COMPLEXITY AS A FORM OF TRANSITION FROM DYNAMICS TO THERMODYNAMICS:APPLICATION TO SOCIOLOGICAL AND BIOLOGICAL PROCESSES

Massimiliano Ignaccolo

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

May 2003

APPROVED:

Paolo Grigolini, Major Professor

Patricia Hamilton, Committee Member

Dan Mauldin, Committee Member

James A. Roberts, Committee Member

Sam Matteson, Chair of the Department of Department of Physics

C. Neal Tate, Dean of the Robert B. Toulouse School of Graduate Studies

Ignaccolo, Massimiliano, <u>Complexity as a form of transition from dynamics to thermodynamics: Application to sociological and biological processes</u>. Doctor of Philosophy (Physics), May 2003, 122 pp., 2 tables, 30 illustrations, references, 64 titles.

This dissertation addresses the delicate problem of establishing the statistical mechanical foundation of complex processes. These processes are characterized by a delicate balance of randomness and order, and a correct paradigm for them seems to be the concept of sporadic randomness. First of all, we have studied if it is possible to establish a foundation of these processes on the basis of a generalized version of thermodynamics, of non-extensive nature. A detailed account of this attempt is reported in Ignaccolo and Grigolini (2001), which shows that this approach leads to inconsistencies. It is shown that there is no need to generalize the Kolmogorov-Sinai entropy by means of a non-extensive indicator, and that the anomaly of these processes does not rest on their non-extensive nature, but rather in the fact that the process of transition from dynamics to thermodynamics, this being still extensive, occurs in an exceptionally extended time scale. Even, when the invariant distribution exists, the time necessary to reach the thermodynamic scaling regime is infinite.   In the case where no invariant distribution exists, the complex system lives forever in a condition intermediate between dynamics and thermodynamics. This discovery has made it possible to create a new method of analysis of non-stationary time series which is currently applied to problems of sociological and physiological interest.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

This disseratation is devoted to the study of complex systems. A widely accepted idea, the one adopted also in this work, of "complexity" is that of a balance between "order" and "randomness" (the first first, alone, being "simple" because totally predictable, and the second, alone, being simple, for the opposite reason: that of being totally unpredictable). Many of the properties that are thought to characterize complex system, such as cooperative behavior, adaptability, ..., are linked to the interplay of order and randomness.

In the present my focus will be on modeling complexity and characterizing (measuring) it. the first part of the original work of this thesis consist in the description and characterization [1, 2, 3] of a model of complexity through the adoption of an intermittent dynamical model based on the Manneville Map. The intermittency characterizing the model proposed is the fact that the phase space of this system is characterized by two regions: one of deterministic motion and one of random motion. A trajectory moving in the phase space will explore both region in an intermittent way, making this dynamical system an ideal candidate for modeling complexity. Moreover I shall use this model to generate a diffusion process and I shall discuss its properties. The second part of the original work of this thesis is dedicated to the characterization of the complexity of a "real" complex system. The one responsible for the teen birth phenomenon, analyzed here under the form of the time series of the daily number of births due to teenage mothers in the state of Texas [4, 5]. This last topic seems disconnected from the first one, but it is not so. In fact since the pioneering work of Hurst [6], the characterization of times series stemming from complex system is done, using the time series itself to generate a diffusion process.

The outline of this work is the following. Chapter 2 is dedicated to the discussion of different kinds of diffusion processes and their derivation within dynamical systems

or within important probabilistic models such as, the Continuous Time Random Walk (CTRW). Chapter 3 is dedicated to the description of the different methods of characterizing complexity. The first part of this chapter is dedicated to important concepts, like the Kolmogorov-Sinai entropy (KS entropy) and the Algorithm Information Content (AIC) that can be used to characterized the complexity of the trajectories of a dynamical system. The second part of Chapter 3 is dedicated to the analysis of complexity of diffusion process: the fundamental notion of "scaling" is discussed and the methodology to detect it, as well. Finally Chapter 4 and Chapter 5 contain the original work of this thesis in the order above described, and Chapter 6 is dedicated to the concluding remarks.

CHAPTER 2

Diffusion Processes and the Continuous Time Random Walk (CTRW)

This chapter is dedicated to the exposition of some fundamental issues in Statistical Mechanics such as Diffusion Processes and CTRW. The exposition will be, by no mean, exhaustive (Diffusion Process and CTRW are the subjects of several books). Only, what is strictly pertinent for understanding the work of this thesis is discussed. I shall start introducing (Sec. 2.1) the Brownian Motion (BM) and its connection (Sec. 2.2) with one of the most important theorem of probability theory, the Central Limit Theorem (CLT). Sec. 2.3 is dedicated to introduce the dynamical approach to diffusion processes, In Sec. 2.4 the Generalized Liouville Equation (GLE) is derived, establishing a connection between the theory of dynamical systems and diffusion processes. This is done using the Projection Method (PM) technique and what is commonly denoted as the density "picture" or Liouville "picture", namely a description of the dynamical system in terms of the motion of a density of trajectories in the phase space of the system. The next three sections are dedicated to anomalous diffusion, namely, any departure from BM, focusing our attention on the variance of the diffusion process (Sec. 2.5) and specifically introducing two important types of anomalous diffusion, the Fractional Brownian Motion (FBM) (Sec. 2.6) and the so called Lévy process (Sec. 2.7). Finally the last section is a brief review of the CTRW and in particular of two kinds of walk denoted as Asymmetric Jump Model (AJM) and Symmetric Velocity Model (SVM) because of their relevance for the results of Chap. 4 .

2.1   Brownian Motion (BM): the Fick's law

Brownian motion provides some of the most spectacular evidence, on the "macroscopic" scale, for the discrete or atomic nature of the matter on the "microscopic" scale. The collisions with the molecules of the fluid in which the Brownian particle is immersed, make the motion of the Brownian particle erratic when looked under a microscope. This kind of motion was observed by the biologist Robert Brown in 1827 in one of his experiment [7]. In 1855 A. Fick [8] on the basis of Fourier's heat conduction equation, derived the diffusion equation (nowadays referred to as Fick's law) and in 1905 A. Einstein [9] connected the diffusion coefficient D of the Fick's law to the atomic properties of the Brownian particles, to the fluid dynamic properties of the medium in which the particles move and to the temperature T of the fluid, as well.

Without entering in the details of Fick's derivation of the diffusion equation, it is interesting to notice that this law can be derived from the law of mass conservation[1]

$$\frac{\partial}{\partial t}\rho(\underline{x},t) = -\nabla \cdot \underline{j}(\underline{x},t), \tag{2.1}$$

where ρ is the fluid density in the position $\underline{x}$ at time t. In fact the assumption that, to a first approximation, the current $\underline{j}$ is proportional to the gradient of the concentration

$$\underline{j}(\underline{x},t) = -\nabla\rho(\underline{x},t), \tag{2.2}$$

leads, immediately, to the Fick's law

$$\frac{\partial}{\partial t}p(\underline{x},t) = D\nabla^2 p(\underline{x},t). \tag{2.3}$$

Notice the change of symbol from ρ to p. This is to indicate the "probabilistic" nature of the assumption of Eq. (2.2). Let us consider for simplicity the one dimensional case of Eq. (2.3)

$$\frac{\partial}{\partial t}p(x,t) = D\frac{\partial^2}{\partial x^2}p(x,t). \tag{2.4}$$

---

[1]Let us imagine, for example, an experiment where a drop of ink diffuses in water.

4

It is straightforward to prove that

$$p(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right),$$ (2.5)

where, without loss of generality the average value of the variable x is fixed to be null. Moreover the standard deviation $\sigma(t)$ obeys to the following equation

$$\sigma(t) = \sqrt{2D}t^{\frac{1}{2}}.$$ (2.6)

Therefore the BM[2] is characterized by a probability density function (pdf) $p(x,t)$ that is a Gaussian function spreading in time. The spreading "speed" can be measured by the variance of the process that, as indicated by Eq.(2.6), increase linearly in time.

2.2 The Central Limit Theorem (CLT)

The *Central Limit Theorem* and its generalization, the *Generalized Central Limit Theorem* described in Sec. 2.7, are the fundamental theorems in probability theory used to develop the theoretical arguments of this work. These theorems are based on the concept of *Infinitely Divisible Variable* (IDV) and *Infinitely Divisible Distributions* (IDD). A stochastic variable Y, is an IDV if for any integer N, it can be represented by a sum of identically distributed stochastic variables X_j $(j = 1, 2, \ldots, N)$

$$Y = \sum_{j=1}^{N} X_j.$$ (2.7)

The distribution function[3] $F_Y(Y)$ of an IDV is an IDD. A distribution function is an IDD if and only if, for any integer N, its characteristic function[4] $f_Y(k)$ is the Nth

---

[2]Here the uni-dimensional case is used for simplicity. In the three dimensional case, using isotropic properties, Eq. (2.4) is obeyed by the variable $R = \sqrt{x^2 + y^2 + z^2}$.

[3]Given a stochastic variable y with probability density function $p(y)$ the corresponding distribution function is defined by $F_y(y) = \int\limits_{-\infty}^{y} p(y^{'})dy^{'}$.

[4]The characteristic function of a stochastic variable y with probability density function $p(y)$ is defined as $f_y(k) = \int\limits_{-\infty}^{+\infty} p(y)e^{iky}dy$.

power of some characteristic function $f_X(k, 1/N)$. Thus $f_Y(k) = [f_X(k, 1/N)]^N$. The importance of the IDDs is that limiting distribution can only belong to this class.

The CLT rests on the following statement;

- **Central Limit Theorem**

  Given the sum Y_N of N independent identically distributed (iid) stochastic variables X_j with finite mean $< X >$ and variance σ_X^2. Let S_N be defined as $S_N = \frac{Y_N - N<X>}{\sigma_X \sqrt{N}}$ and $F_{S_N}(s)$ be the respective distribution function.

  Then $F_S(s) = \lim_{N\to+\infty} F_{S_N}(s)$ is a IDD even if the variables Y_N and X_j may not be.

  Moreover the characteristic function $f_S(k)$ is a Gaussian.

A rigorous proof can be found in [10]. Here I give, just, a simple intuitive explanation. Since

$$S_N = \frac{Y_N - N < X >}{\sigma_X \sqrt{N}}, \tag{2.8}$$

and Y_N is the sum of N iid stochastic variables, S_N itself is a sum of N iid stochastic variable, namely,

$$S_N = \sum_{j=1}^{N} s_j \qquad s_j = \frac{X_j - < X >}{\sigma_X \sqrt{N}}. \tag{2.9}$$

Therefore, it is possible to write for the characteristic function $f_{S_N}(k)$ as follows

$$f_{S_N}(k) = \left(\int_{-\infty}^{+\infty} p_{S_N}(s)e^{iks}ds \right)^N, \tag{2.10}$$

where $p_{S_N}(s)$ is the pdf of the variable s defined in Eq. (2.9).

