



ELSEVIER

Physics Reports 356 (2002) 367–474

PHYSICS REPORTS

www.elsevier.com/locate/physrep

# Predictability: a way to characterize complexity

G. Boffetta<sup>a,b,\*</sup>, M. Cencini<sup>c,d</sup>, M. Falcioni<sup>d,e</sup>, A. Vulpiani<sup>d,e</sup>

<sup>a</sup>*Dipartimento di Fisica Generale, Università di Torino, Via Pietro Giuria 1, I-10125 Torino, Italy*

<sup>b</sup>*Istituto Nazionale Fisica della Materia, Unità dell'Università di Torino, Italy*

<sup>c</sup>*Max-Planck-Institut für Physik komplexer Systeme, Nöthnitzer Str. 38, 01187 Dresden, Germany*

<sup>d</sup>*Dipartimento di Fisica, Università di Roma "la Sapienza", Piazzale Aldo Moro 5, 00185 Roma, Italy*

<sup>e</sup>*Istituto Nazionale Fisica della Materia, Unità di Roma 1, Italy*

Received January 2001; editor: I. Procaccia

This review is dedicated to our masters and to our friends; in particular to Andrei N. Kolmogorov (1903–1987) and Giovanni Paladin (1958–1996) whose influence on our work runs deeper than we can know.

## Contents

1. Introduction	369	4.5. Convective chaos and spatial complexity	398
2. Two points of view	371	4.6. Space–time evolution of localized perturbations	401
2.1. Dynamical systems approach	371	4.7. Macroscopic chaos in globally coupled maps	404
2.2. Information theory approach	374	4.8. Predictability in presence of coherent structures	406
2.3. Algorithmic complexity and Lyapunov exponent	380	5. Predictability in fully developed turbulence	410
3. Limits of the Lyapunov exponent for predictability	381	5.1. Basic facts of turbulence	410
3.1. Characterization of finite-time fluctuations	382	5.2. Reduced model of turbulence	412
3.2. Renyi entropies	385	5.3. Effects of intermittency on predictability of infinitesimal perturbations	413
3.3. The effects of intermittency on predictability	386	5.4. Growth of non-infinitesimal perturbations	415
3.4. Growth of non-infinitesimal perturbations	387	5.5. $\epsilon$ -entropy for turbulent flows	419
3.5. The $\epsilon$ -entropy	390	6. Uncertainty in the evolution equations	424
4. Predictability in extended systems	392	6.1. Uncertainty introduced by numerical computations	428
4.1. Simplified models for extended systems and the thermodynamic limit	393	6.2. Finite resolution effects and parameterization of unresolved scales	428
4.2. Overview on the predictability problem in extended systems	394	6.3. Lyapunov exponents and complexity in dynamical systems with noise	432
4.3. Butterfly effect in coupled map lattices	396	6.4. Random dynamical systems	436
4.4. Comoving and specific Lyapunov exponents	397		

\* Corresponding author.

*E-mail address:* boffetta@to.infn.it (G. Boffetta).

7. Irregular behavior in discrete dynamical systems	440	9. Concluding remarks	457
7.1. Dependence of the period on the number of the states	441	Note added in proof	458
7.2. The transition from discrete to continuous states	442	Acknowledgements	458
7.3. Entropies and Lyapunov exponents in cellular automata	444	Appendix A. On the computation of the finite size Lyapunov exponent	459
8. The characterization of the complexity and system modeling	450	Appendix B. The multifractal model of turbulence	461
8.1. How random is a random number generator?	451	Appendix C. How to compute the $\epsilon$ -entropy with exit times	462
8.2. High-dimensional systems	452	Appendix D. Synthetic signals for turbulence	464
8.3. Diffusion in deterministic systems and Brownian motion	454	D.1. Reproducing the spatial properties or the temporal ones	464
8.4. On the distinction between chaos and noise	455	D.2. Reproducing both the spatial and the temporal properties	466
		References	467

---

## Abstract

Different aspects of the predictability problem in dynamical systems are reviewed. The deep relation among Lyapunov exponents, Kolmogorov–Sinai entropy, Shannon entropy and algorithmic complexity is discussed. In particular, we emphasize how a characterization of the unpredictability of a system gives a measure of its complexity. Adopting this point of view, we review some developments in the characterization of the predictability of systems showing different kinds of complexity: from low-dimensional systems to high-dimensional ones with spatio-temporal chaos and to fully developed turbulence. A special attention is devoted to finite-time and finite-resolution effects on predictability, which can be accounted with suitable generalization of the standard indicators. The problems involved in systems with intrinsic randomness is discussed, with emphasis on the important problems of distinguishing chaos from noise and of modeling the system. The characterization of irregular behavior in systems with discrete phase space is also considered. © 2002 Elsevier Science B.V. All rights reserved.

*PACS:* 05.45.–a; 89.75.Fb

*Keywords:* Predictability; High-dimensional chaos; Algorithmic complexity; Turbulence

---

All the simple systems are simple in the same way, each complex system has its own complexity

(freely inspired by *Anna Karenina* by Lev N. Tolstoy)

## 1. Introduction

The ability to predict the future state of a system, given the present one, stands at the foundations of scientific knowledge with relevant implications from a conceptual and applicative point of view. The knowledge of the evolution law of the system may induce one to conclude that this aim has been attained. This is the classical deterministic point of view as clearly stated by Laplace [134]: once the evolution laws of the system are known, the state at a certain time  $t_0$  completely determines the subsequent states for every time  $t > t_0$ .<sup>1</sup> However it is well established now that this cannot be accomplished in practice.

One limitation occurs in systems with many degrees of freedom, namely the impossibility to manage the huge amount of data required for a detailed description of a single state of a macroscopic body. This aspect, which is not discussed in this review, has led to the development of statistical mechanics.

Another source of difficulty, which arises even in low dimensional systems, is related to the unavoidable uncertainty in the initial condition. As clearly stated by Poincaré, this implies that one can make long-time predictions only if the evolution law does not amplify the initial uncertainty too rapidly. This aspect had a relevant role in the development of the theory of dynamical chaos.

Therefore, from the point of view of predictability, we need to know how an error in the initial state of the system grows in time. In deterministic chaotic systems, i.e., with sensitive dependence on initial condition, one has an exponential growth of errors and, consequently, a severe limitation on the ability to predict the future states. In addition, since the details of the evolution laws are not completely known (or, at least, cannot be specified with an arbitrary accuracy) or some degrees of freedom cannot be resolved, one has another unavoidable source of unpredictability. This is also true for systems with discrete states.

A branch of the theory of dynamical systems has been developed with the aim of formalizing and quantitatively characterizing the sensitivity to initial conditions. The Lyapunov exponent and the Kolmogorov–Sinai entropy are the two indicators for measuring the rate of error growth and information produced by the dynamical system. A complementary approach has been developed in the context of information theory, data compression and algorithmic complexity theory. Nowadays it is rather clear that the latter point of view is closely related to the dynamical systems one. If a system is chaotic then the predictability is limited up to a time which is related to the first Lyapunov exponent, and the time sequence generated from one of its chaotic trajectories cannot be compressed by an arbitrary factor, i.e. is algorithmically complex. On the contrary, a regular trajectory can be easily compressed (e.g., for a periodic trajectory it is sufficient to have the sequence for a period) so it is “simple”.

---

<sup>1</sup> In this review we shall always consider the usual setting where a system is studied by an external observer. In this way one can avoid the problem of the self-prediction [192].

In this review we will discuss how these points of view are related and how they complete each other in giving a quantitative understanding of complexity arising in dynamical systems. In particular, we shall consider the extension of this approach, nowadays well established in the context of low dimensional systems and for asymptotic regimes, to high dimensional systems with attention to situations far from asymptotic (i.e. finite time and finite observational resolution).

It is worth remarking that the approach to complexity here discussed is not able to cover the many aspects of what in recent years has been indicated under this term [13]. Indeed *complexity* has been considered in many different fields of science and its meaning has become (sometimes) vague. A challenging problem in this direction is the definition of indicators which are able to fulfill the intuitive idea of complexity, namely one looks for quantities which give a low complexity value both for pure random sequences and completely regular ones [96]. Even if very interesting this last issue is not addressed here: from the point of view of predictability both a chaotic system and a purely random one are highly complex, i.e. unpredictable.

The review is organized as follows. Section 2 is devoted to the introduction of the basic concepts and ideas of dynamical systems, information theory and algorithmic complexity. In particular, we discuss the relations among Lyapunov exponents, Kolmogorov–Sinai entropy and algorithmic complexity and their relevance for predictability. All these quantities are properly defined only in specific asymptotic limits, that are: very long times and arbitrary accuracy. Since in realistic situations one has to deal with finite accuracy and finite time—as Keynes said, “in the long run we shall all be dead”—it is appealing to treat the predictability problem by taking into account these limitations. This is the subject of Section 3 where, relaxing the request of infinite time, we discuss the relevance of the finite time fluctuations of the “effective” Lyapunov exponent. In addition, relaxing the limit of infinitesimal perturbations, we introduce suitable tools, such as the finite size Lyapunov exponent (FSLE) and the  $\epsilon$ -entropy, for the treatment of non-arbitrary high accuracy, i.e. non-infinitesimal perturbations.

Sections 4 and 5 focus on high dimensional dynamical systems which deserve particular attention. Indeed because of the many degrees of freedom, and its interest in applications (e.g. in weather forecasting), it is necessary to consider the detailed behavior of perturbations and not only the asymptotic features (i.e. long time and infinitesimal amplitudes). Section 5 is devoted to fully developed turbulence (here introduced as an important prototype of high dimensional system) and its finite resolution properties in the inertial range.

In Section 6 we consider the effects of uncertainties on the evolution laws and we discuss systems containing some randomness. In such a situation there are two ways to approach the predictability: by considering either two trajectories generated with the same realization of randomness, or two trajectories evolving with different realizations. Both approaches are physically relevant in different contexts, and the results can be very different in presence of strong intermittency.

For the sake of completeness in Section 7 we discuss dynamical systems with discrete states, e.g., Cellular Automata.

Section 8 is dedicated to a discussion on data analysis. In particular we discuss the use of  $\epsilon$ -entropy and FSLE for a pragmatic classification of signals.

Section 9 reports some concluding remarks. In the Appendices we discuss some more technical details.

## 2. Two points of view

### 2.1. Dynamical systems approach

Two standard—tightly linked—indicators are largely used to quantify the behavior of a dynamical system with respect to the asymptotic evolution of an infinitesimal uncertainty: the largest Lyapunov exponent (LE) and the Kolmogorov–Sinai (or metric) entropy [74].

#### 2.1.1. Characteristic Lyapunov exponents

The characteristic Lyapunov exponents are somehow an extension of the linear stability analysis to the case of aperiodic motions. Roughly speaking, they measure the typical rate of exponential divergence of nearby trajectories. In this sense they give information on the rate of growth of a very small error on the initial state of a system.

Consider a dynamical system with an evolution law given, in the case of continuous time, by the differential equation

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}), \quad (2.1)$$

or, in the case of discrete time, by the map

$$\mathbf{x}(t+1) = \mathbf{G}(\mathbf{x}(t)). \quad (2.2)$$

In both cases, for simplicity, we suppose that a vector  $\mathbf{x} \in \mathbb{R}^d$  uniquely specifies one state of the system. We also assume that  $\mathbf{F}$  and  $\mathbf{G}$  are differentiable functions, that the evolution is well-defined for time intervals of arbitrary extension, and that the motion takes place in a bounded region of the phase space. We intend to study the separation between two trajectories,  $\mathbf{x}(t)$  and  $\mathbf{x}'(t)$ , starting from two close initial conditions,  $\mathbf{x}(0)$  and  $\mathbf{x}'(0) = \mathbf{x}(0) + \delta\mathbf{x}(0)$ , respectively.

As long as the difference between the trajectories,  $\delta\mathbf{x}(t) = \mathbf{x}'(t) - \mathbf{x}(t)$ , remains small (infinitesimal, strictly speaking), it can be regarded as a vector,  $\mathbf{z}(t)$ , in the tangent space. The time evolution of  $\mathbf{z}(t)$  is given by the linearized differential equations:

$$\frac{dz_i(t)}{dt} = \sum_{j=1}^d \left. \frac{\partial F_i}{\partial x_j} \right|_{\mathbf{x}(t)} z_j(t) \quad (2.3)$$

or, in the case of discrete time maps:

$$z_i(t+1) = \sum_{j=1}^d \left. \frac{\partial G_i}{\partial x_j} \right|_{\mathbf{x}(t)} z_j(t). \quad (2.4)$$

Under rather general hypothesis, Oseledec [169] proved that for almost all initial conditions  $\mathbf{x}(0)$ , there exists an orthonormal basis  $\{\mathbf{e}_i\}$  in the tangent space such that, for large times,

$$\mathbf{z}(t) = \sum_{i=1}^d c_i \mathbf{e}_i e^{\lambda_i t}, \quad (2.5)$$

where the coefficients  $\{c_i\}$  depends on  $\mathbf{z}(0)$ . The exponents  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$  are called *characteristic Lyapunov exponents*. If the dynamical system has an ergodic invariant measure, the

spectrum of LEs  $\{\lambda_i\}$  does not depend on the initial condition, except for a set of measure zero with respect to the natural invariant measure.

Loosely speaking, (2.5) tells us that in the phase space, where the motion evolves, a  $d$ -dimensional sphere of small radius  $\epsilon$  centered in  $\mathbf{x}(0)$  is deformed with time into an ellipsoid of semi-axes  $\epsilon_i(t) = \epsilon \exp(\lambda_i t)$ , directed along the  $\mathbf{e}_i$  vectors. Furthermore, for a generic small perturbation  $\delta\mathbf{x}(0)$ , the distance between a trajectory and the perturbed one behaves as

$$|\delta\mathbf{x}(t)| \sim |\delta\mathbf{x}(0)| e^{\lambda_1 t} [1 + O(e^{-(\lambda_1 - \lambda_2)t})]. \quad (2.6)$$

If  $\lambda_1 > 0$  we have a rapid (exponential) amplification of an error on the initial condition. In such a case, the system is chaotic and, de facto, unpredictable on the long times. Indeed, if we put  $\delta_0 = |\delta\mathbf{x}(0)|$  for the initial error, and we want to predict the states of the system with a certain tolerance  $\Delta$  (not too large), then the prediction is possible just up to a *predictability time* given by

$$T_p \sim \frac{1}{\lambda_1} \ln\left(\frac{\Delta}{\delta_0}\right). \quad (2.7)$$

This equation shows that  $T_p$  is basically determined by the largest Lyapunov exponent, since its dependence on  $\delta_0$  and  $\Delta$  is very weak. Because of its preeminent role, very often one simply refers to  $\lambda_1$  as “the Lyapunov exponent”, and one indicates it with  $\lambda$ .

Eq. (2.6) suggests how to numerically compute  $\lambda_1$ . We introduce the response, after a time  $t$ , to a perturbation on  $\mathbf{x}(\tau)$ , defined as follows:

$$R_\tau(t) \equiv \frac{|\mathbf{z}(\tau+t)|}{|\mathbf{z}(\tau)|} = \frac{|\delta\mathbf{x}(\tau+t)|}{|\delta\mathbf{x}(\tau)|}, \quad (2.8)$$

where, again,  $|\delta\mathbf{x}(\tau)|$  and  $|\delta\mathbf{x}(\tau+t)|$  are infinitesimal. The LE  $\lambda_1$  is obtained by averaging the logarithm of the response over the initial conditions or along the trajectory:

$$\lambda_1 = \lim_{t \rightarrow \infty} \frac{1}{t} \langle \ln R_\tau(t) \rangle, \quad (2.9)$$

where  $\langle \cdot \rangle$  denotes the time average  $\lim_{T \rightarrow \infty} (1/T) \int_{\tau_0}^{\tau_0+T} (\cdot) d\tau$ . The Oseledec’s theorem implies that  $(1/t) \ln R_\tau(t)$ , for large  $t$ , is a non-random quantity, i.e. for almost all the initial conditions its value does not depend on the specific initial condition. Therefore, for large times, the average in (2.9) can be neglected.

As the typical growth rate of a generic small segment in phase space is driven by the largest LE, the sum of the first  $n$  ( $\leq d$ ) Lyapunov exponents controls the variations of small  $n$ -dimensional volumes in phase space. This gives us a way to compute the sub-leading Lyapunov exponents. After the selection of  $n \leq d$  non-parallel tangent vectors  $[\mathbf{z}^{(1)}(0), \dots, \mathbf{z}^{(n)}(0)]$ , one introduces the  $n$ th-order response  $R_\tau^{(n)}(t)$  [20]

$$R_\tau^{(n)}(t) \equiv \frac{|\mathbf{z}_1(t+\tau) \times \mathbf{z}_2(t+\tau) \times \dots \times \mathbf{z}_n(t+\tau)|}{|\mathbf{z}_1(\tau) \times \mathbf{z}_2(\tau) \times \dots \times \mathbf{z}_n(\tau)|}. \quad (2.10)$$

Analogously to the LE, it can be shown that

$$\sum_{i=1}^n \lambda_i = \lim_{t \rightarrow \infty} \frac{1}{t} \langle \ln R_\tau^{(n)}(t) \rangle. \quad (2.11)$$

Let us stress that the Lyapunov exponents give information on the typical behaviors along a generic trajectory, followed for infinite time and keeping the perturbation infinitesimally small. In this respect, they are global quantities characterizing fine-grained properties of a system.

### 2.1.2. Kolmogorov–Sinai entropy

The LE,  $\lambda$ , gives a first quantitative information on how rapidly we lose the ability of predicting the evolution of a system. A state, initially determined with an error  $\delta\mathbf{x}(0)$ , after a time enough larger than  $1/\lambda$ , may be found almost everywhere in the region of motion. In this respect, the Kolmogorov–Sinai (KS) entropy,  $h_{\text{KS}}$ , supplies a more refined information. The error on the initial state is due to the maximal resolution we use for observing the system. For simplicity, let us assume the same resolution  $\epsilon$  for each degree of freedom. We build a partition of the phase space with cells of volume  $\epsilon^d$ , so that the state of the system at  $t = t_0$  is found in a region of volume  $V_0 = \epsilon^d$  around  $\mathbf{x}(t_0)$ . Now we consider the trajectories starting from  $V_0$  at  $t_0$  and sampled at discrete times  $t_j = j\tau$  ( $j = 1, 2, 3, \dots, t$ ); in the case of a map one can put  $\tau = 1$ . Since we are considering motions that evolve in a bounded region, all the trajectories visit a finite number of different cells, each one identified by a symbol. In this way a unique sequence of symbols  $\{s(0), s(1), s(2), \dots\}$  is associated with a given trajectory. In a chaotic system, although each evolution  $\mathbf{x}(t)$  is univocally determined by  $\mathbf{x}(t_0)$ , a great number of different symbolic sequences originates by the same initial cell, because of the divergence of nearby trajectories. The total number of the admissible symbolic sequences,  $\tilde{N}(\epsilon, t)$ , increases exponentially with a rate given by the topological entropy

$$h_T = \lim_{\epsilon \rightarrow 0} \lim_{t \rightarrow \infty} \frac{1}{t} \ln \tilde{N}(\epsilon, t). \quad (2.12)$$

However, if we consider only the number of sequences  $N_{\text{eff}}(\epsilon, t) \leq \tilde{N}(\epsilon, t)$  which appear with very high probability in the long time limit—those that can be numerically or experimentally detected and that are associated with the natural measure—we arrive at a more physical quantity called the Kolmogorov–Sinai or metric entropy [74]:

$$h_{\text{KS}} = \lim_{\epsilon \rightarrow 0} \lim_{t \rightarrow \infty} \frac{1}{t} \ln N_{\text{eff}}(\epsilon, t) \leq h_T. \quad (2.13)$$

$h_{\text{KS}}$  quantifies the long time exponential rate of growth of the number of the effective coarse-grained trajectories of a system. This suggests a link with information theory where the Shannon entropy measures the mean asymptotic growth of the number of the typical sequences—the ensemble of which has probability almost 1—emitted by a source. In the following we will discuss in more detail the KS-entropy and its relation with the information theory. Here we obtain, by means of a heuristic reasoning, the relation among  $h_{\text{KS}}$  and Lyapunov exponents.

We may wonder what is the number of cells where, at a time  $t > t_0$ , the points that evolved from  $V_0$  can be found, i.e. we wish to know how big is the coarse-grained volume  $V(\epsilon, t)$ , occupied by the states evolved from  $V_0$ , if the minimum volume we can observe is  $\epsilon^d$ . As stated at the end of the preceding subsection, we have  $V(t) \sim V_0 \exp(t \sum_{i=1}^d \lambda_i)$ . However, this is true only in the limit  $\epsilon \rightarrow 0$ . In this (unrealistic) limit,  $V(t) = V_0$  for a conservative system (where  $\sum_{i=1}^d \lambda_i = 0$ ) and  $V(t) < V_0$  for a dissipative system (where  $\sum_{i=1}^d \lambda_i < 0$ ). As a consequence of limited resolution power, in the evolution of the volume  $V_0 = \epsilon^d$  the effect

of the contracting directions (associated with the negative Lyapunov exponents) is completely lost. We can experience only the effect of the expanding directions, associated with the positive Lyapunov exponents. As a consequence, in the typical case, the coarse grained volume behaves as

$$V(\epsilon, t) \sim V_0 e^{(\sum_{\lambda_i > 0} \lambda_i)t}, \quad (2.14)$$

when  $V_0$  is small enough. Since  $N_{\text{eff}}(\epsilon, t) \propto V(\epsilon, t)/V_0$ , one has

$$h_{\text{KS}} = \sum_{\lambda_i > 0} \lambda_i. \quad (2.15)$$

This argument can be made more rigorous with a proper mathematical definition of the metric entropy. In this case one derives the Pesin relation [179,74]

$$h_{\text{KS}} \leq \sum_{\lambda_i > 0} \lambda_i. \quad (2.16)$$

Because of its relation with the Lyapunov exponents—or by the definition (2.13)—it is clear that also  $h_{\text{KS}}$  is a fine-grained and global characterization of a dynamical system.

The metric entropy is an invariant characteristic quantity of a dynamical system [125,204], i.e. given two systems with invariant measures, their KS-entropies exist and they are equal if the systems are isomorphic [31]. This intrinsic quantity may be properly defined by means of tools borrowed from the mathematical theory of communication.

## 2.2. Information theory approach

In experimental investigations of physical processes, we typically have access to the system only through a measuring device which produces a time record of a certain observable, i.e. a sequence of data. In this regard a system, whether or not chaotic, generates messages and may be regarded as a source of information. This observation opens the possibility to study dynamical systems from a very interesting point of view.

Information has found a proper characterization in the framework of the theory of communication to cope with the practical problem of transmitting a message in the cheapest way without losing information. The characterization of the information contained in a sequence can be approached by two very different points of view. The first one, that of information theory [201], is a statistical approach, i.e., it does not consider the transmission of a specific message (sequence) but refers to the statistical properties of all the messages emitted by the source. Information theory approach characterizes the source of information, so that it gives us a powerful method to characterize chaotic systems.

The second point of view considers the problem of characterizing a single sequence. This latter has led to the theory of algorithmic complexity and algorithmic information theory [53,126,207].

### 2.2.1. Shannon entropy

At the end of 1940s, Shannon [201] introduced rather powerful concepts and techniques for a systematic study of sources emitting sequences of discrete symbols (e.g. binary digit sequences). Originally information theory was introduced in the practical context of electric communications;

nevertheless in a few years it became an important branch of both pure and applied probability theory with strong relations with other fields as computer science, cryptography, biology and physics [230].

For the sake of self-consistency we briefly recall the basic concepts and ideas about the Shannon entropy. Consider a source that can output  $m$  different symbols; denote by  $s(t)$  the symbol emitted by the source at time  $t$  and with  $P(C_N)$  the probability that a given word  $C_N = (s(1), s(2), \dots, s(N))$ , of length  $N$ , is emitted:

$$P(C_N) = P(s(1), s(2), \dots, s(N)) . \quad (2.17)$$

We assume that the source is stationary, so that for the sequences  $\{s(t)\}$  the time translation invariance holds:  $P(s(1), \dots, s(N)) = P(s(t+1), \dots, s(t+N))$ .

Now we introduce the  $N$ -block entropies

$$H_N = - \sum_{\{C_N\}} P(C_N) \ln P(C_N) , \quad (2.18)$$

and the differences

$$h_N = H_{N+1} - H_N , \quad (2.19)$$

whose meaning is the average information supplied by the  $(N+1)$ th symbol, provided the  $N$  previous ones are known. One can also say that  $h_N$  is the average uncertainty about the  $(N+1)$ th symbol, provided the  $N$  previous ones are given. For a stationary source the limits in the following equations exist, are equal and define the Shannon entropy  $h_{\text{Sh}}$ :

$$h_{\text{Sh}} = \lim_{N \rightarrow \infty} h_N = \lim_{N \rightarrow \infty} \frac{H_N}{N} . \quad (2.20)$$

The  $h_N$  are non-increasing quantities:  $h_{N+1} \leq h_N$ ; that is: the knowledge of a longer past history cannot increase the uncertainty on the next outcome. In the case of a  $k$ th-order Markov process  $h_N = h_{\text{Sh}}$  for all  $N \geq k$ . This is because a  $k$ th-order Markov process has the property that the conditional probability to have a given symbol only depends on the results of the last  $k$  times, i.e.

$$P(s(t)|s(t-1), s(t-2), \dots) = P(s(t)|s(t-1), s(t-2), \dots, s(t-k)) . \quad (2.21)$$

The Shannon entropy is a measure of the “surprise” the source emitting the sequences can reserve to us, since it quantifies the richness (or “complexity”) of the source. This can be precisely expressed by the first theorem of Shannon–McMillan [121] that applies to stationary ergodic sources:

If  $N$  is large enough, the ensemble of  $N$ -long subsequences can be partitioned in two classes,  $\Omega_1(N)$  and  $\Omega_0(N)$  such that all the words  $C_N \in \Omega_1(N)$  have the same probability  $P(C_N) \sim \exp(-Nh_{\text{Sh}})$  and

$$\sum_{C_N \in \Omega_1(N)} P(C_N) \rightarrow 1 \quad \text{for } N \rightarrow \infty \quad (2.22)$$

while

$$\sum_{C_N \in \Omega_0(N)} P(C_N) \rightarrow 0 \quad \text{for } N \rightarrow \infty . \quad (2.23)$$

The meaning of this theorem is the following. An  $m$ -states process admits in principle  $m^N$  possible sequences of length  $N$ , but the number of typical sequences  $N_{\text{eff}}(N)$  (those ones in  $\Omega_1(N)$ ) effectively observable is

$$N_{\text{eff}}(N) \sim \exp(Nh_{\text{Sh}}). \quad (2.24)$$

Note that  $N_{\text{eff}} \ll m^N$  if  $h_{\text{Sh}} < \ln m$ . Moreover the entropy per symbol,  $h_{\text{Sh}}$ , is a property of the source. Because of the ergodicity it can be obtained by analyzing just one single sequence in the ensemble of the typical ones, and it can also be viewed as a property of each one of the typical sequences. Therefore, as in the following, one may speak about the Shannon entropy of a sequence.

The above theorem in the case of processes without memory is nothing but the law of large numbers. Let us observe that (2.24) is somehow the equivalent in information theory of the Boltzmann equation in statistical thermodynamics:  $S \propto \ln W$ ,  $W$  being the number of possible microscopic configurations and  $S$  the thermodynamic entropy. This justifies the name “entropy” for  $h_{\text{Sh}}$ . Under rather natural assumptions it is possible to prove that the Shannon entropy, apart from a multiplicative factor, is the unique quantity which characterizes the “surprise” [121].

Let us now mention another important result about the Shannon entropy. It is not difficult to recognize that the quantity  $h_{\text{Sh}}$  sets the maximum compression rate of a sequence  $\{s(1), s(2), s(3), \dots\}$ . Indeed a theorem of Shannon states that, if the length  $T$  of a sequence is large enough, one cannot construct another sequence (always using  $m$  symbols), from which it is possible to reconstruct the original one, whose length is smaller than  $(h_{\text{Sh}}/\ln m)T$  [201]. In other words  $h_{\text{Sh}}/\ln m$  is the maximum allowed compression rate.

The relation between Shannon entropy and the data compression problem is well highlighted by considering the Shannon–Fano code to map  $\mathcal{N}$  objects (e.g. the  $N$ -words  $C_N$ ) into sequences of binary digits  $(0, 1)$  [219]. The main goal in building a code is to define the most efficient coding procedure, i.e. the one which generates the shortest possible (coded) sequence. The Shannon–Fano code is as follows. At first one orders the  $\mathcal{N}$  objects according to their probabilities, in a decreasing way,  $p_1, p_2, \dots, p_{\mathcal{N}}$ . Then the passage from the  $\mathcal{N}$  objects to the symbols  $(0, 1)$  is obtained by defining the coding  $E(r)$ , of binary length  $\ell(E(r))$ , of the  $r$ th object with the requirement that the expected total length,  $\sum_r p_r \ell_r$ , be the minimal one. This can be realized with the following choice:

$$-\ln_2 p_r \leq \ell(E(r)) < 1 - \ln_2 p_r. \quad (2.25)$$

In this way highly probable objects are mapped into short code words while the low probability ones are mapped to longer code words. So that the average code length is bounded by

$$\frac{H_N}{\ln 2} \leq \sum_r p_r \ell(E(r)) \leq \frac{H_N + 1}{\ln 2}, \quad (2.26)$$

and in the limit  $N \rightarrow \infty$  one has

$$\lim_{N \rightarrow \infty} \frac{\langle \ell_N \rangle}{N} = \lim_{N \rightarrow \infty} \frac{\sum_r p_r \ell(E(r))}{N} = \frac{h_{\text{Sh}}}{\ln 2}, \quad (2.27)$$

i.e., in a good coding, the mean length of a  $N$ -word is equal to  $N$  times the Shannon entropy (apart from a multiplicative factor, due to the fact that in the definition (2.20) of  $h_{\text{Sh}}$  we used the natural logarithm and here we want to work with a two symbol code).

An alternative coding method, based on variable length words, is due to Ziv and Lempel [138]. Remarkably it is very efficient for data compression and gives the same asymptotic result of the Shannon–Fano code.

### 2.2.2. Again on the Kolmogorov–Sinai entropy

After the introduction of the Shannon entropy we can give a more precise definition of the KS-entropy. Consider a trajectory,  $\mathbf{x}(t)$ , generated by a deterministic system, sampled at the times  $t_j = j\tau$ , with  $j = 1, 2, 3, \dots$ . Perform a finite partition  $\mathcal{A}$  of the phase space. With the finite number of symbols  $\{s\}_{\mathcal{A}}$  enumerating the cells of the partition, the time-discretized trajectory  $\mathbf{x}(t_j)$  determines a sequence  $\{s(1), s(2), s(3), \dots\}$ , whose meaning is clear: at the time  $t_j$  the trajectory is in the cell labeled by  $s(j)$ . To each subsequence of length  $N \cdot \tau$  one can associate a word of length  $N$ :  $W_j^N(\mathcal{A}) = (s(j), s(j+1), \dots, s(j+(N-1)))$ . If the system is ergodic, as we suppose, from the frequencies of the words one obtains the probabilities by which one calculates the block entropies  $H_N(\mathcal{A})$ :

$$H_N(\mathcal{A}) = - \sum_{\{W^N(\mathcal{A})\}} P(W^N(\mathcal{A})) \ln P(W^N(\mathcal{A})). \quad (2.28)$$

It is important to note that the probabilities  $P(W^N(\mathcal{A}))$ , computed by the frequencies of  $W^N(\mathcal{A})$  along a trajectory, are essentially dependent on the stationary measure selected by the trajectory. This implies a dependence on this measure of all the quantities defined below,  $h_{\text{KS}}$  included. We shall always understand to consider the natural invariant measure and do not indicate this kind of dependence. The entropy per unit time of the trajectory with respect to the partition  $\mathcal{A}$ ,  $h(\mathcal{A})$ , is defined as follows:

$$h_N(\mathcal{A}) = \frac{1}{\tau} [H_{N+1}(\mathcal{A}) - H_N(\mathcal{A})], \quad (2.29)$$

$$h(\mathcal{A}) = \lim_{N \rightarrow \infty} h_N(\mathcal{A}) = \frac{1}{\tau} \lim_{N \rightarrow \infty} \frac{1}{N} H_N(\mathcal{A}). \quad (2.30)$$

Notice that, for the deterministic systems we are considering, the entropy per unit time does not depend on the sampling time  $\tau$  [31]. The KS-entropy, by definition, is the supremum of  $h(\mathcal{A})$  over all possible partitions [31,74]

$$h_{\text{KS}} = \sup_{\mathcal{A}} h(\mathcal{A}). \quad (2.31)$$

It is not simple at all to determine  $h_{\text{KS}}$  according to this definition. A useful tool in this respect would be the Kolmogorov–Sinai theorem, by means of which one is granted that  $h_{\text{KS}} = h(\mathcal{G})$  if  $\mathcal{G}$  is a generating partition. A partition is said to be generating if every infinite sequence  $\{s(n)\}_{n=1, \dots, \infty}$  individuates a single initial point. However the difficulty now is that, with the exception of very simple cases, we do not know how to construct a generating partition. We only know that, according to the Krieger theorem [133], there exists a generating partition with  $k$  elements such that  $e^{h_{\text{KS}}} < k \leq e^{h_{\text{KS}}} + 1$ . Then, a more tractable way to define  $h_{\text{KS}}$  is based upon considering the partition  $\mathcal{A}_\epsilon$  made up by a grid of cubic cells of edge  $\epsilon$ , from which one has

$$h_{\text{KS}} = \lim_{\epsilon \rightarrow 0} h(\mathcal{A}_\epsilon). \quad (2.32)$$

We expect that  $h(\mathcal{A}_\epsilon)$  becomes independent of  $\epsilon$  when  $\mathcal{A}_\epsilon$  is so fine to be “contained” in a generating partition.

For discrete time maps what has been exposed above is still valid, with  $\tau = 1$  (however, Krieger’s theorem only applies to invertible maps).

An important point to note is that, for a truly stochastic (i.e. non-deterministic) system, with continuous states,  $h(\mathcal{A}_\epsilon)$  is not bounded and  $h_{KS} = \infty$ .

### 2.2.3. Algorithmic complexity

We saw that the Shannon entropy puts a limit on how efficiently the ensemble of the messages emitted by a source can be coded. We may wonder about the compressibility properties of a single sequence. This problem can be addressed by means of the notion of algorithmic complexity, that is concerned with the difficulty in reproducing a given string of symbols.

Everybody agrees that the binary digits sequence

$$01110100010111001011010\dots \tag{2.33}$$

is, in some sense, more random than

$$10101010101010101010\dots \tag{2.34}$$

The notion of algorithmic complexity, independently introduced by Kolmogorov [126], Chaitin [53,56] and Solomonov [207], is a way to formalize the intuitive idea of randomness of a sequence.

Consider a binary digit sequence (this does not constitute a limitation) of length  $N$ ,  $(i_1, i_2, \dots, i_N)$ , generated by a certain computer code on some machine  $\mathcal{M}$ . One defines the algorithmic complexity, or algorithmic information content,  $K_{\mathcal{M}}(N)$  of a  $N$ -sequence as the bit length of the shortest computer program able to give the  $N$ -sequence and to stop afterward. Of course, such a length depends not only on the sequence but also on the machine. However, a result of Kolmogorov [126] proves the existence of a universal computer,  $\mathcal{U}$ , that is able to perform the same computation a program  $p$  makes on  $\mathcal{M}$  with a modification of  $p$  that depends only on  $\mathcal{M}$ . This implies that for all finite strings:

$$K_{\mathcal{U}}(N) \leq K_{\mathcal{M}}(N) + C_{\mathcal{M}} \tag{2.35}$$

where  $K_{\mathcal{U}}(N)$  is the complexity with respect to the universal computer and  $C_{\mathcal{M}}$  depends only on the machine  $\mathcal{M}$ . At this point we can consider the algorithmic complexity with respect to a universal computer—and we can drop the machine dependence in the symbol for the algorithmic complexity,  $K(N)$ . The reason is that we are interested in the limit of very long sequences,  $N \rightarrow \infty$ , for which one defines the algorithmic complexity per unit symbol:

$$\mathcal{C} = \lim_{N \rightarrow \infty} \frac{K(N)}{N} \tag{2.36}$$

that, because of (2.35), is an intrinsic quantity, i.e. independent of the machine.

Now coming back to the two  $N$ -sequences (2.33) and (2.34), it is obvious that the second one can be obtained with a small-length minimal program, i.e.

$$\text{“PRINT } 10 \frac{N}{2} \text{ TIMES”} . \tag{2.37}$$

The bit length of the above program is  $O(\ln N)$  and therefore when taking the limit  $N \rightarrow \infty$  in (2.36), one obtains  $\mathcal{C} = 0$ . Of course,  $K(N)$  cannot exceed  $N$ , since the sequence can always be obtained with a trivial program (of bit length  $N$ )

$$\text{“PRINT } i_1, i_2, \dots, i_N \text{”} . \quad (2.38)$$

Therefore, in the case of a very irregular sequence, e.g., (2.33) one expects  $K(N) \propto N$ , i.e.  $\mathcal{C} \neq 0$ . In such a case one calls the sequence complex (i.e. of nonzero algorithmic complexity) or random.

Algorithmic complexity cannot be computed. Since the algorithm which computes  $K(N)$  cannot have less than  $K(N)$  binary digits and since in the case of random sequences  $K(N)$  is not bounded in the limit  $N \rightarrow \infty$  then it cannot be computed in the most interesting cases. The un-computability of  $K(N)$  may be understood in terms of Gödel’s incompleteness theorem [54–56]. Beyond the problem of whether or not  $K(N)$  is computable in a specific case, the concept of algorithmic complexity brings an important improvement to clarify the vague and intuitive notion of randomness. For a systematic treatment of algorithmic complexity, information theory and data compression see [142].

There exists a relation between the Shannon entropy,  $h_{\text{Sh}}$ , and the algorithmic complexity  $\mathcal{C}$ . It is possible to show that

$$\lim_{N \rightarrow \infty} \frac{\langle K(N) \rangle}{H_N} = \frac{1}{\ln 2} , \quad (2.39)$$

where  $\langle K(N) \rangle = \sum_{C_N} P(C_N) K_{C_N}(N)$ , being  $K_{C_N}(N)$  the algorithmic complexity of the  $N$ -words. Therefore the expected complexity  $\langle K(N)/N \rangle$  is asymptotically equal to the Shannon entropy (apart the  $\ln 2$  factor).

Eq. (2.39) stems from the results of the Shannon–McMillan theorem about the two classes of sequences (i.e.  $\Omega_1(N)$  and  $\Omega_0(N)$ ). Indeed in the limit of very large  $N$ , the probability to observe a sequence in  $\Omega_1(N)$  goes to 1, and the entropy of such a sequence as well as its algorithmic complexity equals the Shannon entropy. Apart from the numerical coincidence of the values of  $\mathcal{C}$  and  $h_{\text{Sh}}/\ln 2$  there is a conceptual difference between the information theory and the algorithmic complexity theory. The Shannon entropy essentially refers to the information content in a statistical sense, i.e. it refers to an ensemble of sequences generated by a certain source. On the other hand, the algorithmic complexity defines the information content of an individual sequence [96]. Of course, as noted above, the fact that it is possible to use probabilistic arguments on an individual sequence is a consequence of the ergodicity of the system, which allows to assume good statistical properties of arbitrary long  $N$ -words.

For a dynamical system one can define the notion of algorithmic complexity of the trajectory starting from the point  $\mathbf{x}$ ,  $\mathcal{C}(\mathbf{x})$ . This requires the introduction of finite open coverings of the phase space, the consideration of symbolic sequences thus generated, for the given trajectory, sampled at constant time intervals, and the searching of the supremum of the algorithmic complexity per symbol at varying the coverings [5]. Then it can be shown—Brudno’s and White’s theorems [42,225]—that for almost all  $\mathbf{x}$  (we always mean with respect to the natural invariant

measure) one has

$$\mathcal{C}(\mathbf{x}) = \frac{h_{\text{KS}}}{\ln 2}, \quad (2.40)$$

where, as before, the factor  $\ln 2$  is a conversion factor between natural logarithms and bits.

This result says that the KS-entropy quantifies not only the richness, or surprise, of a dynamical system but also the difficulty of describing (almost) everyone of its typical sequences.

### 2.3. Algorithmic complexity and Lyapunov exponent

Summing up, the theorem of Pesin together with those of Brudno and White show that a chaotic dynamical system may be seen as a source of messages that cannot be described in a concise way, i.e. they are complex. We expose here two examples that may help in understanding the previous conclusion and the relation between the Lyapunov exponent, the KS-entropy and the algorithmic complexity.

Following Ford [79,80], let us consider the shift map

$$x(t+1) = 2x(t) \pmod{1}, \quad (2.41)$$

which has  $\lambda = \ln 2$ . If one writes an initial condition in binary representation, i.e.,  $x(0) = \sum_{j=1}^{\infty} a_j 2^{-j}$ , such that  $a_j = 0$  or  $1$ , it is clearly seen that the action of the map (2.41) on  $x(0)$  is just a shift of the binary coordinates:

$$x(1) = \sum_{j=1}^{\infty} a_{j+1} 2^{-j} \quad \dots \quad x(t) = \sum_{j=1}^{\infty} a_{j+t} 2^{-j}. \quad (2.42)$$

With this observation it is possible to verify that  $K(N) \simeq N$  for almost all the solutions of (2.41). Let us consider  $x(t)$  with accuracy  $2^{-k}$  and  $x(0)$  with accuracy  $2^{-l}$ , of course  $l=t+k$ . This means that, in order to obtain the  $k$  binary digits of the output solution of (2.41), we must use a program of length no less than  $l=t+k$ . Basically one has to specify  $a_1, a_2, \dots, a_l$ . Therefore we are faced with the problem of the algorithmic complexity of the binary sequence  $(a_1, a_2, \dots, a_{\infty})$  which determines the initial condition  $x(0)$ . Martin-Löf [156] proved a remarkable theorem stating that, with respect to the Lebesgue measure, almost all the binary sequences  $(a_1, a_2, \dots, a_{\infty})$ , which represent real numbers in  $[0, 1]$ , have maximum complexity, i.e.  $K(N) \simeq N$ . In practice, no human being will ever be able to distinguish the typical sequence  $(a_1, a_2, \dots, a_{\infty})$  from the output of a fair coin toss.

Finally, let us consider a 1d chaotic map

$$x(t+1) = f(x(t)). \quad (2.43)$$

If one wants to transmit to a friend on Mars the sequence  $\{x(t), t = 1, 2, \dots, T\}$  accepting only errors smaller than a tolerance  $\Delta$ , one can use the following strategy [174]:

- (1) Transmit the rule (2.43): for this task one has to use a number of bits independent of the length of the sequence  $T$ .
- (2) Specify the initial condition  $x(0)$  with a precision  $\delta_0$  using a finite number of bits which is independent of  $T$ .

- (3) Let the system evolve till the first time  $\tau_1$  such that the distance between two trajectories, that was initially  $\delta x(0) = \delta_0$ , equals  $\Delta$  and then specify again the new initial condition  $x(\tau_1)$  with precision  $\delta_0$ .
- (4) Let the system evolve and repeat the procedure (2) and (3), i.e. each time the error acceptance tolerance is reached specify the initial conditions,  $x(\tau_1 + \tau_2)$ ,  $x(\tau_1 + \tau_2 + \tau_3) \dots$ , with precision  $\delta_0$ . The times  $\tau_1, \tau_2, \dots$  are defined as follows: putting  $x'(\tau_1) = x(\tau_1) + \delta_0$ ,  $\tau_2$  is given by the minimum time such that  $|x'(\tau_1 + \tau_2) - x(\tau_1 + \tau_2)| \geq \Delta$  and so on.

By following the steps (1)–(4) the friend on Mars can reconstruct with a precision  $\Delta$  the sequence  $\{x(t)\}$  simply iterating on a computer the system (2.43) between 1 and  $\tau_1 - 1$ ,  $\tau_1$  and  $\tau_1 + \tau_2 - 1$ , and so on.

Let us now compute the amount of bits necessary to implement the above transmission (1)–(4). For simplicity of notation we introduce the quantities

$$\gamma_i = \frac{1}{\tau_i} \ln \frac{\Delta}{\delta_0} \quad (2.44)$$

which can be considered as a sort of *effective* Lyapunov exponents (see Section 3.1). The LE  $\lambda$  can be written in terms of  $\{\gamma_i\}$  as follows:

$$\lambda = \langle \gamma_i \rangle = \frac{\sum_i \tau_i \gamma_i}{\sum_i \tau_i} = \frac{1}{\bar{\tau}} \ln \frac{\Delta}{\delta_0}, \quad (2.45)$$

where

$$\bar{\tau} = \frac{1}{N} \sum \tau_i$$

is the average time after which we have to transmit the new initial condition (let us observe that to obtain  $\lambda$  from the  $\gamma_i$ 's one has to perform the average (2.45) because the transmission time,  $\tau_i$ , is not constant). If  $T$  is large enough the number of transmissions,  $N$ , is  $T/\bar{\tau} \simeq \lambda T / \ln(\Delta/\delta_0)$ . Therefore, noting that in each transmission for the reduction of the error from  $\Delta$  to  $\delta_0$  one needs to use  $\ln_2(\Delta/\delta_0)$  bits, the total amount of bits used in the whole transmission is

$$\frac{T}{\bar{\tau}} \ln_2 \frac{\Delta}{\delta_0} = \frac{\lambda}{\ln 2} T. \quad (2.46)$$

In other words the number of bits for unit time is proportional to  $\lambda$ .

In more than one dimension, we have simply to replace  $\lambda$  with  $h_{KS}$  in (2.46). To intuitively understand this point one has to repeat the above transmission procedure in each of the expanding directions.

In this section, we briefly discussed how the limit of predictability, the fact that a sequence cannot be arbitrarily compressed, and the information contained in a signal are deeply related. In the following we will mainly discuss the dynamical system point of view, i.e., in terms of Lyapunov exponents, Kolmogorov Sinai entropy and their generalizations for less ideal cases.

### 3. Limits of the Lyapunov exponent for predictability

We saw how a first obvious way for quantifying the predictability of a physical system is in terms of the *predictability time*  $T_p$ , i.e. the time interval on which one can typically forecast

the system. A simple argument previously suggested

$$T_p \sim \frac{1}{\lambda} \ln \left( \frac{\Delta}{\delta_0} \right). \quad (3.1)$$

However, in any realistic system, relation (3.1) is too naive to be of actual relevance. Indeed, it does not take into account some basic features of dynamical systems:

- The Lyapunov exponent (2.9) is a global quantity: it measures the average rate of divergence of nearby trajectories. In general, there exist finite-time fluctuations and the probability distribution functions (pdf) of these fluctuations is important for the characterization of predictability. The *generalized Lyapunov exponents* have been introduced with the purpose to take into account these fluctuations [23,85].
- The Lyapunov exponent is defined for the linearized dynamics, i.e., by computing the rate of separation of two infinitesimally close trajectories. On the other hand, for measuring the predictability time (3.1) one is interested in a finite tolerance  $\Delta$ , because the initial error  $\delta_0$  is typically finite. A recent generalization of the Lyapunov exponent to *finite size* errors extends the study of the perturbation growth to the non-linear regime, i.e. both  $\delta_0$  and  $\Delta$  are not infinitesimal [11,12].

