J. Phys. A: Math. Theor. **46** (2013) 254019 (26pp) doi:10.1088/1751-8113/46/25/254019

REVIEW

Finite size Lyapunov exponent: review on applications

Massimo Cencini [1] **and Angelo Vulpiani** [1,2]

[1] Istituto dei Sistemi Complessi, Consiglio Nazionale delle Ricerche, Via dei Taurini 19, I-00185 Rome, Italy
[2] Dipartimento di Fisica, Università 'Sapienza', Piazzale A Moro 2, I-00185 Rome, Italy

E-mail: massimo.cencini@cnr.it and angelo.vulpiani@roma1.infn.it

Received 5 July 2012, in final form 25 September 2012
Published 4 June 2013
Online at stacks.iop.org/JPhysA/46/254019

Abstract

In dynamical systems, the growth of infinitesimal perturbations is well characterized by the Lyapunov exponents. In many situations of interest, however, important phenomena involve finite amplitude perturbations, which are ruled by nonlinear dynamics out of tangent space, and thus cannot be captured by the standard Lyapunov exponents. We review the application of the finite size Lyapunov exponent (FSLE) for the characterization of non-infinitesimal perturbations in a variety of systems. In particular, we illustrate their usage in the context of predictability of systems with multiple spatio-temporal scales of geophysical relevance, in the characterization of nonlinear instabilities, and in some aspects of transport in fluid flows. We also discuss the application of the FSLE to more general aspects such as chaos-noise detection and coarse-grained descriptions of signals.

This article is part of a special issue of *Journal of Physics A: Mathematical and Theoretical* devoted to 'Lyapunov analysis: from dynamical systems theory to applications'.

PACS numbers: 05.45.−a, 05.45.Pq, 47.27.ed

(Some figures may appear in colour only in the online journal)

1. Introduction

The Lyapunov exponents (LEs) $\{\lambda_i\}$ and the Kolmogorov–Sinai (KS) entropy h_{KS} are mathematically well-established quantities [1–3]: they are intrinsic properties of a dynamical system, invariant under change of variables and independent of the used norm. In their definition, the limits of arbitrary resolution and asymptotic times must be imposed, this means that LEs describe the long-time growth of infinitesimally small perturbations. In low-dimensional systems with a single characteristic time, the LE and KS entropy are typically sufficient for characterizing the main aspects of predictability and complexity. However, in high-dimensional systems and in the presence of many characteristic times and scales, such as in fully developed turbulence [4], they provide only a partial description. Such a problem is particularly important for complex systems such as, for instance, those relevant to geophysics. In this respect, Lorenz already realized the necessity to go beyond the understanding of very small perturbations [5]:

Small errors in the finer structure—e.g. the positions of individual clouds—tend to grow much more rapidly, doubling in hours or less [. . .]. Errors in the finer structure, having attained appreciable size, tend to produce errors in the coarser structures [. . .]. Certain special quantities [. . .] may be predictable at a range at which entire weather patterns are not.

The aim of this paper is to review the applications of the finite size Lyapunov exponent (FSLE), originally introduced to characterize predictability in turbulence [6], for the characterization of the growth of small but non-infinitesimal perturbations in a wider context, including signal classification, transport and mixing in fluids. At the basis of the FSLE, and similar approaches [7–9], is the idea of relaxing the request for arbitrarily small perturbations and thus to quantify the growth rate at any given scale reached by the perturbation. The price to pay in this path toward the nonlinear regime is a partial loss of mathematical rigor: the FSLE, unlike the LEs, can depend on the used variables and the used norm. Such dependence is not necessarily negative as it typically reflects some aspects of high-dimensional systems. Some mathematical rigor can be recovered by considering the generalization of the KS entropy to finite resolutions in terms of the ε-entropy [10, 11].

The paper is organized as follows. In section 2, after recalling the definition of LEs, we illustrate, by means of a simple example, the importance to go beyond the regime of infinitesimal perturbations for predictability issues. Then we introduce the FSLE and the ε-entropy. Section 3 is devoted to the characterization of predictability in systems with several characteristic times and scales as found in geophysics, turbulence or high-dimensional chaotic systems. The use of the FSLE and ε-entropy for the classification of signals is discussed in section 4. In section 5, we present a few simple examples of systems characterized by an unusual property, namely the growth rate of finite perturbations is larger than that of infinitesimal ones. Applications to transport and mixing in fluids are reviewed in section 6.

2. Generalization of Lyapunov analysis to non-infinitesimal perturbations

2.1. Lyapunov exponents and infinitesimal perturbations

We start by briefly recalling the basic aspects of the LEs. For the sake of simplicity, we consider a discrete-time dynamical system as defined by the map

$$\boldsymbol{x}(t+1) = \boldsymbol{G}(\boldsymbol{x}(t)) \tag{1}$$

and assume that the motion takes place in a bounded region of $\mathbb{R}^d$. We are interested in the evolution of the separation between two trajectories, $\boldsymbol{x}(t)$ and $\boldsymbol{x}'(t)$, starting from close initial conditions, $\boldsymbol{x}(0)$ and $\boldsymbol{x}'(0) = \boldsymbol{x}(0) + \delta\boldsymbol{x}(0)$, respectively. As long as the difference $\delta\boldsymbol{x}(t) = \boldsymbol{x}'(t) - \boldsymbol{x}(t)$ remains small (strictly speaking, infinitesimally small, $|\delta\boldsymbol{x}(t)| \to 0$), it behaves as a vector $\boldsymbol{z}(t)$ in the tangent space, and evolves as

$$z_i(t+1) = \sum_{j=1}^{d} \left.\frac{\partial G_i}{\partial x_j}\right|_{x(t)} z_j(t). \tag{2}$$

Under a rather general hypothesis, Oseledec [1] proved that, for almost all initial conditions $\boldsymbol{x}(0)$ at every point $\boldsymbol{x}(t)$ of the trajectory, there exists an orthonormal basis $\{\mathbf{e}_i(\boldsymbol{x}(t))\}$ in tangent space such that, for large times, the vector $\boldsymbol{z}$ can be written as

$$\boldsymbol{z}(t) = \sum_{i=1}^{d} c_i \mathbf{e}_i(\boldsymbol{x}(t))\, \mathrm{e}^{\lambda_i t}, \tag{3}$$

where the coefficients $\{c_i\}$ depend on $\mathbf{z}(0)$. The exponents $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_d$ are called the *characteristic LEs*. If the dynamical system has an ergodic invariant measure, the spectrum of LEs $\{\lambda_i\}$ does not depend on the initial condition $\boldsymbol{x}(0)$, except for a set of measure zero with respect to the invariant measure.

The first LE λ_1 has an important role in the issue of predictability: if the initial state is known with accuracy δ (infinitesimal) and we ask for how long the state of the system can be predicted within a tolerance, say Δ (also infinitesimal), then the exponential amplification of the initial error implies for the predictability time

$$T_p(\delta, \Delta) = \frac{1}{\lambda_1} \ln\left(\frac{\Delta}{\delta}\right) \sim \frac{1}{\lambda_1}\,. \tag{4}$$

In other terms, the predictability time of infinitesimal perturbation is essentially given by the inverse of maximal LE, but for a weak logarithmic dependence on the ratio between threshold tolerance and initial error.

2.2. Why beyond the limit of infinitesimal perturbations: a simple example

We stress that in the definition of the LE two limits are involved: the perturbation must remain infinitesimal (this is taken into account using the tangent vector) and time must be arbitrarily long (for Oseledec theorem to apply). What does happen when relaxing these constraints? This question, as illustrated by the following simple example, concerns many situations of physical significance where these limits cannot be realized or where considering infinitesimally small disturbances may not be only unnecessary but also misleading.

Following [12], $\boldsymbol{x} \in \mathbb{R}^2$ and $y \in \mathbb{R}$, we consider the discrete-time map

$$\begin{cases} \boldsymbol{x}(t+1) & = \mathbb{R}[\theta]\,\boldsymbol{x}(t) + c\boldsymbol{f}(y(t)) \\ y(t+1) & = g(y(t)) \end{cases} \tag{5}$$

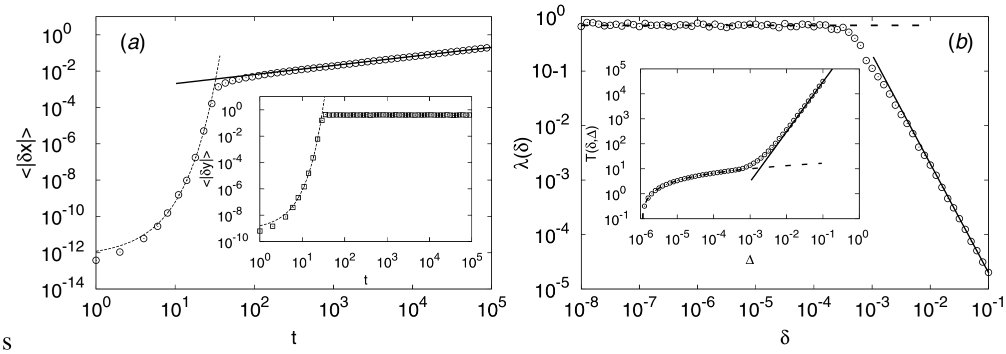
obtained by coupling the rotation $\mathbb{R}[\theta]$ by an arbitrary angle θ on the plane to a chaotic map g via a vector-valued function $\boldsymbol{f}$. For instance, we can take the linear coupling $\boldsymbol{f}(y) = (y, y)$ and the logistic map at the Ulam point $g(y) = 4y(1-y)$. In the absence of coupling, $c=0$, equation (5) describes two independent systems: the predictable and regular $\boldsymbol{x}$-subsystem with $\lambda_{\boldsymbol{x}}(c=0) = 0$ and the chaotic y-subsystem with $\lambda_y = \lambda_1 = \ln 2$. With a small coupling, $0 < c \ll 1$, we have a unique three-dimensional chaotic system with a positive 'global' LE $\lambda_1 = \lambda_y + \mathcal{O}(c)$. In this case, a direct application of equation (4) would imply that the predictability time of the $\boldsymbol{x}$-subsystem is

$$T_p^{(x)} \sim T_p \sim \frac{1}{\lambda_y}\,, \tag{6}$$

according to which the predictability time for $\boldsymbol{x}$ would be basically independent of the coupling strength c, which appears to be at odds with intuition.

We stress that this paradoxical circumstance is not an artifact of the chosen example. For instance, the same happens considering the gravitational three-body problem with one body (the asteroid) of mass m much smaller than the other two (the Sun and one planet). If the gravitational feedback of the asteroid on the two big bodies is neglected (restricted problem), then the result is a chaotic asteroid while the system Sun plus planet is fully predictable. If the feedback is taken into account (i.e. $m > 0$ in the example), the system becomes a fully chaotic non-separable three-body problem. Intuition correctly suggests that if the asteroid has a very small mass ($m \to 0$), then it should be possible to forecast the evolution of the Sun and the planet for very long times in spite of a positive, possibly large, LE for the whole system.

The 'paradox' arises from the misuse of formula (4), which is valid only for the tangent-vector dynamics, i.e. with both δ and Δ infinitesimal. In other words, it stems from the



s

Figure 1. (*a*) Error growth $|\delta\mathbf{x}(t)|$ for the map (5) with $\theta = (\sqrt{5}-1)/2$ and $c = 10^{-3}$. Dashed line $|\delta\mathbf{x}(t)| \sim e^{\lambda_1 t}$, where $\lambda_1 = \ln 2$, solid line $|\delta\mathbf{x}(t)| \sim t^{1/2}$. Inset: evolution of $|\delta y(t)|$, dashed line as in the main figure. The initial error, only on the y variable, is $\delta y = \delta_0 = 10^{-10}$. (*b*) FSLE $\lambda(\delta)$ versus δ for the same system, computed by using the Euclidean norm in the x-subspace. For $\delta \to 0$, $\lambda(\delta) \simeq \lambda_1$ (dashed line). The solid line displays the behavior $\lambda(\delta) \sim \delta^{-2}$. Inset: predictability time $T_p(\delta, \Delta)$ with $\delta = 10^{-6}$ and varying Δ as from equation (13), the dashed line corresponds to equation (4), while the solid line corresponds to the behavior (8). See section 2.3.

application of the correct formula (4) to an improper regime, because as soon as errors become large, the full nonlinear error evolution has to be taken into account (see figure 1(*a*)). The evolution of $\delta\boldsymbol{x}$ is given by

$$\delta\boldsymbol{x}(t+1) = \mathbb{R}[\theta]\delta\boldsymbol{x}(t) + c\,\delta\boldsymbol{f}(y), \tag{7}$$

where, with our choice, $\delta\boldsymbol{f} = (\delta y, \delta y)$. At the beginning, both $|\delta\boldsymbol{x}|$ and $|\delta y|$ grow exponentially. However, the available phase space for y is bounded leading to a saturation of the uncertainty $|\delta y| \sim \mathcal{O}(1)$ in a time $t^\star = \mathcal{O}(1/\lambda_1)$. Therefore, for $t > t^\star$, the two realizations of the y-subsystem are completely uncorrelated and their difference δy acts as deterministic noise in (7), which becomes a sort of discrete-time Langevin equation driven by chaos, instead of noise. As a consequence, the growth of the uncertainty on the $\boldsymbol{x}$-subsystem becomes diffusive with a diffusion coefficient proportional to c^2, i.e. $|\delta\boldsymbol{x}(t)| \sim c\,t^{1/2}$ implying [12]

$$T_p^{(x)} \sim \left(\frac{\Delta}{c}\right)^2, \tag{8}$$

which is much longer than the time expected on the basis of tangent-space error growth (now Δ is not constrained to be infinitesimal).