Using the fact that the stochastic variable s has zero mean and unit standard deviation, Eq.(2.10) can be approximated as

$$f_{S_N}(k) \approx 1 - \frac{k^2}{2N} + o(k^2) \tag{2.11}$$

and using the property

$$\lim_{N \to +\infty} (1 + \frac{x}{N})^N = e^x, \tag{2.12}$$

the characteristic function $f_S(k)$ reads

$$f_S(k) = e^{-\frac{k^2}{2}}. \tag{2.13}$$

From this equation, the pdf $p_S(s)$ is derived making the anti Fourier transform of $f_S(k)$. Finally considering the CLT "valid" for a given $\tilde{N} > N$, the following expression for the pdf of the variable Y_N is obtained:

$$P(Y_N) = \frac{1}{\sqrt{2\pi\sigma_X N}} \exp\left(-\frac{(Y_N - N\langle x \rangle)^2}{2\sigma_X^2 N}\right) \qquad N > \tilde{N}. \tag{2.14}$$

According to this equation the pdf is a Gaussian function whose mean value moves according the law $< Y_N > = N < X >$ and whose standard deviation increases according to the prescription $\sigma_{Y_N} \propto N^{\frac{1}{2}}$. These properties are the same of those characterizing the pdf of the BM.

2.3 Dynamic Approach to Diffusion

The last two sections have been dedicated to the derivation of Fick's Law (with a simple probabilistic argument) and the connection between the BM and the CLT. This section will be dedicated instead to a dynamic approach to diffusion process. In order to connect diffusion processes and dynamical systems, it is convenient to express the result of the CLT can through a differential equation. In fact the sum (Eq. (2.7)) of iid stochastic variables can be seen as the solution of

$$\dot{x} = \xi, \tag{2.15}$$

where ξ is a stochastic variable, delta correlated, whose pdf, $p(\xi)$, has its second moment finite. In fact the solution of Eq. (2.15), selecting for simplicity 0 as initial

time, is

$$x(t) = \int_0^t \xi(t')dt' + x(0), \qquad (2.16)$$

that is the continuous time version, doing the simplification of choosing the initial condition $x(0) = 0$, of the Eq. (2.7).

At this stage the connection with the theory of dynamical systems can be done, imagining an isolated physical system where ξ is one of the variables describing it. The stochastic or stochastic-like behavior of ξ is due to the interaction with the remaining variables describing the system. In this sense, the subsystem described by these variables is called the "bath" and the variables itself, the bath variables. The bath can be composed by a large (an Avogadro number for example) or by a small number of variables. The interactions between the variable ξ and the bath variables and those among the bath variables can be either of deterministic (Hamiltonian system) or random nature. The dynamic of an isolated system described by d variables[5] is, formally, obtained solving the following differential equation

$$\underline{\dot{X}} = \underline{F}(\underline{X}), \qquad (2.17)$$

where $\underline{X}$ is a vector, whose components are $x_1, x_2, \ldots, x_d$, the dot representing the derivative with respect to time and the d-dimensional vector $\underline{F}$ containing the information relative to the mutual interactions. This means that $F_j(x_1, \ldots, x_d)$ describes how the variation in time of the j-th variable depends on itself and on the other variables. In order to establish a connection with Eq. (2.15), I will consider the dynamical system described by the following variables

$$\underline{X} = (x, \xi, b_1, b_2, \ldots, b_n) = (x, \xi, \underline{b}). \qquad (2.18)$$

The variable ξ is coupled to a bath, described by the variables b_i, and the variable x collects the fluctuations of the variable ξ. The dynamics of such a system is described

---

[5]d can be in principle infinite

by the following couple of equations

$$\dot{\underline{Y}} = \underline{F}(\underline{Y}) \tag{2.19}$$

$$\dot{x} = \xi, \tag{2.20}$$

where $\underline{Y} = (\xi, \underline{b})$. The vector $\underline{F}(\underline{Y})$ describes the interactions among the bath variables $\underline{b}$ and between the bath variables and ξ. The two equations above can be written, in a concise way, as

$$\dot{\underline{X}} = \underline{F_{TOT}}(\underline{X}), \tag{2.21}$$

where $\underline{F_{TOT}} = (\xi, \underline{F}(\underline{Y}))$.

2.4 Diffusion Process: the Density Equation

The Fick's law Eq. (2.4) is a Master Equation (ME), an equation for the evolution of a pdf. Many different kinds of master equations are used in literature to explain diffusion processes different from BM (see for example [11] and reference therein). A dynamical approach to diffusion process rests on the derivation of a ME starting from the description of Eq. (2.17). This can be done, using the so called the Liouville Picture.

2.4.1 Liouville Picture

The Liouville Picture (or Density Picture) consists of studying, not the evolution of a single trajectory in the phase space given an initial condition $\underline{X}_0$ at the time $t = t_0$, but rather how an "ensemble" of different initial condition, described by a density function $\rho(\underline{X}, t_0)$, evolves as a function of time in the phase space of the system, if the dynamic of a single trajectory is described by Eq. (2.17). Therefore the goal of this approach is to solve the following differential equation (the Liouville equation)

$$\frac{\partial}{\partial t} \rho(\underline{X}, t) = \mathcal{L} \ \rho(\underline{X}, t), \tag{2.22}$$

where $\mathcal{L}$ is an operator, known as the Liouville operator, acting on the density $\rho(\underline{x}, t)$.

It is a well known result that given $\underline{F}$, of Eq (2.17), the Liouville operator $\mathcal{L}$ is defined by the following relation

$$\mathcal{L}\, \rho(\underline{X}, t) = -\nabla \cdot \left(\underline{F}\, \rho(\underline{X}, t) \right).$$

(2.23)

Moreover, for isolated system, if the dynamic is deterministic (Hamiltonian) then the volume $(V = \int d\underline{X}\, \rho(\underline{x}, t_0))$ of the phase space is conserved under the action of the Liouville operator (Liouville theorem), namely

$$\frac{\partial}{\partial t} \int d\underline{x}\, \rho(\underline{x}, t) = \int d\underline{x}\, \mathcal{L}\, \rho(\underline{x}, t) = 0,$$

(2.24)

where the integration is performed through all the phase space and the symbol $d\underline{x}$ stands for $dx_1 \ldots dx_d$.

2.4.2 PM

Generally, when solving the Liouville Equation (Eq. (2.23)), only the behavior some of the degrees of freedom, those considered relevant or those that can be measured experimentally, is of interest. The PM, introduced independently by H. Mori ([12, 13]) and R. Zwanzig ([14, 15]), permits the extraction of the desired information. In order to illustrate this method I will adopt the following simplified notation:

$$\rho = \rho(\underline{X}, t).$$

(2.25)

The explicit dependence on one of the two parameters $\underline{X}$ or t will be used only if it necessary for a better understanding. Eq. (2.22) of the previous Section, now, reads

$$\frac{\partial}{\partial t} \rho = \mathcal{L}\rho.$$

(2.26)

Then, two projection operators are introduced[6], P and Q, such that $P+Q = I$, where I is the identity operator. Moreover, the following reduced densities are defined

$$\rho_1 = P\rho \tag{2.27}$$

$$\rho_2 = Q\rho = (I - P)\rho. \tag{2.28}$$

Using the projectors P and Q, Eq. (2.26) is split in two equations relative to reduced densities ρ_1 and ρ_2, obtaining

$$\frac{\partial}{\partial t}\rho_1 = P\mathcal{L}(\rho_1 + \rho_2) \tag{2.29}$$

$$\frac{\partial}{\partial t}\rho_2 = Q\mathcal{L}(\rho_1 + \rho_2). \tag{2.30}$$

If ρ_1 is the reduced density relative to the variables considered relevant, then Eq. (2.30) is solved getting

$$\rho_2(t) = \int_0^t e^{Q\mathcal{L}(t-t')}Q\mathcal{L}\rho_1(t')dt' + e^{Q\mathcal{L}t}\rho_2(0) \tag{2.31}$$

and substituting this results in Eq. (2.29), the following differential equation for the evolution of ρ_1 is obtained

$$\frac{\partial}{\partial t}\rho_1(t) = P\mathcal{L}\rho_1(t) + \int_0^t P\mathcal{L}e^{Q\mathcal{L}(t-t')}Q\mathcal{L}\rho_1(t')dt' + P\mathcal{L}e^{Q\mathcal{L}t}\rho_2(0). \tag{2.32}$$

This last equation is what is called the GLE for the reduced density ρ_1. Unlike Eq. (2.3) (for example), Eq. (2.32) is not a "bona fidae" ME, since the evolution of the part of interest ρ_1 depends explicitly from the initial condition of the system through the last term of the r.h.s..

---

[6]an operator P is a projection operator if it satisfies $P^2 = P$.

2.4.3   GLE for a Diffusion Process

It is time to apply the PM formalism to the system described by the variables defined in Eq. (2.18) and subject to the law of motion of Eq. (2.21), to see if it is possible to obtain a result compatible with BM (Fick's law) and the CLT. The corresponding Liouville Equation of the whole system is

$$\frac{\partial}{\partial t}\rho(x,\xi,\underline{b},t) = \mathcal{L}\ \rho(x,\xi,\underline{b},t), \tag{2.33}$$

where

$$\mathcal{L} = \xi\frac{\partial}{\partial x} + \nabla \cdot \underline{F} \tag{2.34}$$

The variable of interest is x, since it is the diffusing variable. Therefore, the PM technique, is used to find the equation driving the evolution of the reduced density

$$\sigma(x,t) = \int d\xi d\underline{b}\ \ \rho(x,\xi,\underline{b},t), \tag{2.35}$$

where $d\underline{b}$ stand for $db_1 \ldots db_n$ and, as initial condition for the density ρ, the following

$$\rho(x,\xi,\underline{b},0) = \sigma(x,0)\eta_{eq.}(\xi.\underline{b}), \tag{2.36}$$

with $\eta_{eq.}(\xi.\underline{b})$ being the equilibrium distribution of the system described by Eq. (2.19). This means that, before the fluctuations of the variable ξ are collected through Eq. (2.20), this system is thermalized to a temperature T and then isolated. The operator that associates to the density ρ the reduced density σ through the integration of Eq. (2.35) is not a projector and it is not possible to identify σ with ρ_1 of Eq. (2.32). In order to use the PM, the following definition of P is adopted

$$P\ \rho(x,\xi,\underline{b},t) = \eta_{eq.}(\xi.\underline{b})\int d\xi d\underline{b}\ \ \rho(x,\xi,\underline{b},t) = \eta_{eq.}(\xi.\underline{b})\sigma(x,t). \tag{2.37}$$