### 3.1. Characterization of finite-time fluctuations

Let us consider the linear response, at a delay  $t$ , to an infinitesimal perturbation  $\delta \mathbf{x}(0)$ :

$$R(t) = \frac{|\delta \mathbf{x}(t)|}{|\delta \mathbf{x}(0)|}, \quad (3.2)$$

from which the LE is computed according to (2.9). In order to take into account the finite-time fluctuations, we can compute the different moments  $\langle R(t)^q \rangle$  and introduce the so-called generalized Lyapunov exponents (of order  $q$ ) [23,85]:

$$L(q) = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \langle R(t)^q \rangle, \quad (3.3)$$

where  $\langle \dots \rangle$  indicates the time average along the trajectory (see Section 2). It is easy to show that

$$\lambda_1 = \left. \frac{dL(q)}{dq} \right|_{q=0}. \quad (3.4)$$

In the absence of fluctuations,  $\lambda_1$  completely characterizes the error growth and we have  $L(q) = \lambda_1 q$ , while in the general case  $L(q)$  is concave in  $q$  [172]. Before discussing the properties of the generalized Lyapunov exponents, let us consider a simple example with a nontrivial  $L(q)$ . The model is the one-dimensional map

$$x(t+1) = \begin{cases} \frac{x(t)}{a} & \text{for } 0 \leq x \leq a, \\ \frac{1-x(t)}{1-a} & \text{for } a < x \leq 1, \end{cases} \quad (3.5)$$

which for  $a = \frac{1}{2}$  reduces to the tent map. For  $a \neq \frac{1}{2}$  the system is characterized by two different growth rates. The presence of different growth rates makes  $L(q)$  non-linear in  $q$ . Since the map

(3.5) is piecewise linear and with a uniform invariant density, the explicit computation of  $L(q)$  is very easy. The moments of the response after a time  $t$  are simply given by

$$\langle R(t)^q \rangle = \left[ a \left( \frac{1}{a} \right)^q + (1-a) \left( \frac{1}{1-a} \right)^q \right]^t . \tag{3.6}$$

From (3.3) and (3.6) we have

$$L(q) = \ln[a^{1-q} + (1-a)^{1-q}] , \tag{3.7}$$

which recovers the non-intermittent limit  $L(q) = q \ln 2$  in the symmetric case  $a = 1/2$ . In the general case, assuming  $0 \leq a < 1/2$ , we have that for  $q \rightarrow +\infty$ ,  $L(q)$  is dominated by the less probable, most unstable contributions and  $L(q)/q \simeq -\ln(a)$ . In the opposite limit,  $q \rightarrow -\infty$ , we obtain  $L(q)/q \simeq -\ln(1-a)$ .

We now show how  $L(q)$  is related to the fluctuations of  $R(t)$  at finite time  $t$ . Define an “effective” Lyapunov exponent  $\gamma(t)$  at time  $t$  by

$$R(t) \sim e^{\gamma(t)t} . \tag{3.8}$$

In the limit  $t \rightarrow \infty$ , the Oseledec theorem [169] assures that, for typical trajectories,  $\gamma(t) = \lambda_1 = -a \ln a - (1-a) \ln(1-a)$ . Therefore, for large  $t$ , the probability density of  $\gamma(t)$  is peaked at the most probable value  $\lambda_1$ . Let us introduce the probability density  $P_t(\gamma)$  of observing a given  $\gamma$  on a trajectory of length  $t$ . Large deviation theory suggests

$$P_t(\gamma) \sim e^{-S(\gamma)t} , \tag{3.9}$$

where  $S(\gamma)$  is the Cramer function [216]. The Oseledec theorem implies that  $\lim_{t \rightarrow \infty} P_t(\gamma) = \delta(\gamma - \lambda_1)$ , this gives a constraint on the Cramer function, i.e.  $S(\gamma = \lambda_1) = 0$  and  $S(\gamma) > 0$  for  $\gamma \neq \lambda_1$ .

The Cramer function  $S(\gamma)$  is related to the generalized Lyapunov exponent  $L(q)$  through a Legendre transform. Indeed, at large  $t$ , one has

$$\langle R(t)^q \rangle = \int d\gamma P_t(\gamma) e^{q\gamma t} \sim e^{L(q)t} , \tag{3.10}$$

by a steepest descent estimation one obtains

$$L(q) = \max_{\gamma} [q\gamma - S(\gamma)] . \tag{3.11}$$

In other words, each value of  $q$  selects a particular  $\gamma^*(q)$  given by

$$\left. \frac{dS(\gamma)}{d\gamma} \right|_{\gamma^*} = q . \tag{3.12}$$

We have already discussed that, for negligible fluctuations of the “effective” Lyapunov exponents, the LE completely characterizes the error growth and  $L(q) = \lambda_1 q$ . In presence of fluctuations, the probability distribution for  $R(t)$  can be approximated by a log-normal distribution. This can be understood assuming weak correlations in the response function so that (3.2) factorizes in several independent contributions and the central limit theorem applies. We can thus write

$$P_t(R) = \frac{1}{R \sqrt{2\pi\mu t}} \exp\left(-\frac{(\ln R - \lambda_1 t)^2}{2\mu t}\right) , \tag{3.13}$$

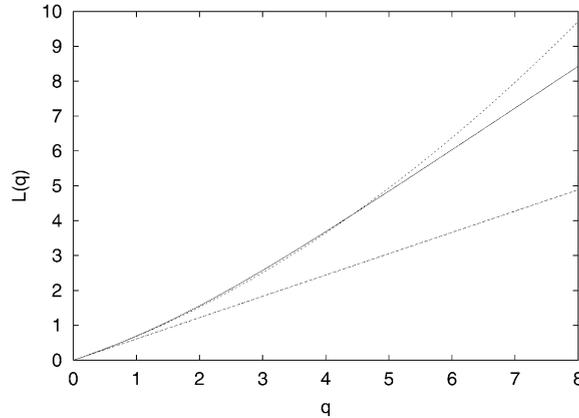


Fig. 1. Generalized Lyapunov exponent,  $L(q)$  for the map (3.5) with  $a = 0.3$  (solid line) compared with the linear non-intermittent approximation,  $\lambda_1 q$  (dashed line), and with the log-normal one (Eq. (3.16), (dotted line)).

where  $\lambda_1$  and  $\mu$  are given by

$$\begin{aligned} \lambda_1 &= \lim_{t \rightarrow \infty} \frac{1}{t} \langle \ln R(t) \rangle, \\ \mu &= \lim_{t \rightarrow \infty} \frac{1}{t} (\langle \ln R(t)^2 \rangle - \langle \ln R(t) \rangle^2). \end{aligned} \tag{3.14}$$

The log-normal distribution for  $R$  corresponds to a Gaussian distribution for  $\gamma$

$$S(\gamma) = \frac{(\gamma - \lambda_1)^2}{2\mu}, \tag{3.15}$$

and to a quadratic in  $q$  generalized Lyapunov exponent:

$$L(q) = \lambda_1 q + \frac{1}{2} \mu q^2. \tag{3.16}$$

Let us remark that, in general, the log-normal distribution (3.13) is a good approximation for non-extreme events, i.e. small fluctuation of  $\gamma$  around  $\lambda_1$ , so that the expression (3.16) is correct only for small  $q$  (see Fig. 1). This is because the moments of the log-normal distribution grow too fast with  $q$  [168]. Indeed from (3.12) we have that the selected  $\gamma^*(q)$  is given by  $\gamma^*(q) = \lambda_1 + \mu q$  and thus  $\gamma^*(q)$  is not finite for  $q \rightarrow \infty$ . This is unphysical because  $\gamma^*(\infty)$  is the fastest error growth rate in the system and, we may reasonably suppose that it is finite.

Let us consider again the map (3.5). In this case we have  $\lambda_1 = L'(0) = -a \ln(a) - (1-a) \ln(1-a)$  and  $\mu = L''(0) = a(1-a)(\ln(a) - \ln(1-a))^2$ , which are nothing but the coefficients of the Taylor expansion of (3.7) around  $q=0$ . For large  $q$  the log-normal approximation gives  $L(q)/q \simeq q\mu/2$  while the correct limit is the constant  $L(q)/q \simeq -\ln(a)$ .

Nevertheless,  $\lambda_1$  and  $\mu$  are the two basic parameters for characterizing the intermittency of a system. To be more specific, let us remark that, assuming (3.13),  $P_t(R)$  reaches its maximum for

$$R_{\max}(t) \cong e^{(\lambda_1 - \mu)t}, \tag{3.17}$$

so that for  $t \rightarrow \infty$ :

$$\begin{aligned} R_{\max} &\rightarrow \infty && \text{if } \mu/\lambda_1 < 1, \\ R_{\max} &\rightarrow 0 && \text{if } \mu/\lambda_1 > 1. \end{aligned} \tag{3.18}$$

Thus in the weak intermittency limit,  $\mu/\lambda_1 < 1$ , the most probable response function  $R_{\max}(t)$  follows the correct behavior (with the corrected exponent  $\lambda_1 - \mu$ ). In the strong intermittent limit,  $\mu/\lambda_1 > 1$ , the most probable estimation breaks down because it predicts an asymptotic stable phase  $R_{\max}(t) \rightarrow 0$  instead of the chaotic exponential growth.

As in the case of the first LE, it is possible to introduce higher order generalized Lyapunov exponents. Starting from the  $n$ th-order response function  $R^{(n)}(t)$  (2.10), we define

$$L^{(n)}(q) = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \langle R^{(n)}(t)^q \rangle, \tag{3.19}$$

where  $L^{(1)}(q) = L(q)$ . From (2.11) we have

$$\sum_{i=1}^n \lambda_i = \left. \frac{dL^{(n)}(q)}{dq} \right|_{q=0}. \tag{3.20}$$

The generalized  $L^{(n)}(q)$  represents the fluctuations of the exponential divergence of a  $n$ -dimensional volume in phase space [172]. The properties of  $L^{(n)}(q)$  are analogous to the properties of  $L(q)$ , i.e.  $L^{(n)}(q)$  is a concave function of  $q$  for any  $n$  and for a non-intermittent behavior they are linear in  $q$ .

### 3.2. Renyi entropies

In Section 2.1.2 we defined the Kolmogorov–Sinai entropy (2.13) and discussed its relation with the Lyapunov exponents by means of the Pesin relation (2.16). Analogously to the generalized LE, it is possible to introduce a generalization of the Kolmogorov–Sinai entropy in order to take into account the intermittency.

Let us recall the definition of Kolmogorov–Sinai entropy

$$h_{\text{KS}} = -\lim_{\tau \rightarrow 0} \lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \frac{1}{N\tau} \sum_{\{W^N(\mathcal{A}_\epsilon)\}} P(W^N(\mathcal{A}_\epsilon)) \ln P(W^N(\mathcal{A}_\epsilon)) \tag{3.21}$$

where  $\mathcal{A}_\epsilon$  is a partition of the phase space in cells of size  $\epsilon$  and  $W^N(\mathcal{A}_\epsilon)$  indicates a sequence of length  $N$  in this partition. The generalized Renyi entropies [171,172],  $K_q$ , can be introduced by observing that (3.21) is nothing but the average of  $-\ln P(W^N)$  with the probability  $P(W^N)$ :

$$K_q = -\lim_{\tau \rightarrow 0} \lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \frac{1}{N\tau(q-1)} \ln \left( \sum_{\{W^N(\mathcal{A}_\epsilon)\}} P(W^N(\mathcal{A}_\epsilon))^q \right). \tag{3.22}$$

As in (3.4) one has  $h_{\text{KS}} = \lim_{q \rightarrow 1} K_q = K_1$ ; in addition from general results of probability theory, one can show that  $K_q$  is monotonically decreasing with  $q$ .

It will not be surprising that the generalized *Renyi entropies* are related to the generalized Lyapunov exponents  $L(q)$ . Introducing the number of non-negative Lyapunov exponents  $n^*$  (i.e.  $\lambda_{n^*} \geq 0, \lambda_{n^*+1} < 0$ ), the Pesin relation (2.16) can be written as

$$h_{KS} = \sum_{i=1}^{n^*} \lambda_i = \left. \frac{dL^{(n^*)}(q)}{dq} \right|_{q=0} . \tag{3.23}$$

Moreover, one has [171]

$$K_{q+1} = \frac{L^{(n^*)}(-q)}{-q} . \tag{3.24}$$

### 3.3. The effects of intermittency on predictability

We have seen that intermittency can be described, at least at a qualitative level, in terms of  $\lambda_1$  and  $\mu$ , which are the two parameters characterizing the log-normal approximation. We discuss now the relevance of the log-normal approximation for the predictability time  $T_p$ .

The predictability time  $T_p$  is defined as the time it takes for the error of initial size  $\delta_0$  to grow up to the tolerance  $\Delta$

$$T_p = \min \left[ t: R(t) \geq \frac{\Delta}{\delta_0} \right] . \tag{3.25}$$

In the framework of the log-normal approximation, we can write

$$\ln R(t) = \lambda_1 t + \sqrt{\mu} w(t) , \tag{3.26}$$

where  $w(t)$  is a Wiener process with  $w(0) = 0, \langle w(t) \rangle = 0$  and  $\langle w(t)w(t') \rangle = \min(t, t')$ . In this case the computation of  $T_p$  is reduced to a first exit problem, which is well known in stochastic process theory [45,78]. The solution gives the pdf of  $T_p$  [66]:

$$p(T_p) = \frac{\ln(\Delta/\delta_0)}{\sqrt{2\pi\mu T_p^3}} \exp \left[ -\frac{(\lambda_1 T_p - \ln(\Delta/\delta_0))^2}{2\mu T_p} \right] . \tag{3.27}$$

Notice that the (3.27) is not normalized, since there is a finite probability of “no exit”. Of course, this is an artifact of the approximation in terms of the stochastic process (3.26). In non-pathological chaotic systems one expects that  $p(T_p)$  is normalized.

In the limit of weak intermittency  $\mu/\lambda_1 \ll 1$ ,  $p(T_p)$  is almost Gaussian and the mean value  $\langle T_p \rangle$  is close to the most probable value of (3.27) corresponding to the naive estimation (3.1). On the contrary, in the strong intermittent limit,  $\mu/\lambda_1 \gg 1$ , the pdf of  $T_p$  shows an asymmetric “triangular shape” and the most probable value is

$$T_p = \frac{1}{3\mu} \ln(\Delta/\delta_0)^2 . \tag{3.28}$$

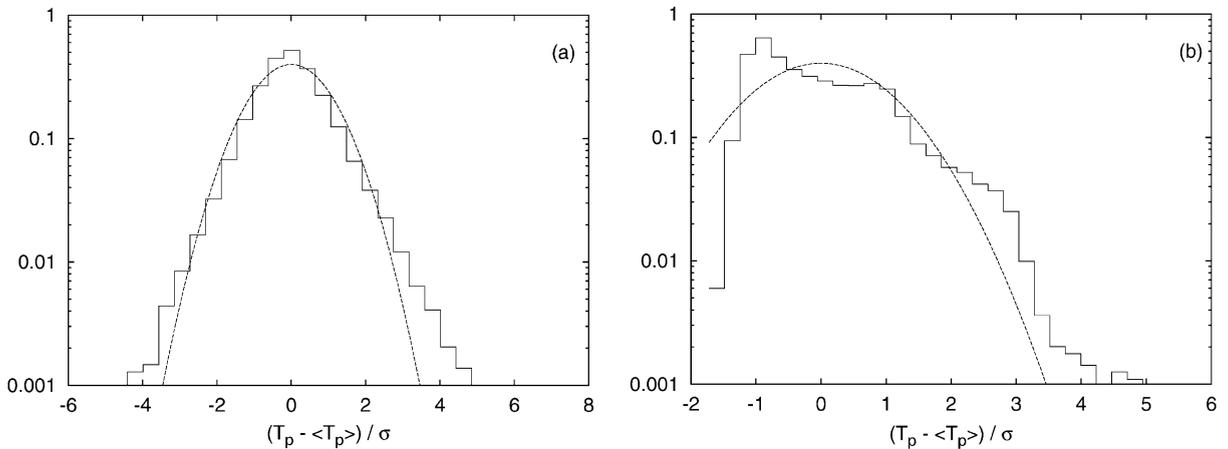


Fig. 2. Rescaled pdf,  $p(T_p)\sigma$ , of the predictability time  $T_p$  for the Lorenz model (3.29): (a) with  $r = 28$  (weak intermittency) the average predictability time is  $\langle T_p \rangle = 10.84$  and its variance is  $\sigma^2 = 3.12$  while  $\lambda = 0.90$ ,  $\mu = 0.06 \pm 0.01$ ; (b) with  $r = 166.3$  (strong intermittency) and  $\langle T_p \rangle = 8.2385$  and  $\sigma^2 = 19.75$ , while  $\lambda = 1.18$  and  $\mu = 3.9 \pm 1$ . The dashed line is the Gaussian distribution.

In order to see the effects of intermittency on the predictability time, let us consider as an example the Lorenz system [145]:

$$\begin{aligned} \frac{dx}{dt} &= \sigma(y - x) , \\ \frac{dy}{dt} &= x(r - z) - y , \\ \frac{dz}{dt} &= xy - bz , \end{aligned} \tag{3.29}$$

with the standard values  $\sigma = 10$  and  $b = \frac{8}{3}$ . For  $r = 28$ , the Lorenz model is very weakly intermittent,  $\mu/\lambda \simeq 7 \times 10^{-2}$ , and the pdf of the predictability time is very close to a Gaussian (Fig. 2). On the contrary, for  $r = 166.3$  the Lorenz model becomes strongly intermittent [191],  $\mu/\lambda \simeq 3.3$  and the pdf of the predictability time displays a long exponential tail responsible for the deviation from (3.1).

This qualitative behavior is typical of intermittent systems. In Section 5.3 we will see a more complex example in the context of turbulence.

### 3.4. Growth of non-infinitesimal perturbations

In realistic situations, the initial condition of a system is known with a limited accuracy. In this case the Lyapunov exponent is of little relevance for the characterization of predictability and new indicators are needed. To clarify the problem, let us consider the following coupled map model:

$$\begin{aligned} \mathbf{x}(t + 1) &= \mathbf{R} \mathbf{x}(t) + \varepsilon \mathbf{h}(y(t)) , \\ y(t + 1) &= G(y(t)) , \end{aligned} \tag{3.30}$$

where  $\mathbf{x} \in \mathbb{R}^2$ ,  $y \in \mathbb{R}^1$ ,  $\mathbf{R}$  is a rotation matrix of arbitrary angle  $\theta$ ,  $\mathbf{h}$  is a vector function and  $G$  is a chaotic map. For simplicity, we consider a linear coupling  $\mathbf{h}(y) = (y, y)$  and the logistic map  $G(y) = 4y(1 - y)$ .

For  $\varepsilon = 0$  we have two independent systems: a regular and a chaotic one. Thus the Lyapunov exponent of the  $\mathbf{x}$  subsystem is  $\lambda_x(\varepsilon = 0) = 0$ , i.e., it is completely predictable. On the contrary, the  $y$  subsystem is chaotic with  $\lambda_y = \lambda_1 = \ln 2$ .

If we now switch on the (small) coupling ( $\varepsilon > 0$ ) we are confronted with a single three-dimensional chaotic system with a positive global Lyapunov exponent

$$\lambda = \lambda_y + O(\varepsilon). \quad (3.31)$$

A direct application of (3.1) would give

$$T_p^{(x)} \sim T_p \sim \frac{1}{\lambda_y}, \quad (3.32)$$

but this result is clearly unacceptable: the predictability time for  $\mathbf{x}$  seems to be independent of the value of the coupling  $\varepsilon$ . Let us underline that this is not due to an artifact of the chosen example. Indeed, one can use the same argument in many physical situations [32]. A well-known example is the gravitational three body problem with one body (asteroid) much smaller than the other two (planets). If one neglects the gravitational feedback of the asteroid on the two planets (restricted problem) one has a chaotic asteroid in the regular field of the planets. As soon as the feedback is taken into account (i.e.  $\varepsilon > 0$  in the example) one has a non-separable three body system with a positive LE. Of course, intuition correctly suggests that it should be possible to forecast the motion of the planets for very long times if the asteroid has a very small mass ( $\varepsilon \rightarrow 0$ ).

The apparent paradox arises from the use of (3.1), which is valid only for the tangent vectors, also in the non-infinitesimal regime. As soon as the errors become large one has to take into account the full nonlinear evolution. The effect is shown for the model (3.30) in Fig. 3. The evolution of  $\delta\mathbf{x}$  is given by

$$\delta\mathbf{x}(t+1) = \mathbf{R}\delta\mathbf{x}(t) + \varepsilon\delta\mathbf{h}(y), \quad (3.33)$$

where, with our choice,  $\delta\mathbf{h} = (\delta y, \delta y)$ . At the beginning, both  $|\delta\mathbf{x}|$  and  $\delta y$  grow exponentially. However, the available phase space for  $y$  is finite and the uncertainty reaches the saturation value  $\delta y \sim O(1)$  in a time  $t \sim 1/\lambda_1$ . At larger times the two realizations of the  $y$  variable are completely uncorrelated and their difference  $\delta y$  in (3.33) acts as a noisy term. As a consequence, the growth of the uncertainty on  $\mathbf{x}$  becomes diffusive with a diffusion coefficient proportional to  $\varepsilon^2$  [32]

$$|\delta\mathbf{x}(t)| \sim \varepsilon t^{1/2} \quad (3.34)$$

so that

$$T_p^{(x)} \sim \varepsilon^{-2}. \quad (3.35)$$

This example shows that, even in simple systems, the Lyapunov exponent can be of little relevance for the characterization of the predictability.

In more complex systems, in which different scales are present, one is typically interested in forecasting the large scale motion, while the LE is related to the small scale dynamics.

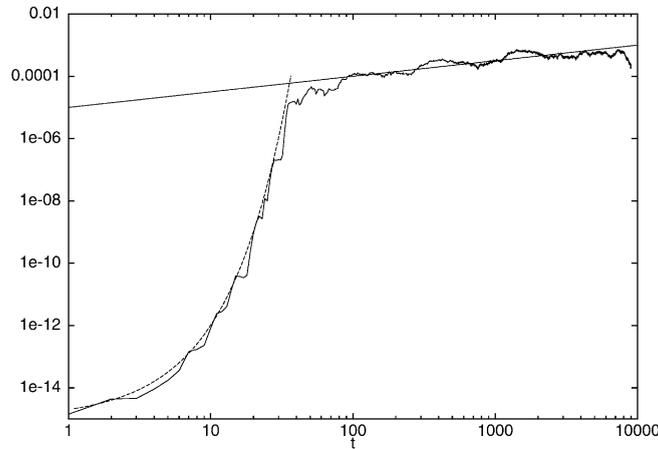


Fig. 3. Growth of error  $|\delta\mathbf{x}(t)|$  for the coupled map (3.30). The rotation angle is  $\theta = 0.82099$ , the coupling strength  $\varepsilon = 10^{-5}$  and the initial error only on the  $y$  variable is  $\delta y = \delta_0 = 10^{-10}$ . Dashed line  $|\delta\mathbf{x}(t)| \sim e^{\lambda_1 t}$  where  $\lambda_1 = \ln 2$ , solid line  $|\delta\mathbf{x}(t)| \sim t^{1/2}$ .

A familiar example is weather forecast: the LE of the atmosphere is indeed rather large due to the small scale convective motion, but (large scale) weather prediction is possible for about 10 days [146,160]. It is thus natural to seek for a generalization of the LE to finite perturbations from which one can obtain a more realistic estimation for the predictability time. It is worth underlining the important fact that finite errors are not confined in the tangent space but are governed by the complete non-linear dynamics. In this sense the extension of the LE to finite errors will give more information on the system.

Aiming to generalize the LE to non-infinitesimal perturbations let us now define the finite size Lyapunov exponent (FSLE) [11,12] (see Appendix A for the computational details). Consider a reference trajectory,  $\mathbf{x}(t)$ , and a perturbed one,  $\mathbf{x}'(t)$ , such that  $|\mathbf{x}'(0) - \mathbf{x}(0)| = \delta$  ( $|\dots|$  is the Euclidean norm but one can also consider other norms). One integrates the two trajectories and computes the time  $\tau_1(\delta, r)$  necessary for the separation  $|\mathbf{x}'(t) - \mathbf{x}(t)|$  to grow from  $\delta$  to  $r\delta$ . At time  $t = \tau_1(\delta, r)$  the distance between the trajectories is rescaled to  $\delta$  and the procedure is repeated in order to compute  $\tau_2(\delta, r), \tau_3(\delta, r) \dots$ .

The threshold ratio  $r$  must be  $r > 1$ , but not too large in order to avoid contributions from different scales in  $\tau(\delta, r)$ . A typical choice is  $r = 2$  (for which  $\tau(\delta, r)$  is properly a “doubling” time) or  $r = \sqrt{2}$ . In the same spirit of the discussion leading to Eqs. (2.44) and (2.45), we may introduce an effective finite size growth rate:

$$\gamma_i(\delta, r) = \frac{1}{\tau_i(\delta, r)} \ln r. \tag{3.36}$$

After having performed  $\mathcal{N}$  error-doubling experiments, we can define the FSLE as

$$\lambda(\delta) = \langle \gamma(\delta, r) \rangle_t = \left\langle \frac{1}{\tau(\delta, r)} \right\rangle_t \ln r = \frac{1}{\langle \tau(\delta, r) \rangle_e} \ln r, \tag{3.37}$$

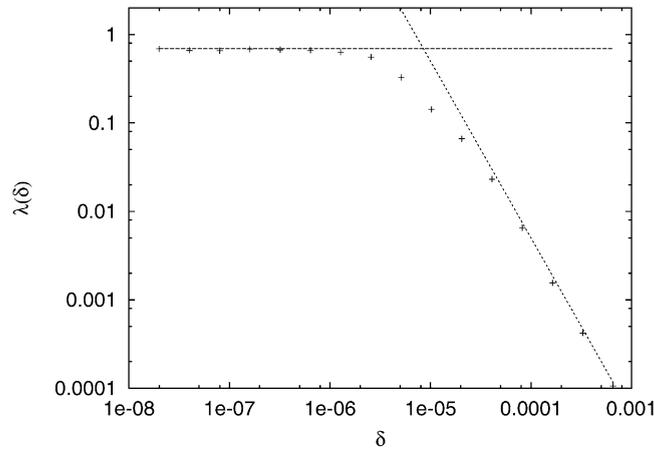


Fig. 4.  $\lambda(\delta)$  as a function of  $\delta$  for the coupled map (3.30) with  $\varepsilon = 10^{-5}$ . The perturbation has been initialized as in Fig. 3. For  $\delta \rightarrow 0, \lambda(\delta) \simeq \lambda_1$  (solid line). The dashed line shows the behavior  $\lambda(\delta) \sim \delta^{-2}$ .

where  $\langle \tau(\delta, r) \rangle_e$  is

$$\langle \tau(\delta, r) \rangle_e = \frac{1}{\mathcal{N}} \sum_{n=1}^{\mathcal{N}} \tau_n(\delta, r). \tag{3.38}$$

(see Appendix A and [12] for details). In the infinitesimal limit, the FSLE reduces to the Lyapunov exponent

$$\lim_{\delta \rightarrow 0} \lambda(\delta) = \lambda_1. \tag{3.39}$$

In practice, this limit means that  $\lambda(\delta)$  displays a constant plateau at  $\lambda_1$  for sufficiently small  $\delta$  (Fig. 3). For finite value of  $\delta$  the behavior of  $\lambda(\delta)$  depends on the details of the non-linear dynamics. For example, in the model (3.30) the diffusive behavior (3.34), by simple dimensional arguments, corresponds to  $\lambda(\delta) \sim \delta^{-2}$ . Since the FSLE measures the rate of divergence of trajectories at finite errors, one might wonder whether it is just another way to look at the average response  $\langle \ln(R(t)) \rangle$  (3.2) as a function of time. A moment of reflection shows that this is not the case. Indeed taking the average at fixed time is not the same as computing the average doubling time at *fixed scale*, as in (3.37). This is particularly clear in the case of strongly intermittent system, in which  $R(t)$  can be very different in each realization. In presence of intermittency averaging over different realizations at fixed times can produce a spurious regime due to the superposition of exponential and diffusive contributions by different samples at the same time [10] (see Fig. 4).

The FSLE method can be easily applied for data analysis [35].

For other approaches to address the problem of non-infinitesimal perturbations see [73,212,115].

### 3.5. The $\epsilon$ -entropy

The Kolmogorov–Sinai entropy,  $h_{KS}$  (2.13), of a system measures the amount of information per unit time necessary to record without ambiguity a generic trajectory of a system. Since

the computation of  $h_{KS}$  involves the limit of arbitrary fine resolution and infinite times (see Section 2.1.2), it turns out that, practically, for most systems it is not possible to compute  $h_{KS}$ . Nevertheless, in the same philosophy of the FSLE, by relaxing the strict requirement of reproducing a trajectory with arbitrary accuracy, one can introduce the  $\epsilon$ -entropy which measures the amount of information for reproducing a trajectory with accuracy  $\epsilon$  in phase-space. Roughly speaking the  $\epsilon$ -entropy can be considered the counterpart, in information theory, of the FSLE (as the KS-entropy is for the Lyapunov exponent). Such a quantity was originally introduced by Shannon [201], and by Kolmogorov [124]. Recently, Gaspard and Wang [89] made use of this concept to characterize a large variety of processes.

We start with a continuous (in time) variable  $\mathbf{x}(t) \in \mathbb{R}^d$ , which represents the state of a  $d$ -dimensional system, we discretized the time by introducing an interval  $\tau$  and we consider the new variable

$$\mathbf{X}^{(m)}(t) = (\mathbf{x}(t), \mathbf{x}(t + \tau), \dots, \mathbf{x}(t + (m - 1)\tau)). \tag{3.40}$$

Of course,  $\mathbf{X}^{(m)}(t) \in \mathbb{R}^{md}$  and it corresponds to the trajectory which lasts for a time  $T = m\tau$ .

In data analysis, the space where the state vectors of the system live is not known. Moreover, usually only a scalar variable  $u(t)$  can be measured. In such a case, one considers vectors  $(u(t), u(t + \tau), \dots, u(t + m\tau - \tau))$ , that live in  $\mathbb{R}^m$  and allow a reconstruction of the original phase space, known as delay embedding in the literature [209,199] (see also [1,2,114,170]), and it is a special case of (3.40).

Introduce now a partition of the phase space  $\mathbb{R}^d$ , using cells of edge  $\epsilon$  in each of the  $d$  directions. Since the region where a bounded motion evolves contains a finite number of cells, each  $\mathbf{X}^{(m)}(t)$  can be coded into a word of length  $m$ , out of a finite alphabet:

$$\mathbf{X}^{(m)}(t) \rightarrow W^m(\epsilon, t) = (i(\epsilon, t), i(\epsilon, t + \tau), \dots, i(\epsilon, t + m\tau - \tau)), \tag{3.41}$$

where  $i(\epsilon, t + j\tau)$  labels the cell in  $\mathbb{R}^d$  containing  $\mathbf{x}(t + j\tau)$ . From the time evolution of  $\mathbf{X}^{(m)}(t)$  one obtains, under the hypothesis of ergodicity, the probabilities  $P(W^m(\epsilon))$  of the admissible words  $\{W^m(\epsilon)\}$ . We can now introduce the  $(\epsilon, \tau)$ -entropy per unit time,  $h(\epsilon, \tau)$  [201]:

$$h_m(\epsilon, \tau) = \frac{1}{\tau} [H_{m+1}(\epsilon, \tau) - H_m(\epsilon, \tau)], \tag{3.42}$$

$$h(\epsilon, \tau) = \lim_{m \rightarrow \infty} h_m(\epsilon, \tau) = \frac{1}{\tau} \lim_{m \rightarrow \infty} \frac{1}{m} H_m(\epsilon, \tau), \tag{3.43}$$

where  $H_m$  is the block entropy of block length  $m$ :

$$H_m(\epsilon, \tau) = - \sum_{\{W^m(\epsilon)\}} P(W^m(\epsilon)) \ln P(W^m(\epsilon)). \tag{3.44}$$

For the sake of simplicity, we ignored the dependence on details of the partition. To make  $h(\epsilon, \tau)$  partition-independent one has to consider a generic partition of the phase space  $\{\mathcal{A}\}$  and to evaluate the Shannon entropy on this partition:  $h_{Sh}(\mathcal{A}, \tau)$ . The  $\epsilon$ -entropy is thus defined as the infimum over all partitions for which the diameter of each cell is less than  $\epsilon$  [89]:

$$h(\epsilon, \tau) = \inf_{\mathcal{A}: \text{diam}(\mathcal{A}) \leq \epsilon} h_{Sh}(\mathcal{A}, \tau). \tag{3.45}$$

Note that the time dependence in (3.45) is trivial for deterministic systems, and that in the limit  $\epsilon \rightarrow 0$  one recovers the Kolmogorov–Sinai entropy

$$h_{\text{KS}} = \lim_{\epsilon \rightarrow 0} h(\epsilon, \tau).$$

The above entropies  $H_m(\epsilon)$  have been introduced by using a partition and the usual Shannon entropy; however it is possible to arrive at the same notion, starting from other entropy-like quantities, that are numerically more convenient. For example, Cohen and Procaccia [61] proposed to estimate  $H_m(\epsilon)$  as follows. Given a signal composed of  $N$  successive records and the embedding variable  $\mathbf{X}^{(m)}$ , let us introduce the quantities:

$$n_j^{(m)} = \frac{1}{N-m} \sum_{i \neq j} \Theta(\epsilon - |\mathbf{X}^{(m)}(i\tau) - \mathbf{X}^{(m)}(j\tau)|), \quad (3.46)$$

then the block entropy  $H_m(\epsilon)$  is given by

$$H_m^{(1)}(\epsilon) = -\frac{1}{(N-m+1)} \sum_j \ln n_j^{(m)}(\epsilon). \quad (3.47)$$

In practice,  $n_j^{(m)}(\epsilon)$  is an approximation of  $P(W^m(\epsilon))$ . From a numerical point of view, correlation entropies [95,210] are sometimes more convenient, so that one studies

$$H_m^{(2)}(\epsilon) = -\ln \left( \frac{1}{N-m+1} \sum_j n_j^{(m)}(\epsilon) \right) \leq H_m^{(1)}(\epsilon). \quad (3.48)$$

This corresponds to approximate the Shannon by the Renyi entropy of order  $q=2$  [114].

The  $(\epsilon, \tau)$ -entropy  $h(\epsilon, \tau)$  is well defined also for stochastic processes. Actually the dependence of  $h(\epsilon, \tau)$  on  $\epsilon$  can give some insight into the underlying stochastic process [89], for instance, in the case of a stationary Gaussian process with spectrum  $S(\omega) \propto \omega^{-2}$  one has [124]

$$\lim_{\tau \rightarrow 0} h(\epsilon, \tau) \sim \frac{1}{\epsilon^2}. \quad (3.49)$$

However, we have to stress that the behavior predicted by Eq. (3.49) may be difficult to be experimentally observed mainly due to problems related to the choice of  $\tau$  [51,3] (see also Appendix C).

#### 4. Predictability in extended systems

Here we consider *extended* dynamical systems, whose degrees of freedom depend on space and time, and which can display unpredictable behaviors both in the time and space evolution, i.e. *spatio-temporal chaos*. The inadequacy of the Lyapunov exponents in characterizing predictability becomes now well evident.

Following Hohenberg and Shraiman [104] (see also [68]) we give a more precise meaning to the terms *spatio-temporal chaos* and *extended* systems. For a generic system of size  $L$ , we can define three characteristic lengths:  $\ell_D$ ,  $\ell_E$ ,  $\xi$  respectively associated to the scales of *dissipation* (i.e. the scale at which dissipation becomes effective, smaller scales can be considered

as inactive), *excitation* (i.e. the scale at which energy is injected in the system) and *correlation* (that we assume can be suitably defined). Now one has two limiting situations.

When all the characteristic lengths are of the same order ( $\ell_D, \ell_E, \xi \sim O(L)$ ) distant regions of the system are strongly correlated. Because of the coherence, the spatial nature is not very important and one speaks of *temporal* chaos, i.e. the system is basically low dimensional [195].

When  $L \gg \xi \gg \ell_D$  distant parts of the system are weakly correlated so that the number of (active) degrees of freedom and consequently the number of positive Lyapunov exponents, the Kolmogorov–Sinai entropy and the attractor dimension,  $D_F$ , are extensive quantities, i.e. they increase with the system size,  $L$ . Here, spatial aspects are crucial and one speaks of *spatio-temporal chaos*, e.g., Rayleigh–Bénard convection for large aspect ratio [153].

The above picture is just an approximative scenario (see [104] for further details) but sufficiently broad to include systems ranging from fluid dynamics to biological and chemical reactions [68,153]. In spite of the fact that turbulent flows fit in this broad definition we shall discuss the predictability problem in turbulence in the next section.

For detailed discussions on different physical and chemical systems which can be included in the above definition see [68,38]. Here we discuss the predictability problem in a restricted class of models, which are relatively simple from a computational and theoretical point of view but, nevertheless, possess the essential phenomenology of spatio-temporal chaos.

#### 4.1. Simplified models for extended systems and the thermodynamic limit

A great simplification in the study of extended systems, usually described by partial differential equations (PDE), can be achieved by considering discrete time and space models, and introducing the *coupled map lattices* (CML) [108], i.e. maps defined on a discrete lattice. A typical 1-dimensional CML (the extension to  $d$ -dimensions is straightforward) can be written in the following way:

$$\mathbf{x}_i(t+1) = (1 - \varepsilon_0)\mathbf{f}_a[\mathbf{x}_i(t)] + \frac{1}{2} \sum_{j=-L/2}^{L/2} \varepsilon_j(\mathbf{f}_a[\mathbf{x}_{i+j}(t)] + \mathbf{f}_a[\mathbf{x}_{i-j}(t)]), \quad (4.1)$$

with  $\varepsilon_0 = \sum_{j=1}^{L/2} \varepsilon_j$ .  $L$  is the lattice size,  $i = -L/2, \dots, L/2$ ,  $\mathbf{x} \in \mathbb{R}^n$  is the state variable which depends on the site and time, and  $\mathbf{f}_a \in \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a non-linear map, which drives the local dynamics and depends on a control parameter  $a$ . Usually, periodic boundary conditions  $\mathbf{x}_{i+L} = \mathbf{x}_i$  are assumed and, for scalar variables ( $n=1$ ), one studies coupled logistic maps,  $f_a(x) = ax(1-x)$  or tent maps,  $f_a(x) = a|\frac{1}{2} - |x - \frac{1}{2}||$ .

The parameters  $\{\varepsilon_i\}$  rule the strength and the form of the coupling and they are chosen according to the physical system under investigation. For example, with  $\varepsilon_j = 0$  for  $j \geq 2$ , i.e. nearest neighbor coupling, one can mimic PDEs describing reaction diffusion processes (indeed formally the equation assumes the structure of a discrete Laplacian). However, it could be misleading to consider CMLs simply as discrete approximation of PDEs. Indeed, since the local map  $\mathbf{f}_a$  is usually chaotic, chaos in CML, differently from PDE, is the result of many interacting chaotic sub-systems. Hence, the correspondence between the instability mechanisms in the two type of models is not straightforward [68].

Other kinds of coupling can be considered to mimic different physical situations, e.g., asymmetric CML (see Section 4.5) for studying convective instabilities [106,39,217], or mean field (globally coupled maps) version of (4.1) (Section 4.3) for studying neural network or population dynamics [118]. Further generalizations are quoted in Ref. [112].

Lyapunov exponents, attractor dimensions and entropies can be defined (and, at least the Lyapunov exponents, numerically computed) also for extended systems. In particular, for  $L < \infty$  the CMLs have finite-dimensional phase space and the above quantities are well defined. In PDEs some difficulties can rise due to the problem of the non-equivalence of the norms [127]: Lyapunov exponents and consequently the characterization of the predictability may depend on the chosen norm. We shall see in Section 4.8 that this is not just a subtle mathematical problem.

In order to build a statistical description of spatio-temporal chaos, as Ruelle pointed out [197], one has to require the existence of a good thermodynamic limit for the Lyapunov spectrum  $\{\lambda_i(L)\}_{i=1,L}$ . This means the existence of the limit

$$\lim_{L \rightarrow \infty} \lambda_i(L) = A(x), \quad (4.2)$$

where  $x = i/L$  is a continuous index in  $[0,1]$ , and  $A(x)$  is a non-increasing function. The function  $A(x)$  can be viewed as a *density* of Lyapunov exponents. If such limit does not exist, the possibility to build a statistical description of spatio-temporal chaos would be hopeless, i.e., the phenomenology of these systems would depend on  $L$ .

Once the existence of a Lyapunov density is proved, one can generalize some results of low-dimensional systems [97,44], namely the Kaplan–Yorke conjecture [117] and the Pesin relation (2.16). For instance, one can generalize (2.16) to

$$\mathcal{H}_{\text{KS}} = \lim_{L \rightarrow \infty} \frac{h_{\text{KS}}}{L} = \int_0^1 dx A(x) \theta(A(x)) \quad (4.3)$$

$\theta(x)$  being the step function. In the same way one can suppose the existence of a *dimension density*  $\mathcal{D}_F$ , that is to say a density of active degrees of freedom, i.e.  $\mathcal{D}_F = \lim_{L \rightarrow \infty} D_F/L$  which by the Kaplan–Yorke [117] conjecture is given by [97]

$$\int_0^{\mathcal{D}_F} dx A(x) = 0. \quad (4.4)$$

The existence of a good thermodynamic limit is supported by numerical simulations [109,144] and some exact results [205]. Recently, Eckmann and Collet [62] have proved the existence of a density of degrees of freedom in the complex Ginzburg–Landau equation. See also Refs. [97,44] and references therein for a discussion on such a problem.

#### 4.2. Overview on the predictability problem in extended systems

In low-dimensional systems, no matter how the initial disturbance is chosen, after a—usually short—relaxation time,  $T_R$ , the eigendirection with the largest growth rate dominates for almost all the initial conditions (this, e.g., helps in the numerical estimates of the Lyapunov exponents [20]). On the contrary, in high-dimensional systems this may not be true [92,183,173,186]. Indeed, in systems with many degrees of freedom there is room for several choices of the initial perturbation according to the specific problem under investigation (e.g., localized on

certain degrees of freedom or homogeneous in all the degrees of freedom), and it is not obvious that for all of them the time  $T_R$  needed to align along the maximally expanding direction is the same.

In general, the situation can be very complicated. For instance, it is known that, also considering initially homogeneous disturbances, the Lyapunov vectors can localize (permanently or slowly wandering) on certain degrees of freedom [109,75,186]. Of course, this will severely affect the prediction of the future evolution of the system. Indeed, regions of large predictability time could coexist with regions of relatively short predictability time. In Ref. [109,112,186] one finds an abundance of examples displaying this phenomenology. A detailed analysis of this problem is far from the aims of this review; we just mention that the behavior of the Lyapunov vectors can range from a strongly localized regime (the origin of which can be understood by the analogy with Anderson localization of the wave function in disordered potential [91]) to localized wandering structures. In particular, in the latter case there is strong numerical evidence [185,186] that for a large class of (1-dimensional) systems the dynamics of the Lyapunov vectors (actually the logarithm of them) falls into the universality class of the 1-dimensional KPZ equation [120].

In these situations the main contribution to the predictability time comes from the time needed for the perturbation to propagate through the system or to align along the maximally expanding direction, which can be of the order of the system size [173,183,186]. As a consequence the predictability time can be much longer than the rough estimation  $T_p \sim 1/\lambda$ .

Moreover, the LE can also be unsatisfactory if one is interested in perturbations with particular space–time shapes. Indeed, these features have led to the introduction of a number of new indicators; for instance, the *temporal* (or *specific*) Lyapunov exponents [189,139], the *spatial* Lyapunov exponents [91,139] (which characterize respectively perturbations exponentially shaped in time and space) or the *comoving* Lyapunov exponents [71] for the characterization of the spatio-temporal evolution of localized perturbation [109] and of the convective instabilities [8,39].

Convectively unstable systems are rather interesting because, even if the LE (computed in the stationary frame) is negative, some features of the motion can be highly unpredictable (see Section 4.6). It is also worth mentioning the existence of systems with exponentially long (in the system size) transients during which the dynamics is essentially unpredictable despite the fact that the LE is negative [69]. This phenomenon, known under the name of *stable chaos* [188], will be briefly discussed in Section 7.3.

In high-dimensional systems one is also interested in predicting the behavior of some average observable to achieve a macroscopic description of the system. The coarse grained (hydrodynamic like) dynamics may be non-trivial, and the largest LE, which is related to the fine grained dynamics, is not relevant for characterizing the predictability at a coarse grained level (see Section 4.7).

We conclude this brief overview by mentioning another interesting feature: in spatially distributed systems coherent structures may appear. They move maintaining for rather long times their shape. In different contexts one can be interested in predicting the evolution of such structures; e.g., cyclonic/anti-cyclonic structures in the atmosphere. A reasonable question could be the prediction of the center and orientation of such structures: limiting to these aspects one can indeed hope to have a rather long predictability time compared to the rough estimate

$T_p \sim O(1/\lambda)$ . However, since usually such phenomena arise in fields, whose evolution is ruled by PDE, the non-equivalence of the norms makes a general approach to the problem unfeasible. Therefore, one has to resort to ad hoc treatments, based on physical intuition to identify the most suitable norm to be used for the particular needs (see Section 4.8).

### 4.3. Butterfly effect in coupled map lattices

In spatially extended systems it is important to understand the way an uncertainty initially localized in some region will spread. Here we study in particular the time needed for a perturbation, initially seeded in the central site of a lattice of coupled maps, to reach a preassigned value at the border of the lattice [173] (see also [109,212,213] and Section 4.6). In other terms we wonder about the “butterfly effect” starting from the center of the lattice and arriving up to the boundary.

We shall discuss the properties of such time by varying the coupling range from local to non local in the 1-dimensional CML (4.1) with periodic boundary conditions. We consider two cases: local coupling, i.e.  $\varepsilon_j = 0$  if  $j \geq 2$ , and non-local coupling, e.g.

$$\varepsilon_1 = C_1 \quad \text{and} \quad \varepsilon_j = \frac{C_2}{j^\alpha} \quad \text{for } j \geq 2, \quad (4.5)$$

where  $\alpha$  measures the strength of non-locality. The initial perturbation is on the central site, i.e.

$$|\delta x_i(0)| = \delta_0 \delta_{i,0}. \quad (4.6)$$

We look at the predictability time  $T_p$  needed for the perturbation to reach a certain threshold  $\delta_{\max}$  on the boundary of the lattice, i.e. the maximum time,  $t$ , such that  $|\delta x_{L/2}(t)| \leq \delta_{\max}$ .

