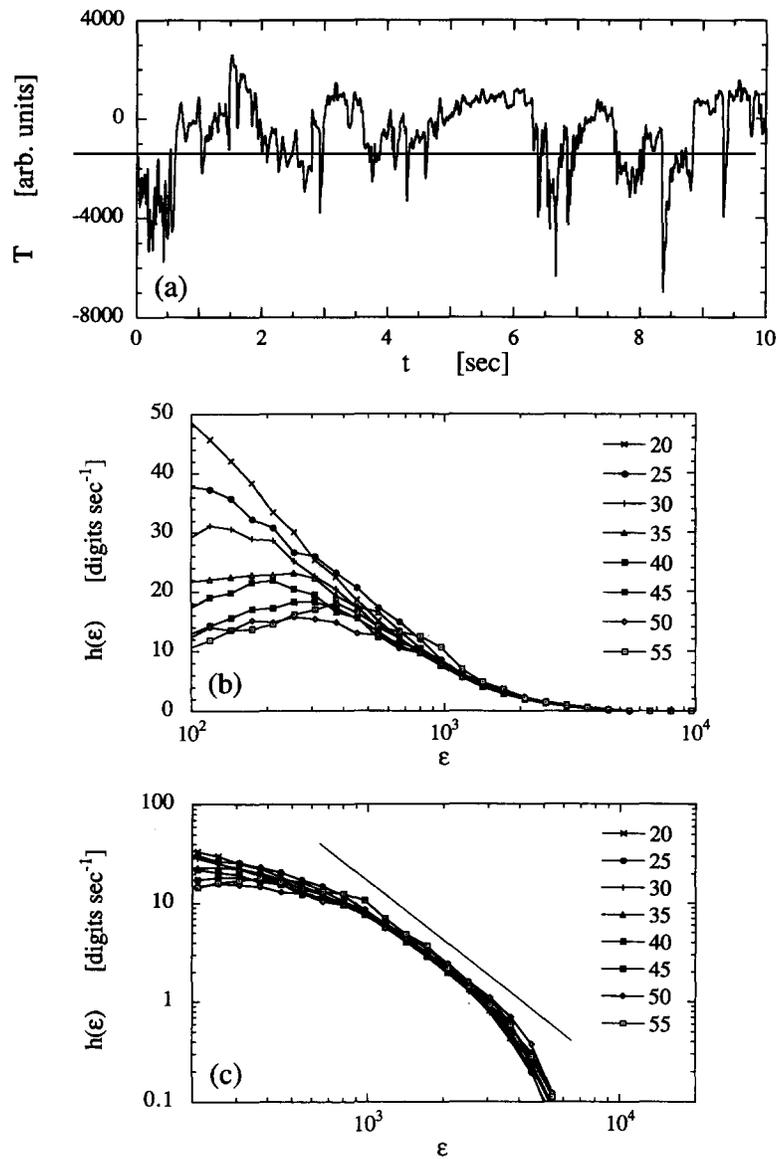


Fig. 22. (a) Temperature versus time in Rayleigh convection in He in a regime of hard turbulence at  $Ra = 1.134 \times 10^{10}$  and inverse diffusivity time  $\kappa/L^2 = 5.127 \times 10^{-5}$  Hz. The total length of the time series is 409 344 points separated by  $\Delta t = 1/640$  s (data of Prof. A. Libchaber [36]). (b) Numerical CP  $(e, \tau)$ -entropy per unit time of the signal (a) calculated over 640 reference points with  $\tau = \Delta t$ . The different curves correspond to the calculation of the entropy for time sequences of lengths  $N = 20, 25, 30, 35, 40, 45, 50, 55$  [cf. (2.37)–(2.41)]. (c) Same as (b) but with a logarithmic scale for the entropy axis. The line has a slope  $-2$  showing the corresponding scaling behavior in this range of values of temperature variations  $e$ , like in fig. 13.

as follows. In order to estimate the  $\varepsilon$ -entropy for a temperature signal,  $\varepsilon$  denotes the size of the temperature fluctuations of a measured time series. In the chaotic regimes at low Rayleigh numbers (figs. 19 and 20), the  $\varepsilon$ -entropy presents a plateau, corresponding to a finite KS entropy

$$h(\varepsilon) \sim h_{\text{KS}} \quad (\text{chaos}). \quad (4.1)$$

In soft turbulence, the plateau has disappeared at high values of the entropy and on fine scales of  $\varepsilon$ . Then, the entropy slowly increases with  $\varepsilon \rightarrow 0$  like

$$h(\varepsilon) \sim [\ln(1/\varepsilon)]^\alpha \quad (\text{soft turbulence}), \quad (4.2)$$

with an exponent  $\alpha \simeq 2 - 3$  (fig. 21).

At still higher Rayleigh numbers in hard turbulence (fig. 22), the  $\varepsilon$ -entropy increases even more rapidly like

$$h(\varepsilon) \sim 1/\varepsilon^2 \quad (\text{hard turbulence}), \quad (4.3)$$

in some intermediate range of values of  $\varepsilon$ . We gave in ref. [37] arguments supporting the behaviors (4.2) and (4.3).

The general increase in the  $\varepsilon$ -entropy shows how the scales where dynamical randomness is active in the fluid changes with the Rayleigh number. If this randomness is small in the chaotic regimes but uniformly distributed on the different scales, it soon becomes much higher in more developed turbulent regimes. A turbulent fluid is characterised by a cascade of dynamical instabilities from large to smaller eddies, the scale-dependent  $(\varepsilon, \tau)$ -entropy can then be used to measure dynamical randomness induced by these instabilities at various scales of the process. Our conclusion ought also to be applicable to other systems of developed fluid turbulence, like the Couette–Taylor cylinder, or open shear flows.

With the advent of the theory of chaos, the KS entropy was introduced which is able to express quantitatively the intuitive concept of dynamical randomness which has been invoked for long in the observation of turbulence. However, with the growing interest in the regimes at higher Rayleigh or Reynolds numbers, the quantitative measure of randomness has once again been overlooked for lack of appropriate concepts. We think that the  $\varepsilon$ -entropy helps us to bridge this gap. It allows to generalize the KS entropy per unit time toward those regimes of higher randomness, while still maintaining the possibility of comparison with regimes of lower degrees of randomness (such as low-dimensional chaos).

## 5. Application to statistical mechanics

In the preceding discussion we viewed turbulence as a deterministic, macroscopic phenomenon of strong spacetime chaos, and did not consider the effects of thermodynamic noise. We are now turning to mesoscopic processes of physics and chemistry, like Brownian motion, where randomness is produced by thermodynamic fluctuations. Here, the situation is profoundly different although it presents similarities with the example of turbulence. Compared with laboratory turbulence, the number of active degrees of freedom is vastly superior in a mole of gas ( $10^{23}$ ). Therefore, although the motion of atoms or molecules in a gas or a liquid is ultimately described by the deterministic Newton equations, the extremely large number of particles renders a stochastic description desirable.