It is straightforward to prove that P is a projector and applying Eq. (2.32), one obtains

$$
\begin{aligned}
\eta_{eq}(\xi,\underline{b})\frac{\partial}{\partial t}\sigma(x,t) &= P\mathcal{L}\left[\sigma(x,t)\eta_{eq}(\xi,\underline{b})\right] \\
&+ \int_0^t P\mathcal{L}e^{Q\mathcal{L}\left(t-t'\right)}Q\mathcal{L}\left[\sigma(x,t')\eta_{eq}(\xi,\underline{b})\right]dt' \qquad (2.38) \\
&+ P\mathcal{L}e^{Q\mathcal{L}t}\left[\rho(x,\xi,\underline{b},0)-\sigma(x,0)\eta_{eq}(\xi,\underline{b})\right].
\end{aligned}
$$

This, seemingly complicated, equation can be simplified. The adoption of the initial condition of Eq.(2.36) makes the last member of the r.h.s. of Eq.(2.38) vanish. Another simplification comes from an explicit evaluation of the first term of the r.h.s. of this equation. Using Eq.(2.34) it is possibel to write

$$
\begin{aligned}
P\mathcal{L}\left[\sigma(x,t)\eta_{eq}(\xi,\underline{b})\right] &= -P\left(\xi\frac{\partial}{\partial x}\right)\left[\sigma(x,t)\eta_{eq}(\xi,\underline{b})\right] \\
&- P\,\nabla\cdot\left(\underline{F}\left[\sigma(x,t)\eta_{eq}(\xi,\underline{b})\right]\right) \qquad (2.39)
\end{aligned}
$$

and using the definition of P (Eq. (2.37))

$$
\begin{aligned}
P\mathcal{L}\left[\sigma(x,t)\eta_{eq}(\xi,\underline{b})\right] &= -\eta_{eq.}(\xi,\underline{b})\int d\xi d\underline{b}\left(\xi\frac{\partial}{\partial x}\right)\left[\sigma(x,t)\eta_{eq}(\xi,\underline{b})\right] \\
&- \eta_{eq.}(\xi,\underline{b})\int d\xi d\underline{b}\,\nabla\cdot\left(\underline{F}\left[\sigma(x,t)\eta_{eq}(\xi,\underline{b})\right]\right). \qquad (2.40)
\end{aligned}
$$

The second integral in the r.h.s. of this equation vanishes because the equilibrium density $\eta_{eq}(\xi,\underline{b})$ is, by definition, invariant under the application of Liouvillian $(\nabla\cdot\underline{F})$ driving the motion of the variables ξ and $\underline{b}$. Therefore Eq.(2.40) becomes

$$
P\mathcal{L}\left[\sigma(x,t)\eta_{eq}(\xi,\underline{b})\right] = \eta_{eq}(\xi,\underline{b})\frac{\partial}{\partial x}\sigma(x,t)\int d\xi d\underline{b}\,\xi\,\eta_{eq}(\xi,\underline{b}). \qquad (2.41)
$$

The integral in the r.h.s of Eq.(2.41) is the definition of mean value $(<\xi>)$ of the variable ξ. Without loosing generality, this mean value can be set equal to zero. In this way also the first term of the r.h.s. of Eq. (2.38) can be set to zero and the

differential equation describing the evolution of the density $\sigma(x,t)$ can be written as

$$\frac{\partial}{\partial t}\sigma(x,t) = \frac{1}{\eta_{eq}(\xi,\underline{b})}\int_0^t P\mathcal{L}e^{Q\mathcal{L}\left(t-t'\right)}Q\mathcal{L}\left[\sigma(x,t')\eta_{eq}(\xi,\underline{b})\right]dt'. \qquad (2.42)$$

2.4.4 Derivation of Fick's law from a Deterministic Picture

Eq. (2.42) is the generalized Liouville equation corresponding to the "trajectory" approach of Eq. (2.17). In deriving it, a "smart" choice of initial condition has been made (Eq. (2.36)) and, without loss of generality, null average for the fluctuation ξ has been assumed. Now, in order to see under which condition the Fick's law can be derived from within an Hamiltonian picture, the assumption that the interaction described by Eq. (2.17) is a deterministic one, has to be made.

As a consequence of the Liouville theorem (Eq. (2.24)), one can substitute the first $\mathcal{L}$ inside the integral of Eq. (2.42) with the operator $-\xi\frac{\partial}{\partial x}$ and use the expression for $\mathcal{L}$ given by Eq.(2.34), to obtain

$$\mathcal{L}\left[\sigma(x,t')\eta_{eq}(\xi,\underline{b})\right] = -\eta_{eq}(\xi,\underline{b})\,\xi\,\frac{\partial}{\partial x}\sigma(x,t'). \qquad (2.43)$$

Therefore Eq. (2.42) can be written as

$$\frac{\partial}{\partial t}\sigma(x,t) = \frac{1}{\eta_{eq}(\xi,\underline{b})}\int_0^t P\xi\frac{\partial}{\partial x}e^{Q\mathcal{L}\left(t-t'\right)}\eta_{eq}(\xi,\underline{b})\,Q\xi\,\frac{\partial}{\partial x}\sigma(x,t')dt'. \qquad (2.44)$$

The second Q in the integral of Eq. (2.44) becomes equivalent to the identity operator because of the earlier assumption that the mean value of the variable ξ vanishes, and of the definition of the projector P (Eq. (2.37), as well. Therefore, using the fact that the operator $\frac{\partial}{\partial x}$ commutes with the Liouvillian $\mathcal{L}$

$$\frac{\partial}{\partial t}\sigma(x,t) = \int d\xi d\underline{b}\int_0^t \xi e^{Q\mathcal{L}\left(t-t'\right)}\xi\,\eta_{eq}(\xi,\underline{b})\,\frac{\partial^2}{\partial x^2}\sigma(x,t')dt'. \qquad (2.45)$$

Finally, defining

$$K(t) = \int d\xi\,d\underline{b}\,\xi\,e^{Q\mathcal{L}t}\,\xi\,\eta_{eq}(\xi,\underline{b}), \qquad (2.46)$$

14

Eq. (2.45) can be written as

$$\frac{\partial}{\partial t}\sigma(x,t) = \int_0^t K(t-t')\,\frac{\partial^2}{\partial x^2}\sigma(x,t')dt'. \tag{2.47}$$

The function $K(t)$ of Eq. (2.46), represents a memory kernel (a correlation), since to know $\sigma(x,t)$ at a given time t (Eq. (2.47)) it is necessary to know what happened at time previous to t. In order to recover Fick's law the memory kernel must be integrable, namely

$$\int_0^{+\infty} K(t)dt \;=\; T \;<\; +\infty. \tag{2.48}$$

In fact, in this case T can be regarded as a measure of the length of the correlations, in the sense that the kernel is a Dirac's delta (the correlation lost) after a time $t \gg T$. With this assumption, one can do the, so called, Markov approximation and write Eq. (2.47) as

$$\frac{\partial}{\partial t}\sigma(x,t) = \left(\int_0^{+\infty} K(t)dt\right)\frac{\partial^2}{\partial x^2}\sigma(x,t). \tag{2.49}$$

This last equation is equivalent to the Fick's law, once the coefficient D of Eq. (2.4) is identified with the integral of the memory kernel $K(t)$

2.4.5 Trajectory vs. Density Approach

The density approach, just described, is considered to be equivalent tot the trajectory approach. In other words, taking an ensemble of trajectories and making each of them evolve with Eq. (2.21) (trajectory approach), or describing the ensemble of trajectories with a density ρ and making it evolve with the using the Liouville equation (density approach) is thought to lead to the same results. In Sec. 4.4.2, I will consider, briefly, how this equivalence is broken if the diffusion process is generated by a dichotomous variable.

2.5   Anomalous Diffusion

The term anomalous diffusion is generally used to indicate any kind of diffusion that depart from BM. Historically since the introduction of the *rescaled range analysis* by Hurst,[6] and [16] a way of monitoring if a given diffusion process was anomalous, was to monitor if the standard deviation of the process grew differently from that of BM, namely

$$\sigma(t) \propto t^H, \tag{2.50}$$

where $H \neq 0.5$. The parameter H is usually denoted as Hurst exponent. In Sec. 3.4 different techniques to assess if a diffusion process is anomalous are examined. In this section the technique, above introduced, of monitoring the standard deviation is adopted. I will show how a departure from $H = 0.5$ is connected to an infinitely extended correlation of the variable ξ (see Eq. (2.15)) generating the diffusion. If this correlation is persistent (positive) then $H \in]0.5, 1[$, while in the case on anti-persistent (negative) correlation $H \in]0, 0.5[$. The first condition is addressed to as *superdiffusion* the last one as *subdiffusion*.

The standard deviation σ of Eq. (2.50) is evaluated, adopting the "trajectory" approach of Eq. (2.15) of Sec. 2.3 and an ensemble of statistically independent trajectories $\xi(t)$ is considered. Therefore, using Eq. (2.16), the average position of a trajectory at time t is

$$< x(t) > = \int_0^t < \xi(t') > dt' \ + \ x(0), \tag{2.51}$$

where the symbol $< \dots >$ denotes the average over the ensemble. In a similar way, the average square displacement reads

$$\begin{aligned} < x^2(t) > \ &= \ \int_0^t dt' \int_0^t dt'' < \xi(t')\xi(t'') > \\ &+ \ 2x(0) \int_0^t dt' < \xi(t') > \ + \ (x(0))^2 . \end{aligned} \tag{2.52}$$

Now, using Eqs. (2.51) and (2.52), is possible to obtain, for the square of the standard

deviation $\sigma(t)$, the following expression

$$\sigma^2(t) = \int_0^t dt' \int_0^t dt'' < \left(\xi(t') - <\xi>\right)\left(\xi(t'') - <\xi>\right) >, \qquad (2.53)$$

which can be reduced, playing with the dominion of integration of the double integral and assuming a stationary condition, the integrand depends only on the absolute value of the difference $t' - t''$, to

$$\sigma^2(t) = 2\int_0^t dt' \int_0^{t'} dt'' < (\xi(0) - <\xi>)\left(\xi(t'' - t') - <\xi>\right) > . \qquad (2.54)$$

Finally, multiplying and dividing the integrand of this equation by the following quantity $(<\xi^2(0)>)^7$, the time dependence of the variance can be written as

$$\sigma^2(t) = 2 < \xi^2(0) > \int_0^t dt' \int_0^{t'} dt'' \Phi_\xi(t''), \qquad (2.55)$$

where $\Phi_\xi(t)$ is the autocorrelation function of the variable ξ.

How the property of the correlation function are connected to those of the standard deviation is the argument of the rest of this section.