For nearest neighbor coupling, one has obviously that  $\delta x_{L/2}(t) = 0$  for  $t < L/2$ . Indeed, by a numerical integration of (4.1) for the short range coupling one observes that  $\delta x_{L/2}(t) = 0$  for times  $t < t^* \propto L$ ; while for  $t > t^*$  the perturbation, due to the (local) chaotic dynamics, grows as  $\delta x_{L/2}(t) \sim \delta_0 \exp[\lambda(t - t^*)]$ . Thus for local interactions, the predictability is mainly determined by the waiting time  $t^*$ , necessary to have  $|\delta x_{L/2}| > \delta_0$ , which is roughly proportional to the system size  $L$ . This is confirmed by Fig. 5, where it is shown that the average predictability time  $\langle T_p \rangle$  as a function of  $L$  goes as

$$\langle T_p \rangle = t_1 + GL, \quad (4.7)$$

where the time  $t_1 \sim \lambda^{-1}$  is due to the exponential error growth after the waiting time and can be neglected in large enough lattices. This agrees with the existence of a finite speed for the perturbation spreading [212]; indeed  $G$  is related to the propagation velocity (see Section 4.6).

The scenario changes for non-local interactions (4.5). Now, due to the long range coupling, the perturbation (4.6) may propagate without any delay. The numerical results show that even for weak non-locality (e.g.  $C_2 \ll C_1$  and rather large  $\alpha$ -values), the waiting time  $t^*$  does not increase (or increases very slowly) with  $L$ , so that

$$\langle T_p \rangle \sim t_1 \sim \lambda^{-1}. \quad (4.8)$$

As shown in Fig. 5, weakly non-local couplings, and mean field interactions ( $\varepsilon_j = C_2/N$ ) have the same qualitative behavior. Very accurate numerical computations have confirmed that the dependence on  $L$  is indeed very weak (only logarithmic), i.e.  $\langle T_p \rangle \sim t_1 + \alpha \ln L/\lambda_1$  [213].

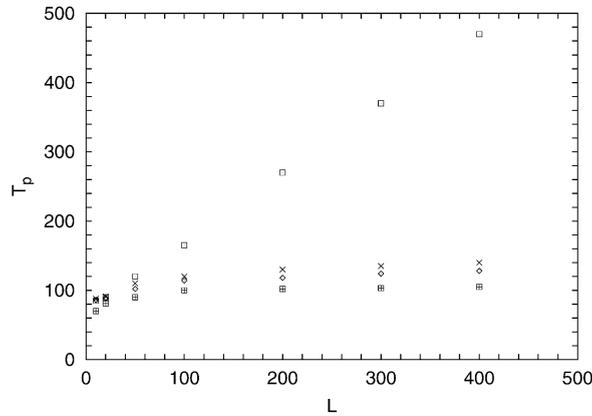


Fig. 5. Average predictability time  $\langle T_p \rangle$  versus  $L$  for a CML of logistic maps  $f_a(x) = ax(1 - x)$  with  $a = 4$ : local coupling  $\epsilon_0 = 0.3$  (squares); non-local coupling (4.5) with  $C_1 = 0.3$ ,  $C_2 = 0.01$   $\alpha = 2$  (crosses) or  $\alpha = 3$  (diamonds); mean field coupling  $\epsilon_i = C_2/L$  with  $C_2 = 0.3$  (crosses squares). The initial perturbation is applied at the center of the lattice (site  $i = 0$ ) and has an amplitude  $10^{-14}$ ; the maximum admitted error is  $\delta_{\max} = 0.1$ .

This example demonstrates that the predictability time is given by two contributions: the waiting time  $t^*$  and the characteristic time  $t_1 \sim \lambda^{-1}$  associated with chaos. For non-local interactions, the waiting time practically does not depend on the system size  $L$ , while for local interactions it is proportional to  $L$ . Let us underline that in these results the non-linear terms in the evolution of  $\delta \mathbf{x}(t)$  are rather important. One numerically observes that the waiting time  $t^*$  is not just the relaxation time  $T_R$  of  $\delta \mathbf{x}$  on the tangent eigenvector. Actually,  $T_R$  is much larger than  $t^*$ .

#### 4.4. Comoving and specific Lyapunov exponents

A general feature of systems evolving in space and time is that a generic perturbation not only grows in time but also propagates in space. Aiming at a quantitative description of such phenomena, Deissler and Kaneko [71] introduced a generalization of the LE to a non-stationary frame of reference: the *comoving* Lyapunov exponent. For the sake of simplicity, we consider again the case of a 1-dimensional CML.

Let us consider an infinitesimally small perturbation initially different from zero only in one site of the lattice (4.6). By looking at the perturbation evolution along the line, defined by  $j(t) = 0 + [vt]$  (where  $[\dots]$  denotes the integer part), one expects

$$|\delta x_j(t)| \approx |\delta x_0(0)| e^{\lambda(v)t}, \tag{4.9}$$

where  $\lambda(v)$  is the largest comoving Lyapunov exponent, i.e.

$$\lambda(v) = \lim_{t \rightarrow \infty} \lim_{L \rightarrow \infty} \lim_{|\delta x_0(0)| \rightarrow 0} \frac{1}{t} \ln \left( \frac{|\delta x_{[vt]}(t)|}{|\delta x_0(0)|} \right). \tag{4.10}$$

In Eq. (4.10) the order of the limits is important to avoid boundary effects. For  $v=0$  one recovers the usual LE. Moreover, one has that  $\lambda(v) = \lambda(-v)$  (and the maximum value is obtained at  $v=0$  [189]) when a privileged direction does not exist, otherwise  $\lambda(v)$  can be asymmetric and the

maximum can be attained at value  $v \neq 0$  (see Section 4.6). By writing the response function (2.8) in the moving frame one can also introduce the generalized comoving Lyapunov exponents  $L_q(v)$  for studying finite time effects [76].

There are other two indicators related to the comoving LE: the local Lyapunov exponent [183] and the specific (or temporal) Lyapunov exponents. Here we only briefly discuss the latter which is indeed nothing but the Legendre transform of  $\lambda(v)$ .

The specific Lyapunov exponent,  $\Lambda(\mu)$ , has been introduced by Politi and Torcini [189] to characterize the growth of exponentially shaped perturbations, i.e.

$$\delta x_i(t) = \Phi_i(t) e^{-\mu|i|}, \quad i = -\frac{L}{2}, \dots, \frac{L}{2}, \quad (4.11)$$

where  $\Phi_i(t)$  gives the fluctuation with respect to the pure exponential shape. One can see that  $\Lambda(\mu)$  is connected through a Legendre transform to  $\lambda(v)$  [189,212]. Indeed, Eq. (4.9) defines a local exponential profile with  $\mu = d\lambda(v)/dv$ , which means that in terms of the specific Lyapunov exponents one expects the perturbation to grow according to

$$\delta x_i(t) \sim \exp\{\Lambda(\mu)t - \mu i\}, \quad i = [vt]. \quad (4.12)$$

Note that for  $\mu = 0$ ,  $\Lambda(\mu)$  reduces to the standard LE. Therefore, the comoving Lyapunov exponent is given by

$$\lambda(v) = \Lambda(\mu) - \mu \frac{d\Lambda(\mu)}{d\mu}. \quad (4.13)$$

The last equation defines a Legendre transform from  $(\lambda(v), v)$  to  $(\Lambda(\mu), \mu)$  [189]. By inverting the transformation (4.13) one obtains  $v = -d\Lambda(\mu)/d\mu$ .

Working in tangent space by using standard algorithms [20], one computes the specific Lyapunov spectrum  $\Lambda_i(\mu)$  with  $i = 1, \dots, L$  for each  $\mu$ . In the limit  $L \rightarrow \infty$  a density of such exponents exists and an interpretation of it is discussed in [139,140].

#### 4.5. Convective chaos and spatial complexity

So far we have considered CMLs with symmetric spatial coupling; however there are many physical systems in which a privileged direction exists, e.g., boundary layers, thermal convection and wind-induced water waves. The term usually employed for denoting such a class of systems is flow systems. See [8,9,39,106] for a discussion of flow systems in different physical contexts.

In recent years it has received much attention a simplified model for flow systems which is able to capture the basic phenomenology [112,198,226]. A minimal model is a chain of maps with unidirectional coupling [9,180,112,217,76]:

$$x_n(t+1) = (1-c)f_a(x_n(t)) + cf_a(x_{n-1}(t)), \quad (4.14)$$

where  $t$  and  $n$  ( $=1, \dots, L$ ) are the discrete time and space respectively; the map  $f_a(x)$  is usually chosen to be the logistic map. One can consider different boundary conditions,  $x_0(t)$ . For instance,  $x_0(t) = x^*$  with  $x^*$  being an unstable fixed point of the map  $f_a(x)$ , or more generic time-dependent boundary conditions where  $x_0(t)$  is equal to a known function of time  $y(t)$ , which can be periodic, quasi-periodic or chaotic. Here, following Pikovsky [180–182], we consider a quasi-periodic boundary condition  $x_0(t) = 0.5 + 0.4 \sin(\omega t)$ , with  $\omega = \pi(\sqrt{5} - 1)$ . However,

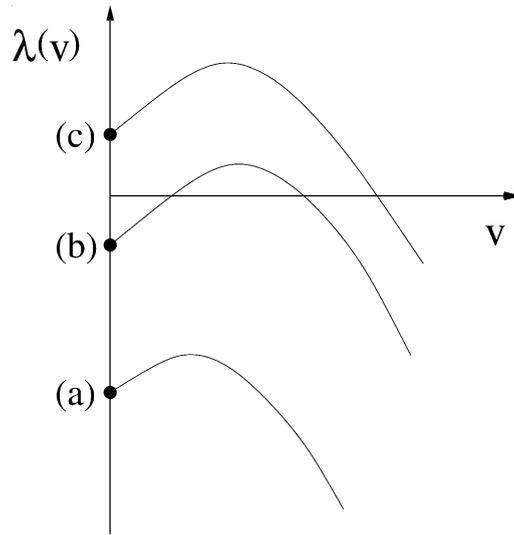


Fig. 6. Sketch of the behavior of  $\lambda(v)$  for (a) an absolutely and convectively stable flow, (b) absolutely stable but convectively unstable flow, and (c) absolutely unstable flow.

the results we are going to discuss do not depend too much on the details of the boundary conditions, i.e. on using  $x_0(t)$  quasi-periodic or chaotic.

A central concept in the study of flow systems is the one of *convective* instability, i.e. when a perturbation grows exponentially along the flow but vanishes locally.

We may give a description of the phenomenology of flow systems in terms of the largest LE and of the comoving LE. The absolute stability is identified by the condition  $\lambda(v) < 0$  for all  $v \geq 0$ ; the convective instability corresponds to  $\lambda_1 = \lambda(v=0) < 0$  and  $\lambda(v) > 0$  for some velocities  $v > 0$  and finally standard chaos (absolute instability) is present when  $\lambda_1 = \lambda(v=0) > 0$ . See Fig. 6 for a sketch of the possible behaviors.

The convective instability is conceptually very interesting, because even if the largest LE is negative the behavior of the system may be very hard to predict, as Fig. 7 suggests.

For this kind of spatial “complexity” there is not a simple and systematic characterization. A first explanation for these features may be found in the sensitivity of convective unstable systems on small perturbations at the beginning of the chain (always present in physical system), which grow exponentially while they are transmitted along the flow. This simple intuition can be made more quantitative defining an indicator which measures the degree of sensitivity on the boundary conditions [217,76]. We wonder how an uncertainty  $|\delta x_0(t)| = \delta_0$  in the knowledge of the boundary conditions will affect the system. We consider only the case of infinitesimal perturbations, i.e.  $\delta x_n(t)$  evolves according to the tangent space dynamics, and, for the moment we do not consider intermittency (i.e. time fluctuations of the comoving Lyapunov exponents).

The uncertainty  $\delta x_n(t)$ , on the determination of the variable at time  $t$  and site  $n$ , is given by the superposition of the evolved  $\delta x_0(t - \tau)$  with  $\tau = n/v$ :

$$\delta x_n(t) \sim \int \delta x_0(t - \tau) e^{\lambda(v)\tau} dv = \delta_0 \int e^{[\lambda(v)/v]n} dv. \quad (4.15)$$

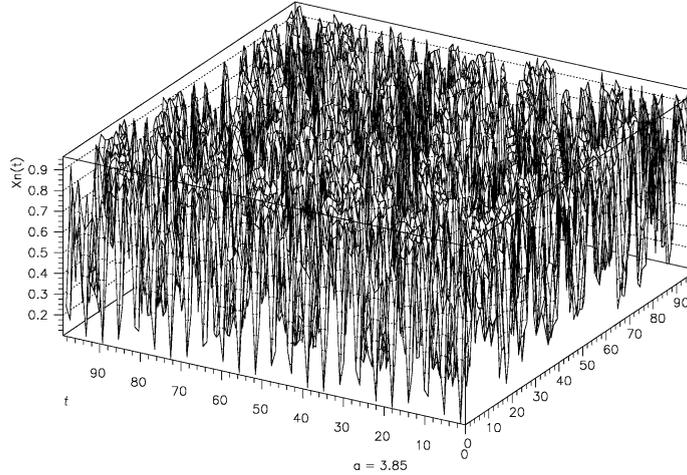


Fig. 7. Evolution of a state of the system (4.14) where  $f_a(x)$  is the logistic maps, the boundary condition is quasi-periodic,  $a = 3.85$  and  $c = 0.7$ : in this case  $\lambda_1 < 0$  but the system is convectively unstable.

Since we are interested in the asymptotic spatial behavior, i.e. large  $n$ , we can write

$$\delta x_n(t) \sim \delta_0 e^{\Gamma n}, \tag{4.16}$$

The quantity  $\Gamma$  can be considered as a sort of spatial-complexity index, an operative definition of which is the following:

$$\Gamma = \lim_{n \rightarrow \infty} \frac{1}{n} \left\langle \ln \frac{|\delta x_n|}{\delta_0} \right\rangle, \tag{4.17}$$

where the brackets mean a time average.

In the particular case of a non-intermittent system, a simple saddle-point estimate of Eq. (4.15) gives

$$\Gamma = \max_v \left[ \frac{\lambda(v)}{v} \right]. \tag{4.18}$$

Eq. (4.18) is a link between the comoving and the “spatial” Lyapunov exponent  $\Gamma$ , i.e. a relation between the convective instability of a system and its sensitivity to the boundary conditions.

Eq. (4.18) holds exactly only in absence of intermittency; in general the relation is more complicated. One can introduce the effective comoving Lyapunov exponent,  $\tilde{\gamma}_t(v)$ , that gives the exponential changing rate of a perturbation, in the frame of reference moving with velocity  $v$ , on a finite time interval  $t$ . According to general arguments (see Section 3.1 and [172]) one has  $\langle \tilde{\gamma}_t(v) \rangle = \lambda(v)$ . Then, instead of (4.15) one obtains

$$\delta x_n(t) \sim \delta_0 \int e^{[\tilde{\gamma}_t(v)/v]n} dv, \tag{4.19}$$

and therefore

$$\Gamma = \lim_{n \rightarrow \infty} \frac{1}{n} \left\langle \ln \frac{|\delta x_n|}{\delta_0} \right\rangle = \lim_{n \rightarrow \infty} \frac{1}{n} \ln \frac{|\delta x_n^{\text{typical}}|}{\delta_0} = \left\langle \max_v \left[ \frac{\tilde{\gamma}_t(v)}{v} \right] \right\rangle. \tag{4.20}$$

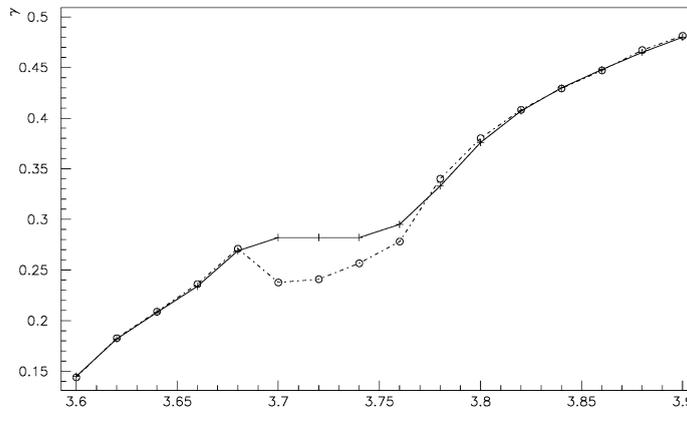


Fig. 8.  $\Gamma(+)$  and  $\Gamma^*$  (○) vs.  $a$  at a fixed value of  $c (=0.7)$ , for the system (4.14) of logistic maps with quasi-periodic boundary conditions (the system is convectively unstable for all the considered values of the parameters).

Therefore, because of the fluctuations, it is not possible to write  $\Gamma$  in terms of  $\lambda(v)$ , although one can obtain a lower bound [217]:

$$\Gamma \geq \max_v \left[ \frac{\langle \tilde{\gamma}_t(v) \rangle}{v} \right] = \max_v \left[ \frac{\lambda(v)}{v} \right] \equiv \Gamma^* . \tag{4.21}$$

In Fig. 8 we show  $\Gamma$  and  $\Gamma^*$  vs.  $a$  for a fixed value of  $c$ . There is a large range of values of the parameter  $a$  for which  $\Gamma$  is rather far from  $\Gamma^*$ . This difference is only due to intermittency, as investigations of the map  $f_a(x) = ax \bmod 1$  or the computation of the generalized spatial Lyapunov exponents  $L_s(q)$  [217] confirm.

Concluding, we underline that the spatial complexity displayed by these systems indicates that the unpredictability of a system cannot be completely reduced to the existence of at least one positive LE.

#### 4.6. Space–time evolution of localized perturbations

In Fig. 9 we show the evolution of a perturbation  $|\delta x_i(t)|$  initialized as (4.6) as a function of space and time  $(i, t)$ , for a 1-dimensional lattice of locally coupled tent maps. The perturbation grows in time and propagates linearly in space creating a sort of predictability “horizon”: this defines a propagation velocity  $V_F$  [109,189,212].

The velocity  $V_F$  is defined in terms of the left and right edges of the disturbance i.e. the first left (right) site for which at time  $t$  the perturbation reaches a preassigned arbitrary threshold. Numerically, it has been found that  $V_F$  is independent either of the amplitude of the initial perturbation,  $\delta_0$ , and of the threshold value, so that it is a well-defined quantity [109].

It is easy to realize that  $V_F$  is nothing but the velocity of a moving frame of reference in which the perturbation is seen neither to grow nor to decrease (i.e. the frame comoving with the edges of the perturbation). Therefore,  $V_F$  is given by [109]

$$\lambda(V_F) = 0 . \tag{4.22}$$

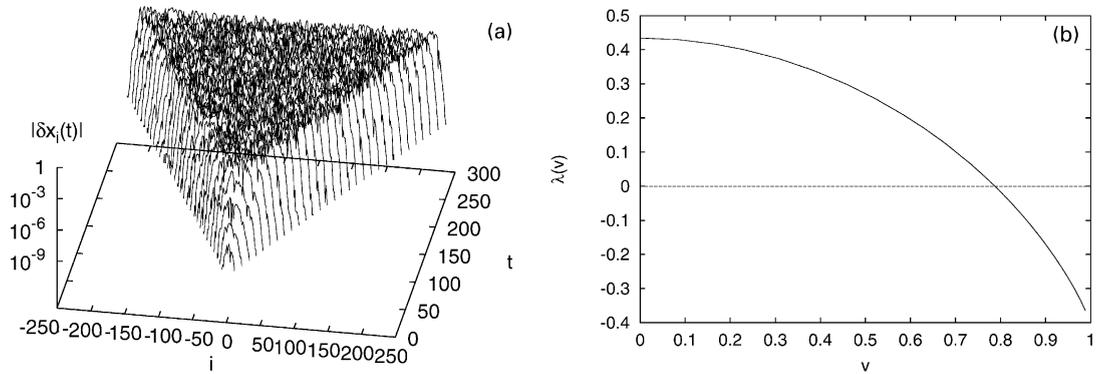


Fig. 9. (a) Space–time evolution of  $|\delta x_i(t)|$  for an initially localized perturbation (4.6) with  $\delta_0 = 10^{-8}$ . We used a CML of tent maps,  $f_a(x) = a(1/2 - |x - 1/2|)$ , with  $a = 2$ ,  $\epsilon = 2/3$  and  $L = 1001$ . (b)  $\lambda(v)$  for  $v > 0$  for the CML of (a). The straight line indicates the zero and the intersection between the curve  $\lambda(v)$  and 0 indicates the perturbation velocity  $V_F \approx 0.78$ .

The interesting point in Eq. (4.22) is that it gives not only a definite prescription to derive the propagation velocity but also a link between the velocity and the stability properties of the system. From this point of view it is instructive to look at the propagation velocity in another way [212].

The perturbation at different times resembles a propagating front, similar to those encountered in reaction–diffusion processes. But while in the latter context the front usually separates two stable phases or a stable from an unstable phase, here one phase is unstable and the other chaotic [212]. Made this analogy one can ask if it is possible to obtain the propagation velocity as in reaction–diffusion phenomena, where we know that the dynamics spontaneously selects the minimum allowed velocity [128].

Torcini et al. [212] have shown that this is indeed the case. They studied the evolution of a perturbation with an exponential profile (4.11) which, according to the definition of the specific Lyapunov exponent, evolves as in Eq. (4.12), i.e.  $\delta x_i(t) \sim \exp\{\Lambda(\mu)t - \mu i\}$ . This last relation tells us that the velocity of a front with shape given by  $\mu$  is  $V(\mu) = \Lambda(\mu)/\mu$ . According to the analogy with reaction–diffusion systems, one expects that a generic localized perturbation develops an exponential decaying shape (at the edges) with a definite exponent  $\mu_0$  (selected by the dynamics) [212]. This means that the propagation velocity  $V_F$  is determined by the relation  $V_F = V(\mu_0)$ . Now the problem is to compute  $\mu_0$ .

From Eq. (4.13), which relates  $\Lambda(\mu)$  and  $\lambda(v)$  through a Legendre transformation, one obtains

$$\frac{dV}{d\mu} = \frac{1}{\mu} \left( \frac{d\lambda}{d\mu} - \frac{\lambda}{\mu} \right) = -\frac{\lambda(v)}{\mu^2}. \tag{4.23}$$

Moreover, since  $\lambda(V_F) = 0$  (4.22) one has that  $dV/d\mu = 0$  at  $\mu_0$  such that  $V(\mu_0) = V_F$ , i.e.  $\mu_0$  selects the minimal velocity. Indeed  $\Lambda(\mu)$  is convex (being a Legendre transform), so that the

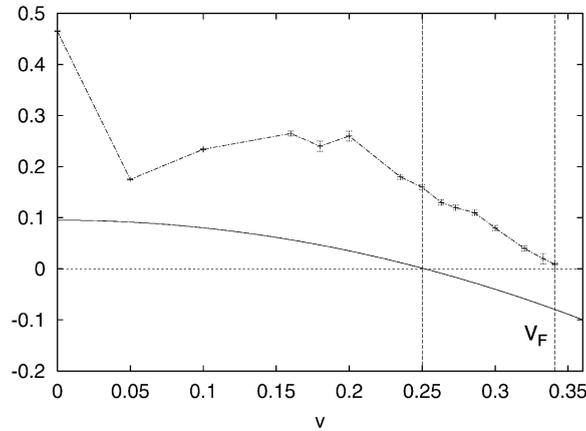


Fig. 10.  $\max_{\delta}[\lambda(\delta, v)]$  (dashed line with points) versus  $v$  for the shift map  $f(x) = rx \bmod 1$  with  $r = 1.1$  and  $\epsilon_0 = \frac{1}{3}$ , compared with  $\lambda(v)$  (continuous line). The two vertical lines indicates the velocity obtained by (4.22) which is about 0.250 and the directly measured one  $V_F \approx 0.342$ . Note that  $\max_{\delta}[\lambda(\delta, v)]$  approaches zero exactly at  $V_F$ .

minimum is unique and

$$V_F = \frac{A(\mu_0)}{\mu_0} = \left. \frac{dA(\mu)}{d\mu} \right|_{\mu=\mu_0}. \quad (4.24)$$

Thus for an infinitesimal perturbation, the selected velocity is the lowest possible one [212].

Summarizing, for short-range coupling the speed of propagation is finite and fully determines the spatio-temporal evolution of the perturbation. The situation becomes different for long-range coupling as (4.5). In this case the velocity of propagation is unbounded [173]. For the sake of completeness, we mention that the long-range coupling case has been also investigated in terms of a specific-like Lyapunov exponent which characterizes power law shaped perturbations [213]. The result of this analysis shows that the perturbation propagates exponentially fast with a rate given by the ratio of the largest LE and the power of the coupling.

We conclude this section by mentioning that there are cases in which the analysis in terms of  $\lambda(v)$  or, equivalently,  $A(\mu)$  fails to give the measured propagation velocity. Indeed, it has been found that  $V_F$  can be larger than the velocity for which  $\lambda(v) = 0$ . A finite propagation velocity has been measured even in systems with  $\lambda < 0$  (the so-called stable chaos phenomenon, see Section 7.3) for which the above analysis predicts  $V_F = 0$  [190].

This failure is related to the presence of strong non linearities. Recently, it has been proposed to generalize (4.22) to the non-linear regime of the perturbation growth by the definition of the finite size comoving Lyapunov exponent [52],  $\lambda(\delta, v)$ . It measures the divergence rate of perturbations of size  $\delta$  (not necessarily infinitesimal) in a moving frame. The algorithm is a generalization of the FSLE (Section 3.4), where now one follows an initially localized perturbation along the line  $[vt]$ . In Fig. 10 we compare  $\lambda(v)$  with  $\lambda(\delta, v)$  for a CML of shift maps, i.e.  $f(x) = rx \bmod 1$ . The latter goes to zero exactly at the directly measured propagation velocity  $V_F$ . Similar results hold for other maps [52]. These results suggest that a generalization of

Eq. (4.22), which is able to take into account also possible non-linear effects, is

$$\max_{\delta} [\lambda(\delta, V_F)] = 0.$$

The numerical evidences also suggest that the condition which should be accomplished in order to have deviation from the linear prediction given by (4.22) and (4.24) is that  $\lambda(\delta, v = 0) > \lambda(0, 0) = \lambda$ , confirming a conjecture done in [212]. However, even if interesting such a behavior seems to be rather non-generic.

#### 4.7. Macroscopic chaos in globally coupled maps

Recently the emergence of non-trivial collective behaviors in high-dimensional dynamical systems has gathered much attention [58,110,111,184]. A limit case of macroscopic coherence is the global synchronization of all the parts of the system. Beyond synchronization there exist other interesting phenomena, among which we just mention: clustering [110,129,67] and collective motion in globally coupled maps (GCM) [113,202,184]. The latter behavior, in the case that we call *macroscopic chaos* [50,203] (see below), is the subject of this section.

Let us consider a globally coupled map (GCM) defined as follows:

$$x_n(t+1) = (1 - \varepsilon)f_a(x_n(t)) + \frac{\varepsilon}{N} \sum_{i=1}^N f_a(x_i(t)), \quad (4.25)$$

where  $N$  is the total number of elements.

The evolution of a macroscopic variable, e.g., the center of mass

$$m(t) = \frac{1}{N} \sum_{i=1}^N x_i(t), \quad (4.26)$$

upon varying  $\varepsilon$  and  $a$  in Eq. (4.25), displays different behaviors [50]:

(a) *Standard chaos*:  $m(t)$  obeys a Gaussian statistics with a standard deviation  $\sigma_N = \sqrt{\langle m(t)^2 \rangle - \langle m(t) \rangle^2} \sim N^{-1/2}$ .

(b) *Macroscopic periodicity*:  $m(t)$  is a superposition of a periodic function and small fluctuations  $O(N^{-1/2})$ .

(c) *Macroscopic chaos*:  $m(t)$  displays an irregular motion as it can be seen by looking at the plot of  $m(t)$  vs.  $m(t-1)$  that appears as a structured function (with thickness  $\sim N^{-1/2}$ ), and suggests a chaotic motion for  $m(t)$ .

Phenomena (a) and (b) also appear in CML with local coupling in high enough dimensional lattices [58], for the interesting case (c), as far as we know, there is not a direct evidence in finite-dimensional CMLs.

In the case of *macroscopic chaos* one can expect that the center of mass evolves with typical times longer than the characteristic time of the full dynamics (i.e. the microscopic dynamics); the order of magnitude of the latter time may be estimated as  $1/\lambda_1$ . Indeed, conceptually, macroscopic chaos for GCM can be thought of as the analogous of the hydro-dynamical chaos for molecular motion, where, in spite of a huge microscopic Lyapunov exponent ( $\lambda_1 \sim 1/\tau_c \sim 10^{11} \text{ s}^{-1}$ ,  $\tau_c$  being the collision time), one can have rather different behaviors at

a hydro-dynamical (coarse grained) level, i.e.: regular motion ( $\lambda_{\text{hydro}} \leq 0$ ) or chaotic motion ( $0 < \lambda_{\text{hydro}} \ll \lambda_1$ ). In principle, if one knows the hydrodynamic equations, it is possible to characterize the macroscopic behavior by means of standard dynamical system techniques. However, in generic CML there are no general systematic methods to build up the macroscopic equations, apart from particular cases [113,184]. Therefore, here we discuss the macroscopic behavior of the system relying upon the full microscopic level of description.

The microscopic Lyapunov exponent cannot give a characterization of the macroscopic motion. To this purpose, recently different approaches have been proposed based on the evaluation of the self-consistent Perron–Frobenius (PF) operator [113,178,184] and on the FSLE [50,203]. Despite the conceptual interest of the former (in some sense the self-consistent PF-operator plays a role similar to that of the Boltzmann equation for gases [50]), here we shall only discuss the latter which seems to us more appropriate to address the predictability problem.

We recall that for chaotic systems, in the limit of infinitesimal perturbations  $\delta \rightarrow 0$ , one has  $\lambda(\delta) \rightarrow \lambda_1$ , i.e.  $\lambda(\delta)$  displays a plateau at the value  $\lambda_1$  for sufficiently small  $\delta$ . However, for non-infinitesimal  $\delta$ , one can expect that the  $\delta$ -dependence of  $\lambda(\delta)$  may give information on the characteristic time-scales governing the system, and, hence, it could be able to characterize the macroscopic motion. In particular, at large scales, i.e.  $\delta \gg 1/\sqrt{N}$ , one expects the (fast) microscopic components to saturate and  $\lambda(\delta) \approx \lambda_M$ , where  $\lambda_M$  can be fairly called the “macroscopic” Lyapunov exponent.

The FSLE has been determined by looking at the evolution of  $|\delta m(t)|$ , which has been initialized at the value  $\delta m(t) = \delta_{\min}$  by shifting all the elements of the unperturbed system by the quantity  $\delta_{\min}$  (i.e.  $x'_i(0) = x_i(0) + \delta_{\min}$ ), for each realization. The computation has been performed by choosing the tent map as local map, but similar results can be obtained for other maps [203,50].

Fig. 11a shows  $\lambda(\delta)$  versus  $\delta$  in the case of macroscopic chaos. One has two plateaus: at small values of  $\delta$  ( $\delta \leq \delta_1$ ), as expected from general considerations,  $\lambda(\delta) = \lambda_1$ ; for  $\delta \geq \delta_2$

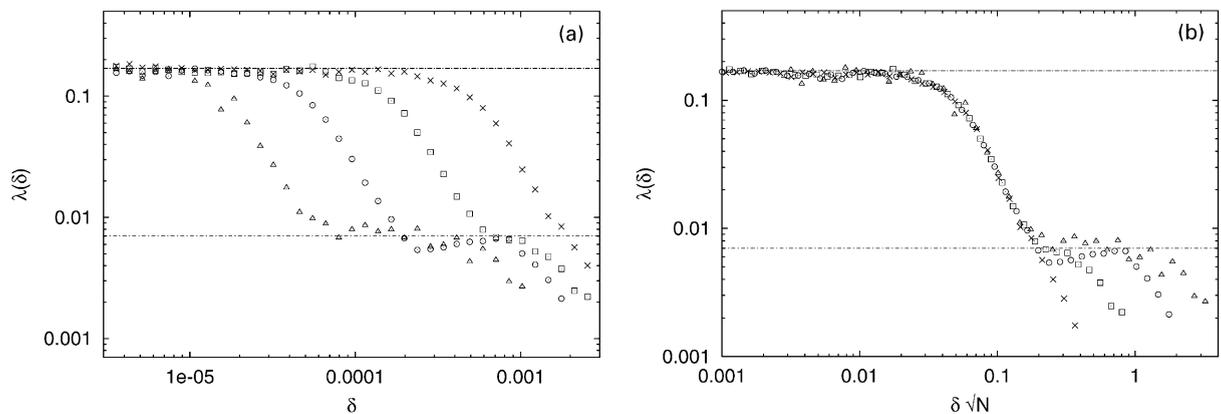


Fig. 11.  $\lambda(\delta)$  versus  $\delta$  for the system (4.25) with  $a = 1.7$ ,  $\epsilon = 0.3$  for  $N = 10^4$  ( $\times$ ),  $N = 10^5$  ( $\square$ ),  $N = 10^6$  ( $\odot$ ) and  $N = 10^7$  ( $\triangle$ ). The first plateau corresponds to the microscopic Lyapunov exponent  $\lambda_{\text{micro}} \approx 0.17$  and the second one to the macroscopic Lyapunov exponent  $\lambda_{\text{macro}} \approx 0.007$ . The average is over  $2 \times 10^3$  realizations for  $N = 10^4, 10^5, 10^6$  and 250 realizations for  $N = 10^7$ . (b) The same as (a) rescaling the  $\delta$ -axis with  $\sqrt{N}$ .

one has another plateau, the “macroscopic” Lyapunov exponent,  $\lambda(\delta) = \lambda_M$ . Moreover,  $\delta_1$  and  $\delta_2$  decrease at increasing  $N$ : indeed, by looking at Fig. 11b one can see that  $\delta_1, \delta_2 \sim 1/\sqrt{N}$ . It is important to observe that the macroscopic plateau, being almost non-existent for  $N = 10^4$ , becomes more and more resolved and extended on large values of  $\delta\sqrt{N}$  at increasing  $N$  up to  $N = 10^7$ . Therefore we can argue that the macroscopic motion is well defined in the limit  $N \rightarrow \infty$  and one can conjecture that in this limit the microscopic signature in the evolution of  $\delta m(t)$  completely disappears in favor of the macroscopic behavior. In the case of standard chaos ( $\lambda_M < 0$ ) one has only the microscopic plateau and then a fast decreasing of  $\lambda(\delta)$  [50].

#### 4.8. Predictability in presence of coherent structures

Here we discuss some problems which arise in characterizing the predictability in continuous systems, described by PDE. In this case the norms are not equivalent [127] and the computation of the LE can give different results. Rather than discussing the problem in general terms, we consider here two-dimensional turbulence as a specific example. The choice of this example is due to several reasons. First of all, two-dimensional turbulence is a continuous system of relevance in atmospheric physics, and it has been extensively investigated in the last years [206,208,176,34]. The statistical theory for two-dimensional turbulence has been developed by Kraichnan and Batchelor [132] on a similar basis to the Kolmogorov theory for three-dimensional turbulence. The main difference is the existence of a second inviscid invariant, the enstrophy (average square vorticity). As a consequence, in the limit of high Reynolds numbers, the energy cannot be dissipated by viscosity and one expects a direct cascade of enstrophy. With an input source at intermediate scales, the energy injected into the system is transferred to large scales by an *inverse* cascade. A large numbers of numerical simulations [206,34] and experiments [176] have demonstrated the universality of the inverse cascade with spectral index very close to the predicted Kolmogorov exponents.

The situation is much less clear for what concerns the direct cascade. The predicted spectral slope (Kraichnan–Batchelor spectrum) is seldom observed and even universality with respect to the forcing or to the form of dissipation is questioned [159]. The freely decaying evolution is characterized by the emergence of coherent structures [159] which eventually dominate the dynamics. Coherent structures are weakly dissipative, rather regular regions of fluids in the turbulent background flow whose interactions can be approximately described by a conservative dynamics [135]. The spontaneous emergence of coherent structures makes two-dimensional turbulence a prototype model for geophysical flows [141] and, most important for our purpose, gives a natural example for illustrating the effects of choosing different error norms.

The equation for describing two-dimensional turbulence is the Navier–Stokes equation written for the scalar vorticity  $\omega = \nabla \times \mathbf{v}$  as [18,159]

$$\frac{\partial \omega}{\partial t} + \frac{\partial(\psi, \omega)}{\partial(x, y)} = (-1)^{p+1} \nu_p \Delta^p \omega \quad (4.27)$$

where  $\psi$  is the stream function such that  $\mathbf{v} = (\partial_y \psi, -\partial_x \psi)$  and  $\Delta \psi = -\omega$ . As customary in direct numerical simulations, the dissipation is modified by employing high order viscosity  $p > 1$  in order to achieve larger Reynolds numbers. The numerical results discussed below are obtained

by integrating (4.27) by means of a standard pseudo-spectral code on a periodic computational domain with resolution  $N \times N$ .

The classical theory of predictability in turbulence [136,137] studies the evolution of a difference (or error) field, defined as

$$\delta\omega(\mathbf{x}, t) = \frac{1}{\sqrt{2}}(\omega'(\mathbf{x}, t) - \omega(\mathbf{x}, t)) \quad (4.28)$$

where  $\omega$  and  $\omega'$  are solutions of (4.27) started from slightly different initial conditions. The “error” is computed from  $\delta\omega$  and measured in terms of a given norm which is the subject of our discussion. Indeed the method used for defining the distance between the reference and perturbed field is a delicate point for continuous systems such as Navier–Stokes equations. Classical norms are based on the invariants of (4.27) in the inviscid limit  $\nu_p = 0$ , i.e. enstrophy and energy norms [137,122]

$$Z_\delta(t) = \frac{1}{2} \int d^2x |\delta\omega(\mathbf{x}, t)|^2 = \int_0^\infty dk Z_\delta(k, t), \quad (4.29)$$

$$E_\delta(t) = \int_0^\infty dk k^{-2} Z_\delta(k, t) = \int_0^\infty dk E_\delta(k, t), \quad (4.30)$$

where we have also introduced the enstrophy ( $Z_\delta$ ) and energy ( $E_\delta$ ) error spectra. It is also natural to introduce the relative errors, defined as

$$r(t) = \frac{E_\delta(t)}{E(t)}, \quad z(t) = \frac{Z_\delta(t)}{Z(t)}, \quad (4.31)$$

where  $E(t) = \frac{1}{2} \int v^2(x) dx$  and  $Z = \frac{1}{2} \int \omega^2(x) dx$ , and the relative error spectrum

$$r(k, t) = \frac{E_\delta(k, t)}{E(k, t)} = \frac{Z_\delta(k, t)}{Z(k, t)}. \quad (4.32)$$

This issue was already addressed in [122] where the infinitesimal (linear) error growth was computed using several “Eulerian norms” as (4.29), (4.30).

We will consider an initial error given by complete uncertainty at small scales, i.e.  $r(k, 0) = 0$  for  $k < k_0$  and  $r(k, 0) = 1$  for  $k > k_0$ . This assumption is physically justified by the finite resolution of any measurement device and/or the numerical simulation scheme. For an infinitesimal perturbation, the error is expected to grow exponentially with the largest LE  $\lambda_1$ . Because we are dealing with a dissipative system which ultimately collapses on the trivial fixed point  $\omega = 0$ ,  $\lambda_1$  is formally negative. However, this is only a formal problem. Indeed in high Reynolds number turbulence the dissipation time scale is much longer than the dynamical time and we can make use of the effective LE  $\gamma(t)$  (3.8). For  $t$  much smaller than the dissipation time, we can consider, from any point of view,  $\gamma(t)$  as the Lyapunov exponent of the decaying turbulence.

The exponential growth regime ends at times much smaller than the dissipative time, as soon as the separation of the two fields cannot be any more considered infinitesimal and the finite error regime sets in. The predictability time is defined by means of the accepted tolerance  $\Delta$  or, which is equivalent, by a threshold for the relative errors (4.31). We will follow the classical prescription for the predictability time  $r(T_p) = \frac{1}{4}$  [137]. In Fig. 12 we plot relative errors (4.31) as functions of time for a  $512^2$  resolution simulation [33]. For small times ( $t < 250$ ) we can see an exponential growth for both  $r(t)$  and  $z(t)$  with effective LE  $\gamma \simeq 0.08$ . At larger times the

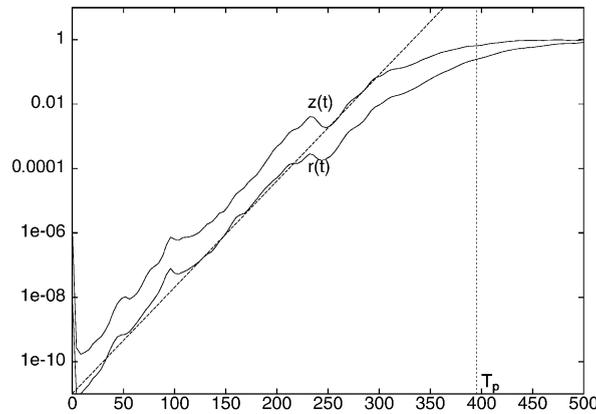


Fig. 12. Relative energy ( $r$ ) and enstrophy ( $z$ ) error growth for a  $512^2$  simulation.  $T_p$  indicate the predictability time defined as  $r(T_p) = \frac{1}{4}$ . The dashed line represents the exponential regime  $r(t) \sim \exp(0.08t)$ .

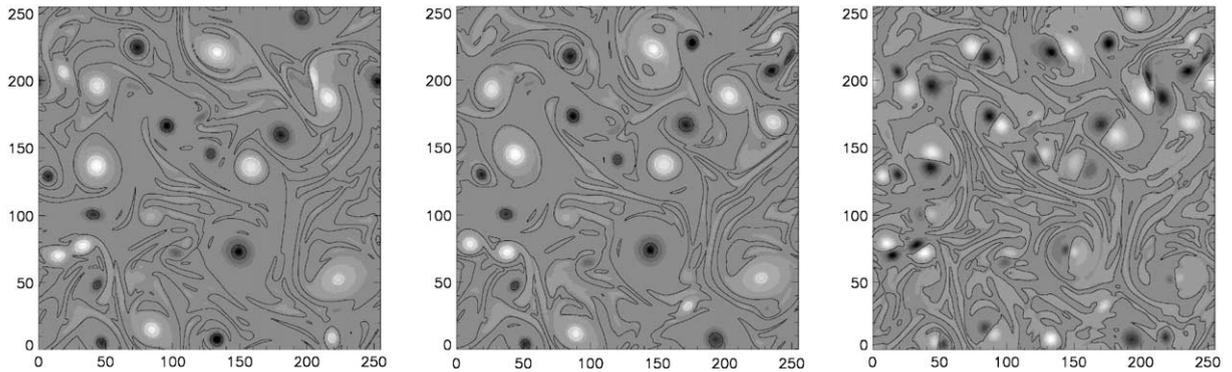


Fig. 13. Gray scale map of the vorticity fields (obtained by a  $256^2$  simulation) at time  $T_p = 177$ . White corresponds to positive vorticity regions, black to negative ones. (a) Reference field  $\omega(\mathbf{x})$ , (b) the perturbed one  $\omega'(\mathbf{x})$ , (c) the error field  $\delta\omega(\mathbf{x})$ .

error curves bend and a predictability time estimation with energy norm gives  $T_p \simeq 395$ . From Fig. 12 we learn at least two lessons. First (but not surprisingly) about half of the predictability time is governed by non-exponential error growth behavior. This is another demonstration of the little relevance of LE for characterizing predictability in realistic complex systems. The second observation is that the different norms  $r(t)$  and  $z(t)$  give qualitatively similar results. Because the error is initially confined to small scales  $k > k_0$ , the vorticity-based norm is always larger than the energy-based norm, but the predictability time is essentially independent of the norm used. It is not difficult to understand that any Eulerian norm would give comparable result. Because the error propagates from small to large scales, a norm which emphasizes small-scale features (as the enstrophy norm) saturates earlier than a large scale based norm (energy, in our example), but the results remain essentially the same.

In Fig. 13 we plot the vorticity field of the reference  $\omega(\mathbf{x}, T_p)$  and perturbed field  $\omega'(\mathbf{x}, T_p)$  at the predictability time  $T_p$ . Although the two fields differ, by definition, by 25% in energy and

about 65% in enstrophy, they look still remarkably similar for what concerns the distribution of vortices. Most of the large coherent structures are almost in the same positions.

In Fig. 13 we also plot the difference field  $\delta\omega(\mathbf{x}, T_p)$ . The typical bipolar configuration, usually observed in simulations [122,164], indicates that the error is concentrated in correspondence of the vortices and that it is essentially due to the different position of the vortex structures in the two realizations.

This result suggests that a Lagrangian measure of the error, based on the vortex positions, would be more suitable for the present system. For example, to emphasize the limits of the Eulerian measure for the error (4.29), (4.30), consider the limiting case of singular point vortices, where an infinitesimal error in the coordinates gives error saturation and hence zero predictability time. In general, we expect that, in presence of vortices, an Eulerian-based measure underestimates the predictability time.

This problem can be overcome by resorting to the natural distance among vortex centers. We use a vortex tracking algorithm which recognizes and follows vortices during the dynamics. First we need a definition of vortex, the one here adopted is: a connected region  $D_\alpha$  in the computational domain with vorticity maximum  $z_\alpha$  larger (in absolute value) than a given threshold and vorticity larger than a fraction (we used 0.2) of the vorticity peak. Given the vortex domains  $D_\alpha$ , all the physical quantities are computed by integrating inside the domains. For example, vortex circulation is defined as  $\Gamma_\alpha = \int_{D_\alpha} d^2x \omega(\mathbf{x})$  and vortex center  $\mathbf{x}_\alpha$  is the center of mass computed from the vorticity field. Finally, vortex trajectories are reconstructed by matching center positions at different times.

A Lagrangian, vortex-based, measure of the error can, e.g., be defined as

$$d^2(t) = \frac{1}{\sum_\alpha |\Gamma_\alpha|} \sum_\alpha |\Gamma_\alpha| |\mathbf{x}'_\alpha - \mathbf{x}_\alpha|^2 \tag{4.33}$$

where  $\mathbf{x}_\alpha$  and  $\mathbf{x}'_\alpha$  are the vortex positions respectively in the reference and perturbed field. In Fig. 14 we plot  $d^2$  obtained from our simulation. We observe that at the classical predictability time  $T_p$  the mean vortex separation is about one-tenth of the saturation level.