The above simple example illustrates the necessity to go beyond the LEs when copying with non-infinitesimal perturbations; we will see in the course of this review that in many circumstances the full characterization of the nonlinear growth regime is necessary to answer physically motivated issues, of which predictability, as here, is one of the most relevant.

2.3. The finite size Lyapunov exponent (FSLE)

Before defining the FSLE, originally introduced in [6, 13] to quantify predictability in turbulence, we mention a few other approaches to the characterization of non-infinitesimal perturbations. Dressler and Farmer [7] introduced a generalization of the LE based on higher order derivatives, relevant to the early stage of the nonlinear regime of perturbation growth. Torcini *et al* [8] were able to analytically compute an indicator that is equivalent to the FSLE for two simple maps showing that there are circumstances in which finite size perturbations grow faster than infinitesimal ones (see section 5); Letz and Kantz [9] extended the analysis

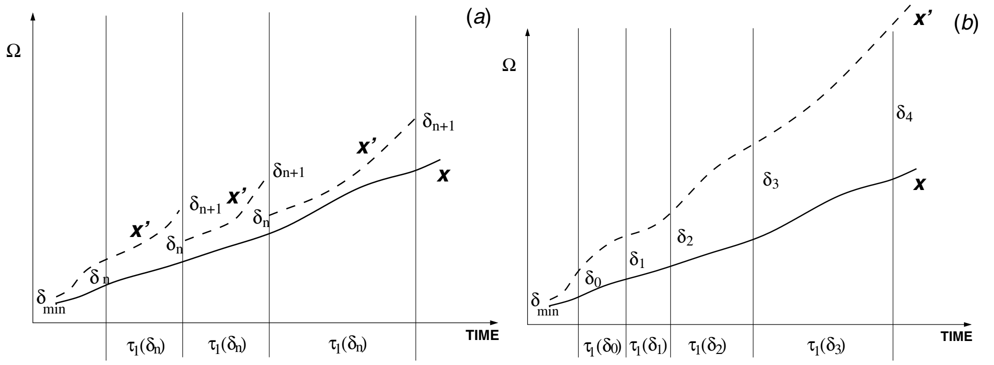


Figure 2. Sketch of the algorithm for computing the FSLE. Panels (*a*) and (*b*) correspond to the first and second algorithms described in the text, respectively. Reproduced with permission from [116]. Copyright 2009 World Scientific.

of these kinds of systems by introducing a scale-dependent stability indicator similar in spirit to the FSLE but based on fixed time analysis instead of fixed scale.

The FSLE quantifies at different observation scales the average growth rate of non-infinitesimal perturbations. Such a quantity has a less firm mathematical ground than the LEs, and we will introduce it operatively through the algorithm for its computation. Assume that a system has been evolved for long enough that the transient dynamics has lapsed, e.g., for dissipative systems the motion has settled onto the attractor. Consider at $t = 0$ a 'reference' trajectory $\boldsymbol{x}(0)$ supposed to be on the attractor, and generate a 'perturbed' trajectory $\boldsymbol{x}'(0) = \boldsymbol{x}(0)+\delta\boldsymbol{x}(0)$. We need the perturbation to be initially very small (essentially infinitesimal) in some chosen norm $\delta(t = 0) = ||\delta\boldsymbol{x}(t=0)|| = \delta_{\min} \ll 1$ (typically in numerical experiments $\delta_{\min} = \mathcal{O}(10^{-6}–10^{-8})$). Then, in order to study the perturbation growth through different scales, we can define a set of thresholds δ_n, e.g., $\delta_n = \delta_0\varrho^n$ with $\delta_{\min} \ll \delta_0 \ll 1$, where δ_0 can still be considered infinitesimal and $n = 0, \ldots, N_s$. To avoid saturation on the maximum allowed separation (i.e. the attractor size) attention should be paid to have $\delta_{N_s} < \langle||\boldsymbol{x} - \boldsymbol{y}||\rangle_\mu$ with $\boldsymbol{x}, \boldsymbol{y}$ being the generic points on the attractor and μ the invariant measure of the system. The numerical factor ϱ should be larger than 1 but not too large in order to avoid interferences of different length scales: typically, $\varrho = 2$ or $\varrho = \sqrt{2}$. Now to measure the perturbation growth rate at scale δ_n, we can proceed as follows. After time t_0, the perturbation has grown from $\delta_{\min}$ up to δ_n, ensuring that the perturbed trajectory relaxes on the attractor and aligns along the maximally expanding direction. Then, we measure the time $\tau_1(\delta_n)$ the error needed to grow up to δ_{n+1}, i.e. the first time such that $\delta(t_0) = ||\delta\boldsymbol{x}(t_0)|| = \delta_n$ and $\delta(t_0+\tau_1(\delta_n)) = \delta_{n+1}$. The perturbation is thus rescaled to δ_n, along the direction $\boldsymbol{x}'-\boldsymbol{x}$ (see figure 2(*a*)). The procedure is repeated $\mathcal{N}_d$ times for each threshold, obtaining the set of the 'doubling' times (strictly speaking the name applies for $\varrho = 2$ only) $\{\tau_i(\delta_n)\}$ for $i = 1, \ldots, \mathcal{N}_d$ error-doubling experiments. Note that $\tau(\delta_n)$ also depends on ϱ. The doubling rate

$$\gamma_i(\delta_n) = \frac{1}{\tau_i(\delta_n)} \ln\varrho\,, \tag{9}$$

when averaged defines the FSLE $\lambda(\delta_n)$ through the relation

$$\lambda(\delta_n) = \langle\gamma(\delta_n)\rangle_t = \frac{1}{T}\int_0^T \mathrm{d}t\ \gamma = \frac{\sum_i \gamma_i\tau_i}{\sum_i \tau_i} = \frac{\ln\varrho}{\langle\tau(\delta_n)\rangle_d}\,, \tag{10}$$

where $\langle\tau(\delta_n)\rangle_d = \sum\tau_i/\mathcal{N}_d$ is the average over the doubling experiments and the total duration of the trajectory is $T = \sum_i \tau_i$. Equation (10) assumes the distance between the two trajectories to be continuous in time. For maps or time-continuous systems sampled at discrete times the method has to be slightly modified defining $\tau(\delta_n)$ as the minimum time such that $\delta(\tau) \geqslant \varrho\delta_n$. In these cases, moreover, $\delta(\tau)$ is a fluctuating quantity, and from (10) we have

$$\lambda(\delta_n) = \frac{1}{\langle\tau(\delta_n)\rangle_d}\left\langle \ln\left(\frac{\delta(\tau(\delta_n))}{\delta_n}\right)\right\rangle_d. \tag{11}$$

When δ_n is infinitesimal, $\lambda(\delta_n)$ recovers the maximal LE

$$\lim_{\delta\to 0}\lambda(\delta) = \lambda_1. \tag{12}$$

We must underline that, unlike the standard LE, when δ is finite $\lambda(\delta)$ depends on the chosen norm. This is not due to an ill-definition of $\lambda(\delta)$ as it applies also to the mathematically well-defined ε-entropy, see section 2.4. The dependence on the norm is a manifestation of the fact that in the nonlinear regime the predictability time depends on the chosen observable.

A possible problem with the above-described algorithm to compute (and define) $\lambda(\delta)$ is the implicit assumption of homogeneity with respect to finite perturbations. Typically, the measure on the attractor is singular and not equally dense at all distances; this may cause an incorrect sampling of the doubling times at large δ_n when rescaling the perturbation. To cure such a problem, the algorithm can be modified to avoid the rescaling at finite δ_n as follows. The thresholds $\{\delta_n\}$ and the initial perturbation ($\delta_{\min} \ll \delta_0$) are chosen as before, but now the perturbation growth is followed from δ_0 to δ_{N_s} without rescaling back the perturbation once the threshold is reached (see figure 2(*b*)). In practice, after the system reaches the first threshold δ_0, we measure the time $\tau_1(\delta_0)$ to reach δ_1, then following the same perturbed trajectory we measure the time $\tau_1(\delta_1)$ to reach δ_2, and so forth up to δ_{N_s}: we thus record the time $\tau(\delta_n)$ for going from δ_n to δ_{n+1} for each value of n. The evolution of the error from the initial value $\delta_{\min}$ to the largest threshold δ_N carries out a single error-doubling experiment, and the FSLE is finally obtained by using (10) or (11), which are accurate also in this case, according to the continuous-time or discrete-time nature of the system, respectively. As finite perturbations are realized by the dynamics (i.e. the perturbed trajectory is on the attractor), the problem of the attractor inhomogeneity is no longer present.

Even though some differences between the two methods are possible for large δ, they should give the same result for $\delta \to 0$. In most cases, numerical computations show that the differences between the two methods are typically very tiny at all scales. It is worth noting that the two above algorithms, being based on doubling times, cannot detect negative LEs. In section 5, we will briefly mention a modification of the first algorithm to be used when $\lambda(\delta) < 0$.

Let us now go back to the example equation (5); figure 1(*b*) displays the functional shape of the FSLE computed with the second method described. For $\delta \ll 1$, a plateau at the value of maximal LE $\lambda_1 \approx \ln 2$ is recovered as from the limit (12). For finite δ, the behavior of $\lambda(\delta)$ in general depends on the details of the nonlinear dynamics; here the diffusive behavior (seen in figure 1(*a*)) implies $\lambda(\delta) \sim \delta^{-2}$ as suggested by dimensional analysis. Now that we have introduced an indicator to quantify the error growth rate as a function of the error amplitude, we can compute the predictability time from an initial error δ to a given tolerance Δ as

$$T_p(\delta, \Delta) = \int_{\delta}^{\Delta} \frac{\mathrm{d}\ln\delta'}{\lambda(\delta')}. \tag{13}$$

Clearly, in the infinitesimal regime, $\lambda(\delta) \approx \lambda_1$ and (13) recovers (4) but, out of the tangent space, $\lambda(\delta)$ depends on the details of the dynamics and T_p can be much longer than expected from the LE. For instance, in this case $\lambda(\delta) \propto \delta^{-2}$, so that from (13) we have $T_p(\delta, \Delta) \propto \Delta^2$ as guessed on dimensional grounds in equation (8); see the inset of figure 1(*b*).

We conclude this section with a final remark on the FSLE. Denoting with $\boldsymbol{x}(t)$ and $\boldsymbol{x}'(t)$ a reference trajectory and a perturbed trajectory of a given dynamical system, respectively, and with $R(t) = |\boldsymbol{x}(t) - \boldsymbol{x}'(t)|$ their separation, we can define a scale-dependent growth rate also using

$$\tilde{\lambda}(\delta) = \frac{1}{2\langle R^2(t)\rangle} \left.\frac{\mathrm{d}\left\langle R^2(t)\right\rangle}{\mathrm{d}t}\right|_{\left\langle R^2\right\rangle=\delta^2} \quad \text{or} \quad \tilde{\lambda}(\delta) = \left.\frac{\mathrm{d}\left\langle \ln R(t)\right\rangle}{\mathrm{d}t}\right|_{\langle \ln R(t)\rangle = \ln\delta}. \tag{14}$$

This way to define a scale-dependent indicator is somewhat similar to that introduced in [9]. We note, however, that $\tilde{\lambda}(\delta)$ is in general different from the FSLE $\lambda(\delta)$, as $\left\langle R^2(t)\right\rangle$ usually depends on $\left\langle R^2(0)\right\rangle$ while $\lambda(\delta)$ depends only on δ. This difference has an important conceptual and practical consequence, for instance, when considering the relative dispersion of two tracer particles in turbulence or geophysical flows [14, 15] (see section 6). However, we must note that the procedure used to define the FSLE cannot account for negative rates, while $\tilde{\lambda}(\delta)$ can also be negative.

2.4. A more rigorous scale-dependent indicator: the ε-entropy

In this section, we briefly discuss the ε-entropy [10, 11] (see also [16]) which measures the amount of information per unit time necessary to record without ambiguity a generic trajectory of a chaotic system with ε-accuracy, and which can be related to the FSLE. Indeed, as the FSLE generalizes the (maximum) LE to the nonlinear regime of perturbation growth, the ε-entropy generalizes the KS entropy [2, 3] to a coarse-grained description. For systems with only one positive LE, we expect that the two quantities should provide essentially equivalent information. However, the ε-entropy, unlike the FSLE, has a mathematical firm ground.