2.5.1 Ordinary Diffusion: $\mathbf{H} = \mathbf{0.5}$

The standard deviation increases like the BM in the following cases. If

$$\Phi_\xi(t) = \delta(t), \qquad (2.56)$$

then using Eq. (2.55) it is straightforward to prove that

$$\sigma(t) = \sqrt{2 < \xi^2(0) >} \times t^{\frac{1}{2}}. \qquad (2.57)$$

If, instead,

$$\Phi_\xi(t) = e^{-\frac{t}{\tau}}, \qquad (2.58)$$

---

[7] Assuming that this quantity is finite.

17

one obtains

$$\sigma(t) = \left[2 < \xi^2(0) > \left[\tau t - \tau^2 + \tau^2 e^{-\frac{t}{\tau}}\right]\right]^{\frac{1}{2}}, \qquad (2.59)$$

and therefore for $t \gg \tau$ only the first term inside the parenthesis counts and therefore $\sigma(t) \propto t^{\frac{1}{2}}$ in this limit.

More in general an increase of the variance like the BM is obtained, whenever there is a finite correlation time, that is, whenever the following condition is satisfied

$$0 < \int_0^{+\infty} \Phi_\xi(t) dt = \tau < +\infty. \qquad (2.60)$$

In fact the derivative with respect to time of Eq. (2.55) is

$$\frac{d}{dt}\sigma^2(t) = 2 < \xi^2(0) > \int_0^t \Phi_\xi(t') dt', \qquad (2.61)$$

which, using Eq. (2.60), becomes, for $t \gg \tau$,

$$\frac{d}{dt}\sigma^2(t) = 2 < \xi^2(0) > \tau \qquad (2.62)$$

2.5.2 Superdiffusion: $\mathbf{0.5 < H < 1}$

An autocorrelation function that is integrable leads to an Hurst exponent of 0.5. What happens when the autocorrelation function is not integrable? For example, an autocorrelation function $\Phi_\xi(t)$ of the type $t^{-\beta}$ with $0 < \beta < 1$. As stated before *superdiffusion* means that

$$\sigma^2(t) \propto t^{2H} \qquad 0.5 < H < 1. \qquad (2.63)$$

Differentiating with respect to time both sides of Eq. (2.61), one obtains

$$\frac{d^2}{dt^2}\sigma^2(t) = 2 < \xi^2(0) > \Phi_\xi(t). \qquad (2.64)$$

A comparison between this result and the second derivative in time of Eq. (2.63), leads to

$$2H\left(2H-1\right)t^{2H-2} \sim t^{-\beta} \tag{2.65}$$

and, therefore,

$$\beta = 2 - 2H. \tag{2.66}$$

This relationship indicates that to a value of β in the interval $]0, 1[$, corresponds a Hurst exponent in the interval $]1, 0.5[$.

2.5.3 Subdiffusion: $0 < H < 0.5$

Eq. (2.65) can be used, in order to see which asymptotic behavior for the autocorrelation function $\Phi_\xi(t)$ is compatible with the subdiffusive regime. With H in the interval $]0, 0.5[$ the possible values of the parameter β are in the interval $]1, 2[$. Therefore the autocorrelation is integrable, but at the same time is negative in value. In fact the factor $2H\left(2H-1\right)$ of Eq. (2.65) is negative in the interval of H considered. Since for $t = 0$ the autocorrelation function has to assume value 1 a possible model for $\Phi_\xi(t)$ can be the following:

$$\Phi_\xi(t) = (1+\epsilon)\, e^{-\gamma t}\, +\, \epsilon \left(\frac{T}{T+t}\right)^\beta. \tag{2.67}$$

Fig. 2.1 shows the typical behavior of a function described by Eq. (2.67). The integral of the correlation function depends on the parameter ϵ in the following way

$$\tau = \int_0^{+\infty} \Phi_\xi(t)dt = \frac{1+\epsilon}{\gamma} - \frac{\epsilon T}{\beta - 1}, \tag{2.68}$$

while the square of the standard deviation is, using Eqs. (2.55) and (2.67),

$$
\begin{aligned}
\sigma^2(t) &= 2 < \xi^2(0) > \left(\frac{1+\epsilon}{\gamma} - \frac{\epsilon T}{\beta - 1}\right) t \\
&+ 2 < \xi^2(0) > \frac{1+\epsilon}{\gamma^2}\left(e^{-\gamma t} - 1\right)
\end{aligned}
$$
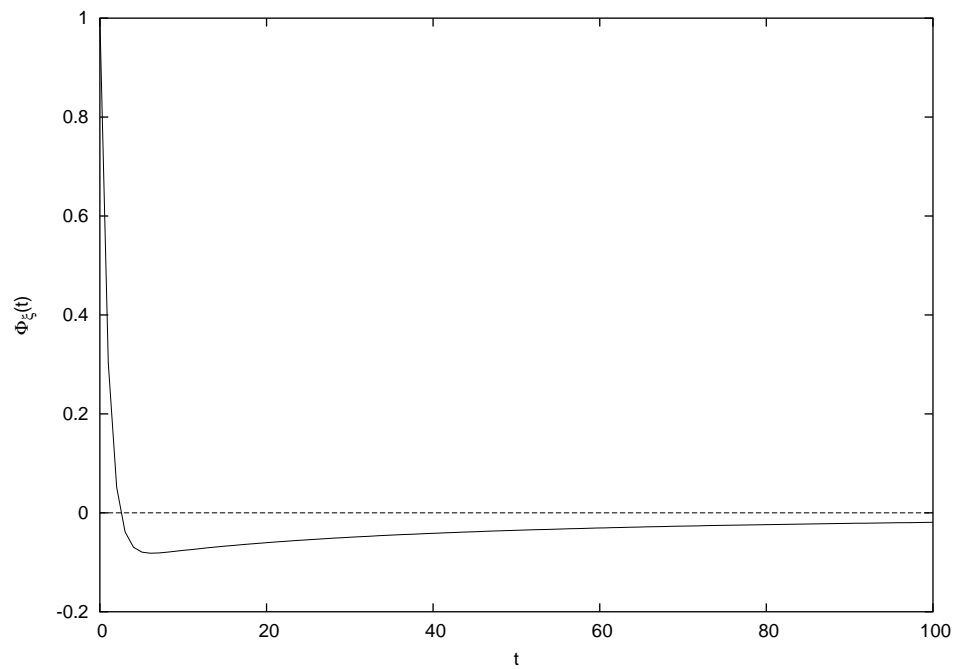
Figure 2.1: The correlation function $\Phi_\xi(t)$ of Eq. (2.67) (full line) and the constant of zero value (dashed line). Here $\epsilon = 0.1$, $\gamma = 1$, $T = 50$ and $\beta = 1.5$.

$$+ \quad 2 < \xi^2(0) > \frac{T^2 \epsilon}{(2-\beta)(\beta-1)} \left[\left(\frac{T+t}{T} \right)^{2-\beta} - 1 \right]. \qquad (2.69)$$

The leading term of Eq. (2.69), as $t \to \infty$ is the linear one, unless the condition

$$\frac{1+\epsilon}{\gamma} = \frac{\epsilon T}{\beta - 1}, \qquad (2.70)$$

applies. This condition implies $\tau = 0$. Therefore if this equation is not satisfied, the standard deviation σ increases as $t^{\frac{1}{2}}$. The subdiffusion grow of the standard deviation is possible only in the case when Eq. (2.70) is valid, since the last term of the r.h.s. of Eq. (2.69) becomes the dominant.

2.6   Fractional Brownian Motion (FBM)

The FBM is a generalization of the ordinary BM, introduced by Mandelbrot [17]. It is a kind of diffusion process that rests on a Gaussian pdf, but, differently, from the BM the scaling parameter is not 0.5: $\delta \in]0, 0.5[$ or $\delta \in]0.5, 1[$. The result of the previous section shows how the FBM is connected with the autocorrelation function of the fluctuating variable.

2.7   Lévy Processes and The Generalize Central Limit Theorem (GCLT)

The GCLT [18, 19], generalizes the CLT for the sum of iid stochastic variables whose pdf has infinite variance. This generalization is due to the French mathematician Paul Lévy, who finds [18] the most general form for the limiting pdf of the sum of iid stochastic variables. For the purpose of this work the content of the GCLT can be expressed as follows:

- **Definition**
  A pdf P(x) is said to be *stable* if, $\forall a_1 > 0, b_1, a_2 > 0, b_2$ there are $a > 0, b$ such that the equation:

$$P(ax + b) = P(a_1 x + b_1) * P(a_2 x + b_2) \qquad (2.71)$$

holds (where the symbol $*$ means convolution).

- **Theorem**

  For the pdf $P(x)$ be a limit distribution for the sum of iid stochastic variables, it is necessary and sufficient to be stable.

- **Theorem**

  For the pdf $P(x)$ to be stable it is necessary and sufficient for its characteristic function, $\hat{P}(k)$, to be of the form

  $$\hat{P}(k) = e^{\imath\gamma k - b|k|^{\alpha}\left[1 + \imath\beta\frac{k}{|k|}\omega(k,\alpha)\right]}, \tag{2.72}$$

  where γ is any real number, $b \geq 0$, $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$ and

  $$\omega(k,\alpha) = \begin{cases} \tan\frac{\pi\alpha}{2} & \text{if } \alpha \neq 1 \\ \frac{2}{\pi}\log[k] & \text{if } \alpha = 1. \end{cases}$$

The index $\gamma \neq 0$ corresponds to a spatial translation (it is equivalent to say that the pdf $P(x)$ has a non vanishing mean), while the index β characterizes the degree of asymmetry of the pdf $P(x)$, since

$$\frac{1-\beta}{1+\beta} = \lim_{L\to+\infty}\frac{R(-L)}{1-R(L)}, \tag{2.73}$$

where

$$R(L) = \int_{-\infty}^{L} P(x)dx. \tag{2.74}$$

Considering, therefore, a symmetric $P(x)$ with zero mean, Eq. (2.72) reduces to

$$\hat{P}(k) = e^{-b|k|^{\alpha}}. \tag{2.75}$$

For $\alpha = 2$, Eq. (2.75) becomes a Gaussian function. The GCLT includes the Gaussian as a special case of admissible limiting distribution of the sum of iid stochastic variables. In fact, the CLT of Sec. 2.2 shows that the Gaussian distribution is the

22

limiting distribution of the sum of iid stochastic variables with finite variance. The case $\alpha \neq 2$ is, instead, the limiting distribution in the case of infinite variance. In particular the parameter α can be connected to the behavior for big values of the pdf $p(x)$ of the iid stochastic variables that are summed, namely if

$$p(\mid x \mid) \sim \frac{1}{\mid x \mid^\mu} \qquad \mid x \mid \to +\infty \qquad \mu \in]1, 3[, \tag{2.76}$$

then $\alpha = \mu - 1$.