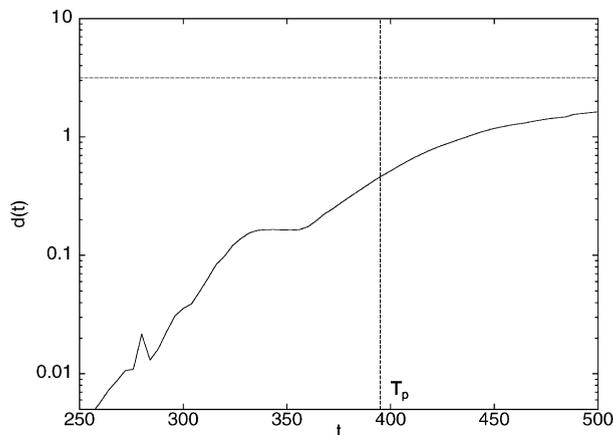


Fig. 14. Mean vortex separation  $d(t)$  at resolution 512<sup>2</sup>. At the classical predictability time  $T_p$ , the mean vortex separation is about one-tenth of the saturation level.

time, the mean vortex separation is  $d(T_p) \simeq 0.5$ , well below the saturation value ( $d_{\max} \sim L/2 = \pi$  in the periodic computational box). This result is a quantitative confirmation of the observations drawn from Fig. 13, i.e. the existence of an intermediate regime in which the (finite) error is ruled by the displacement of the strong coherent structures. If one is interested in predicting, with some tolerance, positions and intensities of coherent structures, it is possible to have a much larger predictability time.

## 5. Predictability in fully developed turbulence

### 5.1. Basic facts of turbulence

Perhaps, *fully developed turbulence* [161,84] is the most important instance of high-dimensional chaotic system. To give an example, let us consider a classical experiment in fluid dynamics: in a wind tunnel, an air mass conveyed by a large fan impinges some obstacles, which perturb significantly the velocity of fluid particles. Big and small whirls appear, and the flow evolves irregularly in time. One could wonder whether the features of the flow depend crucially on the physical properties of the fluid, the size and shape of the obstacle, the mean wind velocity, and so on. It is easy to understand that, with a given geometry the only relevant parameter which characterizes the flow is the Reynolds number  $Re = UL/\nu$ , where  $U$  is the mean wind velocity,  $L$  is the typical size of the obstacle and  $\nu$  is the kinematic viscosity of the fluid. When  $Re$  is very large, i.e., of the order of a thousand or more [19] turbulence is called fully developed. The fundamental physical interest in this regime is motivated by the existence of universal properties with respect to the details of the experimental setup [19,161]. If a velocity probe is placed at some distance past the obstacle, it is possible to record a temporal series that gives us statistical information. If one sits far enough from the obstacle, there the small-scale properties of the flow do not depend sensitively on the precise site and orientation of the probe, that is the turbulence is approximately homogeneous and isotropic. Since the flow is swept across the probe at a mean velocity  $U$ , that largely exceeds the magnitude of the fluctuations, one can expect that the time record essentially amounts to a one-dimensional spatial section of the velocity field. Thus time-scales and length-scales are interchangeable, this is the essence of the *Taylor hypothesis* [161]. Assuming the above hypothesis, we can safely reinterpret temporal variations of the velocity, on an interval  $\tau$ , in a fixed-point of the space as spatial increments on scale  $\ell = U\tau$ , at a fixed time.

The first important result about the expected universality is the behavior of the velocity power spectrum which closely follows a power law decay  $E(k) \propto k^{-5/3}$  on a given range of wave numbers [123,161]. At larger wave number the spectrum falls off with an exponential-like behavior, whereas the form at small  $k$  (i.e. large scales) depends on the mechanism of forcing and/or boundary conditions. For  $k \rightarrow 0$  one often observes a self-similar energy spectrum  $E(k) \sim k^s$  with scaling exponent  $s > 0$ . In incompressible decaying turbulence, there are some arguments indicating that asymptotically  $s \leq 4$  where the limiting value  $s = 4$  is observed if initially the spectrum has  $s > 4$  [19]. A typical turbulence spectrum is shown in Fig. 15.

The two crossovers unveil the presence of two characteristic scales: a large excitation scale  $L \sim k_L^{-1}$ , associated with the energy containing eddies, and a small dissipation scale  $\ell_D \sim k_D^{-1}$ ,

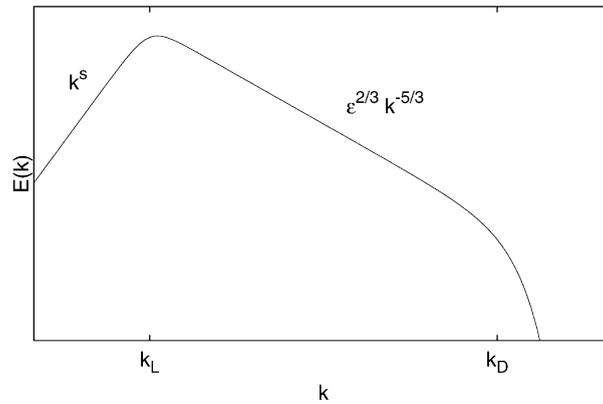


Fig. 15. Typical turbulent energy spectrum,  $k_L^{-1}$  is the energy containing integral scale and  $k_D^{-1}$  the dissipative Kolmogorov scale.

related to the smallest active eddies. The appearance of a power law in between these two extremes unveils that no other characteristic scale is involved.

A simple and elegant explanation of these experimental findings is due to A.N. Kolmogorov [16]: in a nutshell, it is assumed the existence of a range of scales where the energy—injected at the scale  $L$ —flows down (with a cascade process, as remarked by Richardson [194]) to the dissipative scale  $\ell_D$ , where it is dissipated by molecular viscosity. Since, practically, neither injection nor dissipation takes place in this interval, it is called the inertial range. In this range the only relevant quantity is the average energy transfer rate  $\bar{\epsilon}$ : dimensional counting imposes then a power spectrum  $E(k) \propto \bar{\epsilon}^{2/3} k^{-5/3}$  in agreement with the experimental observations discussed above. The scaling for the spectrum is equivalent to a power law dependence for the second-order structure function (SF)

$$S_2(\ell) = \langle \delta v_\parallel^2 \rangle = \langle (v(x + \ell) - v(x))^2 \rangle \sim \ell^{2/3} . \tag{5.1}$$

The original Kolmogorov theory (K41) assumes self-similarity of the turbulent flow. As a consequence, the scaling behavior of higher-order structure functions  $S_p(\ell) = \langle |v(x + \ell) - v(x)|^p \rangle \sim \ell^{\zeta_p}$  is described by a single scaling exponent. The value of the exponent is determined by the so-called “4/5 law”, an exact relation derived by Kolmogorov from the Navier–Stokes equations [123,84], which, under the assumption of stationarity, homogeneity and isotropy states

$$\langle \delta v_\parallel^3(\ell) \rangle = -\frac{4}{3} \bar{\epsilon} \ell , \tag{5.2}$$

where  $\delta v_\parallel(\ell)$  is the longitudinal velocity difference between two points at distance  $\ell$ , and  $\bar{\epsilon}$  is the average rate of energy transfer. The structure function exponent  $\zeta_p$  is thus predicted by Kolmogorov similarity theory to be  $\zeta_p = p/3$ .

Several experimental investigations [7,84] have shown that the Kolmogorov scaling is not exact and  $\zeta_p$  is a non-linear function (with  $\zeta_3=1$  as a consequence of the “4/5 law”). This means a breakdown of the self-similarity in the turbulent cascade. Larger and larger excursions from mean values are observed as one samples smaller and smaller scales. This phenomenon goes under the name of intermittency [84]. A complete theoretical understanding of intermittency in

Navier–Stokes turbulence is still lacking. Nevertheless, there are approaches, as the multifractal model [177], which are able to characterize at a phenomenological level the intermittency.

In brief, the basic idea of the multifractal model [177,172,84] consists in assuming a local scale-invariance for the velocity fluctuations, i.e. one has  $\delta v_\ell \sim \ell^h$ , with a continuous spectrum of (Hölder) exponents  $h$ , each belonging to a given fractal set. In other words, in the inertial range one has

$$\delta v_\ell(x) \sim \ell^h, \quad (5.3)$$

if  $x \in S_h$ , and  $S_h$  is a fractal set with dimension  $D(h)$  and  $h \in (h_{\min}, h_{\max})$ . The probability to observe a given scaling exponent  $h$  at the scale  $\ell$  is thus  $P_\ell(h) \sim \ell^{3-D(h)}$ . In this language the Kolmogorov similarity theory [123,84] corresponds to the case of only one singularity exponent  $h = \frac{1}{3}$  with  $D(h = \frac{1}{3}) = 3$ , see also Appendix B.

### 5.2. Reduced model of turbulence

In numerical simulations of the Navier–Stokes equations in the regime of fully developed turbulence, one has to discretize the original PDE to obtain a set of approximate ODE which must be integrated numerically. This is the direct numerical simulation approach which, in its simplest form, is implemented on a regular 3D grid of  $N^3$  points. Since the dissipative scale (Kolmogorov scale) is related to the Reynolds number as  $\ell_D \sim LRe^{-3/4}$ , an estimate of the number  $\mathcal{N}$  of active spatial degrees of freedom leads to

$$\mathcal{N} \sim (L/\ell_D)^3 \sim Re^{9/4}. \quad (5.4)$$

An obvious consequence of the fast growth of  $\mathcal{N}$  with the Reynolds number is the unfeasibility of a complete turbulent simulations at high  $Re$ . The maximum limit of present computers is about  $N=10^3$  which corresponds to  $Re \simeq 10^4$ . An alternative approach has been introduced with the so-called shell models by the works of Obukhov, Gledzer and Desnyansky and Novikov (see [38] for a detailed discussion). The basic idea, originally motivated in the context of closure theory, is to implement a dynamical cascade with a set of variables  $u_n$  ( $n = 1, \dots, N$ ) each representing the typical magnitude of the velocity fluctuation in a shell of wave-numbers  $k_n < |\mathbf{k}| < k_{n+1}$ . The representative wave-numbers are spaced geometrically,  $k_n = k_0 2^n$ , in this way, assuming locality in the cascade, interactions are confined to neighboring shells only.

We will discuss a specific model, known as GOY model (see [38] for a review), which makes use of complex velocity variables  $u_n$  and for which the equations of motion are

$$\left( \frac{d}{dt} + \nu k_n^2 \right) u_n = i k_n (u_{n+1} u_{n+2} - \frac{1}{4} u_{n-1} u_{n+1} - \frac{1}{8} u_{n-2} u_{n-1})^* + f_n, \quad (5.5)$$

where  $\nu$  is the viscosity and  $f_n$  is a forcing term (typically restricted on the first shells). The coefficients in the nonlinear term (which has the same structure of Navier–Stokes equations) are chosen to conserve energy  $E \equiv 1/2 \sum_n |u_n|^2$  in the unforced, inviscid limit.

Without entering in the details, we recall that shell model (5.5) displays energy cascade à la Kolmogorov from the large scales of forcing to the dissipative scales ( $n \sim N$ ) with a statistical constant energy flux  $\bar{\epsilon}$ . On these inertial range scales, the moments of velocity show power law scaling  $\langle |u_n|^p \rangle \sim k_n^{-\zeta_p}$  with exponents close to those experimentally observed for fully developed turbulence.

The number of shells  $N$  necessary to mimic the cascade mechanism of fully developed turbulence is rather small, due to the geometrical progression in  $k_n$  one has  $N \sim \log_2 Re$ . We have thus a chaotic dynamical system with a reasonably small number of degrees of freedom where standard methods of deterministic chaos can be used in order to relate the “turbulent” statistical description in terms of structure functions and intermittency, and dynamical properties, such as the spectrum of Lyapunov exponents. The absence of any stochastic term in (5.5) makes the shell model a natural model for investigating the predictability problem in turbulence.

### 5.3. Effects of intermittency on predictability of infinitesimal perturbations

The sensitive dependence on initial conditions makes the long term forecasting in turbulent flow practically impossible. For instance, Ruelle [196] remarked that thermal fluctuations in the atmosphere produces observable changes on a scale of centimeters after only few minutes. As a consequence after one or two weeks, the large-scale atmospheric circulation would be completely unpredictable, even if the exact evolution equations were known. This is the so-called *butterfly effect*, in the words of Lorenz: *A butterfly moving its wings in Brazil might cause the formation of a tornado over Texas*. To support this argument, one can observe that the largest LE of fully developed turbulence is roughly proportional to the inverse of the smallest characteristic time of the system, the turn-over time  $\tau_D$  of eddies of the size of the Kolmogorov length  $\ell_D$ . From  $\ell_D \sim LRe^{-3/4}$  one obtains

$$\tau_D \sim \ell_D / \delta v_D \sim \tau_L Re^{-1/2}, \quad (5.6)$$

where  $\tau_L \approx L/U$  is the eddy turn-over time of the energy containing scales. As a consequence, as first pointed out by Ruelle [196], the largest LE scales with  $Re$  like  $\lambda \sim 1/\tau_D \sim Re^{1/2}/\tau_L$ . Thus fully developed turbulence is characterized by a Lyapunov exponent diverging with  $Re$ .

Nevertheless, a large value of the LE does not prevent the possibility of long term prediction, at least if one is interested in predicting the large scales behavior (which is related to finite errors), see Section 5.4.

Remaining in the framework of infinitesimal perturbations, we discuss the effects of intermittency on the predictability time. The multifractal model [177] predicts a spectrum of viscous cut-offs: each singularity exponent  $h$  selects a different damping scale,  $\ell_D(h) \sim LRe^{-1/(1+h)}$ , and hence a spectrum of (dissipative) turn-over times,  $\tau_D(h)$ , such that (5.6) becomes

$$\tau_D(h) \sim \ell_D(h) / \delta v_D \sim \tau_L Re^{-(1-h)/(1+h)}, \quad (5.7)$$

(see Appendix B for details). To obtain the largest Lyapunov exponent now we have to integrate  $\tau_D(h)^{-1}$ , at the scale  $\ell = \ell_D(h)$ , over the  $h$ -distribution  $P_\ell(h) \sim \ell^{3-D(h)}$ :

$$\lambda \sim \int \tau(h)^{-1} P_\ell(h) dh \sim \frac{1}{\tau_L} \int \left( \frac{\ell_D}{L} \right)^{h-D(h)+2} dh. \quad (5.8)$$

Since the viscous cut-off vanishes in the limit  $Re \rightarrow \infty$ , the integral can be estimated by the saddle-point method, i.e.

$$\lambda \sim \frac{1}{\tau_L} Re^\alpha \quad \text{with } \alpha = \max_h \left[ \frac{D(h) - 2 - h}{1 + h} \right]. \quad (5.9)$$

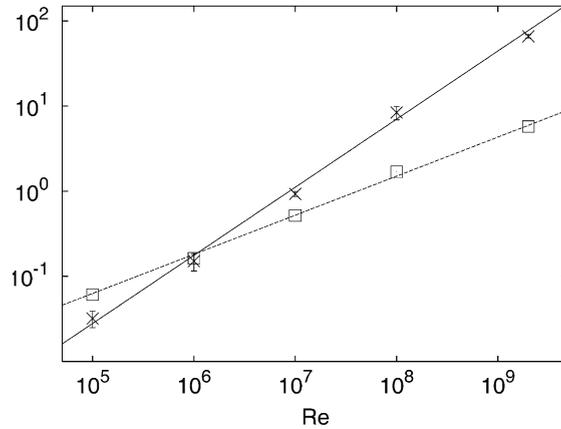


Fig. 16. Lyapunov exponent  $\lambda$  (□) and variance  $\mu$  (×) as a function of the Reynolds number  $Re$  for the shell model (5.5) with  $N = 27$  shells. The dashed line is the multifractal prediction  $\lambda \sim Re^\alpha$  with  $\alpha = 0.459$ , with function  $D(h)$  obtained by the random beta model fit of the  $\zeta_p$  exponents. The full line represents  $\mu \sim Re^w$  with  $w = 0.8$ .

The value of  $\alpha$  depends on the shape of  $D(h)$ . By using the function  $D(h)$  obtained by fitting the exponents  $\zeta_q$  with the random  $\beta$ -model [22] one finds  $\alpha = 0.459\dots$ , slightly smaller than the Ruelle prediction  $\alpha = 0.5$ . This result is confirmed by numerical simulations on the shell model (5.5) (see Fig. 16).

We remind that the fluctuations of the effective Lyapunov exponent  $\gamma(t)$  can be characterized by the ratio of  $\mu/\lambda$  (Section 3.1). The variance  $\mu$  is

$$\mu = \lim_{t \rightarrow \infty} t[\langle \gamma(t)^2 \rangle - \langle \gamma(t) \rangle^2] \sim t_c \langle (\gamma - \lambda)^2 \rangle, \tag{5.10}$$

where in the last expression we have introduced the integral correlation time  $t_c = \int C(t) dt$  of the effective Lyapunov exponent [66,38], where  $C(t)$  is the normalized correlation function of the fluctuation of  $\gamma(t)$  (i.e.  $\gamma(t) - \lambda$ ).

The quantity  $\langle (\gamma - \lambda)^2 \rangle$  can be computed by repeating the argument for  $\lambda$ :

$$\langle \gamma^2 \rangle \sim \langle \tau^{-2} \rangle \sim \frac{1}{\tau_L^2} Re^y. \tag{5.11}$$

An explicit calculation [66] gives  $y = 1$  independently of intermittency. Assuming that the correlation time  $t_c$  vanishes as a power of  $Re$

$$t_c \sim \tau_L Re^{-z} \tag{5.12}$$

one ends with the prediction

$$\mu \sim \frac{1}{\tau_L} Re^w \quad \text{with } w = 1 - z. \tag{5.13}$$

Numerical simulations on the shell model (5.5) give  $w \simeq 0.8$  (see Fig. 16). Because  $w > \alpha$  we obtain that  $\mu/\lambda$  diverges with  $Re$ . From Fig. 16 we see that the strong intermittency regime begins, for the shell model, at  $Re \sim 10^6$ . Let us stress that in the absence of intermittency one would expect that  $t_c \sim \lambda^{-1}$  and thus  $z = \frac{1}{2}$  and  $\mu/\lambda$  constant. The fact that  $z \sim 0.2$  indicates that

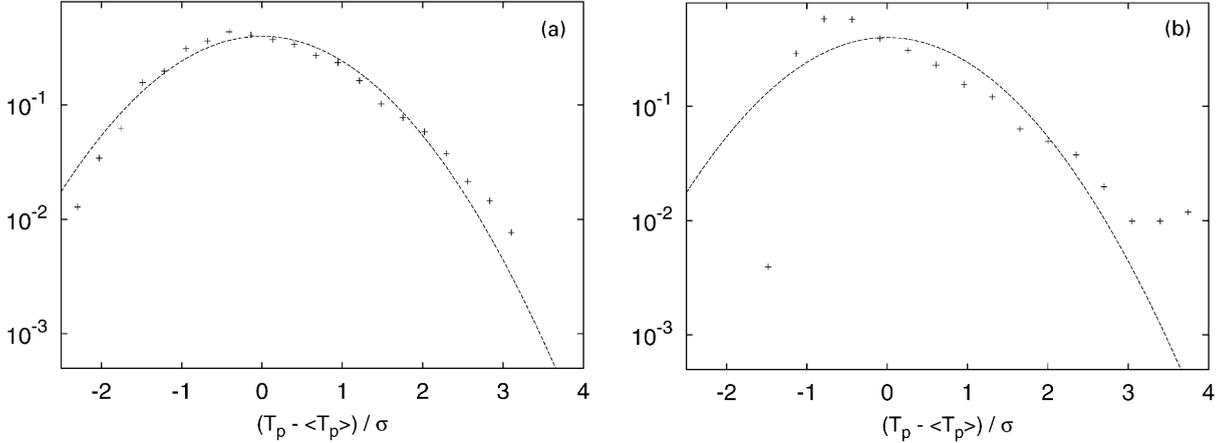


Fig. 17. Rescaled probability distribution functions of the predictability time  $T_p$  for the shell model (5.5) for (a)  $Re = 10^6$  and (b)  $Re = 2 \times 10^9$ . The respective average values are  $\langle T_p \rangle = 84.0$  and  $6.32$  and the standard deviations are  $\sigma(T_p) = 22.2$  and  $3.16$ . The line is the Gaussian.

the presence of quiescent periods in the turbulent activity is much more relevant for the decay rate of time correlations than for the Lyapunov exponent.

We have seen in Section 3.3 that the fluctuations of the effective LE affect the distribution of predictability time, and thus we expect a similar effect in fully developed turbulence. In the shell model one can estimate the predictability time by computing the time  $T_p$  at which the difference  $\delta u_m(t)$  (where  $m$  corresponds to the integral scale) among two realizations of the model becomes larger than the tolerance  $\Delta$ . The initial difference  $\delta_0$  is restricted to the shell  $u_{n^*}$  on the Kolmogorov scale and  $m \ll n^*$ . The predictability time distribution function is computed at two different Reynolds number. At  $Re = 10^6$  we are at the border of the weak intermittent range: the observed PDF (Fig. 17) is indeed close to a Gaussian with mean value

$$\langle T_p \rangle \simeq \frac{1}{\lambda} \ln \left( \frac{\Delta}{\delta_0} \right). \quad (5.14)$$

On the contrary, at  $Re = 2 \times 10^9$ , the PDF exhibits the asymmetric triangular shape and the mean value is ruled by  $\mu$  according to (3.28).

#### 5.4. Growth of non-infinitesimal perturbations

The classical theory of predictability in turbulence has been developed by Lorenz [146] (see also [143]) using physical arguments, and by Leith and Kraichnan [137] on the basis of closure approximations. The fundamental ingredients of the Lorenz approach stem from dimensional arguments on the time evolution of a perturbation in an energy cascade picture. In this framework, it is rather natural to assume that the time  $\tau_\ell$  for a perturbation at scale  $\ell/2$  to induce a complete uncertainty on the velocity field on the scale  $\ell$ , is proportional to the typical eddy turn-over time at scale  $\ell$ :  $\tau_\ell \sim \ell / \delta v_\ell$  where  $\delta v_\ell$  is the typical velocity difference at scale

$\ell$ . Kolmogorov scaling (5.1) gives

$$\tau_\ell \sim \ell^{2/3} . \quad (5.15)$$

Because of the geometric progression (5.15), the predictability time to propagate an uncertainty  $O(\delta v_D)$  from the Kolmogorov scale  $\ell_D$  up to the scale of the energy containing eddies  $L$ , is dominated by the longest time

$$T_p \sim \tau_{\ell_d} + \tau_{2\ell_d} + \dots + \tau_L \sim \tau_L \sim \frac{L}{\delta v_L} . \quad (5.16)$$

Closure approximations, where one still uses dimensional arguments, confirm this result [137,143].

It is important to stress that, in the Lorenz approach, the predictability time is independent of the Reynolds number. This is only in apparent contradiction with the increase of the Lyapunov exponent with  $Re$  (5.9). From the point of view of an observer interested in forecasting the large scales (i.e. not infinitesimal perturbations) the Lyapunov exponent is not physically relevant. Large-scale predictability in turbulence is hence another example where a large LE coexists with a long predictability time. We will see that a coherent description that includes these two features of predictability is given by the finite size Lyapunov exponent (3.37).

It is easy to estimate the scaling behavior of  $\lambda(\delta)$  when the perturbation is in the inertial range  $\delta v_D \leq \delta \leq \delta v_L$ . Following the phenomenological ideas of Lorenz, the doubling time of an error of magnitude  $\delta$  can be identified with the turn-over time  $\tau_\ell$  of an eddy with typical velocity difference  $\delta v_\ell \sim \delta$ . Using the scaling (5.1) one has  $\tau_\ell \sim \tau_L(\ell/L)^{2/3} \sim \tau_L(\delta v_\ell/\delta v_L)^{-2}$ . In conclusion we obtain [11]

$$\lambda(\delta) \sim \delta^{-2} . \quad (5.17)$$

In the dissipative range  $\delta < \delta v_D$ , the error can be considered infinitesimal, implying  $\lambda(\delta) = \lambda$ .

Accounting for intermittency, in the framework of the multifractal approach, one has

$$\lambda(\delta) \sim \tau_L^{-1} \int dh (\delta/\delta v_L)^{[3-D(h)]/h} (\delta/\delta v_L)^{1-1/h} . \quad (5.18)$$

From the basic inequality of the multifractal model  $D(h) \leq 3h + 2$  (see Appendix B), we have

$$\frac{2+h-D(h)}{h} \geq -2 \quad \text{for all } h . \quad (5.19)$$

As a result of the constancy of the energy flux in the inertial range,  $\bar{\varepsilon} = v^3(\ell)/\ell$ , the equality holds for  $h = h^*(3)$ , and gives  $3h^*(3) + 3 - D(h^*(3)) = 1$ . Therefore a saddle-point estimation of (5.18) gives again (5.17). The dimensional scaling of the FSLE in fully developed turbulence  $\lambda(\delta) \sim \delta^{-2}$  is thus not affected by intermittency corrections. This is a direct consequence of the exact result (5.2).

These findings have been numerically tested on the shell model (5.5) for the energy cascade. Fig. 18 shows the scaling of  $\langle 1/\tau(\delta v, r) \rangle_t$  as a function of  $\delta v$  in the GOY model, where  $\tau(\delta v, r)$  is the “doubling time”, i.e. the time necessary for a perturbation of size  $\delta v$  to increase by a factor  $r$  (see Section 3.4 and Appendix A).

For comparison, we also plot the eddy turn-over times  $\tau_\ell^{-1} = \langle |\delta v_\ell|^2 \rangle^{1/2}/\ell$ . We see that below the Kolmogorov scale, the doubling time displays a constant plateau corresponding to the Lyapunov exponent (3.39). At larger errors we observe a good agreement with the prediction (5.17).

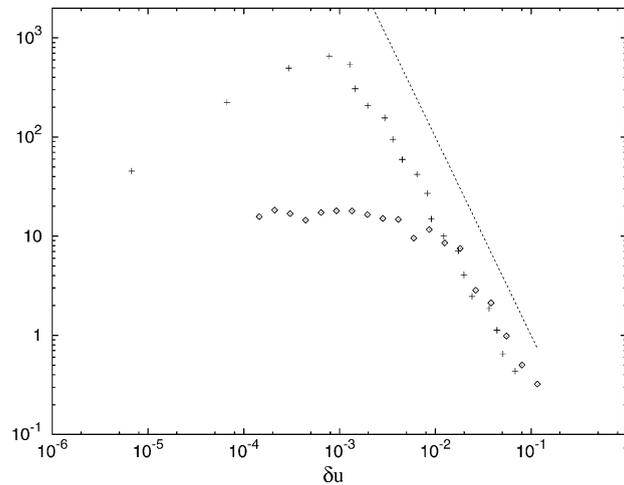


Fig. 18. The inverse of the error doubling times versus  $\delta u$  (diamond) compared with shell turn-over times (plus). Number  $N$  of simulated shells is 27, and Reynolds number  $Re = \nu^{-1} = 10^9$ ,  $k_0 = 0.05$ . The initial perturbation is randomly uniform over all shells in the inertial range, with amplitude of order  $10^{-6}$ . The first threshold is  $\delta_0 = 10^{-4}$  and the error growth rate parameter  $r$  is  $2^{1/2}$ . The number of error doubling experiments is 400. The dashed line has the slope  $-2$ .

Let us observe that, even at this high Reynolds number, the scaling range for the doubling time is rather small.

It is interesting to look at the doubling time as a function of the Reynolds number. For small thresholds the inverse of the doubling time scales as the Lyapunov exponent, i.e. roughly as  $Re^{-1/2}$ . We also observe that the bend away from the infinitesimal growth rate occurs at smaller scales for larger Reynolds numbers. This suggests the following scaling ansatz: times and errors are scaled with the turn-over time and the typical scale of fluctuations at the Kolmogorov scale, that is by  $Re^{-1/2}$  and  $Re^{-1/4}$ , respectively. In Fig. 19 we show the re-scaled data. The data collapse is reasonable, allowing to conclude that small-scale predictability, with small error amplitudes, behaves (apart from intermittency corrections) as predicted by Ruelle [196], whereas large-scale predictability, characterized by large error amplitudes, is well described by Lorenz arguments.

To improve the data collapse, taking into account the multifractal correction as described in Appendix B, one has to make a multiscaling collapse, i.e. to rescale  $\ln\langle 1/\tau(\delta v, r) \rangle$  and  $\ln(\delta v/V_0)$  with  $\ln(Re/Re_0)$  where  $V_0$  and  $Re_0$  are two parameters to be fixed [12]. The result is shown in Fig. 20. The data collapse is clearly improved.

Finite size predictability has been investigated also in two-dimensional turbulence, which is relevant for atmospheric flows. As discussed in Section 4.8, two-dimensional turbulence in the inverse energy cascade regime is characterized by a scaling à la Kolmogorov [132] with no intermittency [34]. As discussed above, the scaling exponent in (5.17) is not affected by intermittency; however intermittency does reduce the scaling range because of the intermediate dissipative range (see Appendix B). The absence of intermittency corrections in 2D turbulence

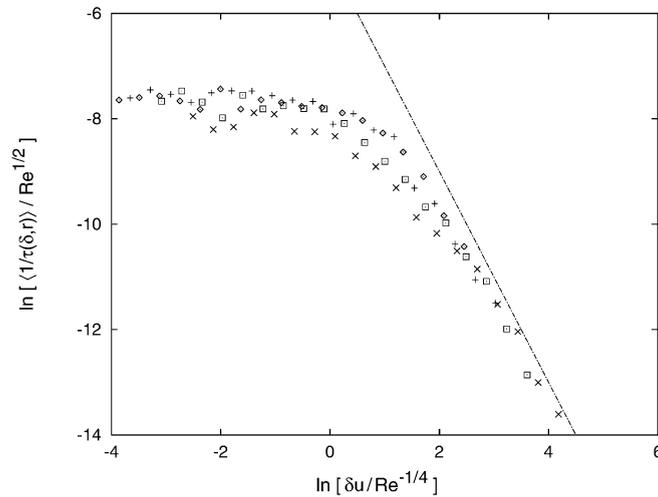


Fig. 19.  $\ln[(1/\tau(\delta, r))/Re^{1/2}]$  versus  $\ln[\delta u/Re^{-1/4}]$  at different Reynolds numbers  $Re=v^{-1}$ . ( $\diamond$ )  $N=24$  and  $Re=10^8$ ; ( $+$ )  $N=27$  and  $Re=10^9$ ; ( $\square$ )  $N=32$  and  $Re=10^{10}$ ; ( $\times$ )  $N=35$  and  $Re=10^{11}$ . The straight line has slope  $-2$ .

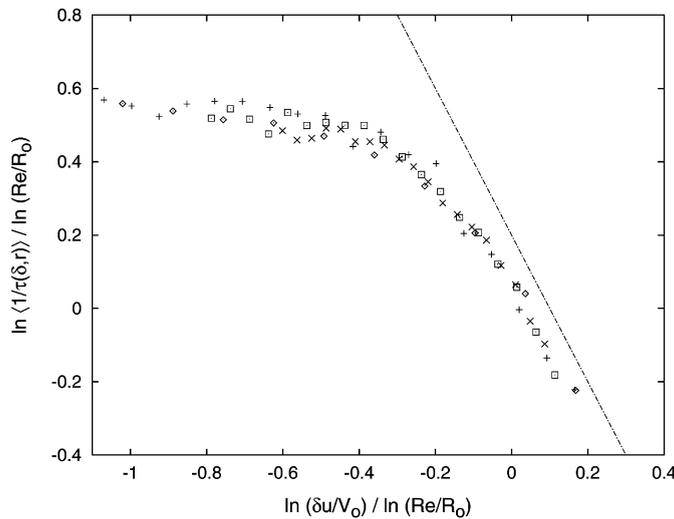


Fig. 20.  $\ln(1/\tau(\delta, r))/\ln(Re/R_o)$  versus  $\ln(\delta u/V_o)/\ln(Re/R_o)$ ; multiscaling data collapse at different Reynolds numbers  $Re=v^{-1}$ . The fitting parameters are  $R_o=6 \times 10^6$ ,  $V_o=5 \times 10^{-2}$ , and  $Re=v^{-1}$ .

suggests that the dimensional scaling (5.17) is observable even in direct numerical simulations at moderate Reynolds number.

Let us consider two realizations of the vorticity field in (4.27) starting from very close initial conditions. The error  $\delta$  is defined, following (4.30), as  $\delta(t) = \sqrt{E_\delta(t)}$ . In Fig. 21 it is shown the FSLE  $\lambda(\delta)$ . It is remarkable the rather wide scaling range for  $\lambda(\delta) \sim \delta^{-2}$  with respect to the shell model simulations (Fig. 18) obtained at much larger  $Re$ . As a consequence of the absence

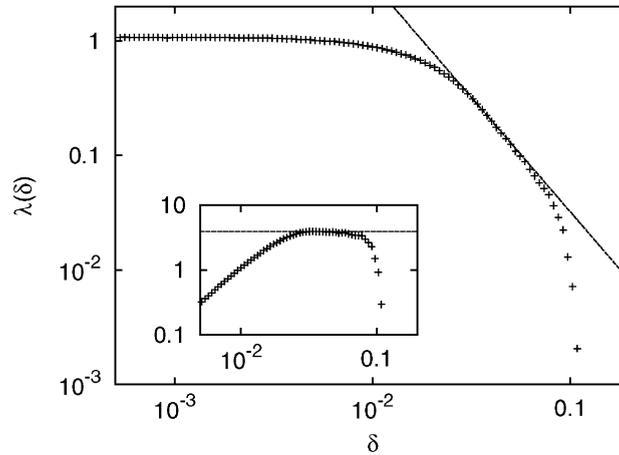


Fig. 21. Finite size Lyapunov exponent  $\lambda(\delta)$  as a function of velocity uncertainty  $\delta$  in a direct numerical simulations with  $1024^2$  grid points of 2D turbulence in the inverse cascade regime. The asymptotic constant value for  $\delta \rightarrow 0$  is the largest Lyapunov exponent of the turbulent flow. The dashed line has slope  $-2$ . In the inset we show the compensated plot  $\lambda(\delta)\delta^2/\bar{\epsilon}$ .

of intermittency, also the crossover from the infinitesimal regime  $\lambda(\delta) = \lambda$  to the inertial range regime (5.17) is sharp.

From a general point of view, it is interesting to observe that even in the absence of intermittency, fixed scale analysis based on the FSLE overpasses fixed time analysis in the characterization of predictability. Dimensional considerations and closure approximations [137] predict a linear growth of the error in the inverse energy cascade as

$$E_\delta(t) = G\bar{\epsilon}t, \quad (5.20)$$

where  $G$  is an adimensional constant. It is easy to realize that (5.20) is equivalent to (5.17),  $\lambda(\delta)$  having the dimension of an inverse time and  $\delta = \sqrt{E_\delta}$ . The result obtained in numerical simulations is shown in Fig. 22, which has to be compared with Fig. 21. The scaling law (5.20) in Fig. 22 is barely visible, making the determination of  $G$  difficult. On the contrary, inverting (5.17) to (5.20) one can measure  $G$  directly from Fig. 21. The result obtained is in close agreement with closure computations [37].

### 5.5. $\epsilon$ -entropy for turbulent flows

A complementary way to look at the predictability of turbulent flows is in terms of its entropy (see Sections 2.1.2 and 3.5).

Unfortunately, a direct measurement of the Kolmogorov–Sinai entropy is practically infeasible. Indeed for  $Re \rightarrow \infty$  due to the huge number of active degrees of freedom, the KS-entropy diverges, so that one needs velocity measurements with an extremely high resolution and lasting for extremely long times, far beyond the actual experimental possibilities. Nevertheless, limiting the analysis to not very high resolution, one can hope to extract some interesting piece of

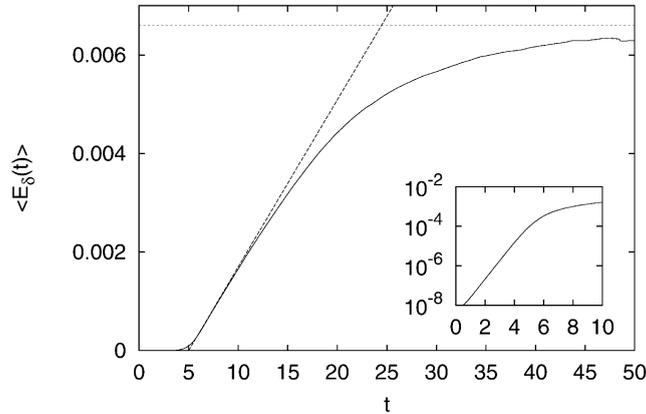


Fig. 22. Average energy error  $\langle E_\delta(t) \rangle$  growth. Dashed line represents linear closure prediction, dotted line is the saturation value  $E$ . The initial exponential growth is emphasized by the lin-log plot in the inset.

information by investigating the behavior of the  $\epsilon$ -entropy,  $h(\epsilon)$ . As far as the  $\epsilon$ -entropy of turbulence is concerned, two questions can be raised.

- (i) Since a direct measurement of the full 3-dimensional velocity field is infeasible, one has usually access just to a time signal measured in one spatial point: which kind of information can we extract from the  $\epsilon$ -entropy per unit time of such a signal?
- (ii) Taking into account (i), can we say something about the  $\epsilon$ -entropy of the full 3-dimensional velocity field?

In (ii) we are referring to the  $\epsilon$ -entropy,  $h^{\text{ST}}(\epsilon)$ , per unit time and volume (the symbol ST means space–time). In other words, we are assuming that the total entropy of a turbulent flow observed for a (very long) time  $T$  on a (very large) volume  $V$  of the 3-dimensional space has the form  $H(V, T, \epsilon) \approx VTh^{\text{ST}}(\epsilon)$ . See Ref. [89] for an introduction of this concept.

Both in (i) and (ii), as we will see, a crucial role is played by the sweeping of the large scales of the flow on the small ones, i.e. the Taylor hypothesis (see Section 5.1).

### 5.5.1. $\epsilon$ -entropy for a time signal of turbulence

In order to estimate the  $\epsilon$ -entropy of a given signal one has to compute the Shannon entropy of the symbolic sequence obtained by making an  $(\epsilon, \tau)$  grid in phase-space (Section 3.5). Unfortunately, this method is rather inefficient for signals in which many scales are excited [3,4,51], e.g., as in turbulence. Therefore, here we resort to a recently proposed method [3] based on the *exit-time* analysis.

In a few words, the idea consists in looking at a sequence of data not at fixed sampling times but at fixed fluctuation (see Appendix C), i.e. when the fluctuation of the signal exceeds a given threshold,  $\epsilon$ . In practice, we code the signal  $v(t)$  of total duration  $T$  in a symbolic sequence  $\Omega^M(\epsilon) = \{t_i(\epsilon), k_i(\epsilon)\}_{i=1}^M$ , where  $t_i(\epsilon)$  is the first time such that  $|v(t_0 + \sum_{k=1}^{i-1} t_k(\epsilon) + t_i(\epsilon)) - v(t_0 + \sum_{k=1}^{i-1} t_k(\epsilon))| \geq \epsilon/2$  (being  $t_0$  a reference time) and  $k_i = \pm 1$  tells us in which direction (up or down with respect to  $v(t_0 + \sum_{k=1}^{i-1} t_k(\epsilon))$ ) the fluctuation has been realized.  $M$  is

the total number of exit events, i.e.  $\sum_{i=1}^M t_i(\epsilon) = T$ . Note that  $\Omega^M(\epsilon)$  is a faithful coding of the signal within the required accuracy  $\epsilon$ . Now the evaluation of the entropy goes as usual through the evaluation of the Shannon entropy,  $h^\Omega(\epsilon)$ , of the sequence  $\Omega^M(\epsilon)$ . Finally, the  $\epsilon$ -entropy per unit time is given by [3]

$$h(\epsilon) \approx \frac{h^\Omega(\epsilon, \tau_r)}{\langle t(\epsilon) \rangle}, \quad (5.21)$$

where a coarse-graining of the possible values assumed by  $t(\epsilon)$  with a resolution time  $\tau_r$  has been considered, and  $\langle t(\epsilon) \rangle$  is the average exit time, i.e.  $\langle t(\epsilon) \rangle = (1/M) \sum_{i=1, M} t_i(\epsilon)$ . Formula (5.22) is exact in the limit  $\tau_r \rightarrow 0$  (in Appendix C one finds the derivation of (5.21) and the details of the method).

This procedure allows a noticeable improvement of the computational possibility to measure the  $\epsilon$ -entropy. In particular, if one is interested in the leading scaling behavior of  $h(\epsilon)$  with  $\epsilon$ , one only needs to estimate the scaling of  $\langle t(\epsilon) \rangle$ . Indeed, the correction induced by  $h^\Omega(\epsilon, \tau_r)$  can be shown to be sub-leading (in particular, logarithmic).

Now, we estimate the average exit time for the velocity signal  $v(t)$ . This can be done assuming the Taylor hypothesis and the multifractal model (see Appendix B). In this framework we can assume that, for  $t$  corresponding to scales  $R = Ut$  in the inertial range, the following relation holds  $|\delta_t v| = |v(t_0 + t) - v(t)| \sim t^h$  and each  $h$  is picked with probability  $P(h) \sim t^{3-D(h)}$ . Since we are interested in the statistics of the first times necessary to observe a fluctuation  $|\delta_t v| \sim \epsilon$ , one can “invert” the above relation [30]:

$$t(\epsilon) \sim \epsilon^{1/h} \quad \text{with } P(h) \sim \epsilon^{(3-D(h))/h}. \quad (5.22)$$

The exit-time moments [30], also called inverse structure functions [107], can be estimated in the multifractal framework as follows

$$\langle\langle t^q(\epsilon) \rangle\rangle \sim \int dh \epsilon^{(q+3-D(h))/h} \sim \epsilon^{\chi(q)}, \quad (5.23)$$

where  $\chi(q)$  may be obtained with a saddle-point estimate in the limit of small  $\epsilon$ :

$$\chi(q) = \min_h \left[ \frac{q+3-D(h)}{h} \right]. \quad (5.24)$$

The average  $\langle[\dots]\rangle$ , obtained by counting the number of exit-time events  $M$ , and the average  $\langle\langle[\dots]\rangle\rangle$ , computed with the uniform time sampling are connected by the relation

$$\langle\langle t^q(\epsilon) \rangle\rangle = \lim_{M \rightarrow \infty} \sum_{i=1}^M t_i^q \frac{t_i}{\sum_{j=1}^M t_j} = \frac{\langle t^{q+1}(\epsilon) \rangle}{\langle t(\epsilon) \rangle}, \quad (5.25)$$

where the term  $t_i / \sum_{j=1}^M t_j$  takes into account the non-uniformity of the exit-time statistics. Therefore the quantity we are looking for, i.e. the mean exit-time, is given by  $\langle t(\epsilon) \rangle = \langle\langle t^{-1}(\epsilon) \rangle\rangle^{-1} \sim (\epsilon)^{-\chi(-1)}$ . By noting that

$$\frac{-1+3-D(h)}{h} = \frac{2-D(h)}{h} \geq -3 \quad \text{for all } h, \quad (5.26)$$

which is nothing but Eq. (5.19), i.e. the  $\frac{4}{5}$  law of turbulence, we finally obtain

$$h(\epsilon) \sim \epsilon^{-3}. \quad (5.27)$$

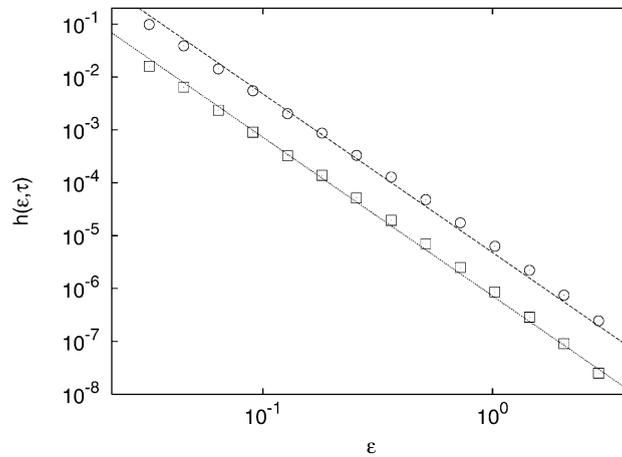


Fig. 23. Numerically computed lower bound ( $\square$ ) and upper bound ( $\circ$ ), with  $\tau = 0.1 \langle t(\epsilon) \rangle$  for the  $(\epsilon, \tau)$ -entropy in the case of a multiaffine signal with  $\zeta(3) = 1$ . The signal has been obtained with the method of Ref. [29] (see also Appendix D) using a  $D(h)$  which fits experimental data at large Reynolds number. The two straight lines show the theoretical scaling  $\epsilon^{-3}$ .

In Fig. 23 we report the evaluation of the upper and lower bounds (see Appendix C) of  $h(\epsilon)$  for a synthetic signal,  $v(t)$ , constructed in such a way as to reproduce the statistical properties of turbulence [29].

Let us now compare the above results with a previous study of the  $\epsilon$ -entropy in turbulence [224], where it was argued that

$$h(\epsilon) \sim \epsilon^{-2}, \tag{5.28}$$

a behavior that differs from the prediction (5.27). The behavior (5.28) has been obtained by assuming that  $h(\epsilon)$ , at scale  $\epsilon$ , is proportional to the inverse of the typical eddy turnover time at that scale: since the typical eddy turnover time for velocity fluctuations of order  $\delta v \sim \epsilon$  is  $\tau(\epsilon) \sim \epsilon^2$ , Eq. (5.28) follows. Indeed this is the argument used to derive (5.17) for the FSLE. The difference between (5.28) and (5.27) can be understood by considering that even if  $\lambda(\delta)$  and  $h(\epsilon)$  are two complementary concepts (the fact that for both the estimate of the scaling behavior reduces to the “4/5 law” is not a coincidence), in the latter case one has to consider the sweeping induced by the large scales. On the contrary, since the former is related to the distance of two realizations which differ in the small scales ( $< \delta$ ) but not on the large scales ( $> \delta$ ), the sweeping of the large scales is not effective.

### 5.5.2. $\epsilon$ -entropy of turbulence and the Taylor hypothesis

Now we study the  $\epsilon$ -entropy per unit time and volume for the velocity field of turbulent flows in  $3 + 1$  dimensions,  $h^{ST}(\epsilon)$ . We will show that, by assuming the usually accepted Taylor hypothesis, one has a spatial correlation which can be quantitatively characterized by an “entropy” dimension  $\mathcal{D} = \frac{8}{3}$ . As already remarked,  $h^{ST}(\epsilon)$  cannot be directly measured so we will discuss its estimation in a theoretical framework by introducing a multi-affine field. For the

sake of simplicity, we neglect intermittency by assuming a pure self-affine field with a unique Hölder exponent  $h = \frac{1}{3}$ .

Let us first introduce a multi-affine field with the proper spatial and temporal scaling [4]. The idea consists in defining the signal as a dyadic 3-dimensional superposition of wavelet-like functions  $\varphi((\mathbf{x} - \mathbf{x}_{n,k}(t))/\ell_n)$  whose centers move due to the sweeping. The coefficients of the decomposition  $a_{n,k}(t)$  are stochastic functions chosen with suitable self-affine scaling properties both in time and in space. A field with spatial Hölder exponent  $h$  in  $d$ -dimensions is (see Appendix D)

$$v(\mathbf{x}, t) = \sum_{n=1}^M \sum_{k=1}^{2^{d(n-1)}} a_{n,k}(t) \varphi\left(\frac{\mathbf{x} - \mathbf{x}_{n,k}(t)}{\ell_n}\right), \tag{5.29}$$

where  $\mathbf{x}_{n,k}$  is the center of the  $k$ th wavelets at the level  $n$ , i.e. for eddies with size  $\ell_n \sim 2^{-n}$ . According to the Richardson–Kolmogorov cascade picture, one assumes that sweeping is present, i.e.,  $\mathbf{x}_{n+1,k} = \mathbf{x}_{n,k'} + \mathbf{r}_{n+1,k}$  where  $(n, k')$  labels the “mother” of the  $(n + 1, k)$ -eddy and  $\mathbf{r}_{n+1,k}$  is a stochastic vector which depends on  $\mathbf{r}_{n,k'}$  and evolves with characteristic time  $\tau_n \propto (\ell_n)^{1-h}$ .