Consider a continuous vector $\boldsymbol{x}(t) \in \mathbb{R}^d$ (with continuous time), representing the state of a d-dimensional system which can be either deterministic or stochastic (as the ε-entropy is well defined also in the latter case). Discretize the time by introducing an interval τ and consider a partition A_ε of the phase space in cells with edges (diameter) $\leqslant \varepsilon$. The partition may be composed of unequal or identical cells: hypercubes of side ε are typically used in practical computations. The partition induces a symbolic dynamics, for which a portion of trajectory

$$\boldsymbol{X}^{(N)}(t) \equiv \{\boldsymbol{x}(t), \boldsymbol{x}(t+\tau), \ldots, \boldsymbol{x}(t+(N-1)\tau)\} \in \mathbb{R}^{Nd} \tag{15}$$

can be coded into a word of length N, from a finite alphabet:

$$\boldsymbol{X}^{(N)}(t) \longrightarrow \mathcal{W}_N(\varepsilon, t) = (s(\varepsilon, t), s(\varepsilon, t+\tau), \ldots, s(\varepsilon, t+(N-1)\tau)), \tag{16}$$

where $s(\varepsilon, t+j\tau)$ labels the cell in $\mathbb{R}^d$ containing $\boldsymbol{x}(t+j\tau)$. The alphabet is finite for bounded motions, which can be covered by a finite number of cells. Assuming ergodicity, from a long time record of $\boldsymbol{X}^{(N)}(t)$ we can estimate the probabilities $P(\mathcal{W}_N(\varepsilon))$ of the admissible words $\{\mathcal{W}_N(\varepsilon)\}$, and thus compute the N-block entropies

$$H_N = -\sum_{\mathcal{W}_N(\varepsilon)} P(\mathcal{W}_N(\varepsilon)) \ln P(\mathcal{W}_N(\varepsilon)). \tag{17}$$

Then, following Shannon [17], the (ε, τ)-entropy per unit time, $h(A_\varepsilon, \tau)$ associated with the partition A_ε is obtained as

$$h_N(A_\varepsilon, \tau) = \frac{1}{\tau}[H_N(A_\varepsilon, \tau) - H_{N-1}(A_\varepsilon, \tau)] \tag{18}$$

$$h(A_\varepsilon, \tau) = \lim_{N\to\infty} h_N(A_\varepsilon, \tau) = \frac{1}{\tau} \lim_{N\to\infty} \frac{H_N(A_\varepsilon, \tau)}{N}. \tag{19}$$

The dependence on τ is retained as in some stochastic systems the ε-entropy may depend on it [16]. Moreover, τ can be important in practical implementations.

The (ε, τ)-entropy can be defined as a partition-independent quantity by taking the infimum over all partitions with cells of diameter smaller than ε [10, 16]:

$$h(\varepsilon, \tau) = \inf_{A:\mathrm{diam}(A)\leqslant\varepsilon} \{h(A_\varepsilon, \tau)\}\,. \tag{20}$$

The infimum in definition (20) is chosen because for continuous stochastic processes, for any ε, $\sup_{A:\mathrm{diam}(A)\leqslant\varepsilon} \{h(A_\varepsilon, \tau)\} = \infty$ as it recovers the Shannon entropy of an infinitely refined partition. It should be remarked that, for $\varepsilon \neq 0$, $h(\varepsilon, \tau)$ depends on the actual definition of diameter, i.e. on the used norm as for the FSLE.

For deterministic systems, the ε-entropy defined by equation (20) can be shown to be independent of τ [18, 19]. Moreover, in the limit $\varepsilon \to 0$, the KS entropy, h_{KS}, is recovered

$$h_{\mathrm{KS}} = \lim_{\varepsilon\to 0} h(\varepsilon, \tau), \tag{21}$$

we recall that thanks to the Pesin relation [20] we have $h_{\mathrm{KS}} \leqslant \sum_{\lambda_i>0} \lambda_i$. Unlike the KS entropy, which is a number, the ε-entropy is a function of the observation resolution ε and its behavior as a function of ε provides information on the dynamical properties of the underlying system [16, 21]. For systems with only one positive LE $h_{\mathrm{KS}} = \lambda_1$ and the limit (21) is equivalent to (12), while at finite ε the ε-entropy $h(\varepsilon)$ and the FSLE $\lambda(\varepsilon)$ provide equivalent information, though differences between the two quantities may be present. For instance, definition (20) implies that $h(\varepsilon)$ is a non-increasing function of ε, while this might not be the case for the FSLE (see, e.g., section 5).

The first possibility of computing the ε-entropy is by using for any fixed ε the symbolic dynamics which results from an equal cell partition. Of course, taking the infimum over all partitions is practically impossible and thus some of the nice properties of the 'mathematically well-defined' ε-entropy will be lost in numerical estimations. Moreover, implementing directly the Shannon definition is sometimes rather time consuming, and faster estimators are necessary [21].

The (ε, τ)-entropy $h(\varepsilon, \tau)$ is well defined also for stochastic processes. Actually, the dependence of $h(\varepsilon, \tau)$ on ε can give some insight into the underlying stochastic process [16]; for instance, in the case of a stationary Gaussian process with the spectrum $S(\omega) \propto \omega^{-(1+2\alpha)}$ with $0 < \alpha < 1$, one has [11]

$$\lim_{\tau\to 0} h(\varepsilon, \tau) \sim \varepsilon^{-1/\alpha}\,. \tag{22}$$

However, the above behavior may be difficult to observe mainly due to problems related to the choice of τ [16, 22].

3. Characterization of predictability in systems with a multiscale structure

3.1. Systems with slow and fast components

We consider here systems with a multiscale structure, where the state of the system $\boldsymbol{x} = (\boldsymbol{X}, \boldsymbol{y})$ can be decomposed into a slow component $\boldsymbol{X}$ which is also the 'largest' one, and a fast component $\boldsymbol{y}$ 'small' with respect to $\boldsymbol{X}$. (This means for example that the typical root mean square (rms) values of the fast variables are smaller than those of the slow ones $y_{\mathrm{rms}} \ll X_{\mathrm{rms}}$.) This kind of situation, though a bit idealized, is common, e.g., in geophysics where several 'subsystems' with different time and spatial scales can be identified, for instance, the coupled system of the ocean and the atmosphere [23]. As a mathematical prototype of such kinds of systems, we discuss here a model introduced by Lorenz in 1996 [24] to study the predictability problem in the atmosphere, where indeed a multiscale structure is present. The model

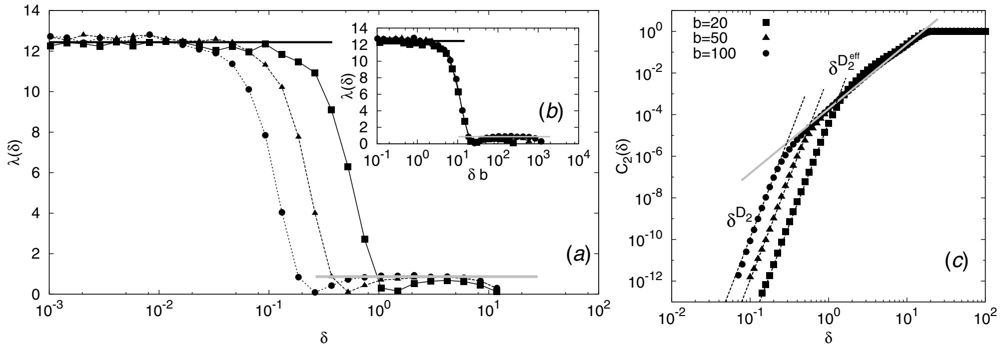


Figure 3. FSLE and correlation dimension computed for the Lorenz96 model ((23)–(24)) with $N = 5$, $K = 10$ forcing $F = 10$, time separation $c = 10$ and coupling $h = 1$ for three values of the amplitude separation $b = 20, 50, 100$. (*a*) $\lambda(\delta)$ versus δ computed using the Euclidean norm in the full $(\mathbf{X}, \mathbf{y})$ phase space with $\delta_{\min} = 10^{-12}$ and $\delta_0 = 10^{-6}$. (*b*) The same as (*a*) but rescaling δ with the scale-separation factor b. The black and gray horizontal lines display the plateaus of $\lambda(\delta)$ to $\lambda_1 = \lambda_{\text{fast}} \approx 12$ and $\lambda_{\text{slow}} \approx 0.9 \approx \lambda_{\text{fast}}/c$, see the text. (*c*) Correlation integral as a function of the scale δ; the straight lines show the measured correlation dimensions $D_2 \approx 9.8$ and $D_2^{\text{slow}} \approx 3.1$, see the text.

reads

$$\frac{\mathrm{d}X_n}{\mathrm{d}t} = X_{n-1}(X_{n+1} - X_{n-2}) - X_n + F - \frac{hc}{b}\sum_{k=1}^{K} y_{k,n} \tag{23}$$

$$\frac{\mathrm{d}y_{k,n}}{\mathrm{d}t} = cb\, y_{k+1,n}(y_{k-1,n} - y_{k+2,n}) - c\, y_{k,n} + \frac{hc}{b}X_n, \tag{24}$$

where $n = 1, \ldots, N$ and $k = 1, \ldots, K$ with the boundary conditions $X_{N\pm n} = X_{\pm n}$, $y_{K+1,n} = y_{1,n+1}$ and $y_{0,n} = y_{K,n-1}$. The above system can be regarded as a one-dimensional caricature of atmospheric motion. The slow variables X_n may be thought of as the values of some atmospheric representative observable along the latitude circle, while the fast variables $\boldsymbol{y}$, which evolve with similar dynamics but are c times faster and b times smaller in amplitude (i.e. $y_{\text{rms}} \approx X_{\text{rms}}/b$), can be seen as representing some convective-scale quantity [24]. The dynamical features of the system ((23)–(24)) are completely determined by the forcing strength F and by the system dimensionality NK; see [25] for a recent study of the model.

As discussed by Lorenz [24] (see also [26] for an early study of the above system with the FSLE), in considering the predictability or, more in general, the error growth in a system such as ((23)–(24)) we should distinguish questions related to the early and the late times of the error evolution, and also specify the size of the errors we are considering. This is well exemplified in figure 3(*a*) where we show the FSLE $\lambda(\delta)$ computed considering an initial error uniformly spread over all the (fast and slow) components, and the error is measured using the Euclidean norm in the full $(\boldsymbol{X}, \boldsymbol{y})$ phase space.

When the perturbation is small, $\delta \ll y_{\text{rms}} \approx X_{\text{rms}}/b$; thanks to (12) the FSLE recovers the maximal LE of the full system that is essentially controlled by the fast subsystem $\lambda_1 = \lambda_{\text{fast}}$. On the other hand, for larger perturbations, $\delta \gtrsim y_{\text{rms}} \approx X_{\text{rms}}/b$, after a fast decreasing related to the saturation of the fast dynamics, the FSLE stabilizes on a second plateau essentially controlled by the slow component, i.e. $\lambda(\delta) \approx \lambda_{\text{slow}} \approx \lambda_1/c$. Figure 3(*b*) displays the same data as a function of $b\delta$; the collapse of the three curves confirms that the slow-component-controlled plateau establishes after the fast dynamics saturates.

We note that one could have considered the initial error present only in the fast degrees of freedom and measured the FSLE using the Euclidean norm only in the slow components.

With the latter procedure (not shown) one would have obtained a slightly different FSLE (as different norms are used). However, remarkably, the two limiting plateaus for small and large δ seem to be independent of the chosen norm. This observation suggests that the evolution of finite-size perturbations, which is fully nonlinear for the fully coupled fast–slow system ((23)–(24)) is actually controlled by the linear dynamics of an effective lower dimensional system. To test this idea we can, for example, measure the correlation dimension of the system on varying the observation resolution as in [27]. The correlation dimension can be obtained by the correlation sum [28] here estimated as

$$C_{m,M}(\delta) = \frac{1}{Mm} \sum_{k=1}^{m} \sum_{j=1}^{M} \Theta(\delta - |\boldsymbol{x}_j - \boldsymbol{x}_\star^{(k)}|), \tag{25}$$

where we consider M points of a long trajectory sampled at discrete-time intervals ($\Delta t = 1$), and m reference points $\{\boldsymbol{x}_\star^{(k)}\}$ on the attractor. Figure 3(c) shows $C_{m,M}$ as a function of δ. The correlation dimension of the attractor defined by the whole system D_2, given by the scaling $C_{m,M}(\delta) \sim \delta^{D_2}$ for $\delta \ll y_{\rm rms} \approx X_{\rm rms}/b$, is rather large ($D_2 \approx 10$). However, for $\delta > y_{\rm rms} \approx X_{\rm rms}/b$, we see a second power law $C_{m,M}(\epsilon) \sim \epsilon^{D_2^{\rm slow}}$ with $D_2^{\rm slow} \approx 3 < D_2$ which defines a sort of 'effective dimension at large scale'.

The above results suggest that the effective dynamics at large scale can be described by fewer degrees of freedom (3 or 4), and can be predicted for longer time than the full dynamics. In particular, the largest LE is expected to be $\lambda_{\rm slow} \approx \lambda_1/c$, as identified by the FSLE in figures 3(a) and (b). It would be interesting to generalize the FSLE to account also for sub-leading FSLEs, and thus to have a spectrum of LEs at large scale. Unfortunately, at present it is not clear if and how this can be done due to some technical difficulties[3]. We also note that the dimensional reduction operating for the large-scale dynamics is somehow suggesting the possibility of building reduced models parametrizing the fast (small-scale) dynamics as expected in systems with different characteristic times. However, the parametrization may be delicate and modeling small/fast degrees of freedom is not, in general, an easy task [29, 30].

The above approach can also be used for separating the fast and slow unstable modes in coupled systems of geophysical relevance with different timescales, e.g. in the El Niño–Southern oscillation as in [31] where the link between the FSLE and the so-called breeding vectors is established. Kalnay and co-workers introduced the breeding method for the study of finite amplitude perturbations [32, 33]. Such a technique, which has many similarities with the FSLE, consists of adding an initial perturbation of size δ to a reference trajectory, integrating forward both the perturbed and unperturbed trajectories and periodically rescaling (every time interval ΔT) the amplitude of the perturbation to the initial value δ. With a proper choice of δ and ΔT (in the physically appropriate scales) it is possible to estimate, e.g., the shape of the baroclinic instabilities [34]. Although in realistic applications, like atmospheric forecasting, it is necessary to face a number of practical aspects, ranging from small scales modeling to the problem of determining the initial state from incomplete observations, dynamical systems tools, such as the FSLE, extending the study of perturbation dynamics to the nonlinear regime have been rather useful (see, e.g., the recent review [35]).