Finally, it is interesting to study the asymptotic properties of $P(x)$. For simplicity, only the case where $\gamma = 0$ and $\beta = 0$ is treated, but the results are valid also in the general case. Therefore, making the inverse Fourier transform of Eq. (2.75) the following expression for $P(x)$ is obtained

$$P(x) = \int\limits_0^{+\infty} dk \cos(kx) \, e^{-bk^\alpha}, \tag{2.77}$$

and, after some algebra:

$$P(x) = \frac{\alpha b}{\pi x^{\alpha+1}} \int\limits_0^{+\infty} dz \, z^{\alpha-1} \sin z \, e^{-b\left(\frac{z}{x}\right)^\alpha}. \tag{2.78}$$

In the limiting case of $\mid x \mid \to +\infty$ the exponential inside the integral of Eq. (2.78) can be considered equal to 1. Therefore

$$P(\mid x \mid) \approx \frac{\alpha b}{\pi \mid x \mid^{\alpha+1}} \int\limits_0^{+\infty} dz \, z^{\alpha-1} \sin z \qquad \mid x \mid \to +\infty. \tag{2.79}$$

Finally [20]
$$P(\mid x \mid) \approx \frac{\alpha b}{\pi \mid x \mid^{\alpha+1}} \sin \frac{\pi \alpha}{2} \Gamma(\alpha) \qquad \mid x \mid \to +\infty, \tag{2.80}$$

where Γ is the well known, Gamma function. Eq. (2.80) shows that symmetric Lévy distribution with zero mean are characterized by slow decaying power-law tails

$$P(\mid x \mid) \propto \frac{1}{\mid x \mid^{\alpha+1}} \qquad \mid x \mid \to +\infty. \tag{2.81}$$

2.8 CTRW

Hereby I will present, for simplicity, and continuity with what done in the previous sections, the one dimensional CTRW. The CTRW is a stochastic process where the length of a given jump made by the walker, as well the waiting time elapsing between two successive jumps are drawn from a pdf $w(x,t)$, referred to as the jump pdf. Starting from $w(x,t)$, the following quantities are defined: the jump length pdf

$$\lambda(x) = \int_0^{+\infty} dt \, w(x,t) \tag{2.82}$$

and the waiting time pdf

$$\psi(t) = \int_{-\infty}^{+\infty} dx \, w(x,t). \tag{2.83}$$

Usually the jump length and the waiting time are independent random variables, [21, 22, 23, 24, 25, 26, 27], and therefore

$$w(x,t) = \lambda(x) \, \psi(t). \tag{2.84}$$

In some cases, [28, 29, 30], the jump length and the waiting time between jumps are coupled and therefore

$$w(x,t) = p(x|t) \, \psi(t) \qquad \text{or} \qquad w(x,t) = p(t|x) \, \lambda(x), \tag{2.85}$$

In a given time span the walker can only travel by a limited distance $(p(x|t))$ or, vice versa, a jump of a certain length involves a time cost $(p(t|x))$. The coupled condition is a more "physical" condition, since it introduce a limitation for the distance that

24

the walker can cover in a given amount of time. This kind of CTRW is called as "walk" in a proper sense, while the decoupled CTRW is know as "flight", in the sense that there is no particular limitation (apart from that given by the jump length pdf $\lambda(x)$) in the distance that the walker can cover in a given amount of time.

2.8.1 Formal Solution of the Flight CTRW

In the case of a "flight", in order to find the pdf, $P(x,t)$, of being in the position x at a time t, it is convenient to introduce two functions. The first one is the function $\Psi(t)$, the probability of not leaving a given position up to a time t. $\Psi(t)$ is related to the jump pdf $w(x,t)$ through the following relation

$$\Psi(t) = \int_t^{+\infty} dt' \int_{-\infty}^{+\infty} dx\, w(x,t`) = \int_t^{+\infty} dt'\, \psi(t'). \tag{2.86}$$

The second function is $Q(x,t)$, the probability of arriving at the position x exactly at the time t and to stop before making another jump. For this function the following recursive relation holds

$$Q(x,t) = \int_{-\infty}^{+\infty} dx' \int_0^t dt'\, Q(x-x',t-t')\, w(x',t') + \delta(t)\delta(r). \tag{2.87}$$

Finally, $P(x,t)$ is connected to $\Psi(t)$ and $Q(x,t)$ by

$$P(x,t) = \int_0^t dt'\, Q(x,t-t')\, \Psi(t'). \tag{2.88}$$

The Fourier $(x \to k)$ and Laplace $(t \to s)$ transform of Eq. (2.88) is

$$P(k,s) = \frac{\Psi(s)}{1 - w(k,s)}, \tag{2.89}$$

that can be written, using Eqs.(2.85) and (2.86), as follows

$$P(k,s) = \frac{1 - \psi(s)}{s} \frac{1}{1 - \lambda(k)\psi(s)}. \tag{2.90}$$

2.8.2 CTRW and Fick's Law

The Fick's law (Eq. (2.4)) can be derived also from within the framework of the CTRW. In fact, considering a waiting time pdf $\psi(t)$ with finite mean (τ) and a symmetric jump length pdf $\lambda(x)$ with finite variance (σ^2) then for large times $(t \to +\infty \leftrightarrow s \to 0)^8$ and large value of the position of the walker $(x \to +\infty \leftrightarrow k \to 0)$, it is possible to write

$$\psi(s) \approx 1 - \tau s \tag{2.91}$$
$$\lambda(k) \approx 1 - \sigma^2 k^2. \tag{2.92}$$

Plugging these results in Eq. (2.90), one gets the following asymptotic expression for $P(k,s)$

$$P(k,s) \approx \frac{1}{s + Dk^2} \qquad D = \frac{\sigma^2}{\tau}, \tag{2.93}$$

which is the Laplace and Fourier transform of (Eq. (2.4)).

2.8.3 The Asymmetric Jump Model (AJM)

The authors of [32] study a case where $\lambda(x) = \delta(x-1)$. In these cases the walker after waiting in a given position for a random time t, drawn from a given pdf $\psi(t)$, moves one step ahead. It can be proven [32] that if the waiting times does not have a finite variance then anomalous diffusion emerge. In fact the asymptotic expression ($x \to +\infty \leftrightarrow k \to 0$ and $t \to +\infty \leftrightarrow s \to 0$) for $P(k,s)$ is compatible with the (Lévy) expression of Eq. (2.72).

This model is important because it is simple and can be applied to many situations.

---

[8]Clues on the behavior at large times of a function $f(t)$ studying the behavior of its Laplace transform in a neighborhood of $s = 0$, can be obtained only in particular circumstances (see Tauberian theorems in [31]).

In fact, if for a ny given event the walker makes a step, then the waiting time pdf characterizes the temporal distance between an event and the next one and $p(x, t)$ can be interpreted as the probability of having x events in a time span t. This model, with, ad hoc, definitions of events, has been successfully applied to model the heartbeat, allowing a distinction between sick[9] and healthy patients [33]. Moreover, it has been connected [34] to the multi fractals properties of the heart beat [35].

2.8.4   The Symmetric Velocity Model (SVM)

Let us now introduce a model, the one used in Chap. 4, where the jump length and the waiting time between jumps are connected. In this model the walker moves by a constant speed in the positive or negative direction for a given amount of time, after which stops and chooses, randomly, a new direction and a new sojourn time. The SVM [36, 37, 38]. is described by the waiting time pdf $\psi(t)$ in a given state of the velocity ($+1$ for the positive direction of the x axis and -1 for the negative direction) and by the conditional pdf

$$p(x \mid t) = \frac{1}{2}\delta(\mid x \mid -t), \tag{2.94}$$

expressing the fact that after the end of a phase of uniform motion in one direction ($\delta(|x| - t)$), the walker decides randomly (the factor $\frac{1}{2}$) a new direction of motion. Using Eq. (2.94) is possible to write for the jump pdf $w(x, t)$ the following expression

$$w(x, t) = \frac{1}{2}\delta(\mid x \mid -t)\,\psi(t). \tag{2.95}$$

Following the same arguments used in Sec. 2.8.1, the function $\Psi(x, t)$ is defined as the probability to pass, starting from the origin, in the location x at a time t in a single motion event (no halts to decide a new random direction). The connection

---

[9]In [33] patients with congestive heart failure are compared to the healthy ones.

between $\Psi(x,t)$ and $w(x,t)$ is the following

$$\Psi(x,t) = \frac{1}{2}\delta(|x|-t) \int\limits_{t}^{+\infty} dt' \ \psi(t').$$

(2.96)

Finally making use of the probability $Q(x,t)$ of Eq. (2.87), the pdf $(p(x,t))$ of being in the position x at a time t is

$$p(x,t) = \int dx' \int\limits_{0}^{t} dt' \ Q(x-x',t-t') \ \Psi(x',t'),$$

(2.97)

whose Fourier and Laplace transform reads

$$P(k,s) = \frac{\Psi(k,s)}{1-w(k,s)}.$$

(2.98)

The SVM has important practical application. DNA sequences ([39, 40, 41], among many others) can be thought as a string of $+$ or $-$ (according if a purine or a pyrimidine is found in the sequence). The diffusion process produced by this string (the so called DNA walk) is just an example of SVM.

CHAPTER 3

Kolmogorov-Sinai Entropy, Algorithmic Complexity and Scaling

This chapter is devoted to discuss the notions of Kolmogorov-Sinai (KS) entropy (Sec. 3.1), that one of Algorithmic Information Content (AIC) and Computable Information Content (CIC) and Scaling for a diffusion process. I will show how KS, IAC and CIC are connected to the theory of dynamical systems and how they are used to to define the concept of Information Complexity (Sec.3.2). Regarding the notion of Scaling of a diffusion process, the focus will be on its detection and the practical application of this concept to the analysis of time series. As for the previous chapter, only the fundamental concepts necessary to understand the original work of this dissertation (Chap. 4 and Chap. 5) will be illustrated.

3.1 KS Entropy, Lyapunov Exponent and Pesin Theorem

The concept of KS entropy Consider a sequence of symbols ω_i of length L ($i = 1, 2, \ldots, L$) drawn from an alphabet $S_1, S_2, \ldots, S_M$ of M letters and select an integer $n < L$ as the length of a window that we move through the sequence of symbols. In principle, having available an infinitely long sequence and a computer with enough memory and computer time, for any sequence of symbols $(\omega_0, \omega_1, \ldots \omega_{n-1})$ possible inside a window of length n, we can evaluate the probability $p(\omega_0, \omega_1, \ldots \omega_{n-1})$ that this sequence occurs in the window. The n-th order Shannon entropy is given by

$$H(n) = - \sum_{\omega_0 \omega_1 \ldots \omega_{n-1}} p(\omega_0, \omega_1, \ldots \omega_{n-1}) \ln(p(\omega_0, \omega_1, \ldots \omega_{n-1})). \tag{3.1}$$

$H(n)$ is a measure [42] of the average "uncertainty" in predicting the $(n+1)$-th symbol following any given portion of n symbols of the sequence ω_i. The KS entropy of the symbolic sequence is defined as the rate at which the "uncertainty" of Eq. (3.1) grows with respect to the size n of the portion of the sequence considered, namely

$$h_{KS} = \lim_{n \to \infty} \frac{H(n)}{n}. \tag{3.2}$$

This quantity represents the complexity of the symbolic sequence: the greater[1] the h_{KS}, the greater the complexity. The problem of the definition of Eq. (3.2), is that h_{KS} is hardly computable, since the possible combinations of M different symbols (letters) inside a window of length n are M^n and, therefore, to have a good numerical evaluation of $H(n)$, a sequence of length $L \gg M^n$ is needed. Nevertheless the KS is very popular in the scientific community, the reason being the Pesin theorem [43, 44] that affords a practicable criterion to evaluate the KS entropy, in the specific case where the symbolic sequence is generated by a well defined dynamic law.