If the coefficients  $\{a_{n,k}\}$  and  $\{\mathbf{r}_{n,k}\}$  have characteristic time  $\tau_n \sim (\ell_n)^{1-h}$  and  $\{a_{n,k}\} \sim (\ell_n)^h$ , it is possible to show (see Appendix D for details) that the field (5.29) has the correct spatio-temporal statistics, i.e.

$$|v(\mathbf{x} + \mathbf{R}, t_0) - v(\mathbf{x}, t_0)| \sim |\mathbf{R}|^h, \tag{5.30}$$

$$|v(\mathbf{x}, t_0 + t) - v(\mathbf{x}, t_0)| \sim t^h. \tag{5.31}$$

In addition the proper Lagrangian sweeping is satisfied. Now we are ready for the  $\epsilon$ -entropy analysis of the field (5.29). If one wants to look at the field  $v$  with a resolution  $\epsilon$ , one has to take  $n$  in (5.29) up to  $N$  given by

$$(\ell_N)^h \sim \epsilon, \tag{5.32}$$

in this way one is sure to consider velocity fluctuations of order  $\epsilon$ . Then the number of terms contributing to (5.29) is

$$\#(\epsilon) \sim (2^d)^N \sim \epsilon^{-d/h}. \tag{5.33}$$

By using a result of Shannon [201] one estimates the  $\epsilon$ -entropy of the single process  $a_{n,k}(t)$  (and also of  $\mathbf{r}_{n,j}$ ) as

$$h_n(\epsilon) \sim \frac{1}{\tau_n} \ln\left(\frac{1}{\epsilon}\right), \tag{5.34}$$

where the above relation is rigorous if the processes  $a_{n,k}(t)$  are Gaussian and with a power spectrum different from zero on a band of frequency  $\sim 1/\tau_n$ . The terms which give the main contribution are those with  $n \sim N$  with  $\tau_N \sim (\ell_N)^{1-h} \sim \epsilon^{((1-h)/h)}$ . Their number is given by (5.33) so that, collecting the above results, one finds

$$h^{\text{ST}}(\epsilon) \sim \frac{\#(\epsilon)}{\tau_N} \sim \epsilon^{-(d-h+1)/h}. \tag{5.35}$$

For the physical case  $d = 3$ ,  $h = \frac{1}{3}$ , one obtains

$$h^{\text{ST}}(\epsilon) \sim \epsilon^{-11}. \tag{5.36}$$

By denoting with  $v_\eta$  the typical velocity at the Kolmogorov scale  $\eta$ , one has that Eq. (5.36) holds in the inertial range, i.e.,  $\epsilon \geq v_\eta \sim Re^{-1/4}$ , while for  $\epsilon \leq v_\eta$ ,  $h^{ST}(\epsilon) = \text{constant} \sim Re^{11/4}$ .

Let us now consider an alternative way to compute the  $\epsilon$ -entropy of the field  $v(\mathbf{x}, t)$ : divide the  $d$ -volume in boxes of edge length  $\ell(\epsilon) \sim \epsilon^{1/h}$  and look at the signals  $v(\mathbf{x}_\alpha, t)$ , where the  $\mathbf{x}_\alpha$  are the centers of the boxes. Denoting with  $h^{(x)}(\epsilon)$  the  $\epsilon$ -entropy of the temporal sequence of the velocity field measured in  $\mathbf{x}_\alpha$ , we have

$$h^{(x)}(\epsilon) \sim \epsilon^{-1/h} \tag{5.37}$$

because of the scaling (5.31). Therefore,  $h^{ST}(\epsilon)$  is obtained summing up all the “independent” contributions (5.37), i.e.

$$h^{ST}(\epsilon) \sim \mathcal{N}(\epsilon)h^{(x)}(\epsilon) \sim \mathcal{N}(\epsilon)\epsilon^{-1/h}, \tag{5.38}$$

where  $\mathcal{N}(\epsilon)$  is the number of independent cells. It is easy to understand that the simplest assumption  $\mathcal{N}(\epsilon) \sim l(\epsilon)^d \sim \epsilon^{d/h}$  gives a wrong result, indeed one obtains

$$h^{ST}(\epsilon) \sim \epsilon^{-(d+1)/h}, \tag{5.39}$$

which is not in agreement with (5.35). In order to obtain the correct result (5.36) it is necessary to assume

$$\mathcal{N}(\epsilon) \sim l(\epsilon)^{\mathcal{D}}, \tag{5.40}$$

with  $\mathcal{D} = d - h$ . In other words, one has that the sweeping implies a non-trivial spatial correlation, quantitatively measured by the exponent  $\mathcal{D}$ , which can be considered as a sort of “entropy” dimension. Incidentally, we note that  $\mathcal{D}$  has the same numerical value of the fractal dimensions of the velocity iso-surfaces [154,218]. From this observation, at first glance, one could conclude that the above result is somehow trivial since it is simply related to a geometrical fact. However, a closer inspection reveals that this is not true. Indeed, one can construct a self-affine field with spatial scaling  $h$  and thus with the fractal dimension of the velocity iso-surfaces given by  $d - h$  for geometrical reasons, while  $\mathcal{D} = d$ . Such a process can be simply obtained by eliminating the sweeping, i.e.,

$$v(\mathbf{x}, t) = \sum_{n=1}^M \sum_{k=1}^{2^{d(n-1)}} a_{n,k}(t) \varphi\left(\frac{\mathbf{x} - \mathbf{x}_{n,k}}{\ell_n}\right), \tag{5.41}$$

where now the  $\mathbf{x}_{n,k}$  are fixed and no longer time-dependent, while  $a_{n,k} \sim (\ell_n)^h$  but  $\tau_n \sim \ell_n$ .

We conclude by noting that it is possible to obtain (see [89]) the scaling (5.35) using Eq. (5.41), i.e. ignoring the sweeping, assuming  $\tau_n \sim (\ell_n)^{1-h}$  and  $a_{n,k} \sim (\ell_n)^h$ ; this corresponds to take separately the proper temporal and spatial spectra. However, this is not satisfactory since one has not the proper scaling in one fixed point (see Eq. (5.37) the only way to obtain this is through the sweeping).

## 6. Uncertainty in the evolution equations

The study of a large class of problems in science (physics, chemistry, biology, etc.) is reduced to the investigation of evolution laws, which describe some aspects of the system. The

assumption that natural processes can be described by mathematical models is at the foundation of this approach [220,70]. The purpose of this section is to discuss how the unavoidable uncertainty in the equation of motion puts limits on the long time forecasting.

To be more concrete, let us consider a system described by a differential equation

$$\frac{d}{dt} \mathbf{x}(t) = \mathbf{f}(\mathbf{x}, t), \quad \mathbf{x}, \mathbf{f} \in \mathbb{R}^n. \quad (6.1)$$

As a matter of fact, we do not know exactly the equations, so we have to devise a model which is different from the true dynamics. In practice, this means that we study

$$\frac{d}{dt} \mathbf{x}(t) = \mathbf{f}_\epsilon(\mathbf{x}, t) \quad \text{where } \mathbf{f}_\epsilon(\mathbf{x}, t) = \mathbf{f}(\mathbf{x}, t) + \epsilon \delta \mathbf{f}(\mathbf{x}, t). \quad (6.2)$$

Therefore, it is natural to wonder about the relation between the true evolution (*reference* or *true* trajectory  $\mathbf{x}_T(t)$ ) given by (6.1) and the one effectively computed (*perturbed* or *model* trajectory  $\mathbf{x}_M(t)$ ) given by (6.2). A typical example is the relation between the true dynamics of the physical system and the one obtained by a computer simulation. This issue is of particular relevance for the study of weather forecast where it is referred to as *predictability of the second kind* [175].

In this context it is particularly relevant the *shadowing lemma* [41] which implies that, for Anosov systems, a computer may not calculate the true orbit but what it does find is nevertheless an approximation of the true one. As a consequence, the statistical properties are well reproduced by an accurate numerical integration [96].

A central point in the discussion of the second kind predictability problem is the issue of *structural stability* [99]: since the evolution laws are known only with finite precision it is highly desirable that at least certain properties were not too sensitive to the details of the equations of motion. For example, in a system with a strange attractor, small generic changes in the evolution laws should not change drastically the statistical properties of the dynamics [74,105].

In order to see that a non-generic perturbation, although very “small” in some sense, can produce dramatic changes in the statistical properties of the dynamics, following Refs. [28,105], we consider the 1-dimensional chaotic map  $x(t+1) = f(x(t))$  with  $f(x) = 4x \bmod 1$ , and a perturbed version of it:

$$f_\epsilon(x) = \begin{cases} 8x - \frac{9}{2}, & x \in \left[ \frac{5}{8}, \frac{247}{384} \right], \\ \frac{1}{2}x + \frac{1}{3}, & x \in \left[ \frac{247}{384}, \frac{265}{384} \right], \\ 8x - \frac{29}{6}, & x \in \left[ \frac{265}{384}, \frac{17}{24} \right], \\ 4x \bmod 1, & \text{otherwise.} \end{cases} \quad (6.3)$$

The perturbed map is identical to the original outside the interval  $\left[ \frac{5}{8}, \frac{17}{24} \right]$ , and the perturbation is small in  $L_2$  norm. Nevertheless, the fixed point  $x = 2/3$ , which is unstable in the original dynamics, becomes stable in the perturbed one, and it is a *global attractor* for  $f_\epsilon(x)$ , i.e. almost every point in  $[0, 1]$  asymptotically approaches  $x = 2/3$  (see Fig. 24).

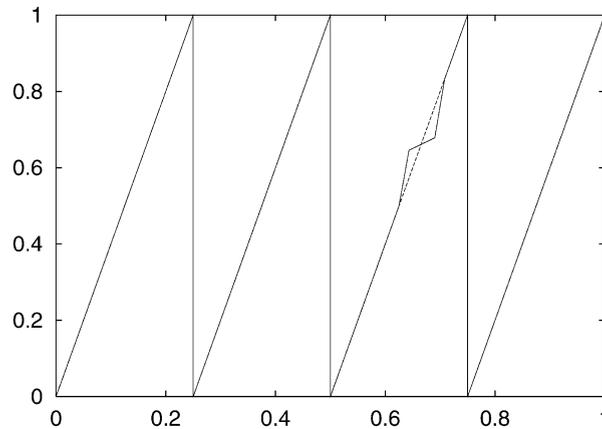


Fig. 24. The map  $f_\epsilon$  of Eq. (6.3) (solid line) and the original chaotic map  $f$  (dashed line).

If one compares the trajectories obtained iterating  $f(x)$  or  $f_\epsilon(x)$  it is not difficult to understand that they may remain identical for a certain time but unavoidably differ utterly in the long time behavior. The transient chaotic behavior of the perturbed orbits can be rendered arbitrarily long by reducing the interval in which the two dynamics differ [28].

As for the problem of predictability with respect to perturbations on the initial conditions, the problem of second kind predictability in the limit of infinitesimal perturbations is essentially understood in terms of the Lyapunov exponents. Indeed, it is possible to show (see below) that a small uncertainty on the evolution laws of chaotic systems has the same effects of an error of the same order of magnitude on the initial conditions. However, also in the case of second kind predictability one has often to deal with errors which are far from being infinitesimal. Moreover, in real systems the size of an uncertainty on the evolution equations is determinable only a posteriori, based on the ability of the model to reproduce some of the features of the phenomenon. Typical examples are systems described by partial differential equations (e.g. turbulence, atmospheric flows). The numerical study of these systems is performed by using a model with unavoidable severe approximations, the most relevant due to the necessity to cut some degrees of freedom off (i.e. the small scale variables). A relevant problem in this case is to quantify the effect of the unresolved scales on the predictability of the resolved ones.

From a general point of view, in the second kind predictability problem we can distinguish three main cases depending on the original dynamics. In particular, Eq. (6.1) may display:

- (i) trivial attractors: asymptotically stable fixed points or attracting periodic orbits;
- (ii) marginally stable fixed points or periodic/quasi-periodic orbits as in integrable Hamiltonian systems;
- (iii) chaotic behavior.

In case (i) small changes in the equations of motion do not modify the qualitative features of the dynamics. Case (ii) is not generic and the outcome strongly depends on the specific perturbation  $\delta\mathbf{f}$ , i.e. it is not structurally stable (see [64] for a discussion on this point). In the

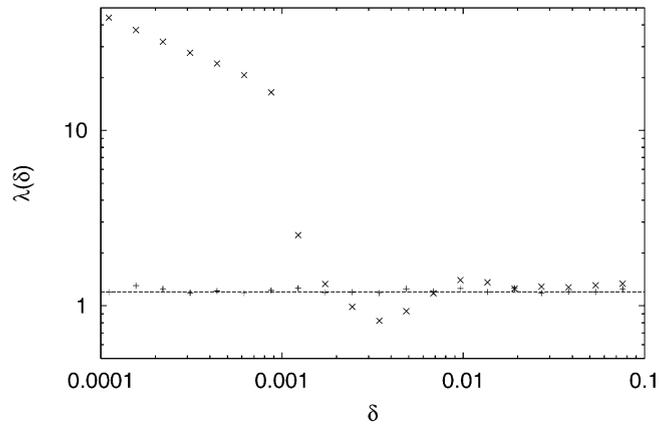


Fig. 25. Finite size Lyapunov exponents  $\lambda_{TT}(\delta)(+)$  and  $\lambda_{TM}(\delta)(\times)$  versus  $\delta$  for the Lorenz model (3.29) with  $\sigma = c = 10$ ,  $b = 83$ ,  $r = 45$  and  $\epsilon = 0.001$ . The dashed line represents the largest Lyapunov exponent for the unperturbed system ( $\lambda_T \approx 1.2$ ). The statistics is over  $10^4$  realizations.

chaotic case (iii) one expects that the perturbed dynamics is still chaotic. In the following we will consider only this latter case.

In chaotic systems, the effects of a small uncertainty on the evolution law is, for many aspects, similar to those due to imperfect knowledge of initial conditions. As an example let us consider the Lorenz system (3.29). In order to mimic an indetermination in the evolution law we assume a small error  $\epsilon$  on the parameter  $r: r \rightarrow r + \epsilon$ . Let us consider the difference  $\delta\mathbf{x}(t) = \mathbf{x}_M(t) - \mathbf{x}_T(t)$ , for simplicity,  $\delta\mathbf{x}(0) = 0$ , i.e. we assume a perfect knowledge of the initial condition. For small  $\epsilon$  one has, with obvious notation:

$$\frac{d\delta\mathbf{x}}{dt} = \mathbf{f}_\epsilon(\mathbf{x}_M) - \mathbf{f}(\mathbf{x}_T) \simeq \frac{\partial\mathbf{f}}{\partial\mathbf{x}} \delta\mathbf{x} + \frac{\partial\mathbf{f}_\epsilon}{\partial r} \epsilon. \quad (6.4)$$

Since at time  $t = 0$  one has  $|\delta\mathbf{x}(0)| = 0$ ,  $|\delta\mathbf{x}(t)|$  initially grows under the effect of the second term in (6.4). At later times, when  $|\delta\mathbf{x}(t)| \approx O(\epsilon)$  the first term becomes the leading one, and we recover the first kind predictability for an initial uncertainty  $\delta_0 \sim \epsilon$ . Therefore, apart from an initial growth, which depends on the specific perturbation, for small enough  $\epsilon$  the evolution of  $\langle \ln(|\delta\mathbf{x}(t)|) \rangle$  follows the usual linear growth with the slope given by the largest LE. Typically the value of the LE computed by using the model dynamics differs from the true one by a small amount of order  $\epsilon$ , i.e.  $\lambda_M = \lambda_T + O(\epsilon)$  [64].

A picture of the error growth, valid also for finite errors, can be obtained by considering the finite size Lyapunov exponent. In addition to the FSLE of the model,  $\lambda_{MM}(\delta)$ , we introduce the FSLE for the true dynamics (6.1)  $\lambda_{TT}(\delta)$  (which cannot be measured in real situations) and  $\lambda_{TM}(\delta)$ , the FSLE computed following the distance between one true trajectory and one model trajectory starting at the same point. In the case of a perfect model  $\lambda_{MM}(\delta) = \lambda_{TT}(\delta)$ . The results of the computation for the Lorenz model (3.29) are shown in Fig. 25.  $\lambda_{TT}(\delta)$  displays the chaotic plateau with  $\lambda \simeq 1.2$ . As discussed above, for  $\delta > \epsilon$  the second term in (6.4) becomes negligible and we observe  $\lambda_{TM}(\delta) \simeq \lambda_{TT}(\delta) \simeq \lambda$ . In this range of errors the model system

recovers the intrinsic predictability of the true system. For very small errors,  $\lambda_{\text{TM}}$  is dominated by the second term in (6.4) and deviates from  $\lambda_{\text{TT}}$ .

### 6.1. Uncertainty introduced by numerical computations

In numerical computations, an unavoidable source of errors is due to the representation of numbers on the computer, as computers work with integers. This has two main consequences: the phase space of the simulated system is necessarily discrete (and finite); and the computation introduces a sort of noise due to the round-off.

A direct consequence of the discreteness in phase space is that any numerical trajectory is periodic. At first sight, this seems a very serious problem, especially when integrating chaotic systems which have non-periodic behavior. However, as discussed in [64], apart from cases in which one uses very low precision, and very low dimensional systems, the period is usually extremely large and one works with an effective continuous phase space dynamical system (see Section 7.1).

The round-off produces on (6.1) and (6.2) a perturbation  $\delta\mathbf{f}(\mathbf{x}, t)$  of order  $\epsilon \sim 10^{-\alpha}$  ( $\alpha$ =number of digits in floating point representation) which depends on  $\mathbf{f}$  and on the software [131]. In general, the round-off error is very small and may have a positive role in selecting the physical probability measure, the so-called *natural measure*, from the set of the admissible invariant ones [74].

In order to show the effect of the numerical precision on the predictability, let us consider again the Lorenz model (3.29). At variance with the previous Section, here we assume to have a perfect knowledge of the model (i.e. of the parameter  $r$ ), and the error is introduced only by the numerical integration, e.g. by different time step  $\Delta t$ . The most precise integration with smallest  $\Delta t$  is taken as the reference trajectory and the other is the perturbed one. The result is shown in Fig. 26 for two different values of  $\Delta t$  for the perturbed integration. In both cases, for small values of the error, the exponential growth rate is given by the largest LE  $\lambda$ . The same behavior is observed by introducing the numerical error in other ways, e.g. by using different precision (single or double) or different integration algorithms [64].

### 6.2. Finite resolution effects and parameterization of unresolved scales

Let us now consider more complex situations, in which many interacting degrees of freedom and different characteristic times are involved [36]. We will consider the particular examples of an extremely simplified model of global circulation [147,148] and the shell model (Section 5.2).

For systems with many different scales usually one is able to represent only the large scale variables. A typical situation is the discretization of partial differential equations. The small scale modes, below the computational grid, are unresolved and are typically parameterized according to some phenomenological prescription (e.g. the eddy viscosity parameterization of the small scales [141,84]). So we consider systems of the following form

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, \mathbf{y}), \quad \frac{d\mathbf{y}}{dt} = \mathbf{g}(\mathbf{x}, \mathbf{y}), \quad (6.5)$$

where  $\mathbf{x} \in \mathbb{R}^n$  represent the large (and typically slow) variables while  $\mathbf{y} \in \mathbb{R}^m$  represent the small (and fast) ones. As explained above, in many practical situations the small variables cannot

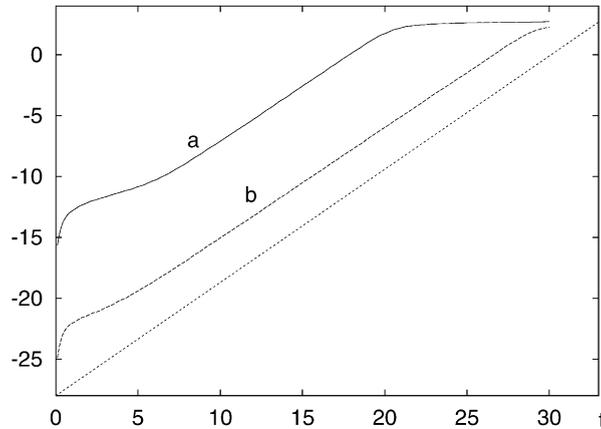


Fig. 26.  $\langle \ln |\delta \mathbf{x}(t)| \rangle$  versus  $t$ , where  $|\delta \mathbf{x}(t)|$  is the Euclidean distance between two trajectories of the Lorenz model (3.29) for  $r=28$ . Curve (a) refers to the comparison between trajectories obtained using a fourth-order Runge–Kutta algorithm with  $\Delta t = 4 \times 10^{-3}$  and  $\Delta t = 4 \times 10^{-5}$ . Curve (b) shows the same quantity obtained with  $\Delta t = 4 \times 10^{-4}$  and  $\Delta t = 4 \times 10^{-5}$ . The dotted line with slope  $\lambda \approx 0.9$  is shown for comparison.

be explicitly resolved. In this framework, a natural question is: how must we parameterize the unresolved modes in order to predict the resolved ones? In this respect, the optimal parameterization is that one for which the predictability on the resolved modes is not worse than the intrinsic predictability of the same variables in the complete system, i.e. in our notation  $\lambda_{\text{TM}} = \lambda_{\text{TT}}$ .

An example in which it is relatively simple to develop a model for the small-scale modes is represented by skew systems, i.e.,  $\mathbf{g}$  depends only on the fast variables  $\mathbf{y}$ . In this case, simply neglecting the fast variables or parameterizing them with a suitable stochastic process does not drastically affect the prediction of the slow variables [32].

On the other hand, in typical cases  $\mathbf{y}$  feels some feedback from  $\mathbf{x}$ , and, therefore, we cannot simply neglect the unresolved modes (see Ref. [36] for details). In practice, one has to construct an effective equation for the resolved variables:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}_M(\mathbf{x}, \mathbf{y}(\mathbf{x})), \tag{6.6}$$

where the functional form of  $\mathbf{y}(\mathbf{x})$  and  $\mathbf{f}_M$  is built by phenomenological arguments and/or by numerical studies of the full dynamics (if available).

Let us now discuss a simplified model for atmosphere circulation [147,148] which includes large scales  $x_k$  (synoptic scales) and small scales  $y_{j,k}$  (convective scales):

$$\begin{aligned} \frac{dx_k}{dt} &= -x_{k-1}(x_{k-2} - x_{k+1}) - vx_k + F - \sum_{j=1}^J y_{j,k}, \\ \frac{dy_{j,k}}{dt} &= -cb y_{j+1,k}(y_{j+2,k} - y_{j-1,k}) - cv y_{j,k} + x_k, \end{aligned} \tag{6.7}$$

where  $k = 1, \dots, K$  and  $j = 1, \dots, J$ . As in [147] we assume periodic boundary conditions on  $k$  ( $x_{K+k} = x_k$ ,  $y_{j,K+k} = y_{j,k}$ ) while for  $j$  we impose  $y_{j+J,k} = y_{j,k+1}$ . The variables  $x_k$  represent some

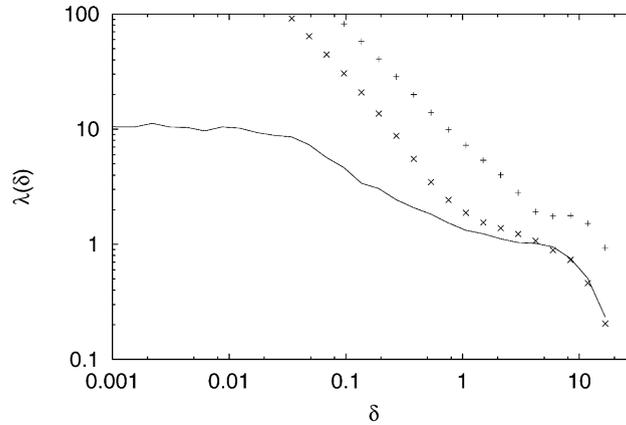


Fig. 27. Finite size Lyapunov exponents for the Lorenz model (6.7)  $\lambda_{TT}(\delta)$  (solid line) and  $\lambda_{TM}(\delta)$  versus  $\delta$  obtained by dropping the fast modes (+) and with eddy viscosity parameterization (x). The parameters are  $F = 10, K = 36, J = 10, \nu = 1$  and  $c = b = 10$ , implying that the typical  $y$  variable is 10 times faster and smaller than the  $x$  variable. The value of the parameter  $\nu_e = 4$  is chosen after a numerical integration of the complete equations as discussed in Ref. [36]. The statistics is over  $10^4$  realizations.

large scale atmospheric quantities in  $K$  sectors extending on a latitude circle, while the  $y_{j,k}$  represent quantities on smaller scales in  $J \cdot K$  sectors. The parameter  $c$  is the ratio between fast and slow characteristic times and  $b$  measures the relative amplitude (both larger than unity). Model (6.7), even if rather crude, contains some nontrivial aspects of the general circulation problem, namely the coupling among variables with very different characteristic times.

Being interested in forecasting the large scale behavior of the atmosphere by using only the slow variables, a natural choice for the model equations is:

$$\frac{dx_k}{dt} = -x_{k-1}(x_{k-2} - x_{k+1}) - \nu x_k + F - G_k(\mathbf{x}), \tag{6.8}$$

where  $G_k(\mathbf{x})$  represents the parameterization of the fast components in (6.7). Following the approach discussed in Ref. [36], a physical reasonable parameterization is

$$G_k(\mathbf{x}) = \nu_e x_k, \tag{6.9}$$

where  $\nu_e$  is a numerically determined parameter.

In Fig. 27 we plot  $\lambda_{TM}(\delta)$  obtained from different choices of  $G_k$ . The simplest possibility is to neglect the fast variable, i.e.  $G_k = 0$ . Also for large errors we have  $\lambda_{TM}(\delta) > \lambda_{TT}(\delta)$  because this crude approximation is not able to capture the intrinsic predictability of the original system. More refined parameterizations in terms of stochastic processes with the correct probability distribution function and correlation time do not improve the forecasting ability. On the contrary, Eq. (6.9) gives the result shown in Fig. 27. At small scales we still observe deviations from  $\lambda_{TT}$  but, at variance with the previous case, we recover intrinsic predictability for error of the size of the resolved scale.

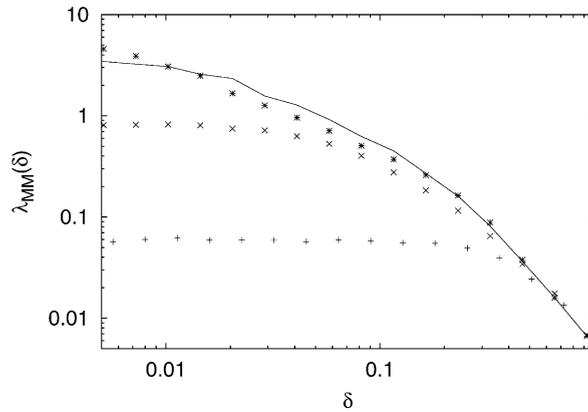


Fig. 28. The FSLE for the eddy-viscosity shell model (6.11), (6.12)  $\lambda_{MM}(\delta)$  at various resolutions  $N_M = 9(+)$ ,  $15(\times)$ ,  $20(*)$ . For comparison it is drawn  $\lambda_{TT}(\delta)$  (continuous line). Here  $\kappa = 0.4$ ,  $k_0 = 0.05$ .

As a more complex example, let us consider a version of the shell model discussed in Section 5.2, more precisely we study [151]

$$\frac{du_n}{dt} = i(k_{n+1}u_{n+1}^*u_{n+2} - \frac{1}{2}k_nu_{n-1}^*u_{n+1} + \frac{1}{2}k_{n-1}u_{n-2}u_{n-1}) - vk_n^2u_n + f_n, \quad (6.10)$$

with  $n = 1, \dots, N$ .

At variance with the previous example, here we have a set of scales  $\ell_n \simeq 1/k_n$  with characteristic times  $\tau_n \sim k_n^{-2/3}$  (see Section 5.4). In order to simulate a finite resolution in the model, we consider a model of (6.10) in terms of an eddy viscosity [24]:

$$\frac{du_n}{dt} = i(k_{n+1}u_{n+1}^*u_{n+2} - \frac{1}{2}k_nu_{n-1}^*u_{n+1} + \frac{1}{2}k_{n-1}u_{n-2}u_{n-1}) - v_n^{(e)}k_n^2u_n + f_n, \quad (6.11)$$

where now  $n = 1, \dots, N_M < N$  and the eddy viscosity, restricted to the last two shells, has the form

$$v_n^{(e)} = \kappa \frac{|u_n|}{k_n} (\delta_{n, N_M-1} + \delta_{n, N_M}), \quad (6.12)$$

where  $\kappa$  is a constant of order 1 [24]. In the model equation  $N_M < N$  the molecular viscosity term is much smaller than the eddy viscosity term and can be simply neglected. Model equations (6.11) and (6.12) are essentially the large eddy simulation for the shell model. Thus, although shell models are not realistic models for large-scale geophysical flows (being nevertheless a good model for small scale turbulent fluctuations), the study of the effect of truncation in term of eddy viscosity is of general interest.

In Fig. 28 we show  $\lambda_{MM}(\delta)$ , i.e. the FSLE computed for the model equations (6.11) with  $N = 24$  at different resolutions  $N_M = 9, 15, 20$ . A plateau is detected for small amplitudes of the error  $\delta$ , corresponding to the LE, which increases with the resolution according to  $\lambda \sim k_{N_M}^{2/3}$ . At larger  $\delta$ , the curves collapse onto the  $\lambda_{TT}(\delta)$ , showing that large-scale statistics of the model is not affected by small-scales resolution.

The ability of the model to predict satisfactorily the features of the “true” dynamics is not anyway determined by  $\lambda_{MM}(\delta)$  but by  $\lambda_{TM}(\delta)$ , which is shown in Fig. 29. Increasing the

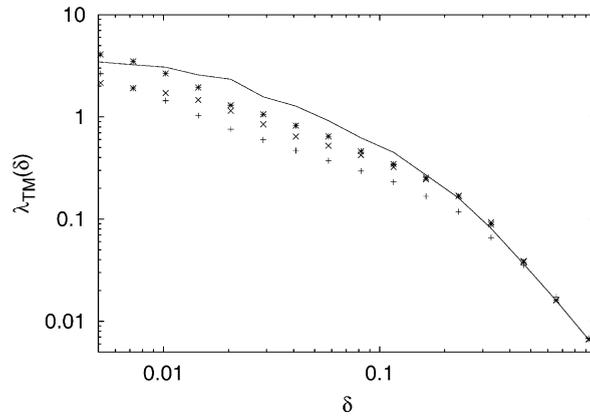


Fig. 29. The FSLE between the eddy-viscosity shell model and the full shell model  $\lambda_{\text{TM}}(\delta)$ , at various resolutions  $N_{\text{M}} = 9(+)$ ,  $15(\times)$ ,  $20(*)$ . For comparison it is drawn  $\lambda_{\text{TT}}(\delta)$  (continuous line). The total number of shell for the complete model is  $N = 24$ , with  $k_0 = 0.05$ ,  $\nu = 10^{-7}$ .

resolution  $N_{\text{M}} = 9, 15, 20$  towards the fully resolved case  $N = 24$  the model improves, in agreement with the expectation that  $\lambda_{\text{TM}}$  approaches  $\lambda_{\text{TT}}$  for a perfect model. At large  $\delta$  the curves practically coincide, showing that the predictability time for large error sizes (associated with large scales) is independent of the details of small-scale modeling.

### 6.3. Lyapunov exponents and complexity in dynamical systems with noise

We saw how in deterministic dynamical systems there exist well established ways to define the complexity of a temporal evolution, either in terms of the Lyapunov exponents and the Kolmogorov–Sinai entropy, or by means of their generalization to non-infinitesimal perturbations, like FSLE and  $\epsilon$ -entropy. The situation is much more ambiguous with random perturbations, which are always present in physical systems as a consequence of thermal fluctuations or hidden changes of control parameters, and, in numerical experiments, because of the roundoff errors [162].

The combined effect of the noise and the deterministic part of the evolution law can produce highly non-trivial behaviors [43,59,73,102,103,157]. Let us mention stochastic resonance, where there is a synchronization of the jumps between two stable points [25,26,163] (for a recent review see [87]), and the phenomena of the so-called noise-induced order [157] and noise-induced instability [43,59].

When facing systems with noise, the simplest possibility is to treat the random term as a time-dependent term, that is to consider the separation of two close trajectories with the same realization of noise. In this way one computes the largest LE,  $\lambda_{\sigma}$ , associated with the separation rate of two nearby trajectories with the same realization of the stochastic term (where  $\sigma$  indicates the noise strength). Although  $\lambda_{\sigma}$  is a well defined quantity, i.e. the Oseledec theorem [169] holds, it is not the most useful characterization of complexity. In addition, a moment of reflection shows that it is practically impossible to extract  $\lambda_{\sigma}$  from experimental data.

We will show how, for noisy and random systems, a more natural indicator of complexity can be obtained by computing the separation rate of nearby trajectories evolving with different noise realizations. This measure of complexity, defined in [174,149], and inspired by the contributions of Shannon [201] and Kolmogorov [126], is related to the mean number of bits per unit time necessary to specify the sequence generated by a random evolution law.

### 6.3.1. The naive approach: noise treated as a standard function of time

The approach in which one treats the random term as an usual time-dependent external force can lead to misleading results, as illustrated in the following example.

Let us consider a one-dimensional Langevin equation

$$\frac{dx}{dt} = -\frac{\partial V(x)}{\partial x} + \sqrt{2\sigma}\eta, \quad (6.13)$$

where  $\eta(t)$  is a white noise and  $V(x)$  diverges for  $|x| \rightarrow \infty$ , like, e.g., the usual double-well potential  $V = -x^2/2 + x^4/4$ .

The Lyapunov exponent  $\lambda_\sigma$ , associated with the separation rate of two nearby trajectories with the same realization of  $\eta(t)$ , is defined as

$$\lambda_\sigma = \lim_{t \rightarrow \infty} \frac{1}{t} \ln |z(t)|, \quad (6.14)$$

where the evolution of the tangent vector is given by

$$\frac{dz}{dt} = -\frac{\partial^2 V(x(t))}{\partial x^2} z(t). \quad (6.15)$$

Since the system is ergodic with invariant probability distribution  $P(x) = Ce^{-V(x)/\sigma}$ , one has

$$\begin{aligned} \lambda_\sigma &= \lim_{t \rightarrow \infty} \frac{1}{t} \ln |z(t)| = -\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \partial_{xx}^2 V(x(t')) dt' \\ &= -C \int \partial_{xx}^2 V(x) e^{-V(x)/\sigma} dx = -\frac{C}{\sigma} \int (\partial_x V(x))^2 e^{-V(x)/\sigma} dx < 0. \end{aligned} \quad (6.16)$$

This result has a rather intuitive meaning: the trajectory  $x(t)$  spends most of the time in one of the “valleys” where  $-\partial_{xx}^2 V(x) < 0$  and only short intervals on the “hills” where  $-\partial_{xx}^2 V(x) > 0$ , so that the distance between two trajectories evolving with the same noise realization decreases on average. Notice that in Ref. [215], supported by a wrong argument, an opposite conclusion has been claimed.

A negative value of  $\lambda_\sigma$  implies a fully predictable process only if the realization of the noise is known. In the case of two initially close trajectories evolving under two different noise realizations, after a certain time  $T_\sigma$ , the two trajectories can be very distant, because they can be in two different valleys. For  $\sigma \rightarrow 0$ , due to the Kramers formula [57], one has  $T_\sigma \sim \exp \Delta V/\sigma$ , where  $\Delta V$  is the difference between the values of  $V$  on the top of the hill and at the bottom of the valley. The result obtained for the one dimensional Langevin equation can easily be generalized to any dimension for gradient systems if the noise is small enough [149].

Another example showing the limitations of this approach is provided by the case of stochastic resonance in chaotic systems. In this case, in fact, one can find the same qualitative behavior both for a positive and a negative LE. We refer to [149] for more details.

### 6.3.2. An information theory approach

The main difficulties in defining the notion of “complexity” of an evolution law with a random perturbation already appears in 1D maps. The generalization to  $N$ -dimensional maps or to ordinary differential equations is straightforward.

Therefore, we consider the model

$$x(t+1) = f[x(t), t] + \sigma w(t), \quad (6.17)$$

where  $t$  is an integer and  $w(t)$  is an uncorrelated random process, e.g.  $w$  are independent random variables uniformly distributed in  $[-\frac{1}{2}, \frac{1}{2}]$ . For the largest LE  $\lambda_\sigma$ , as defined in (6.14), now one has to study the equation

$$z(t+1) = f'[x(t), t] z(t), \quad (6.18)$$

where  $f' = df/dx$ .

Following the approach of Section 2.3, let  $x(t)$  be the trajectory starting at  $x(0)$  and  $x'(t)$  be the trajectory starting from  $x'(0) = x(0) + \delta x(0)$ . Let  $\delta_0 \equiv |\delta x(0)|$  and indicate by  $\tau_1$  the minimum time such that  $|x'(\tau_1) - x(\tau_1)| \geq \Delta$ . Then, we put  $x'(\tau_1) = x(\tau_1) + \delta x(0)$  and define  $\tau_2$  as the time such that  $|x'(\tau_1 + \tau_2) - x(\tau_1 + \tau_2)| > \Delta$  for the first time, and so on. In this way the Lyapunov exponent can be defined as

$$\lambda = \frac{1}{\bar{\tau}} \ln \left( \frac{\Delta}{\delta_0} \right) \quad (6.19)$$

where  $\bar{\tau} = (1/N) \sum \tau_i$  (see also Appendix A). If the above procedure is applied by considering the same noise realization for both trajectories,  $\lambda$  in (6.19) coincides with  $\lambda_\sigma$  (if  $\lambda_\sigma > 0$ ). Differently, by considering two different realizations of the noise for the two trajectories, we have a new quantity

$$K_\sigma = \frac{1}{\bar{\tau}} \ln \left( \frac{\Delta}{\delta_0} \right), \quad (6.20)$$

which naturally arises in the framework of information theory [5] and algorithmic complexity theory. The times  $\tau_1, \tau_2, \dots$  are nothing but the intervals at which it is necessary to repeat the transmission of  $x(t)$ , with a precision  $\delta_0$ , and  $K_\sigma/\ln 2$  is the number of bits *per* unit time one has to specify in order to transmit the sequence. If the fluctuations of the effective Lyapunov exponent  $\gamma(t)$  are very small (i.e. weak intermittency) one has

$$K_\sigma = \lambda + O(\sigma/\Delta). \quad (6.21)$$

The interesting situation happens for strong intermittency when there are alternations of positive and negative  $\gamma$  during long time intervals: this induces a dramatic change for the value of  $K_\sigma$ . This becomes particularly clear when we consider the limiting case of positive  $\gamma^{(1)}$  in an interval  $T_1 \gg 1/|\gamma^{(1)}|$  followed by a negative  $\gamma^{(2)}$  in an interval  $T_2 \gg 1/|\gamma^{(2)}|$ , and again a positive effective LE and so on. During the intervals with positive effective LE the transmission has to be repeated rather often with  $\simeq T_1/(\gamma^{(1)} \ln 2)$  bits at each time, while during the ones with negative effective LE no information has to be sent. Nevertheless, at the end of the contracting intervals one has  $|\delta x| = O(\sigma)$ , so that, at variance with the noiseless case, it is impossible to use them to compensate the expanding ones. This implies that in the limit of very large  $T_i$  only the

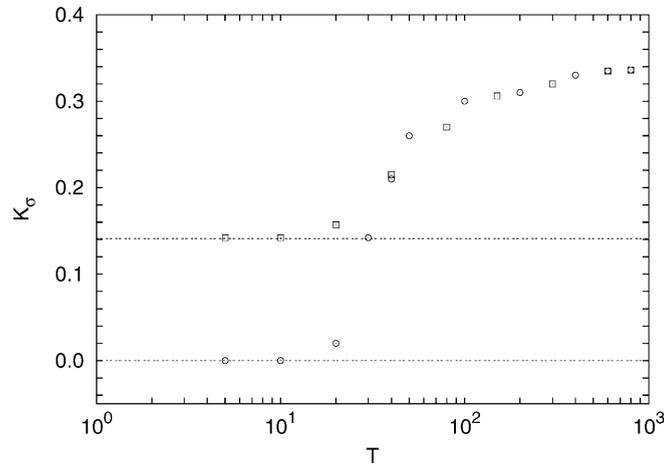


Fig. 30.  $K_\sigma$  versus  $T$  with  $\sigma = 10^{-7}$  for the map (6.23). The parameters of the map are  $a = 2$  and  $b = \frac{2}{3}$  (squares) or  $b = \frac{1}{4}$  (circles). The dashed lines are the noiseless limit of  $K_\sigma$ .

expanding intervals contribute to the evolution of the error  $\delta x(t)$  and  $K_\sigma$  is given by an average of the positive effective Lyapunov exponents:

$$K_\sigma \simeq \langle \gamma \theta(\gamma) \rangle . \tag{6.22}$$

Note that it may happen that  $K_\sigma > 0$  with  $\lambda_\sigma < 0$ . We stress again that (6.22) holds only for strong intermittency, while for uniformly expanding systems or rapid alternations of contracting and expanding behaviors  $K_\sigma \simeq \lambda_\sigma$ .

Note that  $K_\sigma$  is a sort of  $\epsilon$ -entropy (see Section 3.5), indeed, the complexity we consider is defined for  $\delta_0$  not too small ( $\delta_0 \gg \sigma$ ). If  $\delta_0$  and  $\Delta$  are small enough, but still much larger than  $\sigma$ ,  $K_\sigma$  is essentially independent of their values.

The relation  $K_\sigma \simeq \langle \gamma \theta(\gamma) \rangle$  is the time analogous of the Pesin relation (2.15)  $h_{KS} \leq \sum_i \lambda_i \theta(\lambda_i)$ . The latter relation expresses the fact that negative Lyapunov exponents do not decrease the value of  $h_{KS}$ , because the contraction along the corresponding directions cannot be observed for any finite space partition. In the same way the contracting time intervals, if long enough, do not decrease  $K_\sigma$ . Another important remark is that in the usual treatment of the experimental data, where noise is usually present, one practically computes  $K_\sigma$  and the result can be completely different from  $\lambda_\sigma$ .

Let us now briefly discuss some numerical results obtained with two different systems (Figs. 30 and 31). The first example consists in a periodic alternation of two piecewise linear maps of the interval  $[0, 1]$  into itself:

$$f[x, t] = \begin{cases} ax \bmod 1 & \text{if } (2n - 1)T \leq t < 2nT , \\ bx \bmod 1 & \text{if } 2nT \leq t < (2n + 1)T , \end{cases} \tag{6.23}$$

where  $a > 1$  and  $b < 1$ . Note that in the limit of small  $T$ ,  $K_\sigma \rightarrow \max[\lambda_\sigma, 0]$ , because it is a non-negative quantity as shown in Fig. 30.

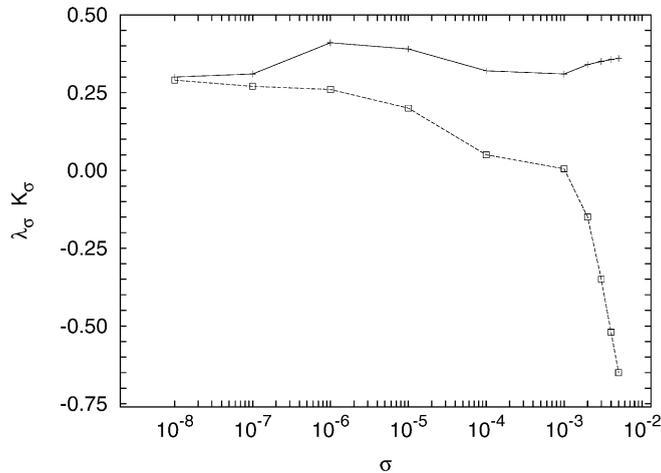


Fig. 31.  $\lambda_\sigma$  (squares) and  $K_\sigma$  (crosses) versus  $\sigma$  for the map (6.24).

The second example (Fig. 31), strongly intermittent without external forcing, is the Beluzov–Zhabotinsky map [103,157], introduced for describing the famous chemical reaction:

$$f(x) = \begin{cases} [(1/8 - x)^{1/3} + a]e^{-x} + b & \text{if } 0 \leq x < 1/8, \\ [(x - 1/8)^{1/3} + a]e^{-x} + b & \text{if } 1/8 \leq x < 3/10, \\ c(10xe^{-10x/3})^{19} + b & \text{if } 3/10 \leq x, \end{cases} \quad (6.24)$$

with  $a=0.50607357, b=0.0232885279, c=0.121205692$ . The map exhibits a chaotic alternation of expanding and contracting time intervals. In Fig. 31, one sees that while  $\lambda_\sigma$  passes from negative to positive values at decreasing  $\sigma$ ,  $K_\sigma$  is not sensitive to this transition [157]. Considering the system with a given realization of noise as the “true” evolution law, one has that  $\lambda_\sigma$  corresponds to  $\lambda_{TT}$  while  $K_\sigma$  corresponds to  $\lambda_{TM}$ .

The previous results show that the same system can be regarded either as regular (i.e.  $\lambda_\sigma < 0$ ), when the same noise realization is considered for two nearby trajectories, or as chaotic (i.e.  $K_\sigma > 0$ ), when two different noise realizations are considered.

#### 6.4. Random dynamical systems

We discuss now dynamical systems where the randomness is not simply given by an additive noise. This kind of systems has been the subject of interest in the last few years in relation to the problems involving disorder [119], such as the characterization of the so-called *on-off intermittency* [187] and to model transport problems in turbulent flows [227,228,86]. In these systems, in general, the random part represents an ensemble of hidden variables believed to be implicated in the dynamics. Random maps exhibit very interesting features ranging from stable or quasi-stable behaviors, to chaotic behaviors and intermittency. In particular, *on-off intermittency* is an aperiodic switching between static, or laminar, behavior and chaotic bursts of oscillation. It can be generated by systems having an unstable invariant manifold, within

which it is possible to find a suitable attractor (i.e. a fixed point). For further details we refer to [187].

A random map can be defined in the following way. Denoting with  $\mathbf{x}(t)$  the state of the system at discrete time  $t$ , the evolution law is given by

$$\mathbf{x}(t+1) = \mathbf{f}(\mathbf{x}(t), J(t)), \quad (6.25)$$

where  $J(t)$  is a random variable.