[3] For example, to compute the first and second LEs in the non-infinitesimal regime, three trajectories are needed (a reference and two perturbed trajectories). This is not a problem in tangent space. But for finite perturbation, one needs all the trajectories be on the attractor, and this is not guaranteed by the standard orthonormalization procedure.

3.2. Large-scale predictability in fully developed turbulence

We start by recalling the basic aspects of predictability in turbulence; we refer to [6, 13, 36] for more elaborate treatments [37]. Essentially, we call turbulent the state of motion of a fluid (of viscosity ν) vigorously stirred by a force acting at a typical scale L, so that the Reynolds number $Re = UL/\nu$ becomes very large (U being the typical large-scale fluid velocity). Under such conditions the fluid velocity field $\boldsymbol{v}(\boldsymbol{x}, t)$ is dominated by the nonlinear terms of the Navier–Stokes equation and behaves chaotically, becoming so irregular and complex that a statistical description is mandatory [38]. The main phenomenological features were established by Kolmogorov in 1941 (K41) who argued that the velocity power spectrum $E_v(k)$ develops a power-law behavior $E_v(k) \sim k^{-5/3}$ from wave numbers corresponding to the scale of excitation L to that of dissipation (the Kolmogorov length ℓ_D)—i.e. in the so-called inertial range. The $-5/3$-spectrum implies that $\langle(\delta_\ell v)^2\rangle \propto \ell^{2/3}$, where $\delta_\ell v = (\boldsymbol{v}(\boldsymbol{x}+\boldsymbol{\ell}, t) - \boldsymbol{v}(\boldsymbol{x}), t) \cdot \boldsymbol{\ell}/\ell$ is the longitudinal velocity fluctuation at scale $\ell = |\boldsymbol{\ell}|$. Moreover, he derived from the Navier–Stokes equations an asymptotically Re exact relation $\langle(\delta_\ell v)^3\rangle = -(4/5)\epsilon r$ (where ϵ is the energy dissipation rate). From these two observations, we can conjecture that, in a statistical sense, fluctuations over a scale $\ell_D \ll \ell \ll L$ behave as

$$\delta_\ell v \sim \ell^{1/3}, \tag{26}$$

for $\ell \ll \ell_D$ one expects $\delta_\ell v \propto \ell$ due to the dissipative smoothing, while for $\ell > L$ correlations disappear so that $\delta_\ell v \approx U$. Equation (26) suggests that $\langle(\delta_\ell v)^p\rangle \propto \ell^{p/3}$, which is close to experimental observations but for small corrections (important at large p) which are a manifestation of intermittency [38, 37], ignored here.

The classical theory of predictability in turbulence was developed by Lorenz [4] using physical arguments confirmed by more refined treatments [39, 40]. Lorenz's approach stems from the assumption that the time needed for a perturbation at scale $\ell/2$ to induce a complete uncertainty on the velocity field at scale ℓ is proportional to the characteristic time τ_ℓ of the scale ℓ. From equation (26), τ_ℓ can be estimated as

$$\tau_\ell \sim \ell/\delta_\ell v \sim \tau_L(\ell/L)^{2/3}, \tag{27}$$

which increases with the scale ℓ. The fastest characteristic time would be that associated with the Kolmogorov length scale τ_{ℓ_D}, which is expected to be of the order of the inverse of the maximal LE λ_1 of turbulence [41, 42].

Because of the algebraic progression (27), the time T_p to propagate an uncertainty from, say, ℓ_D upward to the large scale L is dominated by the slowest timescale τ_L, indeed: $T_p \sim \tau_{\ell_d} + \tau_{2\ell_d} + \cdots + \tau_L \sim \tau_L \sim L/\delta_L v$. Such a result would imply that the predictability time T_p is Reynolds independent, apparently at odds with the fact that the maximum LE increases with Re, i.e. $\lambda_1 \propto \tau_{\ell_D}^{-1} \sim Re^{1/2}$, as predicted by Ruelle [41]. (Actually, small corrections to $1/2$ are expected due to intermittency [42].) However, as argued in the previous sections, the LE plays no role in the growth of large-scale perturbations, and hence there is no contradiction. The above phenomenological considerations can be recast in dynamical systems language using the FSLE $\lambda(\delta)$. (Here $\delta = |\boldsymbol{v} - \boldsymbol{v}'|$ now denotes velocity uncertainties.) In terms of the FSLE, the predictability time $T(\delta, \Delta)$ for an error δ and a given tolerance Δ is obtained summing up the inverse of the error growth rate at all scales as in equation (13), which can be much longer than the naive expectation $T_p \approx \ln(\Delta/\delta)/\lambda_1$ (see equation (4) and the inset of figure 1(*b*)) because, in general, larger errors are characterized by smaller growth rates, i.e.$\lambda(\delta)$ is a decreasing function of δ. In particular, within the phenomenological framework of K41 theory and Lorenz's ideas, we can predict the behavior of $\lambda(\delta)$ when the perturbation is in the inertial range $\delta_{\ell_D} v \ll \delta \ll \delta_L v$. According to Lorenz's argument, the doubling time of an error of amplitude δ is proportional to the turnover time τ_ℓ of an eddy with typical velocity

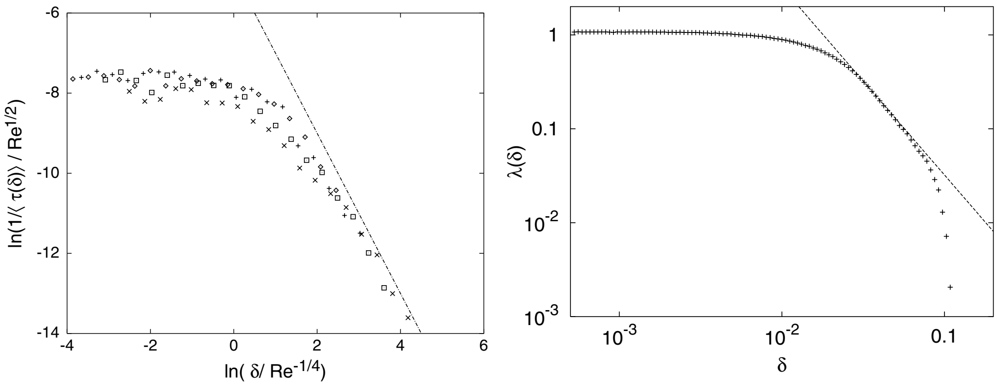


Figure 4. Left: $\ln(1/\langle\tau(\delta)\rangle/Re^{1/2})$ versus $\ln\left(\delta/Re^{-1/4}\right)$ at different Re and N. ($\Diamond$) $N = 24$ and $Re = 10^8$; (+) $N = 27$ and Re $= 10^9$; ($\Box$) $N = 32$ and Re $= 10^{10}$; ($\times$) $N = 35$ and Re $= 10^{11}$. The straight line has slope -2. Right: FSLE $\lambda(\delta)$ versus the velocity uncertainty δ for 2D turbulence in the inverse cascade regime, for details on the numerics see [45]. The asymptotic constant value for $\delta \to 0$ corresponds to λ_1. The dashed line has slope -2. Reproduced with permission from [116]. Copyright 2009 World Scientific.

difference $\delta_\ell v \sim \delta$. Rewriting equation (26) as $\delta_\ell v \sim U(\ell/L)^{1/3}$ and using equation (27) we then obtain $\tau_\ell \sim \tau_L(\ell/L)^{2/3} \sim \tau_L(\delta_\ell v/U)^2$. In the inertial range, the FSLE $\lambda(\delta)$ should be proportional to the inverse of the turnover time associated with velocity uncertainties of size δ; hence,

$$\lambda(\delta) \sim \delta^{-2} \tag{28}$$

remarkably intermittency does not impinge the above scaling behavior as revealed by refining the argument with the multifractal model [6, 13].

Testing the scaling (28) in direct numerical simulations (DNS) of three-dimensional (3D) turbulence presents several difficulties and, moreover, it is currently still difficult to obtain an extended inertial range to verify the δ^{-2} scaling. In figure 4, we show the results obtained in two simpler systems: a set of ordinary differential equations (ODEs), which reproduces most of the phenomenological aspects of turbulence (including intermittency)—namely, the *shell model* [43]—and 2D turbulence in the inverse cascade regime, which can be simulated with an extended inertial range and is essentially well described by the K41 phenomenology, without corrections due to intermittency [44].

The shell model simulated in [6, 13], from whence data are taken, is the so-called GOY model. It consists in a set of ODEs for the complex velocity variables u_n ($n = 1, \ldots, N$) representing velocity fluctuations in a shell of wave numbers $k_n < |\mathbf{k}| < k_{n+1}$. Assuming locality [38], the nonlinear interactions are confined to neighboring shells, and the ODEs read

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} + \frac{k_n^2}{Re}\right) u_n = \mathrm{i}k_n \left(u_{n+1}u_{n+2} - \frac{1}{4}u_{n-1}u_{n+1} - \frac{1}{8}u_{n-2}u_{n-1}\right)^* + f_n\,, \tag{29}$$

where $*$ denotes complex conjugation and f_n is a forcing term (typically restricted to $n = 1$–3). The coefficients in the nonlinear term (having the same structure as Navier–Stokes equations) are chosen to conserve energy $E \equiv 1/2\sum_n |u_n|^2$ in the unforced ($f_n = 0$), inviscid ($Re \to \infty$) limit. The wave numbers are geometrically spaced, $k_n = k_0 2^n$, so that a rather small number of variables, $N \sim \log_2 Re$, are necessary also for high Reynolds numbers. Remarkably, the shell model (29) displays a phenomenology similar to 3D turbulence and, in particular, it reproduces power-law scaling, $\langle|u_n|^p\rangle \sim k_n^{-\zeta_p}$, with exponents close to those experimentally observed for fully developed turbulence, and thus slightly deviating from K41 $\zeta_p \neq p/3$ [43]. Hence, the

GOY model represents a good theoretical laboratory for turbulence where standard methods of deterministic chaos can be used.

The left panel of figure 4 shows the inverse of the doubling time $\langle\tau(\delta)\rangle$, which is proportional to $\lambda(\delta)$, as a function of velocity uncertainty δ for the GOY model (29) [6, 13]. When the perturbation is in the dissipative range, $\delta < \delta v(\ell_D) \sim URe^{-1/4}$, it can be considered infinitesimal so that $\tau(\delta)$ does not depend on δ being proportional to the inverse of the maximal LE, $\tau(\delta) \approx \tau_{\ell_D} \propto 1/\lambda_1 \sim Re^{-1/2}$ [41]. In the inertial range, the scaling (28) is well reproduced. The figure also shows that data obtained with different *Re* are fairly well collapsed onto the same curve when the time $\langle\tau(\delta)\rangle$ and velocity δ are normalized using their Kolmogorov-scale typical values, behaving as $Re^{-1/2}$ and $Re^{-1/4}$, respectively. The collapse can even be improved by accounting for intermittency [6, 13], which however does not change the scaling (28).

The right panel of figure 4 shows the FSLE as obtained in a high-resolution DNS of the 2D Navier–Stokes equation in the inverse cascade regime [45]. As mentioned earlier, 2D inverse cascade well fits the K41 phenomenology with the scaling (26) not modified by corrections due to intermittency [44]. In particular, this means that the algebraic organization of the characteristic times $\tau(\ell) \sim \ell^{2/3}$ (see equation (27)) should apply as well as the Lorenz theory which is based on it. Moreover, as in 2D high *Re* can be reached with numerical simulations, 2D turbulence constitutes a valid framework to further test the result (28) for the FSLE. Indeed, we see from the figure that $\lambda(\delta)$ approaches a constant value for $\delta \to 0$ corresponding to the largest LE of the turbulent flow, while, at inertial range scales, the δ^{-2} scaling behavior is clearly detected. The large δ fall-off is due to the saturation of the error at the largest available scale in the simulation. Remarkably, the scaling range is wider for 2D turbulence than for shell model simulations (figure 4, left), obtained at much larger *Re*. Such discrepancy originates from the absence of intermittency in 2D turbulence, which makes the transition from the infinitesimal regime $\lambda(\delta) \approx \lambda_1$ to the inertial range behavior $\lambda(\delta) \sim \delta^{-2}$ very sharp. Remarkably, the scaling behavior (28) was also observed in the FSLE measured from long records of high-resolution data of atmospheric boundary layer flows by Basu *et al* [46].

3.3. Macroscopic chaos

High-dimensional chaotic systems can give rise to remarkable collective phenomena: for instance, macroscopic (global) observables can display well-defined motions even when the microscopic elements, of which they are made up, behave chaotically and their number N is very large [47–55]. A particularly interesting case is the mean field behavior of globally coupled maps (GCMs) defined by the dynamics

$$u_n(t+1) = (1-\sigma)f(u_n(t)) + \sigma m(t), \qquad m(t) = \frac{1}{N}\sum_{i=1}^{N} f(u_i(t)), \tag{30}$$

where f is the map specifying the local dynamics, N is the number of microscopic elements and σ is the coupling strength. Collective behavior can be detected by looking at the mean field $m(t)$, upon varying the coupling σ and the map $f(x)$, different types of behavior have been found [48–50, 52–55], which can be classified as follows [53, 54].

(a) *Standard chaos.* $m(t)$ fluctuates, essentially according to the Gaussian statistics with the standard deviation $\mathcal{O}(N^{-1/2})$, around a time-independent mean value.
(b) *Macroscopic (quasi-)periodicity.* $m(t)$ displays a periodic (or quasi-periodic) motion with superimposed small fluctuations $\mathcal{O}(N^{-1/2})$.