---

[1]The maximum value possible for a sequence done with an alphabet of M letters is $\ln M$, that corresponds to the fact the all the possible different combinations of the M symbols considered, are equiprobable. This being the situation of maximum uncertainty.

3.1.1   KS Entropy for Dynamical Systems

In the following, the term "dynamical systems" refers to systems whose evolution (the evolution of the variables that describe the system) is given by

$$\underline{x}_{n+1} = \Pi \; \underline{x}_n, \tag{3.3}$$

where the vector $\underline{x}$ is the vector containing all the variables necessary to describe the system, n is the discrete time and $\Pi : X \to X$ (X is the phase space of the system) is the operator (map) responsible for the dynamic. The description of the dynamic in terms of Eq. (3.3) has a general validity, since systems (the Hamiltonian ones, for example) described in terms or differential equations can always put in the form of Eq. (3.3) if integrated at regular interval of times.

In order to define the KS entropy of a dynamic system, it is necessary to express the dynamics via a symbolic sequence. For this purpose a partition $\{A\}$ of the phase space is done, namely the phase space is divided in disjoint subsets. Then to each of these subsets is associated a symbol, so that the evolution of a trajectory starting from a given point $\underline{x}$ is translated into a symbolic sequence, simply recording in which subset of the partition $\{A\}$ the trajectory is at any time step. In this way to every point of the phase space is associated a symbolic sequence. Therefore, Eq. (3.1), defines the nth order Shannon entropy $H(n, \underline{x}, \{A\})$ associated with the trajectory starting from $\underline{x}$ as

$$H(n, \underline{x}, \{A\}) = -\sum_{\underline{\omega}_n} p_{\underline{x},\{A\}}(\underline{\omega}_n) \; \ln(p_{\underline{x},\{A\}}(\underline{\omega}_n)), \tag{3.4}$$

where $\underline{\omega}_n$ denotes the sequence of symbols $\omega_0 \omega_1 \ldots \omega_{n-1}$ and the notation $p_{\underline{x},\{A\}}(\ldots)$ makes explicit the dependence of the probability from the initial condition and the partition chosen. Using this equation and Eq. (3.2), the KS entropy $h_{KS}(\underline{x}, \{A\})$ relative to the trajectory of initial condition $\underline{x}$, given the partition $\{A\}$ is defined as

$$h_{KS}(\underline{x}, \{A\}) = \lim_{n \to \infty} \frac{H(n, \underline{x}, \{A\})}{n}. \tag{3.5}$$

Usually an "ensemble" of initial conditions is chosen, defining, therefore, a measure μ in the phase space. Integrating the function $h_{KS}(\underline{x}, \{A\})$ all over the phase space with respect to the measure μ and taking the supremum over all the possible partitions, one obtains the KS entropy of the map Π with respect to the measure μ:

$$h_{KS}(\mu) = \sup_{\{A\}} h_{KS}(\mu, \{A\}), \qquad (3.6)$$

where

$$h_{KS}(\mu, \{A\}) = \int h_{KS}(\underline{x}, \{A\}) d\mu. \qquad (3.7)$$

The supremum over all the possible partitions can be rather cumbersome to evaluate for practical purposes. An important simplification can be done using the concept of *generating partition*. A *generating partition* is a partition of the phase space X such that the identification between a trajectory of the phase space and the relative symbolic sequence is a one to one application. This means that a symbolic sequence built up with such a partition identify an unique initial point $\underline{x}$ in the phase space. It can be shown [45] that the quantity defined in Eq. (3.7) takes its maximum value for this kind of partition and, therefore, it is not necessary to use all the possible partitions and find the supremum in order to evaluate $h_{KS}(\mu)$.

Finally, if the system admits an invariant natural measure[2] and it is ergodic, then quantity $h_{KS}(\underline{x}, \{A\})$ is independent of $\underline{x}$. Moreover, the existence of an invariant natural measure ($\mu_{eq.}$), serves the purpose of uniquely defining the KS entropy of the map Π as the KS entropy of Eq. (3.6), evaluated with respect to $\mu_{eq.}$. This quantity is the KS entropy of the dynamical system described by a map Π (if the conditions of ergodicity and existence of an invariant measure are satisfied).

---

[2]by natural invariant measure is intended the measure defined by the equilibrium density distribution $\rho_{eq.}(\underline{x})$, that is the distribution reached under the action of Π starting by "almost" any initial distribution. The meaning of "almost", is that distributions having a support in invariant (under the action of Π) sets of zero measure must not be considered (for example we exclude distributions like a δ of Dirac in a fixed point of the map).

3.1.2 Lyapunov Exponents

The Lyapunov Exponent, in the simple case of uni dimensional maps (the multidimensional case can be generalized easily) and for differentiable functions $\Pi(x)$, establish the rate of exponential departure of two nearby trajectories. The absolute distance $\Delta(n)$ of two trajectories, initially infinitesimally close ($\Delta(0) = \epsilon$), after n iterations of the map Π can be written as

$$\Delta(n) = \epsilon \left| \frac{d\Pi^n(x)}{dx} \right|_{x=x_0}. \tag{3.8}$$

The notation $|\,.\,|_{x=x_0}$ indicates that the dericative is evaluated in $x = x_0$ and that, at the same time, the modulus is taken. Using the chain rule, Eq. (3.8) can be written as

$$\Delta(n) = \epsilon \prod_{j=0}^{j=n-1} \left| \frac{d\Pi(x)}{dx} \right|_{x=x_j}, \tag{3.9}$$

where x_j represents the j-th iterated starting from x_0. Finally, supposing an exponential departure, one can write

$$\Delta(n) = \epsilon \prod_{j=0}^{j=n-1} \left| \frac{d\Pi(x)}{dx} \right|_{x=x_j} = \epsilon\, e^{\lambda(x_0,n)n}. \tag{3.10}$$

The coefficient $\lambda(x_0, n)$ is the rate of exponential departure relative to the initial condition x_0, after n iterations of the map. The Lyapunov coefficient is defined as the limiting value, if it exists, of the coefficient $\lambda(x_0, n)$ as the number of iterations goes to infinity, namely

$$\lambda(x_0) = \lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{j=n-1} \ln \left| \frac{d\Pi(x)}{dx} \right|_{x=x_j}. \tag{3.11}$$

If the system is ergodic, then the coefficient $\lambda(x_0)$ is independent of the particular initial condition chosen and an unique Lyapunov coefficient, λ, of the map is defined. A Lyapunov coefficient strictly positive is the signature of chaotic motion, a strictly

negative one represents a exponential convergence and a null one implies a departure (or a convergence) that is less than exponential. In the d-multidimensional case, it is possible to define a set of Lyapunov exponents, one for each, possible, independent direction of motion of a trajectory. A chaotic system is, then, one for which at least a strictly positive Lyapunov exponent exists.

3.1.3 Pesin Theorem

The one dimensional version of the Pesin theorem[3] prescribes that the if the map Π is ergodic and has an invariant natural measure $d\mu_{eq} = \rho_{eq}(x)dx$, then:

$$h_{KS} = \lambda = \int \ln \left| \frac{d\Pi(x)}{dx} \right| d\mu_{eq}, \tag{3.12}$$

where h_{KS} is the KS entropy of the map with respect to the natural invariant measure and λ is the Lyapunov coefficient of the map Π. The multidimensional version, of the same theorem, states that the KS entropy is the sum of all positive Lyapunov coefficients, namely

$$h_{KS} = \sum_{j,\lambda_j \neq 0} \lambda_j. \tag{3.13}$$

3.2 AIC, CIC and the Information Complexity

Another form of complexity measure (the Information Complexity) is discussed and its connection with the KS entropy is explored.

3.2.1 AIC and CIC

Given a set of symbols (alphabet) $\mathcal{A}$ and a finite sequence of such symbols (symbolic sequence) ω, the *quantity of information $I(\omega)$* contained in ω is defined [4] as the length of the smallest binary program p that, given as input to a computer C, gives as output

---

[3]for the exact condition that the map Π has to satisfy in order for the Pesin theorem to apply see [43],[46] and [47].

[4]Here I illustrate the main idea with arguments, intuitive but as close as possible to the formal definition (for further details, see [48] and related references).

$(C(p))$ the sequence itself. The *quantity of information* is, also, called Algorithmic Information Content (AIC) of the sequence ω. Formally,

$$I_{AIC}(\omega, C) = \min\{|p| : C(p) = \omega\}, \tag{3.14}$$

where $|p|$ indicates the length, in bits, of the program p. From this point of view, the shortest program p which outputs the sequence ω is a sort of optimal encoding of the sequence itself. Unfortunately, this coding procedure cannot be performed on a generic sequence by any algorithm: the Algorithmic Information Content is not computable by any algorithm [49].

Another measure of the information content of a finite string can also be defined by a loss-less data compression algorithm Z satisfying some suitable properties (see [48] for details). The *quantity of information* of the symbolic sequence ω is considered to be the length of the compressed sequence $Z(\omega)$, namely,

$$I_Z(\omega) = |Z(\omega)|. \tag{3.15}$$

The advantage of using a compression algorithm lies in the fact that, this way, the information content $I_Z(\omega)$ turns out to be a computable function. For this reason $I_Z(\omega)$ is called Computable Information Content (CIC) and. from now on, the notation $I_{CIC}(\omega)$ will be used for it.

3.2.2 Information Complexity and Entropy of an Information Source

An important quantity that can be defined using the *quantity of information*, according to the AIC or CIC procedure, of a sequence is the complexity K[5] of the sequence itself. Given an infinite symbolic sequence, ω, the complexity $K(\omega)$ is defined as follows

$$K(\omega) = \limsup_{n \to \infty} \frac{I(\omega^n)}{n}, \tag{3.16}$$

---

[5]For the complexity K will be used , the same notation used for the *quantity of information*. Therefore K_{AIC} and K_{CIC} will denote, respectively, the case in which the complexity K is defined using I_{AIC} or I_{CIC}.

where ω^n is the sequence obtained taking the first n elements of ω. The meaning of Eq. (3.16) is straightforward: the complexity $K(\omega)$ of an infinite symbolic sequence ω is the average information I contained in a single digit of ω. Moreover, it is worth noting how Eq. (3.16) is similar to Eq. (3.2).