We gave elsewhere an estimation of the KS entropy per unit time in a gas of interacting particles [38]. We also discussed elsewhere about the  $\varepsilon$ -entropy in the classical gas of free particles and

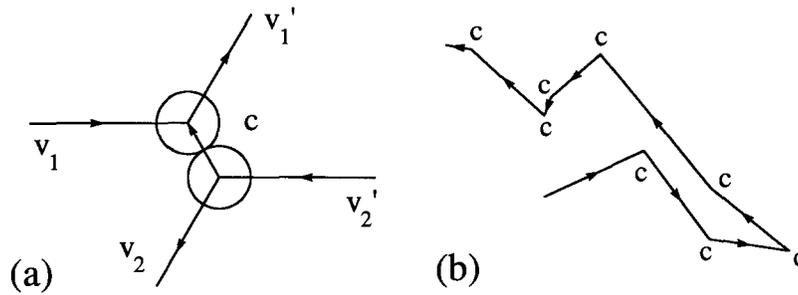


Fig. 23. (a) Schematic representation of a typical collision between two hard spheres of ingoing velocities  $v_1$  and  $v_2$ , and of outgoing velocities  $v_1'$  and  $v_2'$ .  $c$  is the impact unit vector along the line joining the two discontinuities in the trajectories of the particles No. 1 and No. 2. (b) Schematic representation of a typical trajectory of the Boltzmann-Lorentz random process. The trajectory changes its direction at each collision  $c$ . From collision to collision, the process uses six random variables which are continuously distributed: (1) the intercollisional time to which corresponds the infinitesimal  $\tau$ ; (2) the velocity  $v_2$  of the particle No. 2 coming and returning to the thermal bath with a Maxwell distribution and corresponding to the infinitesimal  $\Delta^3 v_2$ ; (3) the angles  $\Omega = (\theta, \varphi)$  of the impact unit vector to which correspond the infinitesimal  $\Delta^2 \Omega$ . Each continuous random variable contributes to the  $(\varepsilon, \tau)$ -entropy by a logarithmic divergence of its corresponding infinitesimal.

how it compares with the entropy per unit time of the quantum gases of free particles [39]. In this paper on the  $(\varepsilon, \tau)$ -entropy, we would like to illustrate our method with the calculation of the  $(\varepsilon, \tau)$ -entropy per unit time for the relaxation of the velocity distribution function toward the equilibrium Maxwellian as described by the Boltzmann equation.

For simplicity, we assume that the relaxation occurs uniformly in space. The linear Boltzmann equation is then [40]

$$\frac{\partial f(v_1)}{\partial t} = \rho \int d^3 v_2 d^2 \Omega |v_1 - v_2| \sigma(\theta, \varphi, v_1, v_2) f_{\text{eq}}(v_2) [f(v_1') - f(v_1)], \tag{5.1}$$

where  $\rho$  is the particle density,  $\sigma$  is the differential cross section of the binary collision, and

$$f_{\text{eq}}(v) = (m\beta/2\pi)^{3/2} \exp(-\frac{1}{2}\beta m v^2). \tag{5.2}$$

Equation (5.1) describes the time evolution of the probability density of the velocity of a test particle No. 1 undergoing multiple collisions with other particles No. 2 in the gas. The outgoing velocities  $v_1'$  and  $v_2'$  after each binary collision is uniquely determined by the velocities  $v_1$  and  $v_2$  of the two particles entering the collision together with the impact unit vector locating the relative positions of the particles No. 1 and No. 2 at the point of the closest approach (fig. 23a).

The Boltzmann-Lorentz equation describes the successive collisions as successive random events where the velocity  $v_2$  of the bath particle as well as the solid angle  $\Omega = (\theta, \varphi)$  of the aforementioned impact unit vector are random variables (fig. 23b). Besides, the Boltzmann-Lorentz equation has the form of a birth-and-death process in the continuous velocity and solid angle variables. Discretizing the velocity and the solid angle, it has the form

$$\frac{d}{dt} p_{v_1} = \sum_{v_1'} W_{v_1 v_1'} p_{v_1'}, \tag{5.3}$$

where

$$p_{v_1} = f(v_1) \Delta^3 v. \tag{5.4}$$

Moreover,

$$W_{v_1, v_1'} = \rho \Delta^3 v \Delta^2 \Omega |v_1 - v_2| \sigma(\Omega, v_1, v_2) f_{\text{eq}}(v_2), \quad (5.5)$$

where  $v_1'$  is a function of  $v_2$  and  $\Omega$ . We shall now estimate the  $(\varepsilon, \tau)$ -entropy per unit time for the stochastic process of the Boltzmann-Lorentz equation in the case of a hard sphere gas. The differential cross section is then constant  $\sigma = a^2/4$  where  $a$  is the radius of the particles. The  $(\varepsilon, \tau)$ -entropy per unit time for a birth-and-death process is given by (3.30). In the present case,  $\Delta^3 v \Delta^2 \Omega$  plays the role of  $\varepsilon$ . We find

$$h(\tau, \Delta^3 v \Delta^2 \Omega) = \left( - \sum_{v_1} p_{v_1}^{\text{eq}} W_{v_1, v_1} \right) \ln(\varepsilon/\tau) - \sum_{v_1 \neq v_1'} p_{v_1}^{\text{eq}} W_{v_1, v_1'} \ln W_{v_1, v_1'} + O(\tau), \quad (5.6)$$

where

$$- \sum_{v_1} p_{v_1}^{\text{eq}} W_{v_1, v_1} = \frac{\rho a^2}{4} \int d^2 \Omega \, d^3 v_1 \, d^3 v_2 |v_1 - v_2| f_{\text{eq}}(v_1) f_{\text{eq}}(v_2), \quad (5.7)$$

while

$$\begin{aligned} \sum_{v_1 \neq v_1'} p_{v_1}^{\text{eq}} W_{v_1, v_1'} \ln W_{v_1, v_1'} &= \frac{\rho a^2}{4} \int d^2 \Omega \, d^3 v_1 \, d^3 v_2 |v_1 - v_2| f_{\text{eq}}(v_1) f_{\text{eq}}(v_2) \\ &\quad \times \ln \left[ \frac{1}{4} \rho a^2 \Delta^3 v \Delta^2 \Omega |v_1 - v_2| f_{\text{eq}}(v_2) \right]. \end{aligned} \quad (5.8)$$

The evaluation of these integrals gives

$$h(\tau \Delta^3 v \Delta^2 \Omega) = 4 \rho a^2 \sqrt{\pi/m\beta} \ln(398.8/\rho a^2 m \beta \tau \Delta^3 v \Delta^2 \Omega) + O(\tau). \quad (5.9)$$

We see that this  $(\varepsilon, \tau)$ -entropy per unit time has the form of the  $\varepsilon$ -entropy per unit time corresponding to the random choice of six continuous random variables at time intervals separated by the mean intercollisional time,  $T_{\text{intercoll}} \sim l/\bar{v}$ .  $l \sim 1/\rho \pi a^2$  is the mean free path while  $\bar{v} \sim (m\beta)^{-1/2}$  is the average velocity. The six continuous random variables are the three velocity components  $v_2$  of the particles of the bath, the two angles  $\Omega = (\theta, \varphi)$  of the impact unit vector, and the random intercollisional time.

The preceding reasoning was done for the linear Boltzmann-Lorentz equation for simplicity but it can also be applied to the Boltzmann equation under the condition that we first map the Boltzmann equation onto a stochastic process. Indeed, the Boltzmann equation cannot be the master equation of a stochastic process because it is nonlinear whereas a master equation must be linear in order to have an interpretation in terms of a random process. Such a master equation for the Boltzmann equation was obtained by several authors, it has the form of a birth-and-death process where the random variables are the numbers  $N_i$  of particles having the velocity  $v_i$ . The average values of  $N_i$  are related to the probability density of the velocity according to  $\bar{N}_i = \rho f(v_i) \Delta^3 v$ . The problem is now mapped onto a birth-and-death process for the binary collisions (see van Kampen [28]). The  $(\varepsilon, \tau)$ -entropy per unit time can then in principle be calculated as before with a similar logarithmic divergence in  $\varepsilon \tau$ .