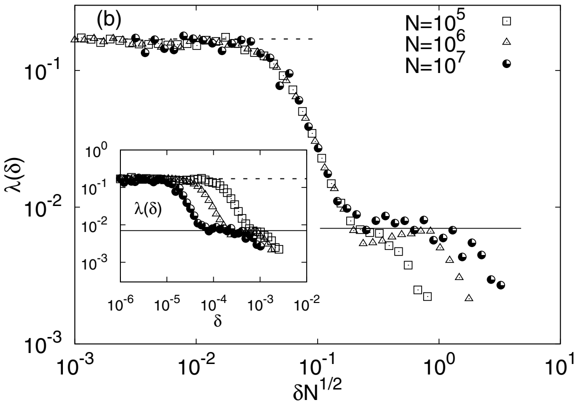


Figure 5. FSLE for the GCM (30) with $f(x) = a(1 - |1/2 - x|)$ with $a = 1.7$, $\sigma = 0.3$ for various N as labeled. The horizontal segments correspond to the microscopic LE $\lambda_{\rm micro} \approx 0.17$ (dashed) and the macroscopic LE $\lambda_{\rm Macro} \approx 0.007$ (solid). The scale δ is multiplied by $\sqrt{N}$ to collapse the curves. Unscaled data are shown in the inset. Data from [54].

(c) *Macroscopic chaos.* $m(t)$ moves erratically with superimposed small $\mathcal{O}(N^{-1/2})$ fluctuations suggesting chaotic collective dynamics.

Standard chaos (a) corresponds to the natural expectation based on the central limit theorem. Behavior such as (b) and (c) is more interesting as it reveals the presence of non-trivial correlations even when many positive LEs are present. Phenomenon (b) has also been observed in diffusively coupled maps in high-dimensional lattices [47, 51].

At least conceptually, macroscopic chaos can be seen as a multiscale phenomenon resembling hydrodynamical chaos emerging from (microscopic) molecular motion of fluids. There, in spite of a huge maximal (microscopic) LE due to molecular collisions ($\lambda_{\rm micro} \sim 1/\tau_c \sim 10^{11}$ s^{-1}, τ_c being the collision time), rather different behavior may appear at the hydrodynamical (coarse-grained) level, regular ($\lambda_{\rm hydro} \leqslant 0$) or chaotic motions ($0 < \lambda_{\rm hydro} \ll \lambda_{\rm micro}$), as in laminar or turbulent flows, respectively. In principle, knowledge of hydrodynamic equations makes the characterization of macroscopic behavior possible by means of standard dynamical system techniques. However, in the generic GCM there are no systematic methods to build up the macroscopic equations, apart from particular cases [48–50].

Whenever macroscopic chaos is an emerging property as described above, we expect that the microscopic LE cannot be straightforwardly used to characterize the macroscopic chaos. A possible strategy, independently proposed by [53] and [54], is to compute the FSLE following the evolution of mean-field perturbations $|\delta m(t)|$. In the limit of infinitesimal perturbations $\delta \to 0$, $\lambda(\delta) \to \lambda_{\rm max} \equiv \lambda_{\rm micro}$, while on scales $\delta \sim \mathcal{O}(N^{-1/2})$ the macroscopic character of mean-field motion should show up. Figure 5 shows $\lambda(\delta)$ versus $\delta N^{1/2}$ for a case with macroscopic chaos [54]. We observe two plateaus: one at small scales where $\lambda(\delta) \to \lambda_{\rm micro}$ and one at large scales $\lambda(\delta) \approx \lambda_{\rm Macro}$ which defines the 'macroscopic' LE. It is important to observe the collapse of the curves (compare with the inset) and that the macroscopic plateau becomes more and more resolved and extended on large values of $\delta\sqrt{N}$ at increasing N. We can thus argue that the macroscopic motion is well defined in the thermodynamics limit $N \to \infty$.

The emerging scenario is that at a coarse-grained level, $\delta \gg 1/\sqrt{N}$, 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 high resolution, $\delta \leqslant \mathcal{O}(N^{1/2})$, providing further support to the picture which emerged in the previous subsections, e.g. recall figure 3. The proper characterization of macroscopic motion thus requires first to perform the thermodynamic limit ($N \to \infty$) and only then to explore the tangent dynamics by considering infinitesimal perturbations [53, 54]. This view is supported by the collapse in figure 5. However, we should mention that recently it has been proposed that collective motion can be completely characterized via the Lyapunov spectrum by looking at those exponents which correspond to Lyapunov vectors spread over all the degrees of freedom; see [55, 56]. It would be very interesting to explore whether these two approaches can be reconciled.

4. Scale-dependent characterization of systems

An interesting issue is to understand whether a given experimental signal, such as the time series of a certain observable, originates from a deterministic chaotic or a stochastic dynamics, of course, without knowing the system that generated it. Despite much effort [22, 57–64], this longstanding problem is still largely unsolved; for a thorough discussion see [65, 66].

In principle, if we were able to measure the maximum LE (λ) and/or the KS entropy (h_{KS}) from a given signal, we could ascertain whether the time series has been generated by a deterministic law ($\lambda, h_{\mathrm{KS}} < \infty$) or a stochastic process ($\lambda, h_{\mathrm{KS}} \to \infty$). Unfortunately, several practical limitations make the determination of h_{KS} and λ [63] problematic. However, part of these restrictions can be, to some extent, circumvented by adopting a scale-dependent description of a given signal in terms of the behavior of quantities such as the (ε, τ)-entropy per unit time, $h(\varepsilon, \tau)$ (see section 2.4), or the FSLE, $\lambda(\varepsilon)$ (for uniformity of notation, within this section the argument of the FSLE is denoted ε instead of δ). When these quantities are properly defined, so that $\lambda = \lim_{\varepsilon\to 0} \lambda(\varepsilon)$ and $h_{\mathrm{KS}} = \lim_{\varepsilon\to 0} h(\varepsilon)$, they can, in principle, be used to answer the question about the deterministic or stochastic character of the dynamical law that generated the signal. In addition, being defined at each observation scale ε, they give us the opportunity to recast the question about the noisy or chaotic character of a signal at each observation scale, as discussed in the following.

For any finite ε, the ε-entropy (for blocks of length m)

$$h_m(\varepsilon, \tau) = \frac{1}{\tau}[(H_{m+1}(\varepsilon, \tau) - H_m(\varepsilon, \tau))] \tag{31}$$

and the FSLE $\lambda(\epsilon)$ are finite and positive for both stochastic and deterministic chaotic signals. To simplify the discussion here we have not considered the problem of reconstructing the phase-space dynamics through the embedding (see, e.g., [66]); for a detailed treatment see [22, 62]. Of course, to ascertain the 'nature' of the signal we should look at the $\varepsilon \to 0$ behavior of the ε-entropy, or equivalently of the FSLE. However, in practical situations, we have a lower resolution cutoff $\varepsilon_1 > 0$, depending on the number of data and the dimensionality of the dynamics [63], below which we are blind on the behavior of these quantities. At any finite scale, including ε_1, entropy and the FSLE are always finite, so that we are unable to decide whether they will extrapolate to infinity or stay constant in the limit $\varepsilon \to 0$.

As proposed in [22], a way to circumvent these difficulties is to classify the character of a signal as deterministic or stochastic according to the following criterion: when in some range of length scales the entropy $h_m(\varepsilon)$ (or the FSLE $\lambda(\varepsilon)$) displays a plateau at a constant value, we can call the signal deterministic on those scales. In contrast, if $h_m(\varepsilon)$ (or $\lambda(\varepsilon)$) increases at decreasing ε, the signal will be considered stochastic on those scales, and the dependence on ε used to characterize it (for instance for bounded independent random variables $h_m(\varepsilon) \sim \ln(1/\varepsilon)$ [67]). Such a definition is free from the necessity to specify a model for the system which generated the signal, so that we are no longer obliged to answer

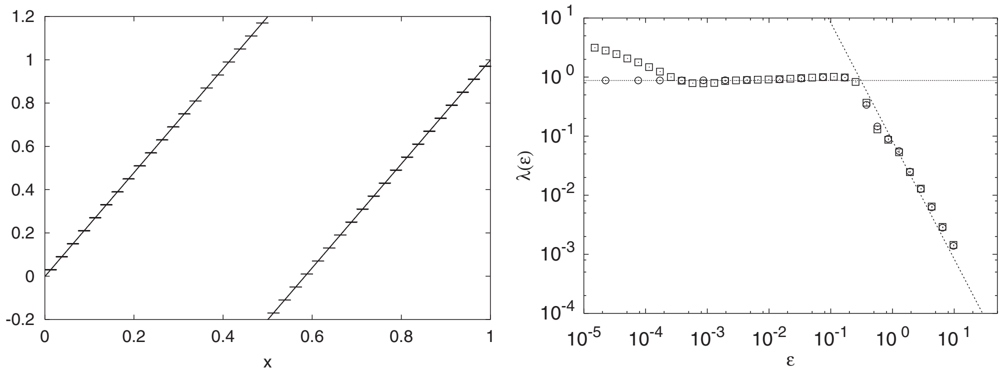


Figure 6. Left: the map $F(x)$ for $\Delta = 0.4$ is shown with the superimposed approximating (regular) map $G(x)$ obtained by using 40 0-slope intervals. Right: $\lambda(\varepsilon)$ versus ε for the chaotic map ((32)–(33)) with $\Delta = 0.4$ ($\circ$) and the noisy map (35) ($\square$), with 10^4 intervals of slope $dG/dy = 0.9$ and noise intensity $\sigma = 10^{-4}$. The straight lines indicate the LE $\lambda = \ln(2.4)$ and the diffusive behavior $\lambda(\varepsilon) \sim \varepsilon^{-2}$. See [22] for details.

the 'metaphysical' question: whether the system which generated the data was deterministic or stochastic?

The distinction between chaos and noise based on (ε, τ)-entropy (or the FSLE) complements previous attempts based on the correlation dimension D_2, where a finite value of D_2 was regarded as a signature for the deterministic nature of the signal [68], which is not completely satisfactory [58]. As an illustration of the above ideas, we briefly discuss the scale-dependent description of signals originating from two systems displaying large-scale diffusion [22]. First, consider the map (figure 6 left)

$$x(t+1) = [x(t)] + F\,(x(t) - [x(t)])\,, \tag{32}$$

where $[u]$ denotes the integer part of u and $F(y)$ is given by

$$F(y) = \begin{cases} (2+\Delta)y & \text{if} \quad y \in [0:1/2[\\ (2+\Delta)y - (1+\Delta) & \text{if} \quad y \in]1/2:1]\,. \end{cases} \tag{33}$$

The above system is chaotic, with LE $\lambda = \ln|F'| = \ln(2+\Delta)$, and generate large-scale diffusion [69], i.e. for large time t the separation between two trajectories diffuses, i.e. $\langle (x(t) - x'(t))^2 \rangle \approx 4Dt$, with D being the diffusion coefficient. As a consequence, the ε-entropy $h(\varepsilon)$ or, equivalently, the FSLE $\lambda(\varepsilon)$ behaves as

$$h(\varepsilon) \approx \begin{cases} \lambda & \text{for} \quad \varepsilon < 1 \\ \dfrac{D}{\varepsilon^2} & \text{for} \quad \varepsilon > 1 \end{cases}. \tag{34}$$

As a second system, consider the noisy map

$$x(t+1) = [x(t)] + G\,(x(t) - [x(t)]) + \sigma\eta_t, \tag{35}$$

where η_t is a time-uncorrelated noise with the uniform distribution in the interval $[-1, 1]$ and σ is a free parameter controlling its intensity. As shown in figure 6 (left), now the deterministic component of the dynamics $G(y)$ is chosen to be a piecewise linear map approximating $F(y)$ in (33). In particular, we can choose $|dG/dy| \leqslant 1$, so that without noise ($\sigma = 0$) the map (35) is non-chaotic.

We can now compare the behavior of $\lambda(\varepsilon)$ in the two systems at varying scale ε. (The ε-entropy provides equivalent information and it is not shown here, see [22].) From a data analysis point of view, one should compute the FSLE by reconstructing the dynamics via embedding [66]; however, being interested only in discussing the resolution effects, one can directly compute it by integrating the evolution equations for two (initially) very close trajectories. In the case of noisy maps, one should use two different realizations of the noise [22]. Figure 6 (right) shows the behavior of $\lambda(\varepsilon)$ for both systems (32) and (35): we can distinguish three ranges of scales with different regimes. On the large length scales, $\varepsilon \gg 1$, we observe diffusive behavior ($\lambda(\varepsilon) \sim \varepsilon^{-2}$) in both models. On intermediate (small) length scales $\sigma < \varepsilon < 1$ both models show chaotic deterministic behavior, because the entropy and the FSLE are independent of ε and larger than zero. Finally, we can see stochastic behavior for the system (35) on the smallest length scales $\varepsilon < \sigma$, while (32) still displays chaotic behavior.

It is thus obvious that if we are asked to define the character of the signal generated by these two systems, our answer would change a lot depending on the smaller cutoff ε_1 (we can reach by analyzing a finite amount of data) being smaller or larger than σ or 1. Adopting a scale-dependent classification scheme gives us the freedom to call deterministic the signal produced by (35) when observed in $\sigma < \varepsilon < 1$, refraining from accounting for its 'true' nature, i.e. its $\varepsilon \to 0$ behavior. Practically, this means that, on these scales, (32) can be considered as an appropriate model for (35). On the other hand, for both systems we can call the behavior at large scales ($\varepsilon > 1$) stochastic.