As for the case of additive noise examined in the previous section, the simplest approach is the introduction of the LE  $\lambda_J$  computed considering the separation of two nearby trajectories evolving with the same realization of the random process  $J(t) = i_1, i_2, \dots, i_t$ . The Lyapunov exponent  $\lambda_J$  generalizes  $\lambda_\sigma$  of Section 6.3.1 and can be computed from the tangent vector evolution:

$$\lambda_J = \lim_{N \rightarrow \infty} \frac{1}{N} \ln |\mathbf{z}(N)|, \quad (6.26)$$

where

$$z_m(t+1) = \sum_n \frac{\partial f_m(\mathbf{x}(t), i_t)}{\partial x_n} z_n(t). \quad (6.27)$$

On the other hand, also for these systems, as in the case of additive noise, it is possible to introduce a measure of complexity,  $K$ , which better accounts for their chaotic properties [174,149]

$$K \simeq h_{\text{Sh}} + \lambda_J \theta(\lambda_J), \quad (6.28)$$

where  $h_{\text{Sh}}$  is the Shannon entropy of the random sequence  $J(t)$ . The meaning of  $K$  is rather clear:  $K/\ln 2$  is the mean number of bits, for each iteration, necessary to specify the sequence  $x_1, \dots, x_t$  with a certain tolerance  $\Delta$ . Note that there are two different contributions to the complexity: (a) one has to specify the sequence  $J(1), J(2), \dots, J(t)$  which implies  $h_{\text{Sh}}/\ln 2$  bits per iteration; (b) if  $\lambda_J$  is positive, one has to specify the initial condition  $x(0)$  with a precision  $\Delta \exp^{-\lambda_J T}$ , where  $T$  is the time length of the evolution. This requires  $\lambda_J/\ln 2$  bits per iteration; if  $\lambda_J$  is negative the initial condition can be specified using a number of bits independent of  $T$ .

#### 6.4.1. A toy model: one-dimensional random maps

Let us discuss a random map which, in spite of its simplicity, captures some basic features of this kind of systems [187,101]:

$$x(t+1) = a_t x(t)(1 - x(t)), \quad (6.29)$$

where  $a_t$  is a random dichotomous variable given by

$$a_t = \begin{cases} 4 & \text{with probability } p, \\ 1/2 & \text{with probability } 1 - p. \end{cases} \quad (6.30)$$

For  $x(t)$  close to zero, we can neglect the non-linear term to obtain

$$x(t) = \prod_{j=0}^{t-1} a_j x(0); \quad (6.31)$$

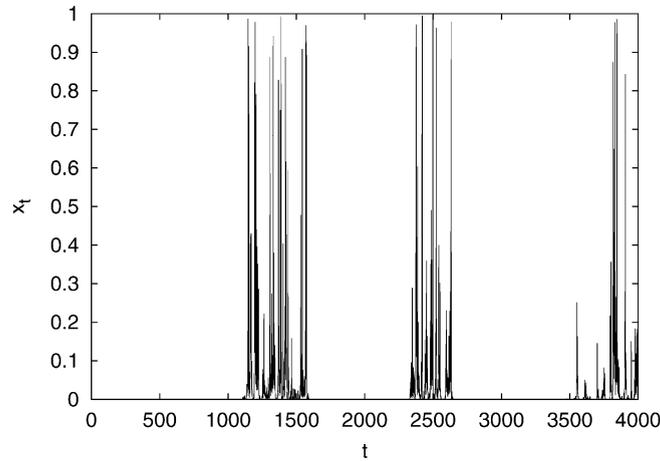


Fig. 32.  $x(t)$  versus  $t$  for the random map (6.29), (6.30), with  $p = 0.35$ .

from the law of large numbers one has that the typical behavior is

$$x(t) \sim x(0)e^{(\ln a)t} . \tag{6.32}$$

Since  $\langle \ln a \rangle = p \ln 4 + (1 - p) \ln 1/2 = (3p - 1) \ln 2$  one has that, for  $p < p_c = 1/3$ ,  $x(t) \rightarrow 0$  for  $t \rightarrow \infty$ . On the contrary for  $p > p_c$  after a certain time  $x(t)$  escapes from the fixed point zero and the non-linear term becomes relevant. Fig. 32 shows a typical *on-off intermittency* behavior for  $p$  slightly larger than  $p_c$ . Note that, in spite of this irregular behavior, numerical computations show that the LE  $\lambda_J$  is negative for  $p < \tilde{p}_c \simeq 0.5$ : this is essentially due to the non-linear terms.

By introducing a finite threshold  $\epsilon$ , in order to discriminate laminar and intermittent phases, we can define a complexity  $K(\epsilon)$ . We denote by  $l_L$  and  $l_J$  the average lifetimes respectively of the laminar and of the intermittent phases for  $p$  close to  $p_c$  ( $l_J \ll l_L$ ). The mean number of bits, per iteration, one has to specify in order to transmit the sequence is [150]

$$\frac{K(\epsilon)}{\ln 2} \simeq \frac{l_J h_{Sh}}{(l_J + l_L) \ln 2} \simeq \frac{l_J}{l_L} \frac{h_{Sh}}{\ln 2} . \tag{6.33}$$

To obtain (6.33) first notice that on an interval  $T$  one has approximatively  $T/(l_J + l_L)$  intermittent bursts and the same number of laminar phases. Then notice that, during a laminar phase, there is not an exponential growth of the distance between two trajectories initially close and computed with the same sequence of  $a_i$ . Since during a laminar phase one has to send a number of bits which does not depend on its duration, one can send all the necessary information simply by giving the sequence of  $a_i$  during the intermittent bursts. Eq. (6.33) has an intuitive interpretation: in systems with a sort of “catastrophic” events, the most important feature is the mean time between two subsequent events.

#### 6.4.2. Sandpile models as random maps

Another example of a system which can be treated in the framework of random maps is represented by the so-called sandpile models [15]. These models are a paradigmatic example

of the self-organized criticality (SOC) [14]. This term refers to the tendency of some large dynamical systems to evolve *spontaneously* toward a critical state characterized by spatial and temporal self-similarity. The original sandpile models are probabilistic cellular automata inspired to the dynamics of avalanches in a pile of sand. Dropping sand slowly, grain by grain on a limited base, one reaches a situation where the pile is critical, i.e. it has a critical slope. This means that a further addition of sand will produce sliding of sand (avalanches) that can be small or cover the entire size of the system. In this case the critical state is characterized by scale-invariant distributions for the size and the lifetime and it is reached without tuning of any critical parameter.

We will refer in particular to the Zhang model [229], a continuous version of the original sandpile model [15], defined on a  $d$ -dimensional lattice. The variable on each site  $x_i$  (interpretable as energy, sand, heat, mechanical stress, etc.) can vary continuously in the range  $[0, 1]$  with the threshold fixed to  $x_c = 1$ . The dynamics is the following:

- (a) one chooses at random a site and adds to it an energy  $\delta e$ ,
- (b) if at a certain time  $t$  the energy in a site, say  $i$ , exceeds the threshold  $x_c$  a relaxation process is triggered defined as

$$\begin{aligned} x_{i+mn} &\rightarrow x_{i+mn} + x_i/2d, \\ x_i &\rightarrow 0, \end{aligned} \tag{6.34}$$

where  $mn$  indicates the  $2d$  nearest neighbors of the site  $i$ ;

- (c) one repeats point (b) until all the sites are relaxed;
- (d) one goes back to point (a).

Let us now discuss the problem of predictability in sandpile models on the basis of the rigorous results [46], which clarify the role of the LE for this class of systems.

In Ref. [46] it has been proved that the LE  $\lambda_J$  is negative. In fact the dynamics of a little difference between two configurations follows the same rules (a)–(d), i.e., the “error” is redistributed to the nearest neighbors site, so that one has

$$\lambda_J \leq -\frac{\text{const}}{R^2}, \tag{6.35}$$

where  $R$  is the diameter of the system.

As for other examples already discussed, the existence of a negative LE does not mean a perfect predictability. This can be understood by looking at the growth of the distance,  $\delta(t)$ , between two initially close trajectories computed with two different realizations of randomness, i.e., by adding sand in different sites. Let us consider the case of the “minimal error”: in the reference realization one adds sand on a site  $i$  chosen at random. In the perturbed realization, instead, one adds a sand grain at one of the nearest sites of  $i$ . In such a case  $\delta(t)$  increases up to a maximal distance in few avalanches [150]. Practically, one has the same kind of phenomenon, already discussed, of the Langevin equation with two noise realizations.

Let us now estimate the complexity  $K$  of this system. An upper bound can be given by using (6.28)  $K = h_{\text{Sh}} + \lambda_J \theta(\lambda_J)$ , where  $h_{\text{Sh}}$  is the entropy of the random sequence of addition

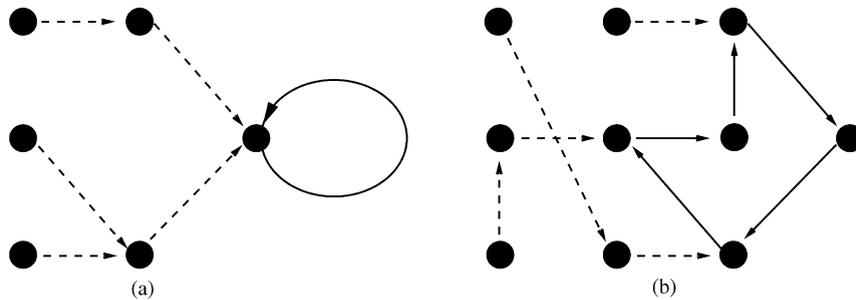


Fig. 33. Schematic representation of the evolution of a deterministic rule with a finite number of states: (a) with a fixed point, (b) with a periodic cycle.

of energy. In sandpile models, since each site has the same probability to be selected, one has  $h_{\text{Sh}} = \ln V$ , where  $V$  is the number of sites of the system. Since the Lyapunov exponent is negative, the complexity is just determined by  $h_{\text{Sh}}$ .

## 7. Irregular behavior in discrete dynamical systems

For the sake of completeness we include in this review a discussion on the characterization of irregular behaviors in systems whose states are discrete. Such systems include Cellular Automata (CA), which have been intensively studied both for their intrinsic interest [223] and for applications as, e.g., to simulate hydrodynamic equations [82] or to study various forms of chemical turbulence [166,167,40]. Other interesting systems with discrete states are the neural networks used for modeling some brain functions [6]. It is also relevant to note that in every simulation with a computer, because of the finite number of digits it can use, one deals with a system with discrete states (see Section 6.1). In addition, the general problem of dynamics of systems with discrete states is important in the debated issue of quantum chaos. Indeed quantum mechanics can be regarded as a discretized version of the classical one, acting on a suitable lattice in phase space, where the number of the possible states is proportional to the inverse of the Planck constant [81,49,60].

If a system consists of  $N$  elements and each element can assume an integer number  $k$  of distinct values,  $\mathcal{N} = k^N$  is the number of states. When these states evolve with a deterministic rule, the dynamics can be depicted in terms of oriented graphs: a set of points, representing the states, are connected by arrows, indicating the time evolution. Of course, each point has one, and only one, outgoing arrow; but different arrows can end at the same point. For any finite system each initial condition evolves to a definite attractor, which can be either a fixed point (as in Fig. 33a), or a periodic orbit (Fig. 33b).

In systems of this kind, obviously, it is not possible to use the previously introduced indicators of chaos, e.g. the Lyapunov exponents or the Kolmogorov–Sinai entropy, whose definitions rely on the continuous character of the system states. Moreover, the asymptotic periodic behavior seems to force the conclusion that discrete states systems are trivial, from an entropic or algorithmic complexity point of view.

The above conclusions, although mathematically correct, are rather unsatisfactory from the physical point of view, indeed from this side the following questions deserve some interest:

- (1) What is the “typical” period,  $\mathcal{T}$ , in a system with  $N$  elements, each assuming  $k$  distinct values?
- (2) When  $\mathcal{T}$  is very large, how can we characterize the (possible) irregular behavior of the trajectories, on times that are large enough but still much smaller than  $\mathcal{T}$ ?
- (3) What does it happen in the transition from discrete to continuous states, i.e. in the limit  $k \rightarrow \infty$ ?

In the next subsections we will deal with the above questions.

### 7.1. Dependence of the period on the number of the states

For deterministic discrete state systems the dependence of the period of the attractor on the number of the states, may be addressed with a statistical approach in terms of random maps [63]. We recall that this problem is important for computer simulations of chaotic systems (see Section 6.1). If  $\mathcal{N} = k^N \gg 1$  is the number of states of the system, the basic result for the average period,  $\mathcal{T}$ , is

$$\mathcal{T}(\mathcal{N}) \sim \sqrt{\mathcal{N}}. \quad (7.1)$$

In the following we give a simple argument, by Coste and Hénon [63].

For simplicity of notation, we consider the case with  $k=2$ , so that the state of the system is a string of  $N$  bits. A deterministic evolution of such a system is given by a map which is one among the possible functions connecting the  $2^N$  states. Let us now assume that all the possible functions can be extracted with the same probability. Denoting by  $I(t)$  the state of the system, for a certain map we have a periodic attractor of period  $m$  if  $I(p+m) = I(p)$  and  $I(p+j) \neq I(p)$ , for  $j < m$ . The probability,  $\omega(m)$ , of this periodic orbit is obtained by specifying that the  $(p+m-1)$ th first successive images of the map are distinct from all the previous ones and the  $(p+m)$ th iterates coincides with the  $p$ th one. Since one has  $I(p+1) \neq I(p+m)$ , with probability  $(1 - 1/\mathcal{N})$ ;  $I(p+2) \neq I(p+m)$ , with probability  $(1 - 2/\mathcal{N})$ ; ...;  $I(p+m-1) \neq I(p+m)$ , with probability  $(1 - (m-1)/\mathcal{N})$ ; and, finally,  $I(p+m) = I(p)$  with probability  $(1/\mathcal{N})$ , one obtains

$$\omega(m) = \left(1 - \frac{1}{\mathcal{N}}\right) \left(1 - \frac{2}{\mathcal{N}}\right) \cdots \left(1 - \frac{(m-1)}{\mathcal{N}}\right) \frac{1}{\mathcal{N}}. \quad (7.2)$$

The average number,  $M(m)$ , of cycles of period  $m$  is

$$M(m) = \frac{\mathcal{N}}{m} \omega(m) \stackrel{(\mathcal{N} \gg 1)}{\approx} \frac{e^{-m^2/2\mathcal{N}}}{m}, \quad (7.3)$$

from which one obtains  $\mathcal{T} \sim \sqrt{\mathcal{N}}$  for the average period.

It is here appropriate to comment on the relevance of Eq. (7.1) for computer-generated orbits of chaotic dynamical systems. Because of the finite number,  $n$ , of digits used in the floating point representation, when one iterates a dynamical system, one basically deals with a discrete system with a finite number  $\mathcal{N}$  of states. If  $d_2$  indicates the correlation dimension of the system

[93,94], one can reasonably assume that  $\mathcal{N} \sim 10^{nd_2}$ , so that, from Eq. (7.1) one has

$$\mathcal{T} \sim 10^{nd_2/2}. \quad (7.4)$$

This estimation gives an upper limit for the typical number of meaningful iterations of a map on a computer. Note that this number, apart from the cases of 1 or 2-dimensional maps with few digits, is very large for almost all practical purposes.

## 7.2. The transition from discrete to continuous states

Following the basic ideas of Ford [79,80], as discussed in Section 2.3, and the results of Section 6—on the predictability in systems whose evolution law is not completely known—we describe now a way to introduce a practical definition of chaos for systems with discrete states. In addition, we deal with the problem of the transition from discrete to continuous states.

Given a system with  $\mathcal{N}$  possible states, denoting by  $\mathbf{I}(t)$  its state at time  $t$  we can write its evolution law as

$$\mathbf{I}(t+1) = \mathbf{F}[\mathbf{I}(t)]. \quad (7.5)$$

A single state  $\mathbf{I}$  is a sequence of (at most)  $\ln_2 \mathcal{N}$  bits, and its time evolution for  $M$  steps can be surely translated in a binary sequence  $\Sigma$  of length  $\ell_\Sigma(M, \mathcal{N}) \leq M \ln_2 \mathcal{N}$ .

Relying on the definition of algorithmic complexity (Section 2.2.3) we can make the following classification: we call regular (compressible) those sequences that can be encoded by a computer program whose length  $\ell_\Sigma(M, \mathcal{N})$  increases less than linearly in  $M$ , when  $M$  ranges over a physically significant interval, at fixed values of  $\mathcal{N}$ . Otherwise the system will be called chaotic or incompressible. Let us call  $\ell_F$  the binary length of the algorithm for one step:  $\ell_F \leq 2\mathcal{N} \ln_2 \mathcal{N}$ . The sequence  $\Sigma$  can be expressed by the record composed by the initial state  $\mathbf{I}(0)$  (specified by  $\ln_2 \mathcal{N}$  bits), the number of steps  $M$  (specified by  $\ln_2 M$  bits) and the rule  $\mathbf{F}$  for one step (specified by  $\ell_F$  bits). Therefore

$$\ell_\Sigma(M, \mathcal{N}) \leq (2\mathcal{N} + 1) \ln_2 \mathcal{N} + \ln_2 M + O(1). \quad (7.6)$$

Let us note that from the above equation one has that – when  $M$  grows indefinitely and  $\mathcal{N}$  is constant –  $\ell_\Sigma$  is logarithmically bounded and hence the sequence appears to be compressible. This is somewhat trivial since, because of the discrete nature of the states, the motion at  $M > \mathcal{N}$  (in practice  $M > \mathcal{T} \sim \sqrt{\mathcal{N}}$ ) is periodic. Therefore it is interesting to consider only  $1 \ll M < \mathcal{T} \ll \mathcal{N}$ . Although the evolution law (7.5) can be carried out, in principle, in exact arithmetic, in practice in real computations one has unavoidable errors due to truncations and approximations. Let us now regard the evolution law (7.5) as a computer program with input  $\mathbf{I}(0)$  and a set  $\mathbf{C}$  of parameters, with  $C$  components, needed to define the algorithm  $\mathbf{F}$ . If these parameters are all known within precision  $O(2^{-q})$ , the binary length of the coding of  $\mathbf{C}$  is  $O(qC)$ .

Consider the following problem: given two identical initial conditions  $\mathbf{I}^{(1)}(0) = \mathbf{I}^{(2)}(0) = \mathbf{I}(0)$ , and two different realizations  $\mathbf{C}^{(1)}$  and  $\mathbf{C}^{(2)}$  of the set of coefficients  $\mathbf{C}$  (with difference  $O(2^{-q})$ ), what is the dependence on  $\epsilon = 2^{-q}$  of the first-error time  $\tilde{M}$  (i.e. the first time for which  $\mathbf{I}^{(1)}(t) \neq \mathbf{I}^{(2)}(t)$ )? Of course, the answer depends on the realizations of the components of  $\mathbf{C}$  and on the

initial conditions  $\mathbf{I}(0)$ . Let us consider  $\mathbf{C}^{(2)}$  as an  $\epsilon$ -perturbation of  $\mathbf{C}^{(1)}$ , i.e. we pose, for each component of the parameter vector:

$$C_i^{(2)} = C_i^{(1)} + \epsilon_i, \tag{7.7}$$

where the random variables  $\epsilon_i$  are uniformly distributed in  $[-2^{(-q-1)}, 2^{(-q-1)}]$ . Let us note that the coding length  $O(Cq) + \ln_2 \mathcal{N}$  is enough to define the sequence  $\Sigma$  up to the first error time  $\tilde{M}$ . Performing an average on the  $\epsilon_i$  and on the initial conditions  $\mathbf{I}(0)$ , one can compute an average first-error time  $\langle \tilde{M}(\epsilon) \rangle$ , and a typical first-error time  $\tilde{M}_{\text{typ}}(\epsilon) = \exp(\ln \tilde{M}(\epsilon))$ . For  $\langle \tilde{M} \rangle$  and  $\tilde{M}_{\text{typ}}$  the dependence on  $\epsilon$  can be of the following type:

- (a)  $\langle \tilde{M} \rangle \sim \epsilon^{-\beta} \sim 2^{q\beta}$ ,
- (b)  $\langle \tilde{M} \rangle \sim \ln_2(1/\epsilon) \sim q$ .

In the case (a) we say that the system is compressible, while if (b) holds one has a chaotic (incompressible) case. The above classification is rather obvious: in case (a) a trajectory of length  $\tilde{M}$  can be coded by a program of length  $O(\ln_2 \mathcal{N}) + O(\ln \tilde{M})$ , while in case (b) one has a length  $O(\ln_2 \mathcal{N}) + O(\tilde{M})$ . For a detailed discussion see Ref. [65].

Let us now discuss in an explicit example the previous approach and the problems in the transition to the continuous state case. We consider a discretized standard map, as obtained by considering lattice points in the torus  $[0, 2\pi]^2$  of the form  $(x, y)$  with  $x = 2\pi Q/L$  and  $y = 2\pi P/L$ , where  $P$  and  $Q$  are integers between 1 and  $L$ . The action of the map is

$$\begin{aligned} Q(t+1) &= \left[ Q(t) + \alpha \frac{L}{2\pi} \sin\left(P(t) \frac{2\pi}{L}\right) \right] \text{mod } L, \\ P(t+1) &= (P(t) + Q(t+1)) \text{mod } L, \end{aligned} \tag{7.8}$$

where  $\alpha$  is the control parameter and  $[\cdot]$  means integer part. From the results of Section 7.1 one has that the typical period for the map (7.8) is  $\mathcal{T}_L \sim L$ ; so if  $L$  is large the periodic motion will be seen only for sufficiently large times. In the system (7.8) one has just one parameter, i.e. the “kick strength”  $\alpha$ . Numerical evidence supports the following picture: at fixed  $L$ , the first-error time is roughly constant for large values of the error  $\epsilon$ , while it goes as  $\epsilon^{-1}$  for small errors  $\epsilon$ . The transition between the two regimes occurs at a critical value  $\epsilon_c(L)$  which scales as  $\epsilon_c \sim 1/L$ . In formulae:

$$\langle \tilde{M}(\epsilon) \rangle \sim \begin{cases} O(1) & \text{for } \epsilon > \epsilon_c(L), \\ 1/\epsilon & \text{for } \epsilon < \epsilon_c(L). \end{cases} \tag{7.9}$$

It is rather easy to give analytical estimates supporting the numerical evidence [65]

$$\langle \tilde{M}(\epsilon) \rangle = \begin{cases} \left( \frac{\pi l}{\epsilon} \right) & \text{for } \epsilon < l, \\ \left( \frac{\epsilon}{\pi l} (1 - \cos \theta) + 1 - \frac{2\theta}{\pi} + \frac{l}{\pi \epsilon} \ln \tan \left( \frac{\theta}{2} \right) \right)^{-1} & \text{for } \epsilon > l. \end{cases} \tag{7.10}$$

where the angle  $\theta$  is defined via  $\sin \theta = l/\epsilon$ , and  $l = 2\pi/L$ . Numerical simulations show that the behavior proposed in Eq. (7.9) holds. To have a comparison with the usual standard map, we have computed the average time  $\langle M_\Delta(\epsilon) \rangle$  required for two trajectories to reach lattice points

farther than a fixed distance  $\Delta$  in the discrete phase space of Eq. (7.8). We found

$$\langle M_\Delta(\epsilon) \rangle \sim \begin{cases} (1/\lambda) \ln(\Delta/\epsilon) & \text{for } \epsilon > \epsilon_c(L), \\ 1/\epsilon & \text{for } \epsilon < \epsilon_c(L). \end{cases} \quad (7.11)$$

We remark that when  $\epsilon < \epsilon_c(L)$ ,  $M_\Delta(\epsilon)$  is weakly dependent, i.e. logarithmically, on  $\Delta$ . This is just another verification of the similarity of the effect of a small disturbance on the equations of motion and of a small error in the initial conditions for a dynamical evolution (see Section 6).

These results unveil the nature of the dynamics of this discrete system: its trajectories are incompressible and therefore chaotic only for large values of  $\epsilon$ , the cutoff value decreasing as  $1/L$ . This helps us also to understand the extent to which the dynamics of the discrete standard map Eq. (7.8) is equivalent to its continuum counterpart. When  $\alpha$  is large, and  $\epsilon > \epsilon_c$ , the two systems possess chaotic trajectories. Simple calculations show that, to the leading order,  $\tilde{M} \sim \log_2(L)$ . After this time, the discrete system appears “regular”, i.e. compressible and predictable. Therefore, continuous and discrete systems are similar (as far as chaos is concerned) only over logarithmically short times. It is important to stress that the system appears “regular” on time scales much smaller than the typical period  $\mathcal{T}_L \sim L^{d_2/2}$  ( $d_2$  being the correlation dimension of the attractor [93,94]).

Recently, Mantica [155] studied the algorithmic complexity in classical polygonal billiards with  $L$  sides. The system, for any finite value of  $L$ , is regular; on the other hand, as  $L \rightarrow \infty$ , it tends to a curved billiard, which can be chaotic. This system is very similar to the discrete dynamical system (7.8) and may be used to study the transition from quantum to classical mechanics and the principle of correspondence. The average complexity of symbolic trajectories in the polygonal billiards has the same scaling behavior (as function of  $L$  and of the precision  $\epsilon$ ) of that one of the system (7.8), i.e. a compressible (regular) regime for  $\epsilon < \epsilon_c \sim 1/L$  and an incompressible (chaotic) one for  $\epsilon > \epsilon_c$ .

It is interesting to note that a similar feature is characteristic of quantum dynamics of systems whose classical behavior is chaotic. Roughly speaking, a quantum system behaves as a system with discrete states whose number is proportional to  $\hbar^{-1}$ . A semi-classical wave function follows a dynamics which is approximately classical up to a time  $t_c \sim (1/\lambda) \ln(I/\hbar)$ , where  $\lambda$  is the Lyapunov exponent of the classical motion, and  $I$  is a typical action of the motion. Over this time, the quantum system has the same complexity of its classical counterpart, while for larger times its quantal (quasi-periodic) nature appears [81,49,60].

### 7.3. Entropies and Lyapunov exponents in cellular automata

Cellular automata (CA) consist of discrete valued state variables,  $\sigma_i(t)$ , defined on a discrete lattice, and updated synchronously at discrete time according to a local rule. They can be defined in any dimensions and for any finite number of possible values for  $\sigma_i(t)$ . For the sake of simplicity we consider Boolean CA, i.e.  $\sigma_i(t) = \{0, 1\}$ , in a 1-dimensional lattice. An evolution rule can be written as

$$\sigma_i(t+1) = F[\sigma_{i-r}(t), \dots, \sigma_i(t), \dots, \sigma_{i+r}(t)], \quad i = 1, \dots, N, \quad (7.12)$$

where  $r$  defines the range of the coupling, i.e. the variable in a site depends on the variables in the  $2r$  neighbor sites. If  $F$  in (7.12) only depends on the sum of the state variables, one

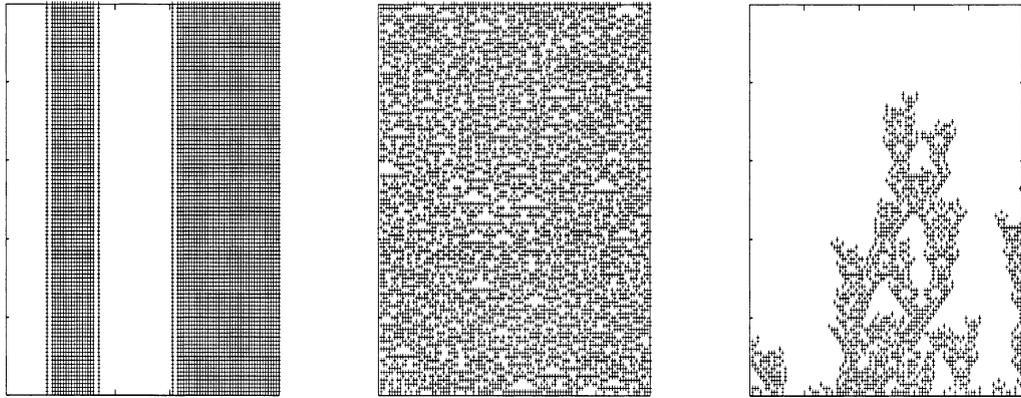


Fig. 34. Typical behavior of (a) regular [Rule 52], (b) chaotic [Rule 22], (c) complex [Rule 20]. We used totalistic  $r = 2$  cellular automata. Time flows from below to above.

speaks of “totalistic” CA. Another usual requirement is to have symmetric rules. For further details we refer to Ref. [223], where the standard scheme for the classification of the possible rules and a number of examples of CA-behavior are discussed.

In the following we refer to 1d Boolean cellular automata with local symmetric rules – as those ones systematically studied by Wolfram [223].

### 7.3.1. Classification of Cellular Automata according to the transient times

For finite lattices with  $N$  sites the number of possible states of CA is finite and equal to  $\mathcal{N} = 2^N$ . As already discussed, this means that, strictly speaking, from an entropic (or algorithmic) point of view CA are trivial. Therefore the problem of the characterization of irregular behaviors in CA has, in principle, some meaning only in the limit  $N \rightarrow \infty$ . In more physical terms, for finite  $N$  one expects the characterization in terms of entropy to be possible for times shorter than the typical period  $\mathcal{T}(N)$  or the typical transient time  $\tilde{\mathcal{T}}(N)$ , provided  $\mathcal{T}(N)$  and  $\tilde{\mathcal{T}}(N)$  are long enough.

Actually, cellular automata behaviors can be classified according to the dependence of  $\mathcal{T}(N)$  and  $\tilde{\mathcal{T}}(N)$  on  $N$ . One has three possible classes of behavior.

*Regular* cellular automata (classes 1 and 2 in Wolfram’s classification [223]) evolve either on homogeneous states both in time and space (the analogous of fixed point in dynamical systems) either to a set of simple stable periodic structures (analogous to limit cycles) which, in general, depend on the initial configuration (Fig. 34a). In these CA,  $\mathcal{T}(N)$  and  $\tilde{\mathcal{T}}(N)$  can range from being almost independent of  $N$  to be, at maximum, proportional to  $N$ .

*Chaotic* cellular automata (class 3 in [223]) yield disordered patterns (Fig. 34b). For any finite  $N$  these CA reach periodic states, but there are rather clear numerical evidences that the transient time  $\tilde{\mathcal{T}}$  increases exponentially with the system size:

$$\tilde{\mathcal{T}}(N) \sim \exp(cN). \quad (7.13)$$

Moreover, also the cycle period shows in most of the cases a similar dependence on  $N$ , this is a reminiscence of what we discussed in Section 7.1.

*Complex* cellular automata (class 4 in [223], Fig. 34c) usually evolve toward complex localized structures (gliders) which interact in a complicate way. For these CA numerical simulations [100] have shown that both the transient time and the cycle period display a non-trivial  $N$ -dependence (i.e. the average, the typical values or the median depend in a different way on  $N$ ). The unpredictability of these system manifests itself in the distribution of these times. In particular, the large variability of these times in dependence of the initial conditions and the lattice size inhibits any forecasting of the duration of the transient.

In the following we limit the discussion to chaotic rules, i.e. class 3 in the Wolfram classification. A detailed characterization of *complex* CA would require the introduction of concepts and tools that are beyond the aim of this review, for further details see Refs. [13,96,100,223].

### 7.3.2. Sensitive dependence on initial conditions

A first intuitively reasonable characterization of irregular behaviors is in terms of sensitive dependence on initial conditions, but in CA it is not possible to have arbitrary small distances between two states. Nevertheless, for large  $N$ , when considering two states with only one different element, one can say that, in some sense (i.e. in an appropriate norm), the difference is small. Denoting with  $R_t$  the number of different elements at time  $t$ , we can define the damage propagation speed as [223]

$$v = \lim_{t \rightarrow \infty} \frac{R_t}{2t} . \quad (7.14)$$

It is not difficult to see that  $v$  is, in a proper space, a Lyapunov exponent (i.e. it measures the rate of divergence of two configurations) [214]. Consider two initial bi-infinite configurations  $\sigma(0) = (\dots, \sigma_{-2}(0), \sigma_{-1}(0), \sigma_1(0), \sigma_2(0), \dots)$  and  $\sigma'(0) = (\dots, \sigma'_{-2}(0), \sigma'_{-1}(0), \sigma'_1(0), \sigma'_2(0), \dots)$ , with  $\sigma_i(0) = \sigma'_i(0)$  for  $|i| < N_0$ , and their evolutions  $\sigma(t)$  and  $\sigma'(t)$ . One can define a distance,  $\|\delta\sigma(t)\|$ , between  $\sigma(t)$  and  $\sigma'(t)$ , as follows:

$$\|\delta\sigma(t)\| = \sum_{n=1}^{\infty} \frac{|\delta\sigma_n(t)| + |\delta\sigma_{-n}(t)|}{2^n} \quad (7.15)$$

where  $\delta\sigma_n = \sigma'_n - \sigma_n$ . With the above norm two systems can be arbitrarily close: one only needs  $N_0 \gg 1$ . At this point it is possible to define the Lyapunov exponent as

$$\lambda = \lim_{t \rightarrow \infty} \lim_{\|\delta\sigma(0)\| \rightarrow 0} \frac{1}{t} \ln \frac{\|\delta\sigma(t)\|}{\|\delta\sigma(0)\|} . \quad (7.16)$$

Note that in (7.16) it has been implicitly taken the limit  $N \rightarrow \infty$ .

Noting that  $\delta\sigma_n(t) = 0$  for  $|n - N_0| > R_t/2 \simeq vt$ , while  $|\delta\sigma_n(t)| = 1$  for  $|n - N_0| < R_t/2 \simeq vt$ , from the definition (7.15) one has

$$\|\delta\sigma(t)\| \sim 2^{-N_0+vt} , \quad (7.17)$$

and therefore

$$\lambda = v \ln 2 . \quad (7.18)$$

In other words, the linear damage spreading in the physical space corresponds to an exponential growth in the norm (7.15). Oono and Yeung [167] stressed a conceptual (and practical) difficulty with the above approach. In systems with continuous states it is clear that by performing an

infinitesimal change on a typical configuration one does not destroy the “typicality”, i.e. the new initial condition will generate a trajectory belonging to the same attractor. On the contrary, it is not obvious that for a, however large, system with discrete states in a typical configuration a change of only one element gives another typical state. For instance, this seemingly innocent change can induce a jump among basins of attraction, so that the perturbed trajectory goes to a different attractor [223,16]. However, taking into account the above criticism, numerically one finds, for most the initial conditions,  $v > 0$  for chaotic CA, and  $v = 0$  for regular CA.

We conclude this subsection mentioning a proposal, by Bagnoli et al. [16], to introduce a Lyapunov exponent for cellular automata, defining it in analogy with continuous states dynamical systems.

In this approach, the equivalent of an infinitesimal perturbation (as for the damage spreading analysis) is the difference between the system and one of its replicas in which one site has been flipped at time  $t = 0$ . Then one formally introduces the Boolean derivatives, a sort of Jacobian of the rule,  $F'_{i,j}$ , the elements of which are 0 or 1. Here, for simplicity, we consider a generic nearest neighbor ( $r = 1$ ) rule so that  $F'_{i,j} = 0$  for  $|i - j| \geq 2$  and

$$F'_{i,i-1} = \frac{\partial \sigma_i(t+1)}{\partial \sigma_{i-1}(t)} \equiv F[\sigma_{i-1}, \sigma_i, \sigma_{i+1}] \text{ XOR } F[\sigma_{i-1} \text{ XOR } 1, \sigma_i, \sigma_{i+1}]$$

where the other non-zero terms are obtained by shifting the XOR operation to  $i$  and  $i + 1$  (respectively). We recall that XOR is the Boolean exclusive operation (i.e.  $0 \text{ XOR } 0 = 0$ ,  $1 \text{ XOR } 1 = 0$ ,  $0 \text{ XOR } 1 = 1$  and  $1 \text{ XOR } 0 = 1$ ). Of course as time goes on the initial perturbation spreads, i.e. new defects appear. As in continuous systems, one needs to maintain the perturbation “infinitesimal”. One introduces a vector  $\mathbf{N}$  (whose components,  $N_i$ , take integer values) which plays the role of the tangent vector. In order to mimic an infinitesimal perturbation at the initial time one assumes  $N_i(0) = \delta_{ij}$ , i.e. only one defect on the site  $j$ . The dynamics of  $N_i$  is ruled by the Boolean derivative, i.e.

$$N_i(t+1) = \sum_j F'_{ij}(t) N_j(t) . \quad (7.19)$$

Finally, putting  $|\mathbf{N}(t)| = \sum_j N_j(t)$ , one can define the “Lyapunov exponent”,  $\lambda_B$ , of the cellular automaton as

$$\lambda_B = \lim_{T \rightarrow \infty} \frac{1}{T} \ln(|\mathbf{N}(T)|) . \quad (7.20)$$

Now, in analogy with continuous systems,  $\lambda_B < 0$  indicates an exponential decrease of the perturbation, while for  $\lambda_B > 0$  the damage spreads. Just to give an example, if one considers the rule 150 of Wolfram classification, i.e. ( $F[0, 0, 1] = F[0, 1, 0] = F[1, 0, 0] = F[1, 1, 1] = 1$  and 0 otherwise) it is easy to see that  $F'_{ij}$  is a tridiagonal matrix with all the elements equal to 1 so that  $\lambda = \ln(3)$ . For a generic rule one has to compute a suitable average over a long trajectory or on many initial configurations (see Fig. 35).

The Lyapunov exponent,  $\lambda_B$ , has been demonstrated to be relevant in the synchronization problem [17] and allows for a qualitative characterization of the cellular automata in agreement with the classification proposed by Wolfram [16,17].

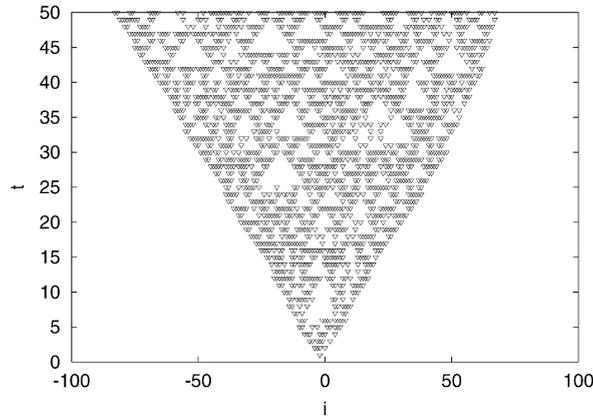


Fig. 35. Damage spreading analysis performed on a totalistic [Rule 10]  $r = 2$  cellular automaton with  $N = 200$ . At time  $t = 0$  a replica is initialized by flipping the value at the center of the lattice.

### 7.3.3. Entropies

For cellular automata one can define a spatial/temporal entropy density by looking at the evolution of the elements in a subset  $\mathcal{L}_L$ , of size  $L$ , of the system. Denoting by  $C(L, T)$  a “word” of spatial size  $L$  and time length  $T$  appearing in the time evolution of the elements in  $\mathcal{L}_L$ , one defines the entropy of the subset  $\mathcal{L}_L$ ,

$$h(L) = \lim_{T \rightarrow \infty} - \frac{1}{T} \sum_{C(L,T)} P(C(L,T)) \ln P(C(L,T)), \tag{7.21}$$

and then the spatio-temporal entropy density as

$$h^{ST} = \lim_{L \rightarrow \infty} \frac{1}{L} h(L). \tag{7.22}$$

This entropy cannot be practically computed. A more accessible quantity is the temporal entropy:

$$h^T = h(1) = \lim_{T \rightarrow \infty} - \frac{1}{T} \sum_{C(1,T)} P(C(1,T)) \ln P(C(1,T)), \tag{7.23}$$

i.e. the Shannon entropy of the time sequence of one element  $(\sigma_n(0), \sigma_n(1), \dots)$ . In principle,  $h^T$  can depend on the site  $n$  and one can classify as non-trivial a system for which the majority of the elements have  $h^T > 0$  [166]. An average measure of the “temporal disorder” is given by the spatial average  $\langle h^T \rangle$ . A systematic study of  $h(1), h(2), h(3), \dots$  – although very difficult in practice – could give, in principle, relevant information on the spatial/temporal behavior. A characterization of the spatial properties can be obtained studying, at a given time  $t$ , the spatial sequences. In practice, one studies  $C(L, 1)$  at increasing  $L$ :

$$h^S = \lim_{L \rightarrow \infty} - \frac{1}{L} \sum_{C(L,1)} P(C(L,1)) \ln P(C(L,1)). \tag{7.24}$$

One can associate to  $h^S$  a sort of “effective” dimension  $d = h^S / \ln 2$  [223]. In a completely disordered cellular automaton configuration one has  $d = 1$ , as expected, while a homogeneous (or spatially periodic) configuration gives  $d = 0$ .

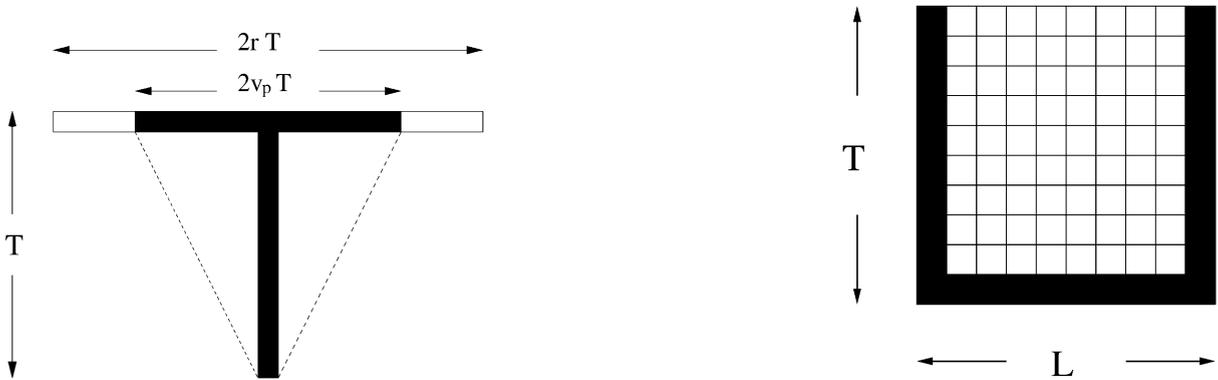


Fig. 36. Sketch of the dependence of temporal sequences on spatial ones.

Fig. 37. The values of the sites in black together with the specification of the rule completely specify the values of the sites in white.

From the definition of cellular automata (7.12) one easily sees that the value of  $\sigma_i(t)$  depends on sites at maximum distance  $r$  from  $i$  at the previous time step. This means that after  $T$  time steps, the value  $\sigma_i(t)$  can depend (at maximum) on sites at distance  $rT$  on both directions, so that the maximum speed for information propagation is  $r$  (i.e. the range of interaction). However, for many CA the actual velocity of information propagation,  $v_p$ , is less than  $r$ , i.e.  $\sigma_i(T)$  depends only on  $v_p T < rT$  sites. By considering a simple construction (see Fig. 36) one can understand that the spatial and temporal entropies are related to each other by the inequality [223]:

$$h^T \leq 2v_p h^S, \tag{7.25}$$

where a good estimate of  $v_p$  can be given in terms of the damage spreading velocity (7.18) [223].

The possible scenario arising from (7.25) can be summarized as follows. One can have “spatial chaos” ( $h^S > 0$ ) in absence of “temporal chaos” ( $h^T = 0$ ), while the existence of “temporal chaos” requires not only a non-zero spatial entropy but also the existence of a finite propagation velocity. This confirms somehow that the classifications of a CA as chaotic in terms of damage spreading velocity and entropy are related to each others.

However, as stressed by Oono and Kohomoto [166], the seemingly natural assumption of calling “turbulent” a cellular automaton for which one has  $h^S > 0$  and  $\langle h^T \rangle > 0$  is not correct in general. This is particularly clear by considering a single direction shift imposed on a “frozen” disordered background. Nevertheless, in spite of this specific counterexample, the attempts based on entropic concepts, for the characterization of the irregular spatial and/or temporal behavior of systems with discrete states, in our opinion, are the most promising ones. In this context Casartelli and coworkers [47,48] introduced the concept of rational partitions in order to define a complexity measure for systems which can be reduced to 1d CA.

Let us conclude this section with a brief discussion and comparison between the unpredictability which characterizes cellular automata evolution with respect to the one encountered in the context of continuous states dynamics, e.g. in coupled map lattices (see Section 4). The latter

indeed seems to be the natural candidate for such a comparison. We limit the discussion to 1-dimensional lattices with  $r = 1$ , i.e. CML and CA with nearest neighbor coupling.

Let us now ask the amount of information we have to specify for knowing all the  $LT$  sites of spatial size  $L (< N)$  and temporal length  $T$ , as shown in Fig. 37. Since both CA and CML are ruled by a local deterministic dynamics one needs to specify the rule of evolution and the values of the  $L + 2(T - 1)$  states at the border, in black in Fig. 37. Basically, one has to specify the initial conditions on the  $L$  sites and the “boundaries”  $\sigma_1(t)$  and  $\sigma_L(t)$  for  $1 < t \leq T$ . But while for CA this specification unambiguously determines the  $LT$  values, for a chaotic CML this is not enough. Indeed, one has to specify the precision,  $\epsilon$ , with which he wants to know the  $LT$  values. Once specified  $\epsilon$ , one knows the necessary initial precision,  $\epsilon_0$ , on the  $L + 2(T - 1)$  sites in black. A conservative estimate gives  $\epsilon_0 \approx \epsilon \exp(-LT \mathcal{H}_{\text{KS}})$ , where  $\mathcal{H}_{\text{KS}}$  is the entropy density defined in Eq. (4.3). This very simple argument suggests that the main difference between CA and continuous systems is the absence of “local” production of information, i.e. in CA the complexity only arises by the spatial propagation of information [97]. Nevertheless, there exist counterexamples in which starting from simple initial configuration complex pattern are generated [222].

From this point of view it is interesting to consider the behavior of certain CMLs which, in spite of their continuous nature, seem to be rather similar to “chaotic” cellular automata. Indeed, it has been found that a class of stable (i.e.  $\lambda < 0$ ) CMLs [69,188] displays an unpredictable dynamics on times exponentially large with the system size. So that in the limit of infinite lattices they are completely unpredictable. Moreover, these CMLs have a finite velocity of propagation for initially localized disturbances (provided that the value of the disturbance was  $O(1)$ ) [188,190]. Recalling the discussion of Section 4.6, we know that this cannot be predicted in terms of the comoving Lyapunov exponents, it is a fully nonlinear phenomenon. The strong analogies with “chaotic” CA have been furtherly explored in Ref. [188], where it has been proposed to classify these CML as large memory cellular automata according to the behavior of their spatial and temporal entropy.

## 8. The characterization of the complexity and system modeling

In the previous sections we discussed the characterization of dynamical behaviors when the evolution laws are known either exactly or with an uncertainty. On the other hand, in experimental investigations only time records of some observable are typically available, and the equation of motions are not known. For the predictability problem, this latter case, at least from a conceptual point of view, can be treated in the same framework of when the evolution laws are known. Indeed, in principle, with the embedding technique one can reconstruct the phase space [209,1,2,114]. Nevertheless, there are rather severe limitations in high dimensional systems [97] and even in low dimensional ones nontrivial features appear in presence of noise [114].