## 6. Spacetime processes

### 6.1. $(\varepsilon, \tau)$ -entropy per unit time and volume

For processes evolving in time and in space, dynamical randomness can occur not only at each time step but also at each space point. Let us consider a simple example of a chain of spins  $1/2$  which evolves in time according to a pure Bernoulli process. Each spin takes the values  $\pm 1/2$  independently of the state of its neighbors and independently of its own previous state. The amount of data produced by this system will be proportional to the time  $T$  of observation but also to the space volume  $V$  under observation. Since the process is of Bernoulli type no compression of data is possible and the record of the spacetime map will require a total of  $TV$  bits. This quantity defines the entropy  $H(T, V)$  of the process over a time  $T$  and a volume  $V$ .

The results of section 2 can be generalized by replacing the group of time translations used in that section by the the group of translations in time and in space. In this way, we can define the entropy  $H(\varepsilon, \tau, T, V)$  for spacetime processes. This entropy grows at most like  $TV$  so that we can define the  $(\varepsilon, \tau)$ -entropy per unit time and volume according to

$$h^{(\text{time,space})}(\varepsilon, \tau) = \lim_{TV \rightarrow \infty} \frac{1}{TV} H(\varepsilon, \tau, T, V), \quad (6.1)$$

as a generalization of  $h^{(\text{time})}(\varepsilon, \tau)$  defined with (2.13) or (2.29). We shall speak of spacetime chaos if  $h^{(\text{time,space})}(\varepsilon, \tau)$  is positive, and bounded as  $\varepsilon$  and  $\tau$  go to zero.

Let us now review several spacetime processes.

### 6.2. Evaluation of $(\varepsilon, \tau)$ -entropy for spacetime processes

#### 6.2.1. Deterministic cellular automata

In these spacetime processes reactualized by Wolfram [41] in the early eighties, the system is defined in discrete states at discrete positions in space. The dynamics is given by logic rules that reset the configuration of the system from one time step to the next. These processes can be viewed as degenerate Markov chains, where states are spatial configurations, and the matrix of transition probabilities (3.9) contains only zeros and ones. As a consequence, no dynamical randomness is produced with time. The dynamics simply propagates the initial configuration of the system which can eventually be random in space. The randomness is due either to the initial conditions or to unknown information coming from outside the observed volume  $V$ .

Let us consider for instance the elementary cellular automata where the coupling is set between nearest neighboring sites. We have  $H(T, V) \sim V \log T$  for Wolfram's rule 132. For other rules like 0, 32, 250, or 254, the initial state is attracted toward a spatially uniform or periodic configuration, in which case the entropy  $H(V, T)$  increases like  $\log(TV)$ . Besides, there is numerical evidence obtained by Grassberger that the rule 22 has a sporadic behavior in time like  $H(T, V) \sim T^{0.82}$  without dependence on  $V \geq 2$  [42]. In all cases, the entropy per unit time and unit volume (6.1) is vanishing. The same conclusion applies to Conway's game of life [43] as well as to the automata of Nowak and May [44].

#### 6.2.2. Lattice gas automata

These processes are discrete in time and in space, where physical quantities like the velocity take discrete amplitudes [45–47]. Contrary to the deterministic cellular automata, however, the lattice gas automata are non-degenerate Markov chains with transition probabilities between zero and one.

Accordingly, randomness is generated at each spacetime point and the entropy is proportional to the spacetime hypervolume  $TV$  as in the example of subsection 6.1. These processes have in general a positive entropy per unit time and volume which does not depend on either  $\varepsilon$  or  $\tau$  due to the discreteness of these models. The same conclusion applies to *probabilistic* cellular automata.

### 6.2.3. Coupled maps

These dynamical systems are defined by deterministic mappings [48, 49]. A continuous variable is assigned to each space position on a lattice. Therefore, the process is discrete in time and in space but have continuous amplitudes. For a finite number of coupled maps, the entropy will then be given by

$$H(T, V) = T \sum_{\lambda_i > 0} \lambda_i, \quad (6.2)$$

according to Pesin's theorem [20]. The number of positive Lyapunov exponents is proportional to the number of degrees of freedom which itself grows proportionally to the volume. As a consequence, we have that  $H(T, V) \sim TV$ . This rapid increase of the entropy is the feature of the dynamical regimes of spacetime chaos.

### 6.2.4. Nonlinear partial differential equations

Such equations define dynamical systems of infinite dimensional phase spaces, with continuous spacetime and observables. Well-known examples are the Navier–Stokes equations [35], the Brusselator reaction-diffusion equations [50], or the Kuramoto–Sivashinsky equation [51]. For PDEs or coupled maps, it is possible to introduce the density  $g(\lambda)$  of Lyapunov exponents as the number of exponents having their value in the range  $(\lambda, \lambda + d\lambda)$  for a system of unit volume. The entropy is then given by

$$H(T, V) = TV \int_0^{\lambda_{\max}} \lambda g(\lambda) d\lambda, \quad (6.3)$$

where  $\lambda_{\max}$  is the largest Lyapunov exponent. This result holds for every deterministic system of arbitrarily large spatial extension. In a regime of spacetime chaos, the entropy per unit time and volume (6.1) is then positive.

### 6.2.5. Stochastic spin dynamics

Glauber and Kawasaki introduced kinetic models describing the stochastic dynamics of spins in a solid [52, 53]. These models are defined as birth-and-death processes of large spatial extension. According to the results of section 3.3, their entropy diverges with  $\tau \rightarrow 0$  according to

$$H(T, V) \sim TV \log(1/\tau). \quad (6.4)$$

The extra randomness has its origin in the assumption that the time is continuous. If we rather adopt a version of the dynamics which is discrete in time the spin dynamics reduce to Markov chains like in lattice gas automata.

### 6.2.6. Spacetime Gaussian fields

These processes are the spacetime generalizations of the Brownian motions of section 3.6. Pinsker and Sofman have generalized the Kolmogorov formula (2.35), (2.36) to those stationary Gaussian

processes [54]. If the spectral density is

$$\Phi(\mathbf{k}, \omega) = \int_{\mathbb{R}^{d+1}} e^{-i(\mathbf{k}\cdot\mathbf{r} + \omega t)} \langle X(\mathbf{r}, t) X(0, 0) \rangle d\mathbf{r} dt, \tag{6.5}$$

the SK  $\varepsilon$ -entropy per unit time and volume is given by

$$\varepsilon^2 = \frac{1}{(2\pi)^{d+1}} \int_{\mathbb{R}^{d+1}} \min[\theta, \Phi(\mathbf{k}, \omega)] d\mathbf{k} d\omega, \tag{6.6}$$

$$h_{\text{SK}}^{(\text{time,space})}(\varepsilon) = \frac{1}{2(2\pi)^{d+1}} \int_{\mathbb{R}^{d+1}} \max[0, \log \Phi(\mathbf{k}, \omega) / \theta] d\mathbf{k} d\omega. \tag{6.7}$$

For instance, let us suppose that the spectral density decreases at large frequencies and wavenumbers like

$$\Phi(\mathbf{k}, \omega) \sim (\omega^\alpha + k^\beta)^{-\nu} \tag{6.8}$$

with  $(1/\alpha + d/\beta) < \nu$ , so that the integral of the power spectrum is bounded. The space power spectrum is

$$\Phi(\mathbf{k}) = \int_{\mathbb{R}} \Phi(\mathbf{k}, \omega) d\omega \sim k^{-\nu_1}, \quad \nu_1 = \beta(\nu - 1/\alpha); \tag{6.9}$$

and the time power spectrum

$$\Phi(\omega) = \int_{\mathbb{R}^d} \Phi(\mathbf{k}, \omega) d\mathbf{k} \sim \omega^{-\nu_2}, \quad \nu_2 = \alpha(\nu - d/\beta). \tag{6.10}$$

The  $\varepsilon$ -entropy then diverges like

$$h_{\text{SK}}^{(\text{time,space})}(\varepsilon) \sim (1/\varepsilon)^\gamma, \quad \gamma = \frac{2(1/\alpha + d/\beta)}{\nu - (1/\alpha + d/\beta)}. \tag{6.11}$$

With  $d = 0$ ,  $\alpha = 2$  and  $\nu = H + 1/2$ , (6.11) is reduced to eq. (3.92) for the Yaglom noise. A possible spacetime generalization of the Yaglom noise is defined with  $\alpha = \beta = 2$ , and  $\nu = H + (1 + d)/2$ , so we have

$$h_{\text{SK}}^{(\text{time,space})}(\varepsilon) \sim (1/\varepsilon)^{(d+1)/H} \tag{6.12}$$

[compare with eq. (3.92)].