5. Linear versus nonlinear instabilities

The stability properties of generic systems are typically controlled by the tangent space dynamics (i.e. by infinitesimal perturbations). In this section, we briefly discuss systems for which, due to higher order corrections, the growth rate associated with finite-size perturbations is larger than that of infinitesimal perturbations (i.e. larger than the standard maximal LE, i.e. $\lambda(\delta) > \lambda_1$ for some δ). When this happens and, furthermore, λ_1 is (vanishing) negative, we have odd situations in which close trajectories diverge from each other despite their (marginal) stability in tangent space. Systems of such kind can generate, especially when spatially coupled, behavior very similar to chaos, even though technically speaking they are non-chaotic ($\lambda_1 \leqslant 0$). We mention for example the phenomenon dubbed *stable chaos* [70–73], namely the presence of irregular transients (characterized by negative LE and thus not to be confused with chaotic transients [73]) that are exponentially long with the number of coupled degrees of freedom. Moreover, spatially extended systems characterized by growth rates of finite-size perturbation larger than the infinitesimal ones are characterized by peculiar propagation [8, 74, 75] and synchronization [76–80] properties which make them behave similarly to cellular automata [73]. Another class of systems where irregular dynamics arises in spite of their marginal stability ($\lambda_1 = 0$) is represented by regular polygonal billiards [81, 82] belonging to the world of so-called *pseudochaos* [83]. Such marginal stable systems [84] can give rise to seemingly chaotic behavior. For instance, Letz and Kantz [9] showed that finite-size instabilities of the type discussed above appear at scales of the order of the inverse of the number of edges in the polygon. In systems of such kind the large-scale properties can be indistinguishable from those of the chaotic ones, and genuine large-scale transport can be observed [85, 86].

The FSLE is a natural candidate to quantify these nonlinear instabilities as it probes the nonlinear stability properties of a system. We start comparing the behavior of two one-dimensional maps, $x_t = f(x_{t-1})$, for which the growth rate of non-infinitesimal perturbations

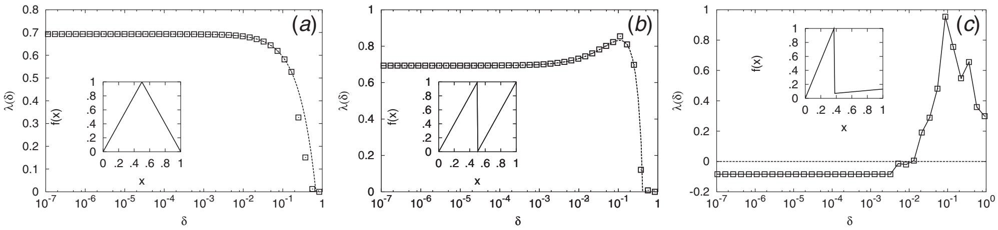


Figure 7. FSLE $\lambda(\delta)$ versus δ for (*a*) the tent map, (*b*) the Bernoulli shift map and (*c*) the stable map (42). The continuous lines in (*a*) and (*b*) are the analytical expressions (39) and (41), respectively. The maps are shown in the insets. Reproduced with permission from [75]. Copyright 2001 American Physical Society.

was analytically estimated in [8]. The first example is the tent map $f(x) = 1 - 2|x - 1/2|$, whose LE can be easily computed as

$$\lambda = \lim_{\delta\to 0}\left\langle \ln\left|\frac{f(x+\delta/2) - f(x-\delta/2)}{\delta}\right|\right\rangle = \int_0^1 \mathrm{d}x\, \rho(x) \ln|f'(x)| = \ln 2, \tag{36}$$

where $\rho(x)$ is the invariant density, which is uniform in the unit interval ($\rho(x) = 1$). Relaxing the request $\delta \to 0$ in (36), we can estimate the FSLE as

$$\lambda(\delta) = \left\langle \ln\left|\frac{f(x+\delta/2) - f(x-\delta/2)}{\delta}\right|\right\rangle = \langle I(x,\delta)\rangle\,, \tag{37}$$

where (for $\delta < 1/2$) $I(x,\delta)$ is given by

$$I(x,\delta) = \begin{cases} \ln 2 & x \in [0:1/2 - \delta/2[\, \cup\,]1/2 + \delta/2:1] \\ \ln \dfrac{|2(2x-1)|}{\delta} & \text{otherwise}\,. \end{cases} \tag{38}$$

The average (37) yields, for $\delta < 1/2$,

$$\lambda(\delta) = \ln 2 - \delta, \tag{39}$$

in agreement with the numerically computed $\lambda(\delta)$ (see figure 7(*a*)). In this case, the error growth rate decreases for finite perturbations, which is somehow the typical behavior. The situation is different for the Bernoulli shift map $f(x) = 2x \bmod 1$ for which, by using the same procedure as before, we easily find that $\lambda = \ln 2$, and for δ not too large

$$I(x,\delta) = \begin{cases} \ln\left[\dfrac{(1-2\delta)}{\delta}\right] & x \in [1/2 - \delta/2, 1/2 + \delta/2] \\ \ln 2 & \text{otherwise.} \end{cases} \tag{40}$$

As the invariant density is uniform, the average of $I(x,\delta)$ gives

$$\lambda(\delta) = (1-\delta)\ln 2 + \delta \ln\left(\frac{1-2\delta}{\delta}\right). \tag{41}$$

In figure 7(*b*), we show the analytic FSLE compared with the numerically evaluated one. In this case, we have that $\lambda(\delta) \geqslant \lambda$ for some $\delta > 0$. Such behavior relates to the discontinuity at $x = 1/2$ which causes trajectories residing on the left (resp.) right of it to experience very different histories irrespective of the original distance between them.

Finally, in figure 7(*c*) we show the FSLE for the map

$$f(x) = \begin{cases} bx & 0 \leqslant x < 1/b \\ 1 - c(1-q)(x - 1/b) & 1/b \leqslant x < \dfrac{b+c}{bc} \\ q + d\left(x - \dfrac{b+c}{bc}\right) & \dfrac{b+c}{bc} \leqslant x \leqslant 1, \end{cases} \tag{42}$$

with $b = 2.7$, $d = 0.1$, $q = 0.07$ and $c = 500$. For $c \to \infty$, the map (42) reduces to the discontinuous map studied in [71]. Interestingly, this map has a negative LE but has a positive growth rate (positive FSLE) for perturbations of finite size (i.e. $\lambda(\delta) > 0 > \lambda_1$ for $\delta \geqslant \delta^* \approx 10^{-2}$). Note for this behavior to be present, the slope c must be large enough [74].

We remark that the FSLE $\lambda(\delta)$ in figure 7(*c*) was measured modifying the first algorithm presented in section 2.3, which cannot be used when $\lambda_1 \leqslant 0$. Here, $\lambda(\delta)$ is measured by taking two trajectories at an initial distance δ_n; after one time step the distance δ between the trajectories is measured. Then one of the two trajectories is rescaled at a distance δ_n from the other, keeping the direction of the perturbation unchanged, and the procedure is repeated several times and for some values of δ_n. Then we averaged $\ln(\delta/\delta_n)$ over many different initial conditions. For $\delta_n \to 0$, this is nothing but the usual algorithm for computing the maximal LE [87]. As already discussed, this method suffers from the problem that when δ_n is finite, we cannot be sure to correctly sample the measure.

We mention that Letz and Kantz [9] considered a system of diffusively coupled maps of type (42) and studied their stability properties in terms of an indicator similar to the FSLE (i.e. able to quantify the growth rate of non-infinitesimal perturbations) and obtained results similar to those of figure 7(*c*). The increased sensitivity to finite perturbations is crucial for establishing (exponentially) long disordered transients in coupled systems of maps such as (42), as reviewed in [73].

We conclude this brief discussion by mentioning that instances of systems with $\lambda(\delta) > \lambda(0)$ have been found also in continuous-time systems (i.e. ODEs). It is worth citing that non-monotonic behavior of $\lambda(\delta)$ similar to that of the Bernoulli map (figure 7(*b*)) has recently been found in a toy model for the climate [88]. In particular, there maxima of the FSLE have been connected to the switching between 'metastable states'. Other examples of systems where finite size instabilities play an important role can be found in the context of neural networks [73, 89].

6. Transport and mixing in fluid flows

The study of transport and mixing properties of small impurities advected by fluid flows is of theoretical interest and great practical importance, e.g., in geophysics and engineering. From a dynamical system point of view investigating the transport of tracer particles—the Lagrangian view of transport—amounts to study the dynamics

$$\frac{\mathrm{d}\boldsymbol{X}(t)}{\mathrm{d}t} = \boldsymbol{u}(\boldsymbol{X}(t), t), \tag{43}$$

where $\boldsymbol{u}(\boldsymbol{x}, t)$ is the Eulerian velocity field; here Brownian fluctuations are ignored for simplicity. Particularly interesting is the study of the relative dispersion [90, 91], i.e. of the separation $\boldsymbol{R}(t) = \boldsymbol{X}^{(2)}(t) - \boldsymbol{X}^{(1)}(t)$ between two tracers whose evolution,

$$\frac{\mathrm{d}\boldsymbol{R}}{\mathrm{d}t} = \boldsymbol{u}(\boldsymbol{X}^{(1)}(t) + \boldsymbol{R}(t), t) - \boldsymbol{u}(\boldsymbol{X}^{(1)}(t), t) = \delta_{\boldsymbol{R}}\boldsymbol{u}(\boldsymbol{X}^{(1)}, t)\,, \tag{44}$$

depends on the properties of the Eulerian velocity field across different scales. For typical velocity fields, in the asymptotics of very small and very large separations (far from any boundaries), the behavior of the relative dispersion is clear. When the initial separation $R(0)$ is very small, tracers evolve in a smooth velocity field and (Lagrangian) chaos will typically be present [92, 93] with exponential amplification of the particle separation. Therefore, at time short enough that $R(t)$ is (infinitesimally) small, we expect

$$\langle R^2(t)\rangle \sim \langle R^2(0)\rangle \mathrm{e}^{L(2)t}, \tag{45}$$

where $L(2) \geqslant 2\lambda_1$ is the generalized LE of order 2 [93], the maximal LE $\lambda_1 \approx \ln[R(t)/R(0)]/t$ controlling the typical behavior. On the other hand, at separation larger than the correlation length of the Eulerian velocity field, the tracers evolve essentially with independent velocities and thus their separation behaves diffusively

$$\langle R^2(t)\rangle \simeq 4Dt, \tag{46}$$

with the diffusion coefficient determined by the large-scale features of the velocity field.

Between these two asymptotic regimes the behavior of $\langle R^2(t)\rangle$ depends on the specific features of the velocity field over the scales. It is rather natural to approach a scale-dependent description in terms of the FSLE and we can indeed define an effective scale-dependent diffusion coefficient as [94]

$$D(R) = \lambda(R)R^2\,. \tag{47}$$

For example, in the above two asymptotics we have $\lambda(R) \approx \lambda_1$ for very small R and $\lambda(R) \propto D/R^2$ at large scales, so that $D(R) \sim \lambda_1 R^2$ and $D(R) \sim D$, respectively.

As shown in the following, this kind of approach to the study of relative dispersion is particularly useful when investigating flows with a multiscale structure, like in turbulence, as with the FSLE the detection of the scale is somehow ‘automatically’ obtained (see the discussion of equation (14)). The proper identification of the scale via the FSLE is also very useful in the presence of boundaries, where observing asymptotic regimes may not be possible, or be spoiled by averaging at fixed times, which may lead to spurious regimes due to interference between different scales.

6.1. Relative dispersion in closed systems

Consider the transport of tracers in a closed domain of size, say, L_B by a velocity field with the typical length scale ℓ_u. In the simple case of laminar (one-scale) flows with $L_B \gg \ell_u$, one expects the following scenario. Until $R \ll \ell_u$, the relative motion is characterized by the exponential regime (45) (with $\lambda(R) \simeq \lambda_1$), then when $\ell_u \ll R \ll L_B$ there should be the diffusive regime (46) (with $\lambda(R) \sim R^{-2}$). Finally, when $R \lesssim L_B$ the average separation must approach a constant value $\langle R^2(t)\rangle \approx$ const, while $\lambda(R)$ quickly decreases toward zero. Actually, by assuming exponential relaxation to the uniform distribution inside the domain, which is expected to be true for a vast class of systems, it is possible to show that [94]

$$\lambda(R) \propto \frac{R^* - R}{R}, \tag{48}$$

where $R^* \approx L_B$ is the average separation between two randomly chosen points.