Finally, considering the set of all infinite sequences Ω, given the alphabet $\mathcal{A}$, and equipping it with a probability measure μ, the couple (Ω, μ) is created. This couple can be viewed as an information source, provided that μ is invariant under the natural shift map σ, which acts on a symbolic sequence $\omega = (\omega_i)$ $(i \in \mathcal{N})$ as follows: $\sigma(\omega) = \tilde{\omega}$ where $\tilde{\omega}_i = \omega_{i-1}$ $\forall i \in \mathcal{N}^+$. The entropy, relative to the measure μ, h_μ of the information source (Ω, μ) can be defined as the expectation value of the complexity:

$$h_\mu = \int_\Omega K(\omega) \, d\mu \ . \tag{3.17}$$

If $I(\omega) = I_{AIC}(\omega)$ or $I(\omega) = I_Z(\omega)$, under suitable assumptions on Z and μ, h_μ turns out to be equal to the Shannon entropy.

3.2.3 Information Complexity and Dynamical Systems

The goal of this section is that of defining the complexity[6] of a trajectory of a dynamical system, starting from a given position $\underline{x}$ of the phase space X. Following the same argument of Sec. 3.1.1, a partition $\{A\}$ of the phase space is done and the trajectory is translated into a symbolic sequence $\omega_{\underline{x}}$. In this way, using Eq. (3.16), the complexity of the trajectory, given the partition, is

$$K(\underline{x}, \{A\}) = \lim_{n \to +\infty} \sup \frac{I(\omega_{\underline{x}}^n)}{n}. \tag{3.18}$$

In order to have a quantity independent of the partition adopted, all the possible partition are considered and

$$K(\underline{x}) = \sup_{\{A\}} K(\underline{x}, \{A\}). \tag{3.19}$$

---

[6]As for the other definitions and concepts of this Chapter, here is reported the main idea in an intuitive way and as close as possible to the rigorous definitions. These can be find in [3]

Then, in analogy with Eq. (3.7), the phase space X is equipped with a measure μ and the complexity of the map Π relative to the measure μ defined as

$$K(\mu) = \int_X K(\underline{x})d\mu. \qquad (3.20)$$

It can be shown [48] that for an ergodic map Π, having an invariant natural measure $\mu_{eq.}$, the complexity $K(\mu_{eq.})$ is equivalent to the KS entropy.

Before concluding this section, some results [50] regarding an one dimensional map in the interval $[0, 1]$ (some of this results will be used in the next Chapter). For this purpose, it is convenient to express the separation of two nearby trajectories (one starting from the position x and another one starting from the position $x + \epsilon$, as defined by Eq. (3.8), in a more general way, as follows

$$\Delta(n) = \epsilon f(x, n). \qquad (3.21)$$

Then, given a partition $\{A\}$ of the phase space, for almost all the points x, the information content (AIC or CIC) is

$$I(\omega_x^n, \{A\}) \sim \ln |f(x, n))| . \qquad (3.22)$$

The following relationships show the behavior of the function $I(\omega_x^n, \{A\})$ for different kind of functions $f(x, n)$

$$
\begin{aligned}
f(x, n) &\sim e^{\lambda n} &\Rightarrow& \quad I(\omega_x^n, \{A\}) \sim n\lambda \\
f(x, n) &\sim n^p &\Rightarrow& \quad I(\omega_x^n, \{A\}) \sim p\ln n \\
f(x, n) &\sim e^{\lambda n^p} &\Rightarrow& \quad I(\omega_x^n, \{A\}) \sim \lambda n^p
\end{aligned}
$$

3.3 Scaling and Complexity in Time Series

Let us go back to the BM of Sec. 2.1 and recall that, in the one dimensional case, the pdf for the walker to be at the position x at a time t is

$$p(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right). \tag{3.23}$$

The above $p(x,t)$ can be written in the form

$$p(x,t) = \frac{1}{t^\delta} F\left(\frac{x}{t^\delta}\right), \tag{3.24}$$

where

$$F(y) = \frac{1}{4\pi D} \exp\left(-y^2\right) \qquad \delta = 0.5. \tag{3.25}$$

Every diffusion process whose pdf $p(x,t)$ that can be put in the form of Eq. (3.24) is said to be a "scaling" diffusion process and the parameter δ is addressed as the "scaling" parameter. A Lévy diffusion process is one where the function $F(y)$ of Eq. (3.24) has a Fourier transform, the one described by Eq. (2.72) and for scaling parameter the value $\delta = \frac{1}{\alpha}$, while for the FBM (Sec. 2.6) the function $F(y)$ is a gaussian but the scaling parameter δ is different from the value 0.5 of ordinary BM.

A diffusion process satisfying the scaling condition is "invariant" under the following "self-affine" transformation acting on the variable x and t:

$$t \mapsto t' = kt \tag{3.26}$$

$$x \mapsto x' = k^\delta x, \tag{3.27}$$

where $k > 0$ is a real valued parameter. In fact, it is straightforward to see that if the scaling condition of Eq. (3.24) applies then the pdf relative to the rescaled (through Eqs. (3.26) and (3.27)) variables satisfies the following condition

$$p(x',t') = \frac{1}{k^\delta} p(x,t). \tag{3.28}$$

Therefore, a part from the factor $\frac{1}{k^\delta}$ due to the conservation of the norm, the "self-affine" transformation, described above, leaves unchanged the pdf of the diffusion process. In other words, the diffusion process described in term of the rescaled variables x' and t' ceases to be a diffusion and becomes a "stationary" process, in the sense that the pdf does not depend anymore explicitly from the time. Thus, the scaling condition has to be considered a "thermodynamic" condition, meaning that the process has reached an equilibrium condition. In general the scaling condition is not a property of the early stage of a diffusion process (it takes time to reach equilibrium), for this reason, the transitional regime in referred as a regime of transition from "dynamics" to "thermodynamics" (using the parallel with isolated system evolving toward equilibrium). For example for the sum of many iid stochastic variables the limiting distribution is reached asymptotically (in the sense that the sum of many variables is needed for the GCLT to be realized). Moreover there are diffusion processes for which the condition of scaling never applies or where the transition to the scaling condition is infinitely extended. This last case is the case of the diffusion process discussed in the next chapter. Let us discuss, now, some of the techniques of analysis used in literature to study a diffusion process and in particular to asses if it satisfies the scaling condition.

A paradigm often used for complexity in the field pf time series analysis, is that a time series stemming from a complex system has to result in in a diffusion process satisfying the scaling condition, just discussed, and departing from BM. Therefore a scaling parameter $\delta \neq 0.5$ is taken as evidence of complexity. The next Section is therefore dedicated to different methods used to detect and measure the scaling of a diffusion process.

3.4   Methods for Scaling Detection in Time Series

By time series is meant a sequence of values $\xi_1, \xi_2, \ldots, \xi_N$ (the data to be analyzed). The connection between diffusion process and time series analysis is a straightforward

one: the values ξ_j can be used to produce, by summation, a diffusion process, namely

$$x(t) = \sum_{j=1}^{t} \xi_j \qquad t = 1, 2, \ldots, N. \tag{3.29}$$

This equation is just the discrete time version of Eq. (2.16). But a statistical approach to a diffusion process is based on the possibility of having an ensemble of statistically equivalent trajectories taken from exact replica of the system examined. How to create this ensemble if one has only a single realization (usually this is the most common case when analyzing a time series)? To solve this problem the overlapping windows technique is introduced. If t is a discrete time in the interval $[1, N]$, $N-t+1$ different diffusion trajectories (number of members in the ensemble) are defined in the following way

$$x_k(t) = \sum_{j=k}^{k+t} \xi_j, \qquad k = 1, 2, \ldots, N - t + 1. \tag{3.30}$$

This corresponds to initiating a "window" (interval) of length t at the data point k, and to aggregating all the data in the sequence ξ_j inside the window. This procedure is shared by all the methods discussed in the following.

Second Moment (SM) Analysis

The SM analysis is based on evaluating the standard deviation of the diffusion process

$$\sigma(t) = \left[< \{x(t) - \bar{x}(t)\}^2 > \right]^{\frac{1}{2}}, \tag{3.31}$$

where $\bar{x}(t) \equiv < x(t) >$ and $< \ldots >$ denotes the average over the ensemble of realizations of diffusion trajectories. If the time series fluctuations have zero mean value, the quantity $\bar{x}(t)$ vanishes and (3.31) can be written as

$$\sigma(t) = \left[\int_{-\infty}^{\infty} x^2 p(x, t) dx \right]^{\frac{1}{2}}, \tag{3.32}$$

which is, in fact, the square-root of the second moment of the distribution.

Therefore the standard deviation of Eq. (3.31) is evaluated, a time domain satisfying the condition

$$\sigma(t) \propto t^{\delta_{sm}} \Leftrightarrow \ln\left[\sigma(t)\right] \propto \delta_{sm} \ln(t). \tag{3.33}$$

is searched. Then, in the time domain for which Eq.(3.33) is fulfilled, the standard deviation "rescales" with the scaling parameter[7] δ_{sm}. The fact that the standard deviation satisfies the condition of Eq. (3.33) does not automatically imply, that the scaling condition is satisfied. Due to the property of the Gaussian function, the equivalence between the scaling condition of the standard deviation and that of the pdf, is true in the case of BM or FBM. In the case of the Lévy processes introduced in Sec. 2.7, instead, the standard deviation is not defined (being infinite), but the pdf satisfies the scaling condition. Finally in the case of the diffusion relative to the SVM, considered in the next Chapter, the scaling condition for the pdf is not satisfied , while the standard deviation rescales.