In fact, one can show that the exponent  $\gamma$  depends only on the scaling properties of the power spectrum  $\Phi(\mathbf{k}, \omega)$ . Let us assume that  $\Phi(\mathbf{k}, \omega)$  has the general form

$$\Phi(\mathbf{k}, \omega) \sim k^{-\gamma} F(\omega/k^z), \tag{6.13}$$

where  $F$  is a scaling function. Then, we have

$$h_{\text{SK}}^{(\text{time,space})}(\varepsilon) \sim (1/\varepsilon)^\gamma, \quad \gamma = 2(d + z) / [y - (d + z)]. \tag{6.14}$$

For instance, let  $d = 3$ ,  $\gamma = 13/3$  and  $z = 2/3$ , so that

$$\Phi(k) = k^{d-1} \Phi(\mathbf{k}) \sim k^{-5/3} \quad \text{and} \quad \Phi(\omega) \sim \omega^{-2}, \quad (6.15)$$

which mimic the well-known power spectra of the homogeneous and isotropic turbulence in three-dimensional fluids in the limit of infinite Reynolds number. Then,

$$h_{\text{SK}}^{(\text{time,space})}(\epsilon) \sim (1/\epsilon)^{11}. \quad (6.16)$$

This result is quite interesting since it is solely determined by the scaling properties of  $\Phi(\mathbf{k}, \omega)$  of the fully developed turbulence. It is hence in this sense unique and universal. Nevertheless, let us emphasize that the result (6.16) is based on the assumption that the process is Gaussian. Non-Gaussian effects (intermittency) might not only lead to a (small) correction to the exponent  $\gamma = 11$ , but also to the non-uniqueness of the exponent. In the case of hard turbulence, the possibility of non-Gaussian behaviors was discussed in ref. [37].

We conclude with the comment that the continuous random fields described in this subsection appear to have a very high degree of randomness compared with the other examples.

### 6.2.7. Spacetime processes of statistical mechanics

Most systems of statistical mechanics have a large spatial extension. There exist different classes of systems (see fig. 24).

The simplest systems are the *ideal gases* where the particles are not interacting with each other. As a consequence, there is no local dynamical randomness generated in these systems and the entropy per unit time and volume vanishes since all the Lyapunov exponents are zero. However, the ideal gases are not completely devoid of randomness since new particles continuously arrive from large distances with arbitrary positions and velocities within the Maxwellian distribution. The entropy  $H(T, V)$  is then proportional to the surface of the observed volume  $V^{(d-1)/d}$  as well as to the time  $T$ . Moreover, each new position and velocity are random variables with continuous distributions so that the entropy has a divergence in  $\log(1/\epsilon)$  where  $\epsilon = \Delta^d x \Delta^d v$  [39, 55–56].

If the particles are interacting either with fixed scatterers like in the *Lorentz gas* or with each others like in the *hard sphere gas* the dynamics becomes locally random and there appears a spectrum of positive Lyapunov exponents like in (6.3). The entropy per unit time and volume may then be positive giving a quantitative measure of the microscopic chaos at the origin of the thermal agitation [38–39, 55].

It is possible to illustrate the transition between the noninteracting and the interacting gases with the Lorentz gas composed of a gas of independent particles which are scattered by a lattice of hard spheres fixed in space and occupying a volume  $V$ . Since the collisional dynamics of the particles is defocusing in the region  $V$  of the scatterer the entropy  $H(T, V)$  contains a term which is proportional to  $TV$  and to the positive Lyapunov exponent of the Lorentz gas. However, particles are continuously coming from large distances and enter the scatterer on the surface of the lattice. This shower of new particles is another source of randomness which is now proportional to the time  $T$ , to the surface  $V^{(d-1)/d}$  of the scatterer, and to a factor with a logarithmic divergence like in the ideal gases. Therefore, the entropy  $H(T, V)$  contains two terms,

$$H(T, V) \sim TV + cTV^{(d-1)/d} \log(1/\Delta^d x \Delta^d p) \quad (6.17)$$

where  $c$  is a positive constant. When the volume is large the first term dominates and the system is the stage of spacetime chaos like in the hard sphere gas.

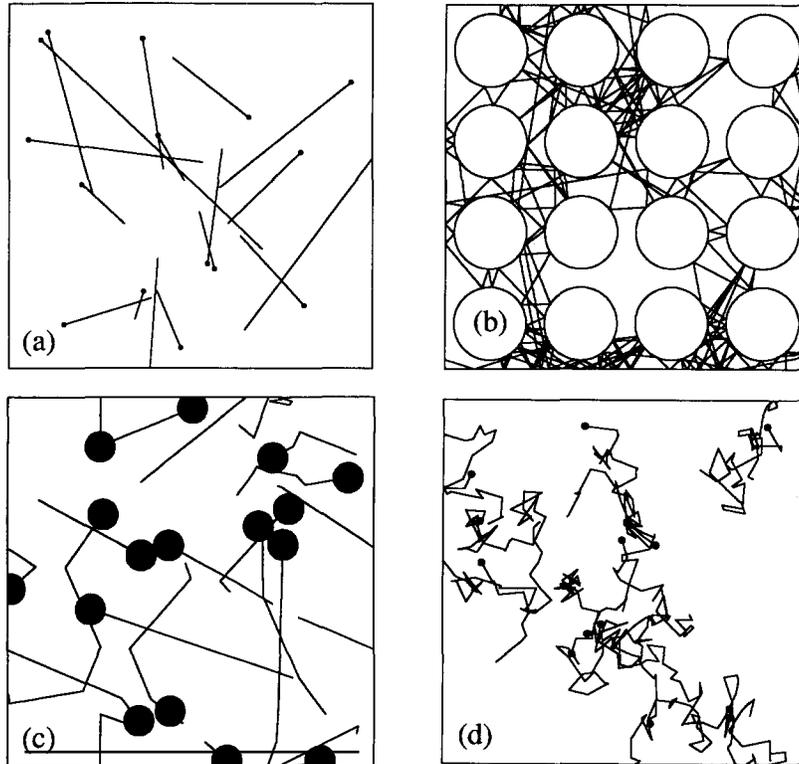


Fig. 24. Schematic representation of the four examples of spacetime processes of statistical mechanics given in subsection 6.2.7: (a) the ideal gas; (b) the finite Lorentz gas; (c) the hard sphere gas; (d) the Boltzmann-Lorentz gas where each collision is a random process like in fig. 23 and contrary to the gases a, b, and c which are deterministic.

Models have been introduced in nonequilibrium statistical mechanics which have a higher degree of dynamical randomness than spacetime chaos. This is the case for the *Boltzmann-Lorentz process with many particles*. Since an extended system has a number of particles which is proportional to the volume, the total entropy of all the particles is obtained by multiplying the single particle entropy (5.9) by the particle density and by the volume. As a consequence, the entropy diverges like (5.9).