In figure 8, we show the behavior of the relative dispersion between tracers evolving in the velocity field generated by $M = 4$ point vortices in a disc (see [14, 94] for details). Tracers are advected by the time-dependent velocity field generated by the vortices and behave chaotically for any $M > 2$. The velocity characteristic scale is not imposed by hand, but depends on M and on the energy of the vortex system. Essentially, with $M = 4$, ℓ_u is of the order of the average inter-vortex distance and thus $\ell_u \sim L_B$, so that the diffusive regime is absent. Moreover, due to intermittency in the particle separation, when considering many particle pairs there are situations in which, at the same time, some pairs may be still in the exponential regime, while others have reached the boundaries. Such intermittency causes the seemingly anomalous superdiffusive behavior $\langle R^2(t)\rangle \propto t^\nu$ (with $\nu > 1$) observed in figure 8(*a*), when we average at fixed times. Conversely, the FSLE, ‘selecting’ the correct scale, displays only the exponential regime (the plateaus at small R) and the saturation (48). Features similar to those shown in figure 8 were observed also in the experimental data in a convective cell [95]. We also

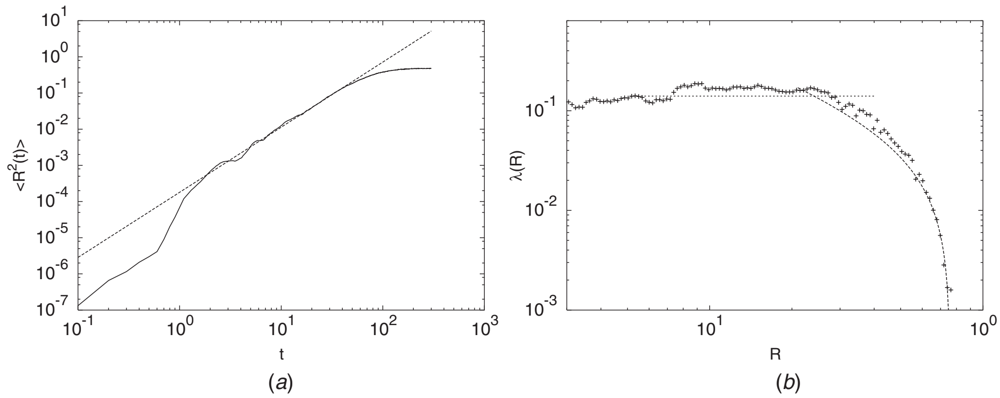


Figure 8. (*a*) $\langle R^2(t)\rangle$ versus t for tracers evolving in a disc with velocity given by four point vortices; see [94, 14] for details. The dashed line is the power law $\langle R^2(t)\rangle \sim t^{1.8}$. (*b*) FSLE for the same model and parameters. The horizontal line indicates the LE ($\lambda \simeq 0.14$) and the dashed curve is the saturation regime (48) with $R^* = 0.76$. Reprinted with permission from [14]. Copyright 2000 American Institute of Physics.

mention that the FSLE has been used to characterize the relative dispersion in experiments on two-dimensional turbulent flows in the Batchelor regime [96] and also to study the Lagrangian trajectories of drifters in the (semi-closed) Adriatic sea [97].

6.2. Relative dispersion in turbulence

We now briefly discuss the application of the FSLE to relative dispersion in turbulence where the classical theory proposed by Richardson predicts [98] (see also [90, 91])

$$\langle R^2(t)\rangle \sim t^3, \tag{49}$$

for R within the inertial range (see section 3.2). It is interesting to note that when Richardson derived (49) he did not know the Kolmogorov scaling (26). Actually, he obtained the law for the relative dispersion using a diffusion equation for the probability distribution of pair separation with a diffusion coefficient $D(R)$ depending on the separation, i.e.

$$\frac{\partial}{\partial t} p(\boldsymbol{R}, t) = \sum_{i=1}^{3} \frac{\partial}{\partial R_i}\left(D(R)\frac{\partial}{\partial R_i} p(\boldsymbol{R}, t)\right). \tag{50}$$

From a collection of a variety of experimental data Richardson proposed $D(R) \sim R^{4/3}$, from which the scaling law for $\langle R^2(t)\rangle$ can be easily obtained. Of course, *a posteriori*, the assumption $D(R) \sim R^{4/3}$ is nothing but a consequence of Kolmogorov scaling (26); indeed, we can estimate $D(R) \sim (\delta_R u)R \sim R^{4/3}$. Richardson's approach to the problem is thus tightly linked to the idea of devising a scale-dependent description of the relative dispersion, in the same spirit we have discussed above within the FSLE framework. In particular, note that, thanks to (47), the Richardson law (49) means $\lambda(R) \sim R^{-2/3}$ for the FSLE.

In spite of the fact that the law (49) was proposed in 1926, still nowadays both experiments and numerical simulations have difficulties in demonstrating it without ambiguity [91]. One of the main sources of difficulties is the presence of intermittency widening the crossovers between the behavior (49) and that expected at small and large scales. Moreover, in the evolution of the separation between two tracers the dependence on the initial separation $R(0)$ typically lasts for very long times, independently of the presence of intermittency.

The comparison between the traditional (fixed time) statistics and the (fixed scale) FSLE reveals interesting features [99–101]. In particular, figure 9(*a*) shows $\langle R^2(t)\rangle$ versus t obtained

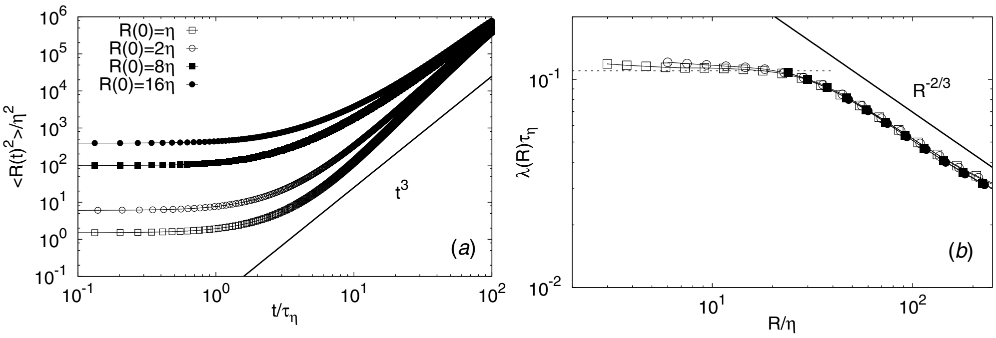


Figure 9. Relative dispersion in numerical simulations of 3D turbulence data from [99]. (*a*) Average separation $\langle R^2(t)\rangle$ normalized with the Kolmogorov length η as a function of the normalized time t/τ_η (where τ_η is the timescale associated with the Kolmogorov scale). Data refer to four different initial separation as labeled; note the dependence on the initial separation and the absence of a clear t^3 range, indicated by the solid line. (*b*) FSLE $\lambda(R)$ versus R for the same data with the dotted and solid lines displaying the exponential $\lambda(R) = \lambda_1 \approx 0.11$ regime and the Richardson prediction $\lambda(R) \propto R^{-2/3}$. Note that the curves obtained by different initial conditions collapse fairly well onto the same curve. Courtesy of G Boffetta and A S Lanotte.

in high-resolution DNS [99]: the dependence on the initial condition is rather evident and the expected t^3 scaling is practically never recovered. Such unpleasant behavior is due to the contamination, at a given t, of different regimes (e.g. exponential in the dissipative range and power law in the inertial range). The above trouble disappears only in the case of an enormous inertial range (say 8–10 decades). However, as shown in figure 9(*b*) the computation of the FSLE is more revealing. First, the dependence on the initial condition essentially disappears as it provides information on the scale at which it is computed. Second, part of the contamination between different regimes is removed so that the $R^{-2/3}$ regime implying the Richardson law (49) can be detected. Clearly, wider inertial ranges, going to higher Reynolds numbers, are mandatory to find more convincing evidence of the Richardson law. We mention that recently exact analytical expressions for the FSLE have been found for some models of tracer dispersion in one and two dimensions [102].

As a historical note we recall that to describe relative dispersion in turbulence, Batchelor proposed an alternative approach [103]. In particular, he replaced the separation-dependent diffusion coefficient ($D(R) \sim R^{4/3}$) with a time-dependent one $D(t) \sim t^2$. As is easily seen by using dimensional analysis, both models predict (49). However, they predict different shapes of the pdf $p(\boldsymbol{R}, t)$. Numerical results suggest the basic validity of the Richardson approximation [99, 101].

We conclude by mentioning that, in [15], the FSLE has been measured on the trajectories of constant-density atmospheric balloons, which were monitored via the satellite EOLE [104]. This study found some evidence of $R^{-2/3}$ scaling as expected for a turbulent atmosphere, which was not clear by the traditional measurement of the relative dispersion.

6.3. Detection of mixing structures in geophysical flows

An interesting application of the FSLE has been put forward for the characterization of mixing in geophysical flows dominated by coherent structures such as the polar vortex [105, 106], and the ocean surface [107–109], where coherent structures can be important for biological activity [110–112].

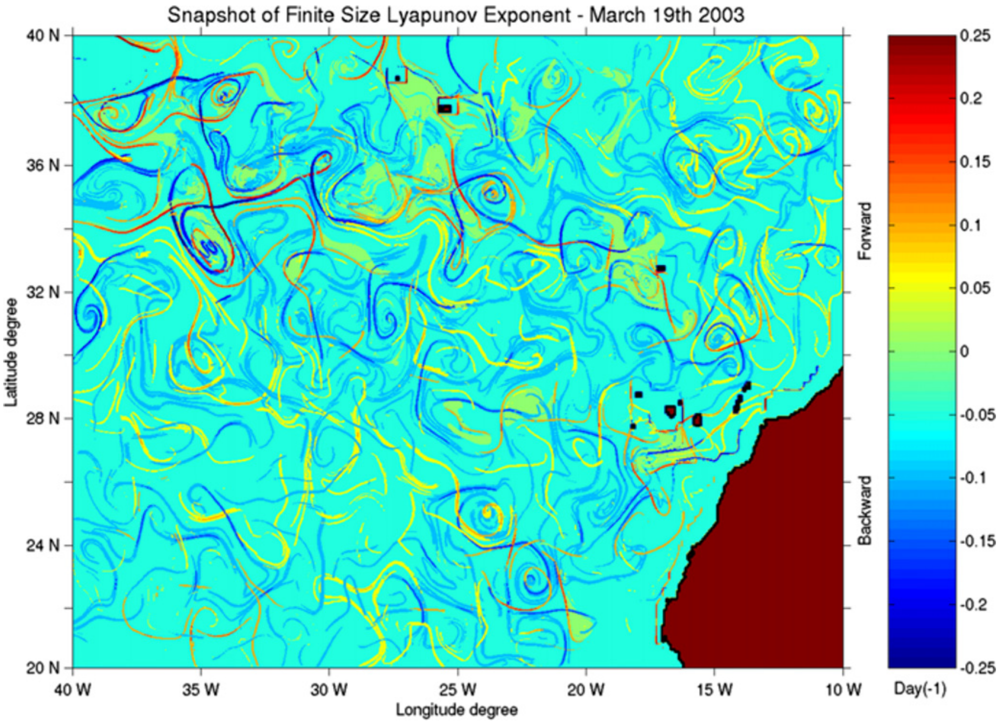


Figure 10. Map of the local FSLE computed backward (negative values) and forward (positive values) in time for an area of the central Atlantic Ocean close to the Canary islands; see [110] for a related study. The maxima and minima of the field, which appear as curves, are good proxies of the coherent structures of the flow. Courtesy of C Lopez.

In quasi-two-dimensional geophysical flows, as occur in the atmosphere and the ocean, it is well known that mixing and stirring are mainly ruled by chaotic advection and, particularly, by hyperbolic lines (i.e. material lines which are locally the most attracting or repelling, see [106, 113]), while the edges of elliptic coherent structures (essentially vortices) constitute barriers to the transport. In general, locating such structures in unsteady flows is not an easy task and traditional criteria based on Eulerian quantities, such as the Okubo–Weiss parameter, do not work properly. It was observed [105, 106] that local measurements of the FSLE (i.e. without averaging) can be used to find 'good' proxies of the hyperbolic lines and thus to map mixing at appropriate length scales. In particular, the idea is to map a given velocity field by measuring, e.g., on a spatial grid defined at an appropriate resolution, the time $\tau_{f,b}$ needed for particle pairs having an initial separation R_0 to reach a separation ϱR_0 both while going forward and backward in time. In this way one can estimate the forward and backward local FSLE at each location $\boldsymbol{x}$ for a given instant of time t:

$$\lambda_{f,b}(\boldsymbol{x}, t, R_0, \varrho) = \frac{1}{\tau_{f,b}} \ln \varrho \,. \tag{51}$$

When ϱ is large enough the extrema of the forward and backward local FSLE (51) are good candidates for the unstable and stable manifolds, respectively, of the flow under consideration (see, e.g., figure 10). Although no rigorous results are at present available, in [105, 106] it was shown that the above strategy can provide good proxies of the Lagrangian coherent structures important for tracer mixing and transport in the polar vortex and also in model flows, where it was compared with more rigorous approaches [113, 114]. Furthermore, the method was applied to data obtained by numerical models and surface data on the Mediterranean sea

[107, 108] and found to provide robust results against measurement noise and resolution details [109]. The use of such an FSLE map was, for instance, very useful for linking the motion of sea birds to Lagrangian structures [111], or establishing the lifetime of phytoplankton niches [112].

The advantage of the FSLE is clearly that it can be easily measured from data. However, a systematic characterization of Lagrangian coherent structures requires more detailed approaches. In this direction, a promising treatment, based on differential geometry, has been recently proposed in [115].

7. Conclusions

In this review, we have discussed the application of the finite size Lyapunov exponent (FSLE), which generalizes the maximal Lyapunov exponent (LE) to the nonlinear regime, that is to the dynamics of perturbation out of the tangent space. The necessity to understand and characterize the nonlinear regime of perturbation growth naturally arises whenever predictability issues concerning realistic systems (involving possibly many different timescales) are considered. In such systems the standard LEs indeed only account for the linear stability and thus are relevant only for the first stages of growth of perturbations which are initially very small (and dominated by the fastest degrees of freedom only). Also, when characterizing experimental signals, practical limitations in the number of data and resolution impose the adoption of a scale-dependent description that can be approached in terms of the FSLE and the ε-entropy, which generalizes the KS entropy to coarser resolutions. The linear stability analysis is sometimes insufficient for understanding properties such as the propagation of perturbations in spatially extended systems, which can be controlled by nonlinear instabilities (see section 5 and [73] for a thorough review). Finally, a scale-dependent indicator such as the FSLE can be very useful in the description of relative dispersion in fluid flows, which is essentially determined by the multiscale structure of the advecting velocity field. Moreover, it provides good proxies for identifying the Lagrangian structures relevant to mixing.