In this section we show that an entropic analysis at different resolution scales allows us for a pragmatic classification of a signal and gives suggestions for modeling of systems. In particular we illustrate, using some examples, how quantities such as the  $\epsilon$ -entropy or the FSLE can display a subtle transition from the large to the small scales. A negative consequence of this is

the difficulty in distinguishing, only from data analysis, a genuine deterministic chaotic system from one with intrinsic randomness [51]. On the other hand, the way the  $\epsilon$ -entropy or finite size Lyapunov exponent depends on the (resolution) scale, allows us for a pragmatic classification of the stochastic or chaotic character of the signal, and this gives some freedom in modeling the system.

### 8.1. How random is a random number generator?

It is rather natural to wonder about the “true character” of the number sequence  $(x_1, x_2, \dots)$  obtained with a (pseudo) random number generator (PRNG) on a computer. One would like to have a sequence with a random character; on the other hand, one is forced to use deterministic algorithms to generate  $(x_1, x_2, \dots)$ . This subsection is mainly based on the paper [116]. A simple and popular PRNG is the multiplicative congruent one [193]:

$$\begin{aligned} z_{n+1} &= N_1 z_n \bmod N_2 \\ x_{n+1} &= z_{n+1}/N_2, \end{aligned} \quad (8.1)$$

with an integer multiplier  $N_1$  and modulus  $N_2$ . The  $\{z_n\}$  are integer numbers and one hopes to generate a sequence of random variables  $\{x_n\}$ , which are uncorrelated and uniformly distributed in the unit interval. A first problem one has to face is the periodic nature of (8.1), because of its discrete character (see Section 7). In practice one wants to fix  $N_1$  and  $N_2$  in such a way to maximize this period. Note that the rule (8.1) can be interpreted as a deterministic dynamical system, i.e.

$$x_{n+1} = N_1 x_n \bmod 1, \quad (8.2)$$

which has a uniform invariant measure and a KS entropy  $h_{\text{KS}} = \lambda = \ln N_1$ . When imposing the integer arithmetics of Eq. (8.1) onto this system, we are, in the language of dynamical systems, considering an unstable periodic orbit of Eq. (8.2), with the particular constraint that, in order to achieve the period  $N_2 - 1$  (i.e. all integers  $< N_2$  should belong to the orbit of Eq. (8.1)) it has to contain all values  $k/N_2$ , with  $k = 1, 2, \dots, N_2 - 1$ . Since the natural invariant measure of Eq. (8.2) is uniform, such an orbit represents the measure of a chaotic solution in an optimal way. Every sequence of a PRNG is characterized by two quantities: its period  $\mathcal{T}$  and its positive Lyapunov exponent  $\lambda$ , which is identical to the entropy of a chaotic orbit of the equivalent dynamical system. Of course a good random number generator has a very large period, and as large as possible entropy.

It is natural to ask how this apparent randomness can be reconciled with the facts that (a) the PRNG is a deterministic dynamical systems (b) it is a discrete state system.

If the period is long enough on shorter times one has to face only point (a). In the following we discuss this point in terms of the behavior of the  $\epsilon$ -entropy,  $h(\epsilon)$  (see Section 3.5). It seems rather reasonable to think that at a high resolution, i.e.  $\epsilon \leq 1/N_1$ , one should realize the true deterministic chaotic nature of the system and, therefore,  $h(\epsilon) \simeq h_{\text{KS}} = \ln N_1$ . On the other hand for  $\epsilon \geq 1/N_1$  one expects to observe the “apparent random” behavior of the system, i.e.  $h(\epsilon) \sim \ln(1/\epsilon)$ .

When the spatial resolution is high enough so that every point of this periodic orbit is characterized by its own symbol, then, for arbitrary block length  $m$ , one has a finite number

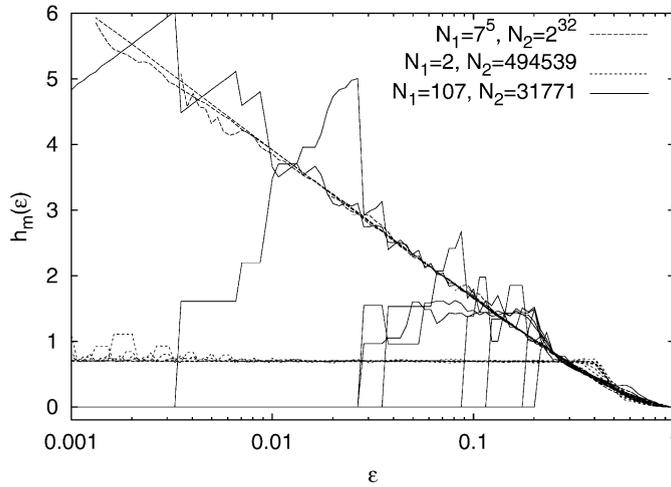


Fig. 38. The  $\epsilon$ -entropies,  $h_m(\epsilon)$ , at varying the embedding dimension  $m$  for the multiplicative congruential random number generator, Eq. (8.1), for different choices of  $N_1$  and  $N_2$ . This figure has been taken from Ref. [116].

of  $m$ -words whose probabilities are different from 0. Therefore, the block entropy  $H_m$  (3.44) is  $m$ -independent and  $h_m = 0$ .

In Fig. 38 it is shown the behavior of  $h_m(\epsilon)$ , computed on sequences of length 60000 of the PRNG with three different pairs  $(N_1, N_2)$  chosen to be  $(7^5, 2^{32})$ ,  $(2, 494539)$ , and  $(107, 31771)$ . The first one is optimal and no deviation from the stochastic behavior is visible. The second one has a small pseudo-entropy, and this is seen by the saturation of all  $h_m(\epsilon)$  at  $\ln N_1 = \ln 2$ , and the last one has large entropy but a rather short period, so that all  $h_m(\epsilon)$  drop to zero for some  $\epsilon_m$ , where  $\epsilon_m$  becomes dramatically larger for increasing  $m$  (strong fluctuations arise from the fact that data are confined to a grid of spacing  $1/31771$ ).

## 8.2. High-dimensional systems

Now we discuss high-dimensional systems that show non-trivial behavior at varying the resolution scales. Olbrich et al. [165] analyzed an open flow system described by unidirectionally coupled map lattice:

$$x_j(t+1) = (1 - \sigma)f(x_{j+1}(t)) + \sigma x_j(t), \quad (8.3)$$

where  $j = 1, \dots, N$  denotes the site of a lattice of size  $N$ ,  $t$  the discrete time and  $\sigma$  the coupling strength. A detailed numerical study (also supported by analytical arguments) of the  $\epsilon$ -entropy  $h_m(\epsilon)$  at different  $\epsilon$ , in the limit of small coupling, gives the following scale-dependent scenario: for  $1 \geq \epsilon \geq \sigma$  there is a plateau  $h(\epsilon) \simeq \lambda_s$  where  $\lambda_s$  is the Lyapunov exponent of the single map  $x(t+1) = f(x(t))$ . For  $\sigma \geq \epsilon \geq \sigma^2$  another plateau appears at  $h(\epsilon) \simeq 2\lambda_s$ , and so on: for  $\sigma^{n-1} \geq \epsilon \geq \sigma^n$  one has  $h(\epsilon) \simeq n\lambda_s$  (see Fig. 39).

Similar results hold for the correlation dimension which increases step by step as the resolution increases, showing that the high-dimensionality of the system becomes evident only as  $\epsilon \rightarrow 0$ .

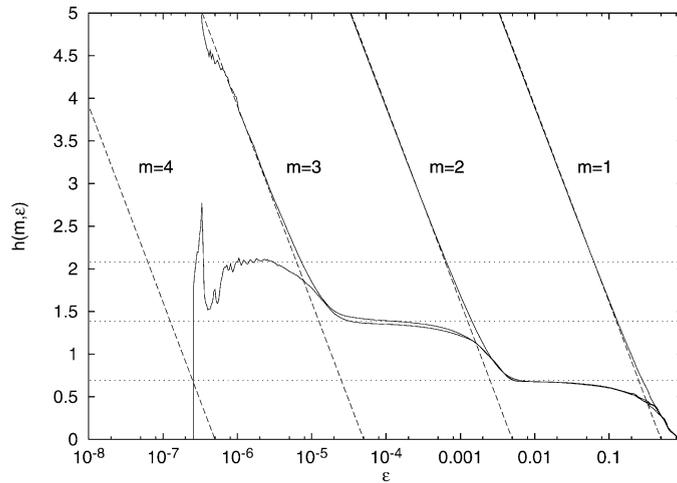


Fig. 39.  $h_m(\epsilon)$  for the system (8.3) where  $f(x)=2|1/2 - |x - 1/2||$  is the tent map and  $\sigma=0.01$ . The horizontal lines indicate the entropy steps which appears at decreasing  $\epsilon$ . The  $\epsilon$ -entropy is computed with the Grassberger–Procaccia method [92]. For further details see Ref. [165].

Therefore one understands that the dynamics at different scales is basically ruled by a hierarchy of low-dimensional systems whose “effective” dimension  $n_{\text{eff}}(\epsilon)$  increases as  $\epsilon$  decreases [165]:

$$n_{\text{eff}}(\epsilon) \sim \left[ \frac{\ln(1/\epsilon)}{\ln(1/\sigma)} \right], \tag{8.4}$$

where [...] indicates the integer part. In addition, for a given resolution  $\epsilon$ , it is possible to find a suitable low-dimensional noisy system (depending on  $\epsilon$ ) which is able to mimic  $x_1(t)$  given by Eq. (8.3). It is interesting to note that, on an extended range of values of  $\epsilon$  ( $\epsilon \geq \sigma^N$ ),  $h(\epsilon)$  can be roughly approximated as log-periodic fluctuations around

$$h(\epsilon) \sim \ln \frac{1}{\epsilon} \tag{8.5}$$

i.e. the typical behavior of a stochastic process. Of course, for  $\epsilon \leq \sigma^N$  one has to realize that the system is deterministic and  $h(\epsilon) = O(N\lambda_s)$ .

Let us now briefly reconsider the issue of the macroscopic chaos, discussed in Section 4.7. The main result can be summarized as follows:

- at small  $\epsilon$  ( $\ll 1/\sqrt{N}$ ), where  $N$  is the number of elements, one recovers the “microscopic” Lyapunov exponent<sup>2</sup>, i.e.  $\lambda(\epsilon) \approx \lambda_{\text{micro}}$
- at large  $\epsilon$  ( $\gg 1/\sqrt{N}$ ) one observes another plateau  $\lambda(\epsilon) \approx \lambda_{\text{macro}}$  which can be much smaller than the microscopic one.

<sup>2</sup>From hereafter we use the same symbol  $\epsilon$  both for the FSLE and the  $\epsilon$ -entropy in order to make a direct comparison between the two quantities.

The emerging scenario is that at a coarse-grained level, i.e.  $\epsilon \gg 1/\sqrt{N}$ , the system can be described by an “effective” hydro-dynamical equation (which in some cases can be low-dimensional), while the “true” high-dimensional character appears only at very high resolution, i.e.

$$\epsilon \leq \epsilon_c = O\left(\frac{1}{\sqrt{N}}\right).$$

### 8.3. Diffusion in deterministic systems and Brownian motion

Consider the following map which generates a diffusive behavior on the large scales [200]:

$$x_{t+1} = [x_t] + F(x_t - [x_t]), \tag{8.6}$$

where  $[x_t]$  indicates the integer part of  $x_t$  and  $F(y)$  is given by:

$$F(y) = \begin{cases} (2 + \alpha)y & \text{if } y \in [0, 1/2], \\ (2 + \alpha)y - (1 + \alpha) & \text{if } y \in [1/2, 1]. \end{cases} \tag{8.7}$$

The largest Lyapunov exponent  $\lambda$  can be obtained immediately:  $\lambda = \ln |F'|$ , with  $F' = dF/dy = 2 + \alpha$ . One expects the following scenario for  $h(\epsilon)$ :

$$h(\epsilon) \approx \lambda \quad \text{for } \epsilon < 1, \tag{8.8}$$

$$h(\epsilon) \propto \frac{D}{\epsilon^2} \quad \text{for } \epsilon > 1, \tag{8.9}$$

where  $D$  is the diffusion coefficient, i.e.

$$\langle (x_t - x_0)^2 \rangle \approx 2Dt \quad \text{for large } t. \tag{8.10}$$

Consider now a stochastic system, namely a noisy map

$$x_{t+1} = [x_t] + G(x_t - [x_t]) + \sigma \eta_t, \tag{8.11}$$

where  $G(y)$ , as shown in Fig. 40, is a piecewise linear map which approximates the map  $F(y)$ , and  $\eta_t$  is a stochastic process uniformly distributed in the interval  $[-1, 1]$ , and no correlation in time. When  $|dG/dy| < 1$ , as is the case we consider, the map (8.11), in the absence of noise, gives a non-chaotic time evolution.

Now we compare the finite size Lyapunov exponent for the chaotic map (8.6) and for the noisy one (8.11). In the latter the FSLE has been computed using two different realizations of the noise. In Fig. 41 we show  $\lambda(\epsilon)$  versus  $\epsilon$  for the two cases. The two curves are practically indistinguishable in the region  $\epsilon > \sigma$ . The differences appear only at very small scales  $\epsilon < \sigma$  where one has a  $\lambda(\epsilon)$  which grows with  $\epsilon$  for the noisy case, remaining at the same value for the chaotic deterministic case.

Both the FSLE and the  $(\epsilon, \tau)$ -entropy analysis show that we can distinguish three different regimes observing the dynamics of (8.11) on different length scales. On the large length scales  $\epsilon > 1$  we observe diffusive behavior in both models. On length scales  $\sigma < \epsilon < 1$  both models show chaotic deterministic behavior, because the entropy and the FSLE are independent of  $\epsilon$

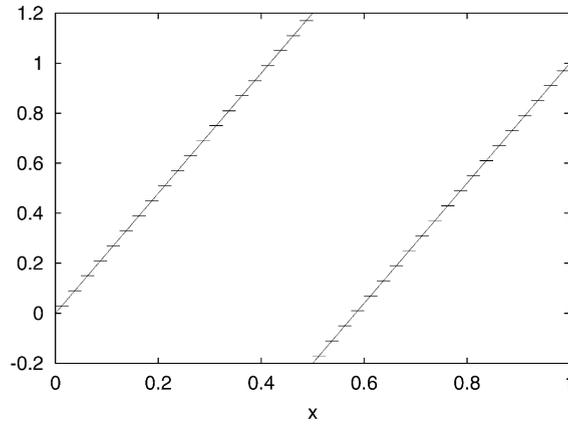


Fig. 40. The map  $F(x)$  (8.7) for  $\alpha = 0.4$  is shown with superimposed the approximating (regular) map  $G(x)$  (8.11) obtained by using 40 intervals of slope 0.

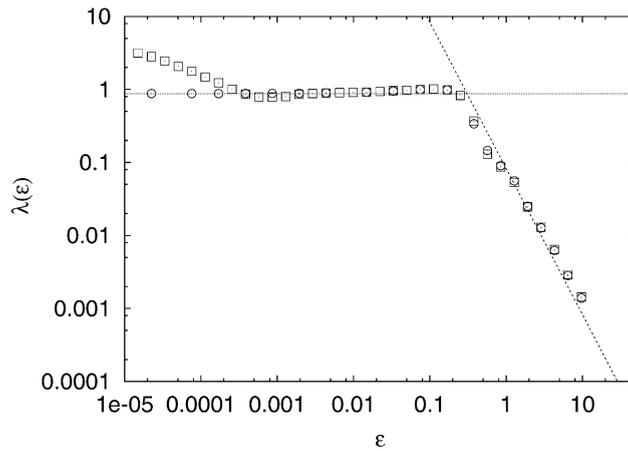


Fig. 41.  $\lambda(\epsilon)$  versus  $\epsilon$  obtained with the map  $F(y)$  (8.7) with  $\alpha = 0.4$  ( $\circ$ ) and with the noisy (regular) map ( $\square$ ) (8.11) with 10 000 intervals of slope 0.9 and  $\sigma = 10^{-4}$ . The straight lines indicates the Lyapunov exponent  $\lambda = \ln 2.4$  and the diffusive behavior  $\lambda(\epsilon) \sim \epsilon^{-2}$ .

and larger than zero. Finally, on the smallest length scales  $\epsilon < \sigma$  we see stochastic behavior for the system (8.11) while the system (8.6) still shows chaotic behavior.

#### 8.4. On the distinction between chaos and noise

The above examples show that the distinction between chaos and noise can be a high non-trivial task, which makes sense only in very peculiar cases, e.g., very low-dimensional systems. Nevertheless, even in this case, the entropic analysis can be unable to recognize the “true” character of the system due to the lack of resolution. Again, the comparison between the

diffusive map (8.6) and the noisy map (8.11) is an example of these difficulties. For  $\sigma \leq \epsilon \leq 1$  both the system (8.6) and (8.11), in spite of their “true” character, will be classified as chaotic, while for  $\epsilon \geq 1$  both can be considered as stochastic.

In high-dimensional chaotic systems, with  $N$  degrees of freedom, one has typically  $h(\epsilon) = h_{KS} \sim O(N)$  for  $\epsilon \leq \epsilon_c$  (where  $\epsilon_c \rightarrow 0$  as  $N \rightarrow \infty$ ) while for  $\epsilon \geq \epsilon_c$ ,  $h(\epsilon)$  decreases, often with a power law [89]. Since also in some stochastic processes the  $\epsilon$ -entropy obeys a power law, this can be a source of confusion.

These kind of problems are not abstract ones, as a recent debate on “microscopic chaos” demonstrates [90,72,98]. The detection of microscopic chaos by data analysis has been recently addressed in a work of Gaspard et al. [90]. These authors, from an entropic analysis of an ingenious experiment on the position of a Brownian particle in a liquid, claim to give an empirical evidence for microscopic chaos. In other words, they state that the diffusive behavior observed for a Brownian particle is the consequence of chaos at a molecular level. Their work can be briefly summarized as follows: from a long ( $\approx 1.5 \times 10^5$  data) record of the position of a Brownian particle they compute the  $\epsilon$ -entropy with the Cohen–Procaccia method [61] (Section 3) from which they obtain

$$h(\epsilon) \sim \frac{D}{\epsilon^2}, \quad (8.12)$$

where  $D$  is the diffusion coefficient. Then, *assuming* that the system is deterministic, and making use of the inequality  $h(\epsilon > 0) \leq h_{KS}$ , they conclude that the system is chaotic. However, their result does not give a direct evidence that the system is deterministic and chaotic. Indeed, the power law (8.12) can be produced with different mechanisms:

- (1) a genuine chaotic system with diffusive behavior, as the map (8.7);
- (2) a nonchaotic system with some noise, as the map (8.11), or a genuine Brownian system;
- (3) a deterministic linear non-chaotic system with many degrees of freedom (see for instance [158]);
- (4) a “complicated” non-chaotic system as the Ehrenfest wind-tree model where a particle diffuses in a plane due to collisions with randomly placed, fixed oriented square scatters, as discussed by Cohen et al. [72] in their comment to Ref. [90].

It seems to us that the weak points of the analysis in Ref. [90] are:

- (a) the explicit assumption that the system is deterministic;
- (b) the limited number of data points and therefore limitations in both resolution and block length.

The point (a) is crucial, without this assumption (even with an enormous data set) it is not possible to distinguish between (1) and (2). One has to say that in the cases (3) and (4) at least in principle it is possible to understand that the systems are “trivial” (i.e. not chaotic) but for this one has to use a huge number of data. For example, Cohen et al. [72] estimated that in order to distinguish between (1) and (4) using realistic parameters of a typical liquid, the number of data points required has to be at least  $\sim 10^{34}$ .

Concluding, we have the apparently paradoxical result that “complexity” helps in the construction of models. Basically, in the case in which one has a variety of behaviors at varying the scale resolution, there is a certain freedom on the choice of the model to adopt. In Section 8.3 one can see that, for some systems, the behavior at large scales can be realized both with chaotic deterministic models or suitable stochastic processes. From a pragmatic point of view, the fact that in certain stochastic processes  $h(\epsilon) \sim \epsilon^{-\alpha}$  can be indeed extremely useful for modeling such high-dimensional systems. Perhaps, the most relevant case in which one can use this freedom in modeling is the fully developed turbulence whose non-infinitesimal (the so-called inertial range) properties can be successfully mimicked in terms of multifractal stochastic process (see Refs. [29,4], Section 5.5 and Appendix D).

## 9. Concluding remarks

The guideline of this review has been how to interpret the different aspects of the predictability of a system as a way to characterize its complexity.

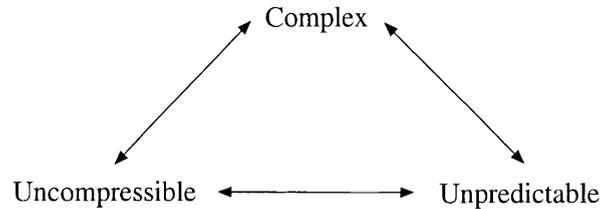
We have discussed the relation between the Kolmogorov–Sinai entropy and the algorithmic complexity (Section 2). As clearly exposed in the seminal works of Alekseev and Yakobson [5] and Ford [79,80], the time sequences generated by a system with sensitive dependence on initial conditions have non-zero algorithmic complexity. A relation exists between the maximal compression of a sequence and its KS-entropy. Therefore, one can give a definition of complexity, without referring to a specific description, as an intrinsic property of the system.

In the presence of intrinsic randomness (Section 6.3), one can introduce two different Lyapunov exponents,  $\lambda_\sigma$  in the case of trajectories with the same realization of noise and  $K_\sigma$  for different realizations. In general,  $\lambda_\sigma$  and  $K_\sigma$  do not coincide and characterize different aspects of the system. Both quantities have their own relevance, the comparison between  $\lambda_\sigma$  and  $K_\sigma$  has shown to be useful in the understanding of apparently intricate phenomena, such as noise-induced order and noise-induced instability.

As an example of system with many degrees of freedom and characteristic times scales, we investigated fully developed turbulence (Section 5). In this case the Lyapunov exponent and the KS-entropy are somehow of limited relevance because they only characterize small scales properties. On the other hand, there exist suitable generalizations—the finite size Lyapunov exponent,  $\lambda(\epsilon)$ , and  $\epsilon$ -entropy,  $h(\epsilon)$ —which characterize the predictability properties at different scales. The scaling of the predictability time with the resolution  $\epsilon$ ,  $\lambda(\epsilon) \sim \epsilon^{-2}$ , has an algorithmic correspondence in the behavior of the  $\epsilon$ -entropy of the signal measured in one point,  $h(\epsilon) \sim \epsilon^{-3}$ . In the words of Lorenz, one can say that the butterfly effect is not so terrible for  $\epsilon$ -resolution in the inertial range.

Complexity in a system can also manifest in the spatial properties as, for example, in open flows with convective chaos but with negative Lyapunov exponents (Section 4). The presence of convective chaos implies a sensitivity on the boundary conditions. An uncertainty,  $\delta x_0$ , on the boundary condition is exponentially amplified with the distance,  $n$ , from the boundary as  $\delta x_n \sim \delta x_0 e^{\Gamma n}$ . The “spatial” Lyapunov exponent  $\Gamma$  is related with the comoving Lyapunov exponent and gives a characterization of the spatial “complexity”.

The study of these different aspects of predictability constitutes a useful method for a quantitative characterization of “complexity”, suggesting the following equivalences:



The above point of view, based on dynamical systems and information theory, quantifies the complexity of a sequence considering each symbol relevant but it does not capture the structural level. Let us clarify this point with the following example. A binary sequence obtained with a coin tossing is, from the point of view adopted in this review, complex since it cannot be compressed (i.e. it is unpredictable). On the other hand such a sequence is somehow trivial, i.e. with low “organizational” complexity. It would be important to introduce a quantitative measure of this intuitive idea. The progresses of the research on this intriguing and difficult issue are still rather slow. We just mention some of the most promising proposals as the logical depth [21] and the sophistication [130], see Ref. [13].

As a final question one can wonder what can one learn by the presented material for practical prediction problems (e.g. weather forecast). The main lesson concerns the framework and limitations about the possible well-posed questions in prediction and modeling. The first relevant fact, now well established, is that in presence of deterministic chaos one cannot hope to reach an arbitrary accuracy in prediction by merely refining the model. A less recognized aspect is that the Lyapunov exponent is usually not sufficient to characterize the limits of predictability in real situations. An appropriate generalization of the Lyapunov exponent is necessary to account for the large-scale properties. Moreover in weather forecast the predictability time, which is typically of 5 days, may be as little as 2 days or as much as 10 days [211]. Thus, simply quoting an average value does not give a satisfactory answer. At a more conceptual level, one has severe limitations in distinguish between deterministic or stochastic nature of systems displaying complex behavior. This implies a certain freedom in the choice of the details of the model, in particular whether to adopt a deterministic or a stochastic model.

### Note added in proof

After completing this review, we became aware of Ref. [231] where a generalization of the Lyapunov exponent and the intermittency are discussed.

### Acknowledgements

We are particularly grateful to A. Celani, for the long collaboration on the subject here discussed and for his contributions at a first stage of this review.

We thank M. Abel, K.H. Andersen, A. Baldassarri, E. Aurell, F. Bagnoli, R. Baviera, A. Bigazzi, R. Benzi, L. Biferale, M. Casartelli, P. Castiglione, A. Crisanti, D.P. Eckmann, U. Frisch, P. Grassberger, K. Kaneko, H. Kantz, M.H. Jensen, G. Lacorata, R. Livi, V. Loreto, R. Mantegna, G. Mantica, U. Marini Bettolo Marconi, A. Mazzino, P. Muratore-Ginanneschi, E. Olbrich, G. Parisi, R. Pasmantier, M. Pasquini, A. Pikovsky, A. Politi, A. Provenzale, A. Puglisi, S. Ruffo, M. Serva, A. Torcini, M. Vergassola, D. Vergni and E. Zambianchi for collaborations, correspondence and discussions in the last years.

We are grateful to F. Benatti, P. Castiglione, A. Politi and A. Torcini for useful remarks and detailed comments on the manuscript.

We acknowledge F. di Carmine, R.T. Lampioni, B. Marani and I. Poeta for their friendly and warm encouragement.

## Appendix A. On the computation of the finite size Lyapunov exponent

This appendix is devoted to the methods for computing the finite size Lyapunov exponent. As we will see there are mainly three possible ways to compute the FSLE. Let us start with a modification of the standard technique for computing the largest Lyapunov exponent [20,221].

Suppose that one has to integrate the equations of motion of a system on a computer. After a long integration time, in order for the motion to settle onto the attractor of the system, we introduce a very small perturbation, i.e. we consider the “reference” trajectory  $\mathbf{x}(0)$ , which is supposed to be on the attractor, and generate a “perturbed” trajectory starting from  $\mathbf{x}'(0) = \mathbf{x}(0) + \delta\mathbf{x}(0)$ . We need the perturbation to be initially very small in some chosen norm  $\delta(t=0) = \|\delta\mathbf{x}(t=0)\| = \delta_{\min} \ll 1$ . Then, in order to study the perturbation growth through different scales, one defines a set of thresholds  $\delta_n$ , e.g.  $\delta_n = \delta_0 r^n$  with  $1 \gg \delta_0 \gg \delta_{\min}$  and  $n = 0, \dots, N$ . To avoid saturation on the maximum allowed separation (i.e. the attractor size) one has to choose  $\delta_N < \langle \|\mathbf{x} - \mathbf{y}\| \rangle_\mu$  with  $\mathbf{x}, \mathbf{y}$  generic points on the attractor. Note that  $r$  should be larger than 1 but not too large in order to avoid interferences of different length scales. Typically, one chooses  $r = 2$  or  $r = \sqrt{2}$ .

In order to measure the perturbation growth rate at scale  $\delta_n$ , one lets the system to evolve from  $\delta_{\min}$  up to the desired scale  $\delta_n$  ensuring the perturbation to be on the attractor and aligned along the maximally expanding direction. After  $\delta_n$  is reached, one computes the first time,  $\tau_1(\delta_n, r)$ , to reach the following threshold,  $\delta_{n+1}$ , and after that the perturbation is rescaled to  $\delta_n$ , keeping the direction  $\mathbf{x}' - \mathbf{x}$  constant. This procedure is repeated  $\mathcal{N}$  times for each thresholds obtaining the set of the doubling times  $\{\tau_i(\delta_n, r)\}$  for  $i = 1, \dots, \mathcal{N}$  error-doubling experiments. Now if we introduce the effective doubling rates

$$\gamma_i(\delta_n, r) = \frac{1}{\tau_i(\delta_n, r)} \ln r, \quad (\text{A.1})$$

we can define their time averages as the effective LEs on the scale  $\delta_n$ . Therefore, we have

$$\lambda(\delta_n) = \langle \gamma(\delta_n, r) \rangle_t = \frac{1}{T} \int_0^T \gamma dt = \frac{\sum_i \gamma_i \tau_i}{\sum_i \tau_i} = \frac{1}{\langle \tau(\delta_n, r) \rangle_e} \ln r, \quad (\text{A.2})$$

where  $\langle \tau(\delta_n, r) \rangle_e = \sum \tau_i / \mathcal{N}$  and  $T = \sum_i \tau_i$ .

To obtain Eq. (A.2) we assumed the distance between the two trajectories to be continuous in time. This is not true for maps or for discrete sampling in time and the method has to be slightly modified. In this case the doubling time,  $\tau(\delta_n, r)$ , is defined as the minimum time such that  $\delta(\tau) \geq r\delta_n$ . Because now  $\delta(\tau)$  is a fluctuating quantity, from (A.2) we have

$$\lambda(\delta_n) = \frac{1}{\langle \tau(\delta_n, r) \rangle_e} \left\langle \ln \left( \frac{\delta(\tau(\delta_n, r))}{\delta_n} \right) \right\rangle_e. \quad (\text{A.3})$$

Let us stress some points.

The computation of the FSLE is not more expensive than the one of the Lyapunov exponent by standard algorithm. One has simply to integrate two copies of the system (or two different systems for second kind predictability) and this can be done without particular problems.

At difference with  $\lambda$ ,  $\lambda(\delta)$  may also depend on the norm one chooses. This fact, apparently disturbing, is however physically reasonable: when one looks at the non-linear regime, for instance, for the predictability problem the answer may depend on the observable considered. A similar problem appears in infinite-dimensional system where the norms are not equivalent [127].

A possible problem with the above-described method is that we have implicitly assumed that the statistically stationary state of the system is homogeneous with respect to finite perturbations. Actually one may plausibly expect the attractor to be fractal, i.e., not at all equally dense at all distances, this may cause an incorrect sampling of the doubling times at large  $\delta_n$ .

A possible way to overcome such a problem is to compute the FSLE avoiding to rescale the perturbation at finite  $\delta_n$ . This can be accomplished by the following modification of the previous method. One defines the thresholds  $\{\delta_n\}$  and initializes the perturbation at  $\delta_{\min} \ll \delta_0$  as before. Then one lets the system to reach the first threshold,  $\delta_0$ . Hence, one starts to measure the doubling time  $\tau(\delta_n, r)$  following the perturbation growth from  $\delta_0$  up to  $\delta_N$ . In practice, one registers the time  $\tau(\delta_n, r)$  for going from  $\delta_n$  to  $\delta_{n+1}$  for each value of  $n$ . The evolution of the error from the initial value  $\delta_{\min}$  to the largest threshold  $\delta_N$  carries out a single error-doubling experiment. When the largest threshold,  $\delta_N$  has been reached the “perturbed” trajectory is rescaled at the initial distance,  $\delta_{\min}$ , with respect to the “reference” trajectory and one starts another experiment measuring a second set of doubling times,  $\{\tau_2(\delta_n, r)\}$ . The procedure is then repeated  $\mathcal{N}$  times to have statistics. In this way one obtains the set of the doubling times  $\{\tau_i(\delta_n, r)\}$  for  $i = 1, \dots, \mathcal{N}$ . The FSLE is finally obtained by using Eq. (A.2) or (A.3), which are accurate also in this case, according to the continuous time and discrete time nature of the system, respectively. One understands that with this method, since finite perturbations are realized by the dynamics (i.e. the perturbed trajectory is on the attractor) and not imposed by hand, the problems related to the attractor inhomogeneity are not present.

In any case, in most numerical experiments, one does not find significant differences between the two numerical methods.

A further possibility to compute the FSLE is to remove the threshold condition used for defining  $\tau(\delta_n, r)$  and simply compute the average error growth rate at every time step. In other words, at every time step  $\Delta t$  in the integration, the perturbed trajectory  $\mathbf{x}'(t)$  is rescaled to the original distance  $\delta$ , keeping the direction  $\mathbf{x} - \mathbf{x}'$  constant. The FSLE is then obtained by the

average of the one-step exponential divergence:

$$\lambda(\delta) = \frac{1}{\Delta t} \left\langle \ln \left( \frac{\|\delta \mathbf{x}(t + \Delta t)\|}{\|\delta \mathbf{x}(t)\|} \right) \right\rangle_t, \tag{A.4}$$

which, if non-negative, is equivalent to the definition (A.2). Let us note that the above procedure is nothing but the finite scale version of the usual algorithm of Benettin et al. [20] for the LE. The one-step method (A.4) can be, in principle, generalized to compute the sub-leading finite-size Lyapunov exponent following the standard ortho-normalization method [20]. One introduces  $k$  perturbed trajectories  $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(k)}$  each at distance  $\delta$  from  $\mathbf{x}$  such that  $\mathbf{x}^{(k)} - \mathbf{x}$  are orthogonal to each others. At every time step, any difference  $\mathbf{x}^{(k)} - \mathbf{x}$  is rescaled at the original value and orthogonalized, while the corresponding finite-size Lyapunov exponent is accumulated according to (A.4).

Here we have again the problem of the implicitly assumed homogeneity of the attractor, and also a problem of isotropy when we re-orthogonalize the perturbations. We note that this could be a more serious problem.

### Appendix B. The multifractal model of turbulence

The multifractal model of turbulence [177,172,84] assumes that the velocity has a local scale-invariance, i.e. it does not have a unique scaling exponent  $h$  such that  $\delta v_\ell \sim \ell^h$ , but a continuous spectrum of exponents, each of which belonging to a given fractal set. In other words, in the inertial range one has

$$\delta v_\ell(x) \sim \ell^h, \tag{B.1}$$

if  $x \in S_h$ , and  $S_h$  is a fractal set with dimension  $D(h)$  and  $h \in (h_{\min}, h_{\max})$ . The probability to observe a given scaling exponent  $h$  at the scale  $\ell$  is thus  $P_\ell(h) \sim \ell^{3-D(h)}$ , so the scaling of the structure function assumes the form

$$S_p(\ell) = \langle \delta v_\ell^p \rangle \sim \int_{h_{\min}}^{h_{\max}} \ell^{hp} \ell^{3-D(h)} dh \sim \ell^{\zeta_p}, \tag{B.2}$$

where in the last equality, being  $\ell \ll 1$ , a steepest descent estimation gives

$$\zeta_p = \min_h \{hp + 3 - D(h)\} = h^*p + 3 - D(h^*), \tag{B.3}$$

where  $h^* = h^*(p)$  is the solution of the equation  $D'(h^*(p)) = p$ . The Kolmogorov “ $\frac{4}{5}$ ” law [84]

$$S_3(\ell) = -\frac{4}{5}\epsilon\ell \tag{B.4}$$

imposes  $\zeta_3 = 1$  which implies that

$$3h + 2 - D(h) \leq 0, \tag{B.5}$$

the equality is realized by  $h^*(3)$ . The Kolmogorov similarity theory corresponds to the case of only one singularity exponent  $h = \frac{1}{3}$  with  $D(h = \frac{1}{3}) = 3$ .

A non-trivial consequence of the intermittency in the turbulent cascade is the fluctuations of the dissipative scale which implies the existence of an intermediate region between the inertial

and dissipative range [83]. The local dissipative scale  $\ell_D$  is determined by imposing the effective Reynolds number to be of order unity:

$$Re(\ell_D) = \frac{\delta v_D \ell_D}{\nu} \sim 1, \quad (\text{B.6})$$

therefore the dependence of  $\ell_D$  on  $h$  is thus

$$\ell_D(h) \sim LRe^{-1/(1+h)} \quad (\text{B.7})$$

where  $Re = Re(L)$  is the large-scale Reynolds number. The fluctuations of  $\ell_D$  modifies the evaluation of the structure functions (B.2): for a given  $\ell$ , the saddle-point evaluation of (B.2) remains unchanged if, for the selected exponent  $h^*(p)$ , one has  $\ell_D(h^*(p)) < \ell$ . If, on the contrary, the selected exponent is such that  $\ell_D(h^*(p)) > \ell$  the saddle-point evaluation is not consistent, because at scale  $\ell$  the power-law scaling (B.1) is no longer valid. In this intermediate dissipation range [83] the integral in (B.2) is dominated by the smallest acceptable scaling exponent  $h(\ell)$  given by inverting (B.7), and the structure function of order  $p$  a pseudo-algebraic behavior, i.e. a power law with exponent  $ph(\ell) + 3 - D(h(\ell))$  which depends on the scale  $\ell$ . Taking into account the fluctuations of the dissipative range, one has for the structure functions

$$S_p(\ell) \sim \begin{cases} \ell^{\zeta_p} & \text{if } \ell_D(h^*(p)) < \ell, \\ \ell^{h(\ell)p+3-D(h(\ell))} & \text{if } \ell_D(h_{\min}) < \ell < \ell_D(h^*(p)). \end{cases} \quad (\text{B.8})$$

A simple calculation [83,84] shows that it is possible to find a universal description valid both in the inertial and in the intermediate dissipative ranges. Let us discuss this point for the energy spectrum  $E(k)$ . Introducing the rescaled variables

$$F(\theta) = \frac{\ln E(k)}{\ln Re} \quad \text{and} \quad \theta = \frac{\ln k}{\ln Re} \quad (\text{B.9})$$

one obtains the following behavior:

$$F(\theta) = \begin{cases} -(1 + \zeta_2)\theta & \text{for } \theta < \frac{1}{1 + h^*(2)}, \\ -2 - 2\theta + \theta D(\theta^{-1} - 1) & \text{for } \frac{1}{1 + h^*(2)} < \theta < \frac{1}{1 + h_{\min}}. \end{cases} \quad (\text{B.10})$$

The prediction of the multifractal model is that  $\ln E(k)/\ln Re$  is an universal function of  $\ln k/\ln Re$ . This is in contrast with the usual scaling hypothesis according which  $\ln E(k)$  should be a universal function of  $\ln(k/k_D)$ . The multifractal universality has been tested by collapsing energy spectra obtained from turbulent flow in a wide range of  $Re$  [88], see also [30].

### Appendix C. How to compute the $\epsilon$ -entropy with exit times

The approach based on exit times differs from the usual one (see Section 3.5) in the procedure to construct the coding sequence of the signal at a given level of accuracy [3]. Indeed an efficient coding procedure reduces the redundancy and improves the quality of the results. The method here discussed is particularly suited for computing the  $\epsilon$ -entropy in processes in which many scales are excited as, e.g., in turbulence [3,4].

The coding of a signal,  $x(t)$ , by the exit-time approach is the following. Given a reference starting time  $t = t_0$ , one measures the first exit time,  $t_1$ , from a cell of size  $\epsilon$ , i.e. the first time necessary to have  $|x(t_0 + t_1) - x(t_0)| \geq \epsilon/2$ . Then one restarts from the time  $t = t_0 + t_1$  to look for the next exit time  $t_2$ , i.e., the first time such that  $|x(t_0 + t_1 + t_2) - x(t_0 + t_1)| \geq \epsilon/2$  and so on. Finally one obtains a sequence of exit times,  $\{t_i(\epsilon)\}$ , and one also records the labels  $k_i = \pm 1$ , which distinguish the direction of the exit (up or down out of a cell).

At the end of this construction, the trajectory is coded without ambiguity, with the required accuracy  $\epsilon$ , by the sequence  $\{(t_i, k_i), i = 1, \dots, M\}$ , where  $M$  is the total number of exit-time events observed during the total time  $T$ . Now, performing a coarse-graining of the possible values assumed by  $t(\epsilon)$  with a resolution time  $\tau_r$ , we accomplish the goal of obtaining a symbolic sequence. One now studies the “exit-time words”,  $\Omega_i^N$ , of various lengths  $n$ :  $\Omega_i^N(\epsilon, \tau_r) = ((\eta_i, k_i), (\eta_{i+1}, k_{i+1}), \dots, (\eta_{i+N-1}, k_{i+N-1}))$ , where  $\eta_j$  labels the cell (of width  $\tau_r$ ) containing the exit time  $t_j$ . From the probabilities of these words one evaluates the block entropies (2.18) at the given time resolution,  $H_N^\Omega(\epsilon, \tau_r)$ , and then the exit time  $(\epsilon, \tau_r)$ -entropies:

$$h^\Omega(\epsilon, \tau_r) = \lim_{N \rightarrow \infty} H_{N+1}^\Omega(\epsilon, \tau_r) - H_N^\Omega(\epsilon, \tau_r). \tag{C.1}$$

The limit of infinite time-resolution gives us the  $\epsilon$ -entropy *per exit*, i.e.

$$h^\Omega(\epsilon) = \lim_{\tau_r \rightarrow 0} h^\Omega(\epsilon, \tau_r). \tag{C.2}$$

The link between  $h^\Omega(\epsilon)$  and the  $\epsilon$ -entropy (3.45) can be obtained as follows. We note that there is a one-to-one correspondence between the (exit-time)-histories and the  $(\epsilon, \tau)$ -histories (in the limit  $\tau \rightarrow 0$ ) originating from a given  $\epsilon$ -cell. The Shannon–McMillan theorem [121] assures that the number of the typical  $(\epsilon, \tau)$ -histories of length  $N$ ,  $\mathcal{N}(\epsilon, N)$ , is such that  $\ln \mathcal{N}(\epsilon, N) \simeq h(\epsilon)N\tau = h(\epsilon)T$ . For the number of typical (exit-time)-histories of length  $M$ ,  $\mathcal{M}(\epsilon, M)$ , we have  $\ln \mathcal{M}(\epsilon, M) \simeq h^\Omega(\epsilon)M$ . If we consider  $T = M \langle t(\epsilon) \rangle$ , where  $\langle t(\epsilon) \rangle = 1/M \sum_{i=1}^M t_i$ , we must obtain the same number of (very long) histories. Therefore, from the relation  $M = T / \langle t(\epsilon) \rangle$  we obtain finally for the  $\epsilon$ -entropy per unit time:

$$h(\epsilon) = \frac{M h^\Omega(\epsilon)}{T} = \frac{h^\Omega(\epsilon)}{\langle t(\epsilon) \rangle} \simeq \frac{h^\Omega(\epsilon, \tau_r)}{\langle t(\epsilon) \rangle}, \tag{C.3}$$

where the last equality is valid at least for small enough  $\tau_r$  [3]. In most of the cases, the leading  $\epsilon$ -contribution to  $h(\epsilon)$  in (C.3) is given by the mean exit time  $\langle t(\epsilon) \rangle$  and not by  $h^\Omega(\epsilon, \tau_r)$ . Anyhow, the computation of  $h^\Omega(\epsilon, \tau_r)$  is compulsory in order to recover, e.g., a zero entropy for regular (e.g. periodic) signals.

One can easily estimate an upper and a lower bound for  $h(\epsilon)$  which can be computed in the exit-time scheme [3]. We use the following notation: for given  $\epsilon$  and  $\tau_r$ ,  $h^\Omega(\epsilon, \tau_r) \equiv h^\Omega(\{\eta_i, k_i\})$ , and we indicate with  $h^\Omega(\{k_i\})$  and  $h^\Omega(\{\eta_i\})$  respectively the Shannon entropy of the sequence  $\{k_i\}$  and  $\{\eta_i\}$ . By applying standard results of information theory [201] one obtains the inequalities (see [3,4] for more details):

$$h^\Omega(\{k_i\}) \leq h^\Omega(\{\eta_i, k_i\}) \leq h^\Omega(\{\eta_i\}) + h^\Omega(\{k_i\}). \tag{C.4}$$

Moreover,  $h^\Omega(\{\eta_i\}) \leq H_1^\Omega(\{\eta_i\})$ , where  $H_1^\Omega(\{\eta_i\})$  is the one-symbol entropy of  $\{\eta_i\}$  (i.e. the entropy of the probability distribution of the exit times measured on the scale  $\tau_r$ ) which

can be written as

$$H_1^\Omega(\{\eta_i\}) = c(\epsilon) + \ln\left(\frac{\langle t(\epsilon) \rangle}{\tau_r}\right),$$

where  $c(\epsilon) = -\int p(z) \ln p(z) dz$ , and  $p(z)$  is the probability distribution function of the rescaled exit time  $z(\epsilon) = t(\epsilon)/\langle t(\epsilon) \rangle$ . Finally, using the previous relations, one obtains the following bounds for the  $\epsilon$ -entropy:

$$\frac{h^\Omega(\{k_i\})}{\langle t(\epsilon) \rangle} \leq h(\epsilon) \leq \frac{h^\Omega(\{k_i\}) + c(\epsilon) + \ln(\langle t(\epsilon) \rangle/\tau_r)}{\langle t(\epsilon) \rangle}. \quad (\text{C.5})$$

Note that such bounds are relatively easy to compute and give a good estimate of  $h(\epsilon)$ . In particular, as far as the scaling behavior of  $h(\epsilon)$  is concerned, the exit-time method allows for very efficient and good estimates of the scaling exponent. The reason is that at fixed  $\epsilon$ ,  $\langle t(\epsilon) \rangle$  automatically selects the typical time at that scale. Consequently, it is not necessary to reach very large block sizes—at least if  $\epsilon$  is not too small. So that the leading contribution is given by  $\langle t(\epsilon) \rangle$ , and  $h^\Omega(\epsilon, \tau_r)$  introduces, at worst, a sub-leading logarithmic contribution  $h^\Omega(\epsilon, \tau_r) \sim \ln(\langle t(\epsilon) \rangle/\tau_r)$  (see Eq. (C.5)).

In Refs. [3,4] one can find the details of the derivation and some applications.

## Appendix D. Synthetic signals for turbulence

In this appendix we recall some recently introduced methods for generating multi-affine stochastic signals [27,29], whose scaling properties are fully under control. The first step consists in generating a 1-dimensional signal, and the second in decorating it such as to build the most general  $(d + 1)$ -dimensional process,  $v(\mathbf{x}, t)$ , with given scaling properties in time and in space.

For the 1-dimensional case there are at least two different kind of algorithms. One is based on a dyadic decomposition of the signal in a wavelet basis with a suitable assigned series of stochastic coefficients [27]. The second is based on a multiplication of sequential Langevin-processes with a hierarchy of different characteristic times [29].

The first procedure suits particularly appealing for modeling of spatial turbulent fluctuations, because of the natural identification between wavelets and eddies in the physical space. The second one looks more appropriate for mimicking the turbulent time evolution in a fixed point of the space.