## 7. Discussions and conclusions

### 7.1. Classification of the random processes

In the preceding sections, we have obtained the  $(\varepsilon, \tau)$ -entropy per unit time for a variety of stochastic processes (and per unit volume for spacetime processes). We observed that there exist two ways to classify the processes using the general  $(\varepsilon, \tau)$ -entropy  $H(\varepsilon, \tau, T)$  over a time interval  $T$ .

First, there is the dependence in the time  $T$ . We have seen that  $H(\varepsilon, \tau, T)$  is extensive in  $T$  for most processes, a property we called regular randomness. In contrast, for some processes

Table 1  
Periodic and sporadically random processes.

Process	$H(\varepsilon, \tau, T)$
periodic	$\log T$
Feigenbaum attractor	$\log T$
intermittent maps ( $z \geq 2$ )	$T^\alpha$ ( $\alpha < 1$ )
intermittent maps ( $z = 2$ )	$T/\log T$
English or German texts	$T^\alpha$ ( $\alpha \simeq 1/4 - 1/2$ )
DNA sequences	$T^\alpha$ ( $\alpha < 1$ )
Lévy flights ( $\alpha < 1$ )	$T^\alpha d \log(1/\varepsilon)$
Lévy flights ( $\alpha = 1$ )	$(T/\log T)d \log(1/\varepsilon)$

that we called sporadically random,  $H(\varepsilon, \tau, T)$  grows more slowly than the time  $T$  according to  $H(\varepsilon, \tau, T) \sim T^\alpha$  where  $0 < \alpha < 1$  is the exponent of sporadicity. Recently, Ebeling, Nicolis and others have shown that such sporadic behavior is a feature of language texts [57–59] as well as of some DNA sequences [60, 61]. This phenomenon can be interpreted by the presence of new broader comprehensive schemes or Kantian syntheses emerging on longer and longer periods of time in human or biological information sequences.

Secondly, there is the dependence of  $H(\varepsilon, \tau, T)$  on  $\varepsilon$  and  $\tau$ , which constitutes the main contribution of the present paper. This dependence disappears for the chaotic deterministic processes because their trajectories are smooth. On the contrary, the dependence remains for processes which have non-smooth trajectories. For instance, the trajectories  $N_i(t)$  of a birth-and-death process are discontinuous in time; the Ornstein–Uhlenbeck, the Yaglom, and the Brownian processes are continuous but nondifferentiable so that the dependence in  $\varepsilon$  remains. Hence, all these noises have an  $(\varepsilon, \tau)$ -entropy per unit time which diverges as  $\varepsilon$  or  $\tau \rightarrow 0$ . As a consequence, they are much more random than the chaotic processes. Furthermore, we can compare one process with another, as we do with Yaglom noises of different exponents  $H$ , to show that some of them are more random than others. In the same way, we can establish equivalences between the degrees of randomness of the processes.

In summary, we have compiled in table 1 the entropies  $H(\varepsilon, \tau, T)$  for various periodic and sporadic processes. In table 2, we gathered the entropies per unit time  $h(\varepsilon, \tau)$  for the random processes considered in sections 3–5. The entropies  $H(\varepsilon, \tau, T, V)$  for spacetime processes discussed in section 6 are listed in table 3. Since we have only considered a limited number of known random processes, the list is far from being exhaustive.

## 7.2. The principle of entropy invariance and some consequences

The principle of entropy invariance by Kolmogorov, Sinai [16], Ornstein [23], and others states that the KS entropy per unit time is invariant under an isomorphism which maps a process onto another. For instance, in section 3, we constructed an isomorphism between a time-discrete Markov chain and a deterministic chaotic mapping. The KS entropy is the same for both processes showing that information on the time evolution of the processes is strictly preserved going from one description to the other.

The invariance of the KS entropy generalizes to the  $(\varepsilon, \tau)$ -entropy per unit time if we admit some differences due to the different definitions of  $(\varepsilon, \tau)$  in the two processes between which we want to establish an isomorphism. This general idea is of application in kinetic theory where we

Table 2  
Time random processes.

Process	$h(\epsilon, \tau)$
periodic	0
Feigenbaum attractor	0
intermittent maps	0
English or German texts	0
DNA sequences	0
Lévy flights ( $\alpha \leq 1$ )	0
Lévy flights ( $\alpha > 1$ )	$d \log(1/\epsilon)$
deterministic chaos	$h_{KS}$
Bernoulli and Markov	$h_{KS}$
birth-and-death	$\log(1/\tau)$
time-discrete, amplitude-continuous	$d \log(1/\epsilon)$
Ornstein-Uhlenbeck	$(1/\epsilon)^2$
Yaglom	$(1/\epsilon)^{1/H}$
Brownian	$(1/\epsilon)^2$
fractional Brownian	$(1/\epsilon)^{1/H}$
white noise	$(1/\tau) \log(1/\epsilon)$
soft turbulence (exp. range)	$(\log \epsilon^{-1})^3$
hard turbulence (exp. range)	$(1/\epsilon)^2$
Boltzmann-Lorentz (1 particle)	$\log(1/\tau \Delta^3 v \Delta^2 \Omega)$

want to map the deterministic description by Newton's equations onto a probabilistic description that we hope to be faithful, if not for all, at least for a restricted set of physical observables. The description remains faithful for the physical observables of interest if the  $(\epsilon, \tau)$ -entropy per unit time is essentially preserved during the change of description.

The  $(\epsilon, \tau)$ -entropy is then able to check if the dynamical randomness of the kinetic model takes the correct values in the range of  $(\epsilon, \tau)$  considered. If the  $(\epsilon, \tau)$ -entropy of the kinetic model overestimates the  $(\epsilon, \tau)$ -entropy of the deterministic mode, a program which simulates the process according to the kinetic model will call the pseudorandom number generator more often than necessary. On the other hand, if the  $(\epsilon, \tau)$ -entropy of the kinetic model underestimates the true

Table 3  
Spacetime random processes.

Process	$H(\epsilon, \tau, T, V)$
cellular automata	$< TV$
Conway's game of life	$< TV$
lattice gas automata	$TV$
coupled maps in spacetime chaos	$TV$
nonlinear PDE's in spacetime chaos	$TV$
Glauber or Kawasaki spin dynamics	$TV \log(1/\tau)$
$(d + 1)$ -dimensional Yaglom fields	$TV (1/\epsilon)^{(d+1)/H}$
$d$ -dimensional ideal gas (many particles)	$TV^{(d-1)/d} \log(1/\Delta^d x \Delta^d p)$
Lorentz gas (fixed scatterers)	$TV + cTV^{(d-1)/d} \log(1/\Delta^d x \Delta^d p)$
Hard sphere gas	$TV$
Boltzmann-Lorentz (many particles)	$TV \log(1/\tau \Delta^3 v \Delta^2 \Omega)$

value, the signal of the simulation will be more regular than it should be and the conclusion will be that some source of randomness has been overlooked.

Let us remark that the limitations of the kinetic models have always been discussed in the literature along lines that are parallel with ours. Both discussions have in common their use of the concepts of mean free path and of mean intercollisional time. These quantities give the scales of  $\varepsilon$  or of  $\tau$  where the kinetic model may need to be amended because of the deterministic character of the molecular dynamics. It is also the scale where differences may appear between the actual  $(\varepsilon, \tau)$ -entropy calculated from Newton's equations and the  $(\varepsilon, \tau)$ -entropy of the kinetic model. An example of amendment which has been considered is to include memory effects [62]. The consequence of such modifications on the  $\varepsilon$ -entropy is that the  $\varepsilon$ -entropy would grow more slowly at small  $\varepsilon$ . Finally, if the deterministic Newton equations are used without stochastic assumptions but only chaotic assumptions, the  $\varepsilon$ -entropy would saturate at the value of the KS entropy.

Let us mention a few examples to illustrate our point.