We hope that this discussion will stimulate further research on the characterization of the nonlinear stability analysis. In particular, it would be important to establish mathematical rigor for FSLE, possibly generalizing it beyond the maximal growth rate, in such a way as to be able to define an FSLE spectrum.

Acknowledgments

We wish to thank M Abel, E Aurell, L Biferale, G Boffetta, F Cecconi, A Celani, A Crisanti, M Falcioni, F Ginelli, P Giuliani, K Kaneko, H Kantz, G Lacorata, A S Lanotte, C Lopez, S Musacchio, E Olbrich, A Pikovsky, A Politi, A Torcini, F Toschi and D Vergni for collaborations and discussions on the subject over the years. We thank G Boffetta, A S Lanotte, C Lopez and S Musacchio for providing us some of the figures.

References

[1] Oseledec V I 1968 *Trans. Moscow Math. Soc.* **19** 197
[2] Kolmogorov A N 1958 *Dokl. Akad. Nauk SSSR* **119** 861
[3] Sinai Y G 1959 *Dokl. Akad. Nauk SSSR* **124** 768
[4] Lorenz E N 1969 *Tellus* **21** 289
[5] Lorenz E 1995 *The Essence of Chaos* (Seattle: University of Washington Press)
[6] Aurell E, Boffetta G, Crisanti A, Paladin G and Vulpiani A 1996 *Phys. Rev. Lett.* **77** 1262
[7] Dressler U and Farmer J D 1992 *Physica* D **59** 365
[8] Torcini A, Grassberger P and Politi A 1995 *J. Phys. A: Math. Gen.* **28** 4533

[9] Letz T and Kantz H 2000 *Phys. Rev.* E **61** 2533
[10] Kolmogorov A N 1956 *IRE Trans. Inform. Theor.* **IT-2** 102
[11] Kolmogorov A N and Tikhomirov V M 1959 *Usp. Math. Nauk* **14** 3
[12] Boffetta G, Paladin G and Vulpiani A 1996 *J. Phys. A: Math. Gen.* **29** 2291
[13] Aurell E, Boffetta G, Crisanti A, Paladin G and Vulpiani A 1997 *J. Phys. A: Math. Gen.* **30** 1
[14] Boffetta G, Celani A, Cencini M, Lacorata G and Vulpiani A 2000 *Chaos* **10** 50
[15] Lacorata G, Aurell E, Legras B and Vulpiani A 2004 *J. Atmos. Sci.* **61** 2936
[16] Gaspard P and Wang X J 1993 *Phys. Rep.* **235** 291
[17] Shannon C E 1948 *Bell Syst. Tech. J.* **27** 379
Shannon C E 1948 *Bell Syst. Tech. J.* **27** 623
[18] Billingsley P 1965 *Ergodic Theory and Information* (New York: Wiley)
[19] Eckmann J P and Ruelle D 1985 *Rev. Mod. Phys.* **57** 617
[20] Pesin Y B 1977 *Russ. Math. Surv.* **32** 55
[21] Abel M, Biferale L, Cencini M, Falcioni M, Vergni D and Vulpiani A 2000 *Physica* D **147** 12
[22] Cencini M, Falcioni M, Kantz H, Olbrich E and Vulpiani A 2000 *Phys Rev.* E **62** 427
[23] Palmer T 1996 Predictability of the atmosphere and oceans: from days to decades *Decadal Climate Variability: Dynamics and Predictability (NATO ASI Series* vol 144*)* (Berlin: Springer) p 83
[24] Lorenz E N 1996 Predictability—a problem partly solved *Predictability: ECMWF Seminar Proc. (4–5 Sept. 1995)* (Reading: ECMWF) p 1
[25] Herrera S, Pazó D, Fernández J and Rodríguez M 2011 *Tellus* A **63** 978
[26] Boffetta G, Giuliani P, Paladin G and Vulpiani A 1998 *J. Atmos. Sci.* **55** 3409
[27] Olbrich E, Hegger R and Kantz H 1998 *Phys. Lett.* A **244** 538
[28] Grassberger P and Procaccia I 1983 *Phys. Rev. Lett.* **50** 346
[29] Boffetta G, Celani A, Cencini M, Lacorata G and Vulpiani A 2000 *J. Phys. A: Math. Gen.* **33** 1313
[30] Palmer T 2001 *Quart. J. R. Meteor. Soc.* **127** 279
[31] Peña M and Kalnay E 2004 *Nonlinear Process. Geophys.* **11** 319
[32] Toth Z and Kalnay E 1997 *Mon. Weather Rev.* **125** 3297
[33] Kalnay E 2002 *Atmospheric Modeling, Data Assimilation and Predictability* (Cambridge: Cambridge University Press)
[34] Toth A and Kalnay E 1993 *Bull. Am. Meteorol. Soc.* **74** 2317
[35] Yoden S 2007 *J. Meteorol. Soc. Japan* B **85** 77
[36] Boffetta G, Cencini M, Falcioni M and Vulpiani A 2002 *Phys. Rep.* **356** 367
[37] Boffetta G, Mazzino A and Vulpiani A 2008 *J. Phys. A: Math. Theor.* **41** 363001
[38] Frisch U 1995 *Turbulence: The Legacy of A N Kolmogorov* (Cambridge: Cambridge University Press)
[39] Leith C E 1971 *J. Atmos. Sci.* **28** 144
[40] Leith C E and Kraichnan R H 1972 *J. Atmos. Sci.* **29** 1041
[41] Ruelle D 1979 *Phys. Lett.* A **72** 81
[42] Crisanti A, Jensen M H, Paladin G and Vulpiani A 1993 *Phys. Rev. Lett.* **70** 166
[43] Biferale L 2003 *Annu. Rev. Fluid Mech.* **35** 441
[44] Boffetta G and Ecke R 2012 *Annu. Rev. Fluid Mech.* **44** 427
[45] Boffetta G and Musacchio S 2000 *Phys. Fluids* A **13** 1060
[46] Basu S, Foufoula-Georgiou E and Porté-Agel F 2002 *Geophys. Res. Lett.* **29** 2038
[47] Chaté H and Manneville P 1992 *Prog. Theor. Phys.* **87** 1
[48] Perez G and Cerdeira H A 1992 *Phys. Rev.* A **46** 7492
[49] Pikovsky A S and Kurths J 1994 *Phys. Rev. Lett.* **72** 1644
[50] Kaneko K 1995 *Physica* D **86** 158
[51] Chaté H and Manneville P 1996 *Physica* A **224** 348
[52] Losson J, Vannitsem S and Nicolis G 1998 *Phys. Rev.* E **57** 4921
[53] Shibata T and Kaneko K 1998 *Phys. Rev. Lett.* **81** 4116
[54] Cencini M, Falcioni M, Vergni D and Vulpiani A 1999 *Physica* D **130** 58
[55] Takeuchi K, Ginelli F and Chaté H 2009 *Phys. Rev. Lett.* **103** 154103
[56] Olmi S, Politi A and Torcini A 2010 *Europhys. Lett.* **92** 60007
[57] Nicolis C and Nicolis G 1984 *Nature* **311** 529
[58] Osborne A R and Provenzale A 1989 *Physica* D **35** 357
[59] Sugihara G and May R 1990 *Nature* **344** 734
[60] Casdagli M J 1991 *J. R. Stat. Soc.* B **54** 303
[61] Kaplan D T and Glass L 1992 *Phys. Rev. Lett.* **68** 427
[62] Kubin G 1995 What is a chaotic signal? *IEEE Workshop on Nonlinear Signal and Image Processing (Halkidiki, Greece)* p 141

[63] Olbrich E and Kantz H 1997 *Phys. Lett.* A **232** 63
[64] Gao J B, Hu J, Tung W W and Cao Y H 2006 *Phys. Rev.* E **74** 066204
[65] Abarbanel H D I 1996 *Analysis of Observed Chaotic Data* (New York: Springer-Verlag)
[66] Kantz H and Schreiber T 1997 *Nonlinear Time Series Analysis* (Cambridge: Cambridge University Press)
[67] Gelfand I M, Kolmogorov A N and Yaglom A M 1958 Amount of information and entropy for continuous distributions *Selected Works of A N Kolmogorov* vol 3 ed A N Shiryeyev (Dordrecht: Kluwer) p 33
[68] Grassberger P and Procaccia I 1983 *Physica* D **9** 189
[69] Schell M, Fraser S and Kapral R 1982 *Phys. Rev.* A **26** 504
[70] Crutchfield J P and Kaneko K 1988 *Phys. Rev. Lett.* **60** 2715
[71] Politi A, Livi R, Oppo G L and Kapral R 1993 *Europhys. Lett.* **22** 571
[72] Ginelli F, Livi R and Politi A 2002 *J. Phys. A: Math. Gen.* **35** 499
[73] Politi A and Torcini A 2010 Stable chaos *Nonlinear Dynamics and Chaos: Advances and Perspectives* ed M Thiel *et al* (Berlin: Springer) p 103
[74] Politi A and Torcini A 1994 *Europhys. Lett.* **28** 545
[75] Cencini M and Torcini A 2001 *Phys. Rev.* E **63** 056201
[76] Bagnoli F and Cecconi F 2000 *Phys. Lett.* A **282** 9
[77] Baroni L, Livi R and Torcini A 2001 *Phys. Rev.* E **63** 036226
[78] Ahlers V and Pikovsky A 2002 *Phys. Rev. Lett.* **88** 254101
[79] Ginelli F, Livi R, Politi A and Torcini A 2003 *Phys. Rev.* E **67** 046217
[80] Ginelli F, Cencini M and Torcini A 2009 *J. Stat. Mech.* **2009** P12018
[81] Vega J L, Uzer T and Ford J 1993 *Phys. Rev.* E **48** 3414
[82] Mantica G 2000 *Phys. Rev.* E **61** 6434
[83] Chirikov B and Vivaldi F 1999 *Physica* D **129** 223
[84] Gutkin E 1996 *J. Stat. Phys.* **83** 7
[85] Cecconi F, del Castillo-Negrete D, Falcioni M and Vulpiani A 2003 *Physica* D **180** 129
[86] Jepps O G and Rondoni L 2006 *J. Phys. A: Math. Gen.* **39** 1311
[87] Benettin G, Galgani L, Giorgilli A and Strelcyn J M 1980 *Meccanica* **15** 9
[88] Mitchell L and Gottwald G 2012 *Chaos* **22** 023115
[89] Zillmer R, Livi R, Politi A and Torcini A 2006 *Phys. Rev.* E **74** 036203
[90] Sawford B 2001 *Annu. Rev. Fluid Mech.* **33** 289–317
[91] Salazar J and Collins L 2009 *Annu. Rev. Fluid Mech.* **41** 405–32
[92] Ottino J M 1990 *Annu. Rev. Fluid Mech.* **22** 207
[93] Crisanti A, Falcioni M, Paladin G and Vulpiani A 1991 *Riv. Nuovo Cimento* **14** 1
[94] Artale V, Boffetta G, Celani A, Cencini M and Vulpiani A 1997 *Phys. Fluids* **9** 3162
[95] Boffetta G, Cencini M, Espa S and Querzoli G 2000 *Phys. Fluids* **12** 3160
[96] Jullien M 2003 *Phys. Fluids* **15** 2228
[97] Lacorata G, Aurell E and Vulpiani A 2001 *Ann. Geophys.* **19** 121
[98] Richardson L F 1926 *Proc. R. Soc. Lond.* A **110** 709
[99] Biferale L, Boffetta G, Celani A, Devenish B, Lanotte A and Toschi F 2005 *Phys. Fluids* **17** 115101
[100] Boffetta G, Celani A, Crisanti A and Vulpiani A 1999 *Europhys. Lett.* **46** 177
[101] Boffetta G and Sokolov I M 2002 *Phys. Rev. Lett.* **88** 094501
[102] Piterbarg L I 2012 *Appl. Math. Model.* **36** 3464
[103] Batchelor G 1952 *Math. Proc. Camb. Phil. Soc.* **48** 345
[104] Morel P and Larceveque M 1974 *J. Atmos. Sci.* **31** 2189
[105] Joseph B and Legras B 2002 *J. Atmos. Sci.* **59** 1198
[106] Koh T and Legras B 2002 *Chaos* **12** 382
[107] d'Ovidio F, Fernández V, Hernández-García E and López C 2004 *Geophys. Res. Lett.* **31** 17203
[108] d'Ovidio F, Isern-Fontanet J, López C, Hernández-García E and García-Ladona E 2009 *Deep-Sea Res.* I **56** 15
[109] Hernández-Carrasco I, López C, Hernández-García E and Turiel A 2011 *Ocean Modelling* **36** 208
[110] Rossi V, López C, Sudre J, Hernández-García E and Garçon V 2008 *Geophys. Res. Lett.* **35** L11602
[111] Tew Kai E, Rossi V, Sudre J, Weimerskirch H, López C, Hernández-García E, Marsac F and Garçon V 2009 *Proc. Natl Acad. Sci. USA* **106** 8245
[112] d'Ovidio F, Monte S D, Alvain S, Dandonneau Y and Lévy M 2010 *Proc. Natl Acad. Sci. USA* **107** 18366
[113] Haller G 2001 *Phys. Fluids* **13** 3365
[114] Haller G and Yuan G 2000 *Physica* D **147** 352
[115] Haller G and Beron-Vera F 2012 *Physica* D **241** 1680
[116] Cencini M, Cecconi F and Vulpiani A 2009 *Chaos: From Simple Models to Complex Systems* (Singapore: World Scientific)