### D.1. Reproducing the spatial properties or the temporal ones

A non-sequential algorithm for 1-dimensional multi-affine signal in  $[0, 1]$ ,  $v(x)$ , can be introduced as [27]

$$v(x) = \sum_{n=1}^N \sum_{k=1}^{2^{(n-1)}} a_{n,k} \varphi\left(\frac{x - x_{n,k}}{\ell_n}\right), \quad (\text{D.1})$$

where we have a set of reference scales  $\ell_n = 2^{-n}$  and  $\varphi(x)$  is a wavelet-like function [77], i.e. of zero mean and rapidly decaying in both real space and Fourier-space. The signal  $v(x)$  is built in

terms of a superposition of fluctuations,  $\varphi((x - x_{n,k})/\ell_n)$  of characteristic width  $\ell_n$  and centered in different points of  $[0, 1]$ ,  $x_{n,k} = (2k + 1)/2^{n+1}$ . In [29] it has been proved that provided the coefficients  $a_{n,k}$  are chosen by a random multiplicative process, i.e. the daughter is given in terms of the mother by a random process,  $a_{n+1,k'} = X a_{n,k}$  with  $X$  a random number identical, independent distributed for any  $\{n, k\}$ , then the result of the superposition is a multi-affine function with given scaling exponents, namely

$$\langle\langle |v(x + R) - v(x)|^p \rangle\rangle \sim R^{\zeta(p)},$$

with  $\zeta(p) = -p/2 - \log_2 \langle X^p \rangle$  and  $\ell_N \leq R \leq 1$ . In this appendix,  $\langle \cdot \rangle$  indicates the average over the probability distribution of the multiplicative process.

Besides the rigorous proof, the rationale for the previous result is simply that due to the hierarchical organization of the fluctuations, one may easily see that the term dominating the expression of a velocity fluctuation at scale  $R$ , in (D.1) is given by the couple of indices  $\{n, k\}$  such that  $n \sim \log_2(R)$  and  $x \sim x_{n,k}$ , i.e.  $v(x + R) - v(x) \sim a_{n,k}$ . The generalization (D.1) to  $d$ -dimension is given by

$$v(\mathbf{x}) = \sum_{n=1}^N \sum_{k=1}^{2^{d(n-1)}} a_{n,k} \varphi\left(\frac{\mathbf{x} - \mathbf{x}_{n,k}}{\ell_n}\right),$$

where now the coefficients  $\{a_{n,k}\}$  are given in terms of a  $d$ -dimensional dyadic multiplicative process.

On the other hand, as previously said, sequential algorithms look more suitable for mimicking temporal fluctuations. Let us now discuss how to construct these stochastic multi-affine fields. With the application to time fluctuations in mind, we will denote now the stochastic 1-dimensional functions with  $u(t)$ . The signal  $u(t)$  is obtained by a superposition of functions with different characteristic times, representing eddies of various sizes [29]:

$$u(t) = \sum_{n=1}^N u_n(t). \tag{D.2}$$

The functions  $u_n(t)$  are defined by the multiplicative process

$$u_n(t) = g_n(t) x_1(t) x_2(t) \dots x_n(t), \tag{D.3}$$

where the  $g_n(t)$  are independent stationary random processes, whose correlation times are supposed to be  $\tau_n = (\ell_n)^\alpha$ , where  $\alpha = 1 - h$  (i.e.  $\tau_n$  are the eddy-turn-over time at scale  $\ell_n$ ) in the quasi-Lagrangian frame (Ref. [152]) and  $\alpha = 1$  if one considers  $u(t)$  as the time signal in a given point, and  $\langle g_n^2 \rangle = (\ell_n)^{2h}$ , where  $h$  is the Hölder exponent. For a signal mimicking a turbulent flow, ignoring intermittency, we would have  $h = \frac{1}{3}$ . Scaling will appear for all time delays larger than the UV cutoff  $\tau_N$  and smaller than the IR cutoff  $\tau_1$ . The  $x_j(t)$  are independent, positive defined, identical distributed random processes whose time correlation decays with the characteristic time  $\tau_j$ . The probability distribution of  $x_j$  determines the intermittency of the process.

The origin of (D.3) is fairly clear in the context of fully developed turbulence. Indeed we can identify  $u_n$  with the velocity difference at scale  $\ell_n$  and  $x_j$  with  $(\varepsilon_j/\varepsilon_{j-1})^{1/3}$ , where  $\varepsilon_j$  is the energy dissipation at scale  $\ell_j$  [29].

The following arguments show that the process defined according to (D.2) and (D.3) is multi-affine. Because of the fast decrease of the correlation times  $\tau_j = (\ell_j)^\alpha$ , the characteristic time of  $u_n(t)$  is of the order of the shortest one, i.e.,  $\tau_n = (\ell_n)^\alpha$ . Therefore, the leading contribution to the structure function  $\tilde{S}_q(\tau) = \langle\langle |u(t + \tau) - u(t)|^q \rangle\rangle$  with  $\tau \sim \tau_n$  stems from the  $n$ th term in (D.2). This can be understood as that in  $u(t + \tau) - u(t) = \sum_{k=1}^N [u_k(t + \tau) - u_k(t)]$  the terms with  $k \leq n$  are negligible because  $u_k(t + \tau) \simeq u_k(t)$  and the terms with  $k \geq n$  are sub-leading. Thus one has

$$\tilde{S}_q(\tau_n) \sim \langle |u_n|^q \rangle \sim \langle |g_n|^q \rangle \langle x^q \rangle^n \sim \tau_n^{hq/\alpha - \log_2 \langle x^q \rangle / \alpha} \tag{D.4}$$

and therefore for the scaling exponents:

$$\zeta_q = \frac{hq}{\alpha} - \frac{\log_2 \langle x^q \rangle}{\alpha} . \tag{D.5}$$

The limit of an affine function can be obtained when all the  $x_j$  are equal to 1. A proper proof of these result can be found in [29].

Let us notice at this stage that the previous “temporal” signal for  $\alpha = 1 - h$  is a good candidate for a velocity measurements in a Lagrangian, co-moving frame of Ref. [152]. Indeed, in such a reference frame the temporal decorrelation properties at scale  $\ell_n$  are given by the eddy-turn-over times  $\tau_n = (\ell_n)^{1-h}$ . On the other hand, in the laboratory reference frame the sweeping dominates the time evolution in a fixed point of the space and we must use as characteristic times of the processes  $x_n(t)$  the sweeping times  $\tau_n^{(s)} = \ell_n$ , i.e.,  $\alpha = 1$ .

### D.2. Reproducing both the spatial and the temporal properties

We have now all the ingredients to perform a merging of temporal and spatial properties of a turbulent signal in order to define stochastic processes able to reproduce in a realistic way both spatial and temporal fluctuations (5.30) and (5.31). We just have to merge in a proper way the two previous algorithms.

For example, for a  $d$ -dimensional multi-affine field such as, say, one of the three components of a turbulent field in a Lagrangian reference frame we can use the following model:

$$v_L(\mathbf{x}, t) = \sum_{n=1}^N \sum_{k=1}^{2^{d(n-1)}} a_{n,k}(t) \varphi \left( \frac{\mathbf{x} - \mathbf{x}_{n,k}}{\ell_n} \right) , \tag{D.6}$$

where the temporal dependency of  $a_{n,k}(t)$  is chosen following the sequential algorithm while its spatial part are given by the dyadic structure of the non-sequential algorithm. In (D.6) we have used the notation  $v_L(\mathbf{x}, t)$  in order to stress the typical Lagrangian character of such a field.

We are now able to guess a good candidate for the same field measured in the laboratory-reference frame, i.e. where the time properties are dominated by the sweeping of small scales by large scales. Indeed, in order to reproduce the sweeping effects one needs that the centers  $\{\mathbf{x}_{n,k}\}$  of the wavelets-like functions move according a swept dynamics.

To do so, let us define the Eulerian model

$$v_E(\mathbf{x}, t) = \sum_{n=1}^N \sum_{k=1}^{2^{d(n-1)}} a_{n,k}(t) \varphi \left( \frac{\mathbf{x} - \mathbf{x}_{n,k}(t)}{\ell_n} \right) , \tag{D.7}$$

where the difference with the previous definition is in the temporal dependency of the centers of the wavelets,  $\mathbf{x}_{n,k}(t)$ . According to the Richardson–Kolmogorov cascade picture, one assumes that sweeping is present, i.e.,  $\mathbf{x}_{n,k} = \mathbf{x}_{n-1,k'} + \mathbf{r}_{n,k}$  where  $(n, k')$  labels the “mother” of the  $(n, k)$ -eddy and  $\mathbf{r}_{n,k}$  is a stochastic vector which depends on  $\mathbf{r}_{n-1,k'}$  and evolves with characteristic time  $\tau_n \propto (\ell_n)^{1-h}$ . Furthermore, its norm is  $O(\ell_n)$ :  $c_1 < |\mathbf{r}_{n,k}|/\ell_n < c_2$  where  $c_1$  and  $c_2$  are constants of order 1.

We now see that if we measure in one fixed spatial point a fluctuation over a time delay  $\Delta t$ , is like to measure a simultaneous fluctuations at scale separation  $R = U\Delta t$ , i.e. due to the sweeping the main contribution to the sum will be given by the terms with scale-index  $n = \log_2(R = U\Delta t)$  while the temporal dependency of the coefficients  $\{a_{n,k}(t)\}$  will be practically frozen on that time scale. Therefore, one has the proper spatial and temporal statistics, see Ref. [4] for details. This happens because in presence of sweeping the main contribution is given by the displacement of the center at large scale, i.e.  $\delta r_0 = |\mathbf{r}_0(t + \Delta t) - \mathbf{r}_0(t)| \sim U\Delta t$ , and the eddy turnover time at scale  $\ell_n$  is  $O((\ell_n)^{1-h})$  always large that the sweeping time  $O(\ell_n)$  at the same scale.

In the previous discussion for the sake of simplicity we did not consider the incompressibility condition. However one can take into account this constraint by the projection on the solenoidal space.

In conclusion we have a way to build up a synthetic signal with the proper Eulerian (laboratory) properties, i.e. with sweeping, and also with the proper Lagrangian properties.

## References

- [1] H.D.I. Abarbanel, R. Brown, J.L. Sidorowich, L.Sh. Tsimring, The analysis of observed chaotic data in physical systems, *Rev. Mod. Phys.* 65 (1993) 1331.
- [2] H.D.I. Abarbanel, *Analysis of Observed Chaotic Data*, Springer, New York, 1996.
- [3] M. Abel, L. Biferale, M. Cencini, M. Falcioni, D. Vergni, A. Vulpiani, Exit time approach to  $(\epsilon, \tau)$ -entropy, *Phys. Rev. Lett.* 84 (2000) 6002.
- [4] M. Abel, L. Biferale, M. Cencini, M. Falcioni, D. Vergni, A. Vulpiani, Exit-times and  $\epsilon$ -entropy for dynamical systems, *Stochastic Processes, and Turbulence, Physica D* 147 (2000) 12.
- [5] V.M. Alekseev, M.V. Yakobson, Symbolic dynamics and hyperbolic dynamic-systems, *Phys. Rep.* 75 (1981) 287.
- [6] D.J. Amit, *Modeling Brain Function*, Cambridge University Press, Cambridge, 1992.
- [7] F. Anselmet, Y. Gagne, E.J. Hopfinger, R.A. Antonia, High order velocity structure functions in turbulent shear flow, *J. Fluid. Mech.* 140 (1984) 63.
- [8] I.S. Arason, A.V. Gaponov-Grekhov, M. Rabinovich, I.P. Starobinets, Strange attractors and the spatial development of turbulence in flow systems, *Sov. Phys. JEPT* 63 (1986) 1000.
- [9] I.S. Arason, A.V. Gaponov-Grekhov, M. Rabinovich, The onset and spatial development of turbulence in flow systems, *Physica D* 33 (1988) 1.
- [10] V. Artale, G. Boffetta, A. Celani, M. Cencini, A. Vulpiani, Dispersion of passive tracers in closed basins: Beyond the diffusion coefficient, *Phys. Fluids A* 9 (1997) 3162.
- [11] E. Aurell, G. Boffetta, A. Crisanti, G. Paladin, A. Vulpiani, Growth of non-infinitesimal perturbations in turbulence, *Phys. Rev. Lett.* 77 (1996) 1262.
- [12] E. Aurell, G. Boffetta, A. Crisanti, G. Paladin, A. Vulpiani, Predictability in the large: an extension of the concept of Lyapunov exponent, *J. Phys. A* 30 (1997) 1.
- [13] R. Badii, A. Politi, *Complexity. Hierarchical Structures and Scaling in Physics*, Cambridge University Press, Cambridge, UK, 1997.
- [14] P. Bak, C. Tang, K. Wiesenfeld, Self-organized criticality an explanation of  $1/f$  noise, *Phys. Rev. Lett.* 59 (1987) 381.

- [15] P. Bak, C. Tang, K. Wiesenfeld, Self-organized criticality, *Phys. Rev. A* 38 (1988) 364.
- [16] F. Bagnoli, R. Rechtman, S. Ruffo, Damage spreading and Lyapunov exponent in cellular automata, *Phys. Lett. A* 172 (1992) 34.
- [17] F. Bagnoli, R. Rechtman, Synchronization and maximum Lyapunov exponent in cellular automata, *Phys. Rev. E* 59 (1999) R1307.
- [18] C. Basdevant, B. Legras, R. Sadourny, M. Beland, A study of barotropic model flows: intermittency, waves and predictability, *J. Atmos. Sci.* 38 (1981) 2305.
- [19] G.K. Batchelor, *The Theory of Homogeneous Turbulence*, Cambridge University Press, Cambridge, 1953.
- [20] G. Benettin, L. Galgani, A. Giorgilli, J.M. Strelcyn, Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems: a method for computing all of them, *Meccanica* 15 (1980) 9.
- [21] C.H. Bennet, How to define complexity in physics and why, in: W.H. Zurek (Ed.), *Complexity, Entropy and the Physics of Information*, Addison-Wesley, Redwood City, 1990, p. 137.
- [22] R. Benzi, G. Paladin, G. Parisi, A. Vulpiani, On the multifractal nature of fully developed turbulence and chaotic systems, *J. Phys. A* 17 (1984) 3521.
- [23] R. Benzi, G. Paladin, G. Parisi, A. Vulpiani, Characterization of intermittency in chaotic systems, *J. Phys. A* 18 (1985) 2157.
- [24] R. Benzi, L. Biferale, S. Succi, F. Toschi, Intermittency and eddy-viscosities in dynamical models of turbulence, *Phys. Fluids* 11 (1998) 1221.
- [25] R. Benzi, A. Sutera, A. Vulpiani, The mechanism of the stochastic resonance, *J. Phys. A* 14 (1981) L453.
- [26] R. Benzi, G. Parisi, A. Sutera, A. Vulpiani, Stochastic resonances in climatic change, *Tellus* 34 (1982) 10.
- [27] R. Benzi, L. Biferale, A. Crisanti, G. Paladin, M. Vergassola, A. Vulpiani, A random process for the construction of multifractal fields, *Physica D* 65 (1993) 352.
- [28] G. Berkooz, An observation probability density equations, or, when simulations reproduce statistics?, *Nonlinearity* 7 (1994) 313.
- [29] L. Biferale, G. Boffetta, A. Celani, A. Crisanti, A. Vulpiani, Mimicking a turbulent signal: sequential multifractal processes, *Phys. Rev. E* 57 (1998) R6261.
- [30] L. Biferale, M. Cencini, D. Vergni, A. Vulpiani, Exit time of turbulent signals: a way to detect the Intermediate Dissipative Range, *Phys. Rev. E* 60 (1999) R6295.
- [31] P. Billingsley, *Ergodic Theory and Information*, Wiley, New York, 1965.
- [32] G. Boffetta, G. Paladin, A. Vulpiani, Strong chaos without butterfly effect in dynamical systems with feedback, *J. Phys. A* 29 (1996) 2291.
- [33] G. Boffetta, A. Celani, A. Crisanti, A. Vulpiani, Predictability in two dimensional decaying turbulence, *Phys. Fluids A* 9 (1997) 724.
- [34] G. Boffetta, A. Celani, M. Vergassola, Inverse energy cascade in two-dimensional turbulence: deviations from Gaussian behavior, *Phys. Rev. E* 61 (2000) R29.
- [35] G. Boffetta, A. Crisanti, F. Paparella, A. Provenzale, A. Vulpiani, Slow and fast dynamics in coupled systems: a time series analysis view, *Physica D* 116 (1998) 301.
- [36] G. Boffetta, A. Celani, M. Cencini, G. Lacorata, A. Vulpiani, The predictability problem in systems with an uncertainty in the evolution law, *J. Phys. A* 33 (2000) 1313.
- [37] G. Boffetta, S. Musacchio, Predictability of the energy cascade in 2D turbulence, *Phys. Fluids* 13 (2000) 1060.
- [38] T. Bohr, M.H. Jensen, G. Paladin, A. Vulpiani, *Dynamical Systems Approach to Turbulence*, Cambridge University Press, Cambridge, 1998.
- [39] T. Bohr, D. Rand, A mechanism for localized turbulence, *Physica D* 52 (1991) 532.
- [40] J.P. Boon, D. Dab, R. Kapral, A. Lawniczak, Lattice gas automata for reactive systems, *Phys. Rep.* 273 (1996) 55.
- [41] R. Bowen, Markov partitions for Axiom-A diffeomorphisms, *Amer. J. Math.* 92 (1970) 725.
- [42] A.A. Brudno, Entropy and the complexity of the trajectories of a dynamical system, *Trans. Moscow Math. Soc.* 44 (1983) 127.
- [43] A.R. Bulsara, E.W. Jacobs, W.C. Schieve, Noise effects in a non-linear dynamic system—the RF—superconducting quantum interference device, *Phys. Rev. A* 42 (1990) 4614.

- [44] L.A. Bunimovich, G. Sinai, Statistical mechanics of coupled map lattices, in: K. Kaneko (Ed.), *Theory and Applications of Coupled Map Lattices*, Wiley, New York, 1993, p. 169.
- [45] J.M. Burgers, *The Nonlinear Diffusion Equation*, D. Reidel Publishing Company, Dordrecht, 1974.
- [46] E. Caglioti, V. Loreto, Dynamical properties and predictability of a class of self-organized critical models, *Phys. Rev. E* 53 (1996) 2953.
- [47] M. Casartelli, Partitions, rational partitions, and characterization of complexity, *Complex Systems* 4 (1990) 491.
- [48] M. Casartelli, M. Zerbini, Metric features of self-organized criticality states in sandpile models, *J. Phys. A* 33 (2000) 863.
- [49] G. Casati, B.V. Chirikov (Eds.), *Quantum Chaos: Between Order and Disorder*, Cambridge University Press, Cambridge, UK, 1995.
- [50] M. Cencini, M. Falcioni, D. Vergni, A. Vulpiani, Macroscopic chaos in globally coupled maps, *Physica D* 130 (1999) 58.
- [51] M. Cencini, M. Falcioni, H. Kantz, E. Olbrich, A. Vulpiani, Chaos or Noise—Sense and Nonsense of a Distinction, *Phys. Rev. E* 62 (2000) 427.
- [52] M. Cencini, A. Torcini, A non-linear marginal stability criterion for information propagation, *Phys. Rev. E* 63 (2001), in press.
- [53] G.J. Chaitin, On the length of programs for computing finite binary sequences, *J. Assoc. Comp. Mach.* 13 (1966) 547.
- [54] G.J. Chaitin, Information-theoretic limitations of formal systems, *J. Assoc. Comp. Mach.* 21 (1974) 403.
- [55] G.J. Chaitin, Gödel theorem and information, *Int. J. Theor. Phys.* 22 (1982) 941.
- [56] G.J. Chaitin, *Information, Randomness and Incompleteness*, 2nd Edition, World Scientific, Singapore, 1990.
- [57] S. Chandrasekhar, Stochastic problems in physics and astronomy, *Rev. Mod. Phys.* 15 (1943) 1.
- [58] H. Chaté, P. Manneville, Collective behaviors in spatially extended systems with local interactions and synchronous updating, *Prog. Theor. Phys.* 87 (1992) 1.
- [59] Z.Y. Chen, Noise-induced instability, *Phys. Rev. A* 42 (1990) 5837.
- [60] B.V. Chirikov, Natural Laws and Human Prediction, in: P. Weingartner, G. Schurz (Eds.), *Laws and Prediction in the Light of Chaos Research*, Springer, Berlin, 1996, p. 10.
- [61] A. Cohen, I. Procaccia, Computing the Kolmogorov entropy from time signals of dissipative and conservative dynamical systems, *Phys. Rev. A* 31 (1985) 1872.
- [62] P. Collet, J.P. Eckmann, The definition and measurement of the topological entropy per unit volume in parabolic PDEs, *Nonlinearity* 12 (1999) 451.
- [63] J. Coste, M. Hénon, Invariant cycles in the random mapping of  $N$  integers onto themselves. Comparison with Kauffman binary network, in: E. Bienenstock et al. (Eds.), *Disordered Systems and Biological Organization*, Springer, Berlin, 1986, p. 360.
- [64] A. Crisanti, M. Falcioni, A. Vulpiani, On the effects of an uncertainty on the evolution law in dynamical systems, *Physica A* 160 (1989) 482.
- [65] A. Crisanti, M. Falcioni, G. Mantica, A. Vulpiani, Applying algorithmic complexity to define chaos in the motion of complex systems, *Phys. Rev. E* 50 (1994) 1959.
- [66] A. Crisanti, M.H. Jensen, G. Paladin, A. Vulpiani, Intermittency and predictability in turbulence, *Phys. Rev. Lett.* 70 (1993) 166.
- [67] A. Crisanti, M. Falcioni, A. Vulpiani, Broken ergodicity and glassy behavior in a deterministic chaotic map, *Phys. Rev. Lett.* 76 (1996) 612.
- [68] M.C. Cross, P.H. Hohenberg, Pattern formation outside of equilibrium, *Rev. Mod. Phys.* 65 (1993) 851.
- [69] J.P. Crutchfield, K. Kaneko, Are attractors relevant for turbulence?, *Phys. Rev. Lett.* 60 (1988) 2715.
- [70] P.J. Davies, R. Hersh, *The Mathematical Experience*, Penguin Books, New York, 1984.
- [71] R.J. Deissler, K. Kaneko, Velocity-dependent Lyapunov exponents as a measure of chaos for open-flow systems, *Phys. Lett. A* 119 (1987) 397.
- [72] C. Dettman, E. Cohen, H. van Beijeren, Microscopic chaos from Brownian motion? *Nature* 401 (1999) 875.
- [73] U. Dressler, J.D. Farmer, Generalized Lyapunov exponents corresponding to higher derivatives, *Physica D* 59 (1992) 365.
- [74] J.P. Eckmann, D. Ruelle, Ergodic theory of chaos and strange attractors, *Rev. Mod. Phys.* 57 (1985) 617.

- [75] M. Falcioni, U.M. Bettolo Marconi, A. Vulpiani, Ergodic properties of high-dimensional symplectic maps, *Phys. Rev. A* 44 (1991) 2263.
- [76] M. Falcioni, D. Vergni, A. Vulpiani, Characterization of the spatial complex behavior and transition to chaos in flow systems, *Physica D* 125 (1999) 65.
- [77] M. Farge, Wavelet transforms and their applications to turbulence, *Ann. Rev. Fluid Mech.* 24 (1992) 395.
- [78] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. 1, 3rd Edition, Wiley, New York, 1970.
- [79] J. Ford, How random is a coin tossing?, *Phys. Today* 36 (1983) 40.
- [80] J. Ford, Chaos: Solving the Unsolvable, Predicting the Unpredictable, in: M.F. Barnaley, S. Demko (Eds.), *Chaotic Dynamics and Fractals*, Academic Press, New York, 1986.
- [81] J. Ford, G. Mantica, G.H. Ristow, The Arnold's Cat: Failure of the Correspondence Principle, *Physica D* 50 (1991) 1493.
- [82] U. Frisch, B. Hasslacher, Y. Pomeau, Lattice-gas automata for the Navier–Stokes equation, *Phys. Rev. Lett.* 56 (1986) 1505.
- [83] U. Frisch, M. Vergassola, A prediction of the multifractal model—The Intermediate Dissipation Range, *Europhys. Lett.* 14 (1991) 439.
- [84] U. Frisch, *Turbulence: the Legacy of A. N. Kolmogorov*, Cambridge University Press, Cambridge, 1995.
- [85] H. Fujisaka, Statistical dynamics generated by fluctuations of local Lyapunov exponents, *Prog. Theor. Phys.* 70 (1983) 1264.
- [86] S. Galluccio, A. Vulpiani, Stretching of material lines and surfaces in systems with Lagrangian Chaos, *Physica A* 212 (1994) 75.
- [87] L. Gammaitoni, P. Hanggi, P. Jung, F. Marchesoni, Stochastic resonance, *Rev. Mod. Phys.* 70 (1998) 223.
- [88] Y. Gagne, B. Castaing, Une représentation universelle sans invariance globale d'échelle des spectres d'énergie en turbulence développée, *C. R. Acad. Sci. Paris* 312 (1991) 441.
- [89] P. Gaspard, X.J. Wang, Noise, chaos, and  $(\epsilon, \tau)$ -entropy per unit time, *Phys. Rep.* 235 (1993) 291.
- [90] P. Gaspard, M.E. Briggs, M.K. Francis, J.V. Sengers, R.W. Gammon, J.R. Dorfman, R.V. Calabrese, Experimental evidence for microscopic chaos, *Nature* 394 (1998) 865.
- [91] G. Giacomelli, A. Politi, Spatio-temporal chaos and localization, *Europhys. Lett.* 15 (1991) 387.
- [92] I. Goldhirsh, P.L. Sulem, S.A. Orszag, Stability and Lyapunov stability of dynamical systems: a differential approach and numerical method, *Physica D* 27 (1987) 311.
- [93] P. Grassberger, I. Procaccia, Characterization of strange attractors, *Phys. Rev. Lett.* 50 (1983) 346.
- [94] P. Grassberger, I. Procaccia, Measuring the strangeness of strange attractors, *Physica D* 9 (1983) 189.
- [95] P. Grassberger, I. Procaccia, Estimation of the Kolmogorov-entropy from a chaotic signal, *Phys. Rev. A* 28 (1983) 2591.
- [96] P. Grassberger, Toward a quantitative theory of self-generated complexity, *Int. J. Theor. Phys.* 25 (1986) 907.
- [97] P. Grassberger, Information content and predictability of lumped and distributed dynamical systems, *Physica Scripta* 40 (1989) 346.
- [98] P. Grassberger, T. Schreiber, Statistical mechanics—Microscopic chaos from Brownian motion?, *Nature* 401 (1999) 875.
- [99] J. Guckenheimer, P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer, Berlin, 1986.
- [100] H.A. Gutowitz, Transient, cycles and complexity in cellular automata, *Phys. Rev. A* 44 (1991) 7881.
- [101] J.F. Heagy, N. Platt, S.M. Hammel, Characterization of on-off intermittency, *Phys. Rev. E* 49 (1994) 1140.
- [102] H. Herzog, W. Ebeling, Th. Schulmeister, Nonuniform Chaotic dynamics and effects of noise in biochemical systems, *Z. Naturforsch.* A 42 (1987) 136.
- [103] H. Herzog, B. Pompe, Effects of noise on a non-uniform chaotic map, *Phys. Lett. A* 122 (1987) 121.
- [104] P.C. Hohenberg, B.I. Shraiman, Chaotic behavior of an extended system, *Physica D* 37 (1988) 109.
- [105] P.J. Holmes, J.L. Lumley, G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge University Press, Cambridge, 1996.
- [106] M.H. Jensen, Fluctuations and scaling in a model for boundary-layer-induced turbulence, *Phys. Rev. Lett.* 62 (1989) 1361.

- [107] M.H. Jensen, Multiscaling and structure functions in turbulence: an alternative approach, *Phys. Rev. Lett.* 83 (1999) 76.
- [108] K. Kaneko, Period-doubling of kink-antikink patterns, quasi-periodicity in antiferro-like structures and spatial intermittency in coupled map lattices—toward a prelude to a field theory of chaos, *Prog. Theor. Phys.* 72 (1984) 480.
- [109] K. Kaneko, Lyapunov analysis and information flow in coupled map lattices, *Physica D* 23 (1986) 436.
- [110] K. Kaneko, Chaotic but regular posi-nega switch among coded attractors by cluster-size variation, *Phys. Rev. Lett.* 63 (1989) 219.
- [111] K. Kaneko, Globally coupled chaos violates the law of large numbers but not the central-limit theorem, *Phys. Rev. Lett.* 65 (1990) 1391.
- [112] K. Kaneko, The coupled map lattice, in: K. Kaneko (Ed.), *Theory and Applications of Coupled Map Lattices*, Wiley, New York, 1993, p. 1.
- [113] K. Kaneko, Remarks on the mean field dynamics of networks of chaotic elements, *Physica D* 86 (1995) 158.
- [114] H. Kantz, T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, Cambridge, 1997.
- [115] H. Kantz, T. Letz, Quasi-chaos and quasi-regularity—the breakdown of linear stability analysis, *Phys. Rev. E* 61 (2000) 2533.
- [116] H. Kantz, E. Olbrich, Coarse grained dynamical entropies: investigation of high entropic dynamical systems, *Physica A* 280 (2000) 34.
- [117] J.L. Kaplan, J.A. Yorke, Pre turbulence: a regime observed in a fluid model of Lorenz, *Comm. Math. Phys.* 67 (1979) 93.
- [118] R. Kapral, Chemical waves and coupled map lattices, in: K. Kaneko (Ed.), *Theory and Applications of Coupled Map Lattices*, Wiley, New York, 1993, p. 135.
- [119] R. Kapral, S.J. Fraser, Dynamics of oscillations with periodic dichotomic noise, *J. Stat. Phys.* 70 (1993) 61.
- [120] M. Kardar, G. Parisi, Y.C. Zhang, Dynamic scaling of growing interfaces, *Phys. Rev. Lett.* 56 (1986) 889.
- [121] A.I. Khinchin, *Mathematical Foundations of Information Theory*, Dover, New York, 1957.
- [122] S. Kida, M. Yamada, K. Ohkitani, Error growth in a decaying two-dimensional turbulence, *J. Phys. Soc. Jpn.* 59 (1990) 90.
- [123] A.N. Kolmogorov, The local structure of turbulence in incompressible viscous fluid for very large Reynold number, *Dokl. Akad. Nauk. SSSR* 30 (1941) 299 (reprinted in *Proc. R. Soc. Lond. A* 434 (1991) 9).
- [124] A.N. Kolmogorov, On the Shannon theory of information transmission in the case of continuous signals, *IRE Trans. Inform. Theory* 1 (1956) 102.
- [125] A.N. Kolmogorov, New metric invariant of transitive dynamical systems and auto-morphism of Lebesgue spaces, *Dokl. Akad. Nauk SSSR* 119 (1958) 861.
- [126] A.N. Kolmogorov, Three approaches to the quantitative definition of information, *Prob. Inform. Trans.* 1 (1965) 1.
- [127] A. Kolmogorov, S. Fomine, *Éléments de la théorie des fonctions et de l'analyse fonctionnelle*, Editions MIR, Moscow, 1976 (Chapter III).
- [128] A.N. Kolmogorov, I. Petrovsky, N. Piskunov, Study of the diffusion equation with growth of the quantity of matter and its application to a biology problem, *Bull. Univ. Moscow, Ser. Int. A* 1 (1937) 1.
- [129] T. Konishi, K. Kaneko, Clustered motion in symplectic coupled maps system, *J. Phys. A* 25 (1992) 6283.
- [130] M. Koppel, H. Atlan, An almost machine-independent theory of program length complexity, sophistication and induction, *Inform. Sci.* 56 (1991) 23.
- [131] D.E. Knuth, *The Art of Computer Programming*, Vol. 2, Addison-Wesley, Reading, MA, 1969.
- [132] R.H. Kraichnan, D. Montgomery, Two-dimensional turbulence, *Rep. Prog. Phys.* 43 (1980) 547.
- [133] W. Krieger, On entropy and generators of measure preserving transformations, *Trans. Am. Math. Soc.* 149 (1970) 453.
- [134] S. Laplace, *Essai philosophique sur les probabilités*, Courcier, Paris, 1814.
- [135] B. Legras, P. Santangelo, R. Benzi, High-resolution numerical experiments for forced two-dimensional turbulence, *Europhys. Lett.* 5 (1988) 37.
- [136] C.E. Leith, Atmospheric predictability and two-dimensional turbulence, *J. Atmos. Sci.* 28 (1971) 144.
- [137] C.E. Leith, R.H. Kraichnan, Predictability of turbulent flows, *J. Atmos. Sci.* 29 (1972) 1041.
- [138] A. Lempel, J. Ziv, On the complexity of finite sequences, *IEEE Trans. Inform. Theory* 22 (1976) 75.

- [139] S. Lepri, A. Politi, A. Torcini, Chronotropic Lyapunov Analysis: (I) a comprehensive characterization of 1D systems, *J. Stat. Phys.* 82 (1996) 1429.
- [140] S. Lepri, A. Politi, A. Torcini, Chronotropic Lyapunov analysis: (II) towards a unified approach, *J. Stat. Phys.* 88 (1997) 31.
- [141] M. Lesieur, *Turbulence in Fluids*, 2nd Edition, Kluwer Academic Publishers, London, 1990.
- [142] M. Li, P. Vitanyi, *An Introduction to Kolmogorov Complexity and its Applications*, Springer, Berlin, 1997.
- [143] D.K. Lilly, Lectures in sub-synoptic scales of motion and two-dimensional turbulence, in: P. Morel (Ed.), *Dynamic Meteorology*, Riedel, Boston, 1973.
- [144] R. Livi, A. Politi, S. Ruffo, Distribution of characteristic exponents in the thermodynamic limit, *J. Phys. A* 19 (1986) 2033.
- [145] E.N. Lorenz, Deterministic non-periodic flow, *J. Atmos. Sci.* 20 (1963) 130.
- [146] E.N. Lorenz, The predictability of a flow which possesses many scales of motion, *Tellus* 21 (1969) 3.
- [147] E.N. Lorenz, Predictability—a problem partly solved, in: *Predictability*, Vol. I, Proceedings of a Seminar, ECMWF, Reading, 4–8 September 1995, 1996, p. 1.
- [148] E.N. Lorenz, K.A. Emanuel, Optimal sites for supplementary weather observations: simulation with a small model, *J. Atmos. Sci.* 55 (1998) 399.
- [149] V. Loreto, G. Paladin, A. Vulpiani, On the concept of complexity for random dynamical systems, *Phys. Rev. E* 53 (1996) 2087.
- [150] V. Loreto, G. Paladin, M. Pasquini, A. Vulpiani, Characterization of Chaos in random maps, *Physica A* 232 (1996) 189.
- [151] V.S. L'vov, E. Podivilov, A. Pomyalov, I. Procaccia, D. Vandembroucq, Improved shell model of turbulence, *Phys. Rev. E* 58 (1998) 1811.
- [152] V.S. L'vov, E. Podivilov, I. Procaccia, Temporal multiscaling in hydrodynamic turbulence, *Phys. Rev. E* 55 (1997) 7030.
- [153] P. Manneville, *Dissipative Structures and Weak Turbulence*, Academic Press, Boston, 1990.
- [154] B. Mandelbrot, On the geometry of homogeneous turbulence, with stress on the fractal dimension of the iso-surfaces of scalars, *J. Fluid. Mech.* 72 (1975) 401.
- [155] G. Mantica, Quantum algorithmic integrability: the metaphor of classical polygonal billiards, *Phys. Rev. E* 61 (2000) 6434.
- [156] P. Martin-Löf, The definition of random sequences, *Inform. Control* 9 (1966) 602.
- [157] K. Matsumoto, I. Tsuda, Noise-induced order, *J. Stat. Phys.* 31 (1983) 87.
- [158] P. Mazur, E. Montroll, Poincare cycles, ergodicity and irreversibility in assemblies of coupled harmonic oscillators, *J. Math. Phys.* 1 (1960) 70.
- [159] J.C. McWilliams, The emergence of isolated coherent vortices in turbulent flows, *J. Fluid. Mech.* 146 (1984) 21.
- [160] A. Monin, *Weather Prediction as a Problem in Physics*, MIT Press, Moscow, 1973.
- [161] A. Monin, A. Yaglom, *Statistical Fluid Dynamics*, Vols. I and II, MIT Press, Cambridge, 1971, 1975.
- [162] F. Moss, P.W.E. McClintock (Eds.), *Noise in Non-Linear Dynamical Systems*, Vols. I–III, Cambridge University Press, Cambridge, 1989.
- [163] C. Nicolis, Stochastic aspects of climate transitions-response to a periodic forcing, *Tellus* 34 (1982) 1.
- [164] K. Ohkitani, Wave number space dynamics of enstrophy cascade in a forced two-dimensional turbulence, *Phys. Fluids A* 3 (1991) 1598.
- [165] E. Olbrich, R. Hegger, H. Kantz, Analyzing local observations of weakly coupled maps, *Phys. Lett. A* 244 (1998) 538.
- [166] Y. Oono, M. Kohomoto, Discrete model of chemical turbulence, *Phys. Rev. Lett.* 55 (1985) 2927.
- [167] Y. Oono, C. Yeung, A cell dynamical system model of chemical turbulence, *J. Stat. Phys.* 48 (1987) 593.
- [168] S.A. Orszag, Indeterminacy of the moment problem for intermittent turbulence, *Phys. Fluids* 13 (1970) 2211.
- [169] V.I. Oseledec, A multiplicative ergodic theorem: Lyapunov characteristic numbers for dynamical systems, *Trans. Moscow Math. Soc.* 19 (1968) 197.
- [170] E. Ott, T. Sauer, J.A. Yorke (Eds.), *Coping With Chaos: Analysis of Chaotic Data and the Exploitation of Chaotic Systems*, Wiley Series in Nonlinear Science, Wiley, New York, 1994.
- [171] G. Paladin, A. Vulpiani, Intermittency in chaotic systems and Renyi entropies, *J. Phys. A* 19 (1986) 997.

- [172] G. Paladin, A. Vulpiani, Anomalous scaling laws in multifractal objects, *Phys. Rep.* 156 (1987) 147.
- [173] G. Paladin, A. Vulpiani, Predictability in spatially extended systems, *J. Phys. A* 27 (1994) 4911.
- [174] G. Paladin, M. Serva, A. Vulpiani, Complexity in dynamical systems with noise, *Phys. Rev. Lett.* 74 (1995) 66.
- [175] T.N. Palmer, Predictability of the atmosphere and oceans: from days to decades, in: *Predictability*, Vol. I, Proceedings of a Seminar, ECMWF, Reading, 4–8 September, 1995, p. 93.
- [176] J. Paret, P. Tabeling, Experimental Observation of the Two-Dimensional Inverse Energy Cascade, *Phys. Rev. Lett.* 79 (1997) 4162.
- [177] G. Parisi, U. Frisch, On the singularity structure of fully developed turbulence, in: M. Ghil et al. (Eds.), *Turbulence and Predictability of Geophysical Fluid Dynamics*, North-Holland, Amsterdam, 1985, p. 84.
- [178] G. Perez, H.A. Cerdeira, Instabilities and non-statistical behavior in globally coupled systems, *Phys. Rev. A* 46 (1992) 7492.
- [179] Y.B. Pesin, Lyapunov characteristic exponent and ergodic properties of smooth dynamical systems with an invariant measure, *Sov. Math. Dokl.* 17 (1976) 196.
- [180] A.S. Pikovsky, Spatial development of Chaos in non-linear media, *Phys. Lett. A* 137 (1989) 121.
- [181] A.S. Pikovsky, Does dimension grow in flow systems?, in: T. Riste, D. Sherrington (Eds.), *Spontaneous formation of Space-Time Structures and Criticality*, Kluwer Academic Publishers, Dordrecht, 1991.
- [182] A.S. Pikovsky, Spatial development of chaos, in: S. Vohra et al. (Eds.), *Proceeding of the first Experimental Chaos Conference*, World Scientific, Singapore, 1992.
- [183] A.S. Pikovsky, Local Lyapunov exponents for spatiotemporal chaos, *Chaos* 3 (1993) 225.
- [184] A.S. Pikovsky, J. Kurths, Collective behavior in ensembles of globally coupled maps, *Physica D* 76 (1994) 411.
- [185] A.S. Pikovsky, J. Kurths, Roughening interfaces in the dynamics of perturbations of spatiotemporal chaos, *Phys. Rev. E* 49 (1994) 898.
- [186] A.S. Pikovsky, A. Politi, Dynamic localization of Lyapunov vectors in space-time chaos, *Nonlinearity* 11 (1998) 1049.
- [187] N. Platt, E.A. Spiegel, C. Tresser, On-off intermittency—a mechanism for bursting, *Phys. Rev. Lett.* 70 (1993) 279.
- [188] A. Politi, R. Livi, G.L. Oppo, R. Kapral, Unpredictable Behavior in Stable Systems, *Europhys. Lett.* 22 (1993) 571.
- [189] A. Politi, A. Torcini, Periodic orbits in coupled Hénon maps: Lyapunov and multifractal analysis, *Chaos* 2 (1992) 293.
- [190] A. Politi, A. Torcini, Linear and Non-Linear mechanisms of Information propagation, *Europhys. Lett.* 28 (1994) 545.
- [191] Y. Pomeau, P. Mannerville, Intermittent transition to turbulence in dissipative dynamical systems, *Comm. Math. Phys.* 74 (1980) 189.
- [192] K.R. Popper, *The Open Universe: an Argument for Indeterminism*, Hutchinson, London, 1982.
- [193] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes*, Cambridge University Press, Cambridge, 1986.
- [194] L.F. Richardson, *Weather Prediction by Numerical Processes*, Cambridge University Press, Cambridge, 1922.
- [195] D. Ruelle, F. Takens, On the nature of turbulence, *Commun. Math. Phys.* 20 (1971) 167.
- [196] D. Ruelle, Microscopic fluctuations and turbulence, *Phys. Lett. A* 72 (1979) 81.
- [197] D. Ruelle, Large volume limit of the distribution of characteristic exponents in turbulence, *Commun. Math. Phys.* 87 (1982) 287.
- [198] O. Rudzick, A. Pikovsky, Unidirectionally coupled map lattices as a model for open flow systems, *Phys. Rev. E* 54 (1996) 5107.
- [199] T. Sauer, J.A. Yorke, M. Casdagli, *Embedology*, *J. Stat. Phys.* 65 (1991) 579.
- [200] M. Schell, S. Fraser, R. Kapral, Diffusive dynamics in systems with translational symmetry—a one-dimensional-map model, *Phys. Rev. A* 26 (1982) 504.
- [201] C.E. Shannon, A mathematical theory of communication, *The Bell System Technical J.* 27 (1948) 623; 27 (1948) 379.

- [202] T. Shibata, K. Kaneko, Heterogeneity induced order in globally coupled systems, *Europhys. Lett.* 38 (1997) 417.
- [203] T. Shibata, K. Kaneko, Collective chaos, *Phys. Rev. Lett.* 81 (1998) 4116.
- [204] Y.G. Sinai, On the concept of entropy for a dynamic system, *Dokl. Akad. Nauk. SSSR* 124 (1959) 768.
- [205] Y.G. Sinai, A remark concerning the thermodynamic limit of the Lyapunov spectrum, *Int. J. Bif. Chaos* 6 (1996) 1137.
- [206] L.M. Smith, V. Yakhot, Bose condensation and small-scale structure generation in a random force driven 2D turbulence, *Phys. Rev. Lett.* 71 (1993) 352.
- [207] R.J. Solomonov, A formal theory of inductive inference, *Inform. Control* 7 (1964) 1; 7 (1964) 224.
- [208] P. Tabeling, Experiments on 2D turbulence, in: P. Tabeling, A. Cardoso (Eds.), *Turbulence a Tentative Dictionary*, Plenum Press, New York, 1994, p. 13.
- [209] F. Takens, Detecting strange attractors in turbulence, in: D.A. Rand, L.-S. Young (Eds.), *Dynamical Systems and Turbulence* (Warwick 1980), *Lecture Notes in Mathematics*, Vol. 898, Springer, Berlin, 1980, p. 366.
- [210] F. Takens, E. Verbitski, Generalized entropies: Renyi and correlation integral approach, *Nonlinearity* 11 (1998) 771.
- [211] H. Tennekes, Karl Popper and the accountability of numerical forecasting, in: *New Development in Predictability, Proceedings of a Seminar*, ECMWF, 1992, p. 21.
- [212] A. Torcini, P. Grassberger, A. Politi, Error propagation in extended chaotic systems, *J. Phys. A* 27 (1995) 4533.
- [213] A. Torcini, S. Lepri, Disturbance propagation in extended systems with long range-coupling, *Phys. Rev. E* 55 (1997) R3805.
- [214] J. Urias, R. Rechtman, A. Enciso, Sensitive dependence on initial conditions for cellular automata, *Chaos* 7 (1997) 688.
- [215] C. van der Broeck, G. Nicolis, Noise-induced sensitivity to the initial conditions in stochastic dynamical systems, *Phys. Rev. E* 48 (1993) 4845.
- [216] S.R.S. Varadhan, *Large Deviations and Applications*, SIAM, Philadelphia, PA, 1984.
- [217] D. Vergni, M. Falcioni, A. Vulpiani, Spatial complex behavior in non-chaotic flow systems, *Phys. Rev. E* 56 (1997) 6170.
- [218] R.F. Voss, Random fractals: self-affinity in noise, music, mountains and clouds, *Physica D* 38 (1989) 362.
- [219] D. Welsh, *Codes and Cryptography*, Clarendon Press, Oxford, 1989.
- [220] E. Wigner, The unreasonable effectiveness of mathematics in natural sciences, *Comm. Pure Appl. Math.* 13 (1960) 1.
- [221] A. Wolf, J.B. Swift, H.L. Swinney, J.A. Vastano, Determining Lyapunov exponents from a time series, *Physica D* 16 (1985) 285.
- [222] S. Wolfram, Origins of randomness in physical systems, *Phys. Rev. Lett.* 55 (1985) 449.
- [223] S. Wolfram, *Theory and Applications of Cellular Automata*, World Scientific, Singapore, 1986.
- [224] X.-J. Wang, P. Gaspard,  $\epsilon$ -entropy for a time series of thermal turbulence, *Phys. Rev. A* 46 (1992) R3000.
- [225] H. White, Algorithmic complexity of Points in Dynamical Systems, *Erg. Theory Dyn. Systems* 13 (1993) 807.
- [226] F.H. Willeboordse, K. Kaneko, Bifurcations and spatial chaos in an open flow model, *Phys. Rev. Lett.* 74 (1994) 533.
- [227] L. Yu, E. Ott, Q. Chen, Transition to chaos for random dynamic-systems, *Phys. Rev. Lett.* 65 (1990) 2935.
- [228] L. Yu, E. Ott, Q. Chen, Fractal distribution of floaters on a fluid surface and the transition to chaos for random maps, *Physica D* 53 (1991) 102.
- [229] Y.-C. Zhang, Scaling theory of self-organized criticality, *Phys. Rev. Lett.* 63 (1989) 470.
- [230] W.H. Zurek (Ed.), *Complexity, Entropy and Physics of Information*, Addison-Wesley, Redwood City, 1990.
- [231] J.S. Nicolis, G. Mayer-Kress, G. Haubs, Non-uniform chaotic dynamics with implications to information processing, *Z. Naturf.* 38A (1983) 1157.