### 7.2.1. Equilibrium states

Suppose a classical statistical mechanical system is described an equilibrium probability density which is Gaussian, such as the Maxwellian velocity distribution of a gas. If its microscopic state is monitored at sampling times much longer than the relaxation time, then the observation of this system consists of a discrete sequence of independent Gaussian random variables, therefore its  $\varepsilon$ -entropy per sampling time is equal to  $h(\varepsilon)$  given by (3.36). It diverges for the reason explained in subsection (3.4.1). This divergence is at the origin of the famous Gibbs paradox in classical equilibrium statistical mechanics, where the thermodynamic entropy is not fixed in absolute value but only in relative value with respect to the entropy of a reference thermodynamic state. The divergent term is in general omitted and only the next term – which is often called the differential entropy in the mathematical literature – is considered in equilibrium thermodynamics. Contrary to the  $\varepsilon$ -entropy, the differential entropy is empty of any operational interpretation in terms of randomness. In quantum statistical mechanics, the size  $\varepsilon^d$  of the cells in phase space cannot be smaller than  $(2\pi\hbar)^d$ , since the quantum states are discrete and fully determined below this limit by the wavefunctions. As a consequence, the constant of the thermodynamic entropy appears to be fixed by quantum mechanics, which is expressed by Nernst's third law of thermodynamics.

### 7.2.2. Kinetic equations

A dilute hard sphere gas is governed by deterministic Newton's equations of interacting particles. The dynamical chaos produced by these interactions is characterized by an  $(\varepsilon, \tau)$ -entropy per unit time and per particle of the gas, and provides a justification for the randomness aspect of Boltzmann's fundamental *stosszahlansatz* of the kinetic theory of gases. An estimation of the maximal Lyapunov exponent per particle of the gas is given by the finite quantity [39]

$$\lambda \sim \frac{1}{T_{\text{intercoll}}} \ln(l/a), \quad (7.1)$$

where  $a$  is the radius of the particles,  $l$  the mean free path, and  $T_{\text{intercoll}}$  the mean intercollision time.

In section 5, we have calculated directly the  $(\varepsilon, \tau)$ -entropy per unit time for the Boltzmann–Lorentz equation describing the collision process of one particle in a dilute hard sphere gas. We obtained a result similar to (7.1), except that, for the Boltzmann–Lorentz process, some infinitesimal  $\varepsilon\tau$  appears in the  $(\varepsilon, \tau)$ -entropy per unit time. As we discussed before, the Boltzmann–Lorentz process is stochastic and, as a consequence, has a degree of randomness superior to the

chaotic process of the deterministic collisions in a gas. Since the  $(\varepsilon, \tau)$ -entropy diverges as  $\varepsilon$  and  $\tau$  becomes arbitrarily small, there must be a limiting value where the randomness of the stochastic process starts to overestimate the randomness of the gas particle motion. (7.1) is equal to the  $(\varepsilon, \tau)$ -entropy per unit time (5.9) for small  $\tau^*$  and  $\varepsilon = \Delta^3 v \Delta^2 \Omega$  when

$$\tau^* \sim a \bar{v}^2 / \Delta^3 v \Delta^2 \Omega, \quad (7.2)$$

which is the limit of applicability of the stochastic process when compared with the chaotic process of the Newtonian dynamics.

### 7.2.3. Langevin equations

A similar discussion applies to the Langevin equations, which provide a very good description of the Brownian motion of particles in a fluid or of other related processes. These processes are characterized by an  $\varepsilon$ -entropy per unit time which grows like (cf. subsection 3.6.2)

$$h(\varepsilon) \sim 1/\varepsilon^2. \quad (7.3)$$

Ultimately, however, this  $\varepsilon$ -entropy should saturate at a finite value as soon as  $\varepsilon$  is small enough to resolve the deterministic motion of the molecules in collisions. Therefore, for  $\varepsilon$  small enough, the  $\varepsilon$ -entropy of the Langevin equation which assumes the irregularity of the signal on arbitrarily small scales will overestimate the actual  $\varepsilon$ -entropy of the process. In that regime, the stochastic assumptions of the kinetic model would have to be revised.

That is the case in models of deterministic diffusion with only two degrees of freedom where the saturation is rapid. As an example, we consider the following map of the real line [63]

$$X_{t+1} = X_t + p \sin 2\pi X_t, \quad (7.4)$$

which presents diffusion, in particular, at the value  $p = 0.8$  with a diffusion coefficient  $D = 0.18$  [63]. Figure 25b depicts the  $\varepsilon$ -entropy per unit time calculated numerically where the saturation toward the KS entropy  $h_{\text{KS}} = \lambda = 0.49$  digits/iteration is visible at small values of  $\varepsilon$ .

We see that the  $\varepsilon$ -entropy is thus able to reveal subtle distinctions in the stochastic assumptions of a particular kinetic model.

### 7.2.4. Master equations

Master equations like (3.23) are often used to model rate processes in physical chemistry. One example is the transport in the phase space of chaotic systems with a finite number of degrees of freedom [64]. Another example is provided by macroscopic reaction-diffusion in far-from-equilibrium systems with a very large number of degrees of freedom [27]. In both cases, the statistical assumptions of the kinetic model usually contains a time scale  $\tau^*$  under which the kinetic model is not valid anymore. This time scale  $\tau^*$  is some recurrence time like the crossing time of the Poincaré surface of section in chaotic systems or the intercollisional time in macroscopic far-from-equilibrium systems. On this small time scale, the underlying deterministic dynamics start to show its effects. Below this time scale, the dynamical randomness comes from the chaotic nature of the motion so that the  $\tau$ -entropy per unit time is always bounded by the value of the KS entropy per unit time of the full deterministic dynamics

$$h(\tau) \leq h_{\text{KS}}. \quad (7.5)$$

In the regime where  $\tau \gg \tau^*$ , the  $\tau$ -entropy per unit time calculated using the deterministic dynamics together with a partition of its phase space into the states  $\{\alpha\}$  should take the same value as the

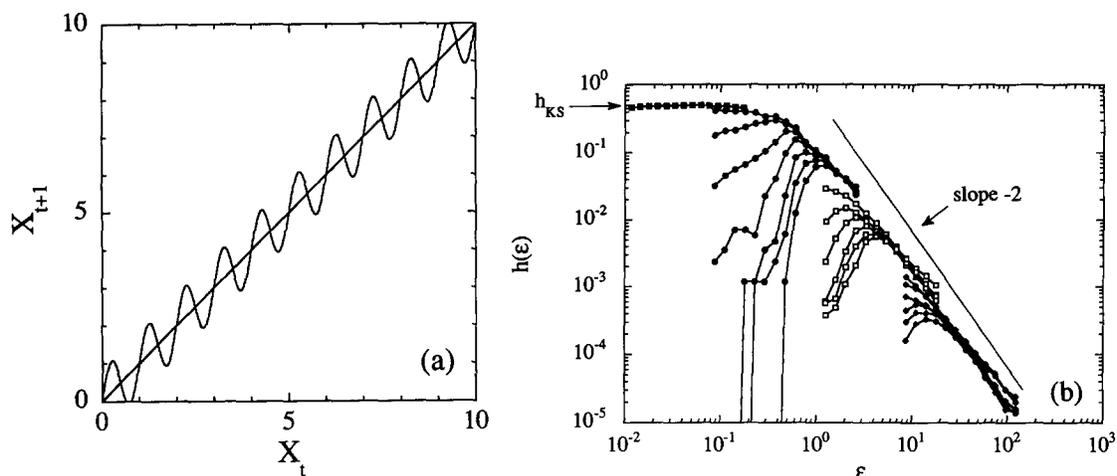


Fig. 25. (a) The map (7.4) as a model of deterministic diffusion. (b) Numerical evaluation of the  $(\epsilon, \tau)$ -entropy of the nonstationary process on the real line for different values of  $\tau = 1$  (filled circles),  $\tau = 10$  (squares), and  $\tau = 100$  (filled diamonds). For each value of  $\tau$ , the different curves correspond to time sequences of lengths  $N = 1, 4, 7, 10, 13, 16, 19$  [cf. (2.37)–(2.41)]. The crossed squares give the entropy calculated with periodic boundary condition over 50 lattice cells ( $0 \leq X < 50$ ) which coincides with the value of the mean Lyapunov exponent  $h_{KS} = \lambda = 0.49$  digits/iteration. The envelope of the different curves decreases like  $(1/\epsilon)^2$  at large  $\epsilon$  but saturates at the value of the KS entropy at small  $\epsilon$ .

$\tau$ -entropy of the master equation, provided that the master equation gives a faithful description of the time evolution of the states  $\{\alpha\}$  (see fig. 26).

This statement is general and constitutes an extension of the principle of invariance of the entropy per unit time under an isomorphism by Kolmogorov, Sinai, and Ornstein [16, 23].

### 7.2.5. Turbulence

In developed turbulence, the application of the preceding reasoning leads to a similar situation. Although turbulence is described by the deterministic Navier–Stokes equations, it turns out that the regimes at large Reynolds numbers can effectively be described by noisy stochastic processes on large and intermediate scales. In this case, the aim is the obtention of an effective isomorphism

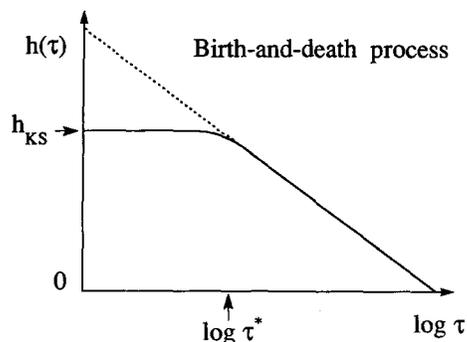


Fig. 26. Schematic behavior of the actual  $\tau$ -entropy per unit time (solid line) compared with the  $\tau$ -entropy of the birth-and-death processes. We see the crossover around  $\tau = \tau^*$  and the saturation to the KS entropy [cf. eq. (7.5)].

between the dynamical system of the Navier-Stokes equations and some stochastic process.

### 7.3. Perspectives

Our purpose in this paper has been to present this theory of the  $(\varepsilon, \tau)$ -entropy per unit time in a systematic way and in its modern physico-chemical context where a lot of different random processes are considered. We think that it is particularly important to compare their degree of randomness in order to understand their limit of validity and their origin in the deterministic chaotic dynamics maybe at remotely small scales. About this question, the equivalence principle by Kolmogorov, Sinai [16], Ornstein [23], and others is a powerful guiding principle to relate the stochastic processes usually considered in nonequilibrium statistical mechanics to the underlying deterministic chaotic dynamics of these many-body systems. On the other hand, the stochastic models may be taken for what they are to model experimental results. It is then important to control the transitions that may occur in their degrees of randomness like in fluid turbulence.

Moreover, we have shown that the concept of entropy over a time interval and a space volume is uniquely able to establish a classification of the different spacetime random processes according to their degree of dynamical randomness. We believe that the very broad perspective that is so provided has far reaching consequences in the natural sciences. First of all, it gives a quantitative measure of the randomness of a system. For instance, a chaotic model is suggested if the  $\varepsilon$ -entropy per unit time presents a plateau. In this case, we find [65]

$$H(\varepsilon, \tau, T) \simeq Th_{KS} + d_1 \log(1/\varepsilon), \quad (7.6)$$

and the divergence with  $\varepsilon \rightarrow 0$  in the term which is independent of the time  $T$  provides us with the information dimension of the chaotic attractor.

However, if the  $\varepsilon$ -entropy per unit time diverges as  $\varepsilon \rightarrow 0$ , a stochastic model will be more appropriate

$$H(\varepsilon, \tau, T) \simeq Th(\varepsilon). \quad (7.7)$$

In that case, the type of the divergence may point toward the kind of noise to consider according to table 2. We remark that the divergence appears in the term which is proportional to the time contrary to (7.6) so that the dependence on  $\varepsilon$  of the next term is hidden in (7.7).

Furthermore, in fields like statistical mechanics where there exist chains of models or theories, the  $\varepsilon$ -entropy can detect the degree of dynamical randomness which is introduced with the assumptions used to go from one model or theory to another. In particular, we have shown here, that the Boltzmann–Lorentz kinetic model assumes a higher dynamical randomness than present at the level of the chaotic Newtonian dynamics, which restricts the domain of validity of the kinetic model to the mesoscopic scales.

There are also important applications in biology or artificial intelligence, for instance, to the stochastic gating dynamics of the ion channels [66] or to the activity of neural networks [67]. Other applications exist in communication problems [5].

As we enter into the dynamics of microscopic systems, we think that it is important to extend the preceding considerations to quantum mechanics. The reasons are twofold. First, the quantum effects are important in many physico-chemical random processes, specially, at low temperature and in quantum optics. In particular, the characterization of the degree of randomness of these processes considered as source of information is a major preoccupation in quantum electronics and telecommunication [68, 69]. Moreover, the origin of randomness in quantum mechanics is

a central problem and we think that a quantitative measure of the dynamical randomness like the entropy per unit time provides a unique opportunity to make progress on this fundamental question. We have already obtained elsewhere significant results about dynamical randomness in quantum systems which shed a new light on this problem [39, 55].

### Acknowledgements

We are grateful to Professors L. Kadanoff and G. Nicolis for support and encouragement in this research; to Professor A. Libchaber who kindly provided us with his data on turbulence in He and his support in their analysis. We also thank Professors L. Kadanoff and W. E for a careful reading of the manuscript. We are also grateful to Y. Elskens for fruitful discussions at the early stages of this work back to 1985. P.G. thanks the National Fund for Scientific Research (F.N.R.S. Belgium) for financial support. X.-J.W. is partly supported by the Office of Naval Research under the contract No. N00014-90J-1194.

### References

- [1] P. Gaspard and G. Nicolis, *J. Stat. Phys.* 31 (1983) 499.
- [2] N. Wax, ed., *Selected papers on noise and stochastic processes* (Dover, New York, 1954).
- [3] Hao Bai-Lin, Ed., *Chaos*, vol. I (World Scientific, Singapore, 1987); vol. II (World Scientific, Singapore, 1990).
- [4] J.-P. Eckmann and D. Ruelle, *Rev. Mod. Phys.* 57 (1985) 617.
- [5] C.E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (The University of Illinois Press, Urbana, 1949).
- [6] A.N. Kolmogorov, *IRE Trans. Inf. Theory* 1 (1956) 102.
- [7] C. E. Shannon, in: *Information and Decision processes*, ed R. Machol (McGraw-Hill, New York, 1960) p. 93.
- [8] A.N. Kolmogorov and V.M. Tikhomirov, *Uspekhi Mat. Nauk* 14 (1959) 3.
- [9] V.M. Tikhomirov, *Russ. Math. Surveys* 18 (1963) 51.
- [10] T. Berger, *Rate Distortion Theory* (Prentice-Hall, Englewood Cliffs, 1971).
- [11] A.N. Kolmogorov, *Russ. Math. Survey* 38 (1983) 29.
- [12] G.J. Chaitin, *Algorithmic Information Theory* (Cambridge University Press, Cambridge, 1987).
- [13] P. Grassberger, *J. Theor. Phys.* 25 (1986) 907.
- [14] P. Gaspard and X.-J. Wang, *Proc. Natl. Acad. Sci. (USA)* 85 (1988) 4591.
- [15] P. Manneville, *J. de Phys.* 41 (1980) 1235; Y. Pomeau and P. Manneville, *Commun. Math. Phys.* 74 (1980) 189.
- [16] I.P. Cornfeld, S.V. Fomin and Ya.G. Sinai, *Ergodic Theory* (Springer, Berlin, 1982).
- [17] A.K. Zvonkin and L.A. Levin, *Russian Math. Surveys* 25 (1970) 83.
- [18] A. Cohen and I. Procaccia, *Phys. Rev. A* 31 (1985) 1872. A variant of the entropy when the Renyi parameter  $q = 2$  was considered by P. Grassberger and I. Procaccia, *Phys. Rev. A* 28 (1983) 259.
- [19] D. Ruelle, *Ann. N.Y. Acad. Sci.* 316, 408 (1979); Yu.I. Kifer, *Math. SSSR Izv.* 8 (1974) 1083.
- [20] Ya.B. Pesin, *Math. USSR Izv.* 10(6) (1976) 1261; *Russ. Math. Surveys* 32(4) (1977) 55.
- [21] M. Hénon, *Commun. Math. Phys.* 50 (1976) 69.
- [22] P. Billingsley, *Ergodic Theory and Information* (Wiley, New York, 1965).
- [23] D.S. Ornstein, *Ergodic Theory, Randomness, and Dynamical Systems* (Yale University Press, New Haven, 1974).
- [24] R.L. Adler and B. Weiss, *Mem. Am. Math. Soc.* 98 (Am. Math. Soc., Providence, 1970).
- [25] J. Schnakenberg, *Rev. Mod. Phys.* 48 (1976) 571.
- [26] G. Nicolis and C. Nicolis, *Phys. Rev. A* 38 (1988) 427.
- [27] G. Nicolis and I. Prigogine, *Self-Organization in Nonequilibrium Systems* (Wiley, New York, 1977).
- [28] N.G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1981).
- [29] A.M. Yaglom, *Correlation Theory of Stationary and related Random Functions* (Springer, Berlin, 1987).
- [30] K. Lindenberg, K.E. Shuler, J. Freeman and T.L. Lie, *J. Stat. Phys.* 12 (1975) 217.
- [31] B.B. Mandelbrot and J.W. Van Ness, *SIAM Review* 10 (1968) 422.
- [32] R.J. Barton and H.V. Poor, *IEEE Trans. Inf. Theory* IT-34 (1988) 943.

- [33] T. Berger, *IEEE Trans. Inf. Theor.* IT-16 (1970) 134.
- [34] X.-J. Wang, *Phys. Rev. A* 45 (1992) 8407.
- [35] D. Ruelle, *Commun. Math. Phys.* 87 (1982) 287.
- [36] F. Heslot, B. Castaing and A. Libchaber, *Phys. Rev. A* 36 (1987) 5870;  
B. Castaing et al., *J. Fluid Mech.* 204 (1989) 1;  
X.-Z. Wu, L. Kadanoff, A. Libchaber and M. Sano, *Phys. Rev. Lett.* 64 (1990) 2140.
- [37] X.-J. Wang and P. Gaspard, *Phys. Rev. A* 46 (1992) R3000.
- [38] P. Gaspard, in: *Solitons and Chaos*, eds I. Antoniou and F.J. Lambert (Springer, Berlin, 1991) pp. 46–57.
- [39] P. Gaspard, in: *Quantum Chaos–Quantum Measurement*, eds P. Cvitanović, I. Percival and A. Wirzba (Kluwer, Dordrecht, 1992) pp. 19–42.
- [40] P. Résibois and M. De Leener, *Classical Kinetic Theory of Fluids* (Wiley, New York, 1977).
- [41] S. Wolfram, *Rev. Mod. Phys.* 55 (1983) 601; *Physica D* 10 (1984) 1.
- [42] P. Grassberger, *J. Stat. Phys.* 45 (1986) 27.
- [43] J.H. Conway, in: *Winning Ways*, Vol. 2, eds E. Berkelamp, J. Conway and R. Guy (Academic Press, New York, 1982).
- [44] M.A. Nowak and R. May, *Nature* 359 (1992) 826.
- [45] D. Dab, J.-P. Boon and Y.-X. Li, *Phys. Rev. Lett.* 66 (1991) 2535.
- [46] A. Lawniczak, D. Dab, R. Kapral and J.-P. Boon, *Physica D* 47 (1991) 132.
- [47] R. Kapral, A. Lawniczak and P. Masiar, *J. Chem. Phys.* 96 (1992) 2762.
- [48] G.-L. Oppo and R. Kapral, *Phys. Rev. A* 36 (1987) 5820;  
R. Kapral, *J. Math. Chem.* 6 (1991) 113.
- [49] J. Crutchfield and K. Kaneko, in: *Chaos*, vol. I, ed. Hao Bai-Lin (World Scientific, Singapore, 1987) p. 272;  
K. Kaneko, in: *Formation, Dynamics and Statics of Patterns*, ed. K. Kawasaki (World Scientific, Singapore, 1989).
- [50] A. DeWit, G. Dewel, P. Borckmans and D. Walgraef, *Physica D* 61 (1992) 289.
- [51] Y. Pomeau, A. Pumir and P. Pelce, *J. Stat. Phys.* 37 (1984) 39;  
P. Manneville, in: *Macroscopic Modeling of Turbulent Flows*, ed. O. Pironneau, *Lecture Notes in Physics* Vol. 230 (Springer, New York, 1985) p. 319.
- [52] R.J. Glauber, *J. Math. Phys.* 4 (1963) 294.
- [53] K. Kawasaki, *Phys. Rev.* 145 (1966) 224.
- [54] M.S. Pinsky and L.B. Sofman, *Problems of Inf. Transm.* 19(3) (1983) 214.
- [55] P. Gaspard, in: *Quantum Chaos*, eds H.A. Cerdeira, M.C. Gutzwiller, R. Ramaswamy and G. Casati (World Scientific, Singapore, 1991).
- [56] T. Hudetz and P. Gaspard, in preparation.
- [57] W. Ebeling and G. Nicolis, *Europhys. Lett.* 14 (1991) 191.
- [58] P. Grassberger, *IEEE Trans. Inf. Theory* 35 (1989) 669.
- [59] W. Ebeling and G. Nicolis, *Chaos, Solitons & Fractals* 2(3) (1992).
- [60] W. Ebeling, R. Feistel and H. Herzel, *Physica Scripta* 35 (1987) 761.
- [61] W. Li and K. Kaneko, *Europhys. Lett.* 17 (1992) 655.
- [62] J.-P. Boon and S. Yip, *Molecular Hydrodynamics* (Dover, New York, 1980).
- [63] M. Schell, S. Fraser and R. Kapral, *Phys. Rev. A* 26 (1982) 504.
- [64] R.S. MacKay, J.D. Meiss and I.C. Percival, *Physica D* 13 (1984) 55;  
J.D. Meiss, *Rev. Mod. Phys.* 64 (1992) 795.
- [65] P. Grassberger, T. Schreiber and C. Schaffrath, *Int. J. Bifurcation Chaos* 1 (1991) 521.
- [66] D. Colquhoun and A.G. Hawkes, *Proc. R. Soc. Lond. B* 199 (1977) 231; 211 (1981) 205.
- [67] X.-J. Wang, in: *Self-Organization, Emergent Properties, and Learning*, ed. A. Babloyantz (Plenum, New York, 1991) p. 155.
- [68] Y. Yamamoto and H.A. Haus, *Rev. Mod. Phys.* 58 (1986) 1001.
- [69] A. Yariv, *Theory and Applications of Quantum Mechanics* (Wiley, New York, 1982).