Multiscaling (MS) Analysis

The MS analysis is a generalization of the SM analysis. In fact (3.31) is replaced by the expression

$$\sigma_q(t) = \left[< |x(t) - \bar{x}(t)|^q >\right]^{\frac{1}{q}}, \tag{3.34}$$

where q is a real number. It is evident that (3.31) is recovered from (3.34) by setting $q = 2$. In case of fluctuations with zero mean value, Eq. (3.34) reads:

$$\sigma_q(t) = \left[\int_{-\infty}^{\infty} |x|^q p(x,t) dx\right]^{\frac{1}{q}}. \tag{3.35}$$

If the scaling condition of (3.24) holds, it is possible to express (3.35) as

$$\sigma_q(t) = B_q t^{\delta}, \tag{3.36}$$

---

[7]Here for uniformity of notation with the following sections, the symbol H is not adopted to indicate the scaling parameter of the standard deviation.

where

$$B_q = \left[\int_{-\infty}^{\infty} |y|^q F(y) dy \right]^{\frac{1}{q}}. \tag{3.37}$$

When the scaling condition applies: $\zeta(q) = \delta q$, for all the values of q for which the corresponding fractional standard deviation $\sigma_q(t)$ is finite. In the case of FBM, ad example,

$$\sigma_q(t) \propto t^{\delta} \qquad \forall q \in]-1, +\infty), \tag{3.38}$$

while for a Lévy process, characterized by the index α (see Sec. 2.7)

$$\sigma_q(t) \propto t^{\frac{1}{\alpha}} \qquad \forall q \in]-1, \alpha[. \tag{3.39}$$

The way this analysis is applied to a generic time series, as for the SM analysis, rests on looking for a time domain where the q-th fractional standard deviation $\sigma_q(t)$ rescales with exponent $\zeta(q)$, or equivalently

$$[\sigma_q(t)]^q \propto t^{q\zeta(q)} \Leftrightarrow \ln [\sigma_q(t)]^q \propto q\zeta(q) \ln t. \tag{3.40}$$

The quantity $\ln [\sigma_q(t)]^q$ as a function of the time t is plotted in a log-log scale and a check is done to establish if a straight line results. In this case, the slope (δ_{ms}) "can" indicate the presence of a scaling conditon with parameter $\delta = \delta_{ms}$. In fact if the scaling condition of Eq. (3.24) sussists, then the MS analysis is expected to results in a straight line, while the reverse is not necessarily true.

In the ideal case where divergent moments are involved, as pointed out earlier, the analysis should be limited to values of q smaller than a given q_{max}. In practice, the calculation could be performed for the entire range of values from $q = -1$ to $q = \infty$[8], since real data are of finite length and, therefore, have no divergent moments.

---

[8]In reality the behavior of $\sigma_q(t)$ for very high value of q, $q > 10$ for example, cannot be trusted because they are dominated by rare events and therefore subject to the problem of lack of statistics. Therefore, only the fractional standard deviations relative to the range of value from $q = -1$ to $q \approx 10$ is used, when dealing with real data.

Diffusion Entropy (DE) Analysis

The rationale behind the adoption of DE analysis is that, if the pdf of the diffusion process satisfies (3.24), then, regardless of the pdf shape, the Shannon entropy of this density satisfies the following relationship

$$S(t) = -\int_{-\infty}^{\infty} p(x, t) \ln [p(x, t)] \, dx = A + \delta \ln(t), \tag{3.41}$$

where A is a constant defined by

$$A = -\int_{-\infty}^{\infty} F(y) \ln [F(y)] \, dy, \qquad y = \frac{x}{t^{\delta}}. \tag{3.42}$$

Therefore the scaling condition of Eq. (3.24) can be detected by searching for the linear dependence of the entropy of Eq.(3.41) on a logarithmic time scale.

In practice, the numerical procedure necessary to evaluate the pdf requires that the x-axis be divided into cells of a given size, that, in principle, might also depend on the cell position on the x-axis. The criterion of assigning to each cell the same size is adopted, but the size is allowed to be time-dependent, with the symbol $\Delta(t)$ denoting it at a time t. Finally, $\Delta(t)$ must be chosen so as to lead to a fair approximation of $p(x, t)$, through the histogram relation

$$p(x_j, t) \approx \frac{P_j}{\Delta(t)}. \tag{3.43}$$

Here $p(x_j, t)$ is the histogram at the value x_j, the midpoint of the j-th cell, at time t and P_j is the fraction of the total number of trajectories found in this cell at time t. The rationale behind the choice of a time-dependent size for the cell has to do with obtaining a good estimate of the pdf from the histogram. At early times, when the trajectories are close together, a constant Δ is adequate to estimate the ratio P_j/Δ. As the trajectories diffuse apart, however, either more trajectories are needed, or a larger Δ is needed, to provide a reasonable estimate for this ratio. To ensure that Eq. (3.43) is satisfied at any time, the standard deviation of the diffusion process, $\sigma(t)$, is evaluated and the cell size is chosen to be a fraction of the standard deviation,

$\Delta\left(t\right) = \epsilon\sigma\left(t\right)$ where $0.1 \leq \epsilon \leq 0.2$. There is some sensitivity to the choice of ϵ, but in the proper range of values in which Eq. (3.43) is satisfied, the diffusion entropy

$$S\left(t\right) = -\int_{-\infty}^{\infty} p(x,t) ln\left[p(x,t)\right] dx \approx -\sum_{j} P_j \ln P_j + \ln \Delta\left(t\right) \qquad (3.44)$$

is insensitive to the particular fraction of the standard deviation adopted.

As stated before for the MS analysis, the validity of Eq. (3.24) implies a linear growth of the DE in a logarithmic time scale, but the vice versa is not guaranteed. Therefore, it is convenient to denote the slope of the linear dependence on the logarithm of time of the DE, with the symbol δ_{de}.

Direct Assessment Scaling (DAS)

The DAS analysis is based of the invariance under the "self-affine" transformation described by Eqs. (3.26) and (3.27) for a pdf satisfying the condition of scaling given by (3.24). In fact, given any two times t_2 and t_1, with $t_2 > t_1$, the pdf at time t_2 coincides with the pdf at time t_1 if the following procedure is adopted: the scale of the variable x is "squeezed" by a factor $R = [t_1/t_2]^\delta$, simultaneously the scale of the distribution intensity is "enhanced" by the factor $1/R$. The DAS analysis consists exactly of this procedure of "squeezing" and "enhancing" aimed at establishing the invariance of the pdf by a scaling transformation. Since this method uses directly the meaning of scaling, a positive results for the DAS is a proof that the such condition is satisfied.

The reader might wonder why even apply the three procedure described before (SM, MS and DE analysis), since the DAS analysis is the only that can afford an unambiguous result for detecting the scaling parameter. The reason is that the detection of the scaling parameter, if it exists, through the DAS analysis, would necessitate many trials before its discovering. The adoption of the DAS becomes useful after the DE and MS methods are applied since these method can give a "fast" clue on the possible candidate as scaling parameter.

CHAPTER 4

Complexity as a Balance between Order and Randomness

In this chapter I will discuss a model of "complexity" that has been adopted by our research group as a result of mine and my colleagues efforts. In particular the connections between this "idea" of complexity and the Information complexity of Sec. 3.2.2 will be explored. Moreover this model will be used to generate a diffusion process, whose "complexity" is studied in terms of the notion of scaling of Sec. 3.3.

A complex system is one whose dynamics shows an interplay of determinism and randomness. The best models to describe our notion are the, so called, intermittent models like the Manneville map, where, as described in Sec. 4.1, a trajectory moving in the phase space will experience, in an intermittent way, deterministic and random motion. The balance between this two kinds of motion (the balance between order and randomness) can be measured and can be tuned in one way or in the other, changing a parameter. In Sec. 4.2 will be discussed how quantities like the natural invariant measure and the Lyapunov exponent change as the balance between order and randomness is changed. In Sec. 4.3, the Information complexity (see Sec. 3.2.2) of the Manneville Map is evaluated. In particular, it will be shown how depending from the balance between order and randomness, the quantity of information (see Sec. 3.2.3) contained by a trajectory can show a linear increase, or a less than linear increase. The latter property being connected to the absence of an invariant measure and to a balance between order and randomness, in favor of the former: a condition that the authors of [51] call "sporadic randomness". Finally, in Sec. 4.4, a generalized version of the Manneville map will be used to generate, dynamically, the SVM diffusion process of Sec. 2.8.4. It will be demonstrated that, also the cases in which the randomness is not sporadic (according to the definition of [51]) can lead to an anomalous diffusion and characterized by a departure from the scaling condition. Finally, the connection between this lack of scaling and the break down of the equivalence between the trajectory and density approach will be, briefly, addressed.

4.1 The Manneville Map as Prototype of Complexity
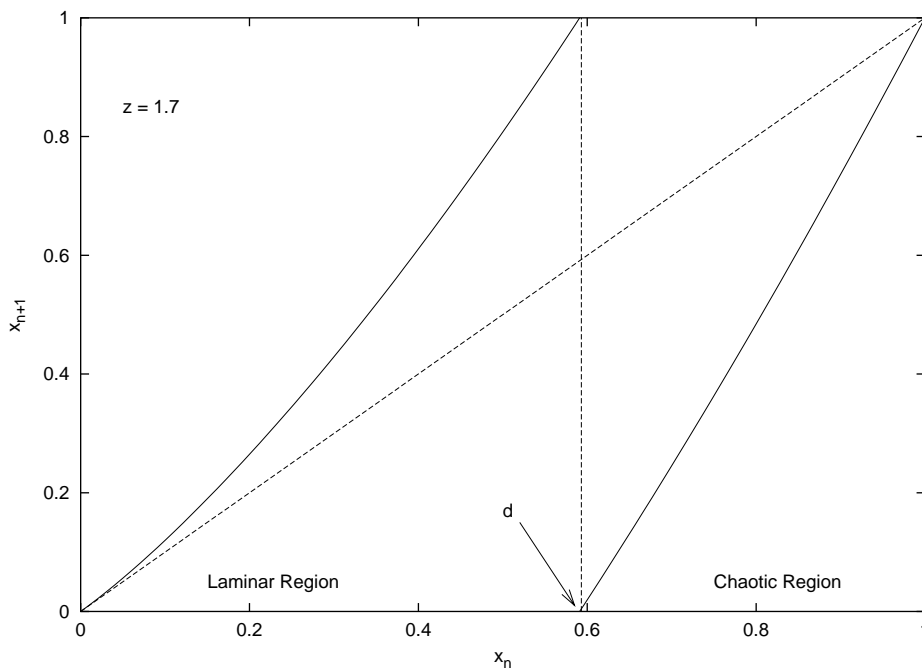
Figure 4.1: The Manneville (full line) map for $z = 1.7$.. The vertical dashed line at $d \approx 0.59$ divides the phase space in Laminar and Chaotic region. Shown in the picture also the Identity transformation (oblique dashed line).

The Manneville Map [52] , see Fig. (4.1), reads

$$x_{n+1} = \Phi(x_n) \qquad \Phi(x) = x + x^z \mod 1, \qquad (4.1)$$

where $z \geq 1$ is a real number. The phase space of the Manneville Map is the interval $[0, 1]$, the fixed points are 0 and 1 and both of them are repeller. The parameter d satisfying

$$d + d^z = 1, \qquad (4.2)$$

divides the phase space in two parts (vertical dashed line of Fig. (4.1)). The "Laminar

region" the "Chaotic region", $[0, d[$ and $]d, 1]$ respectively.

The motion of a trajectory[1] in the phase space is an "intermittent" one. In fact the trajectory will spend some time in the Laminar region of the phase space before jumping in the Chaotic one and after some time going back to the Laminar region, and so on. The rationale for the terms Laminar and Chaotic is the following. The motion of a trajectory inside the Laminar region is a "slow" one, since the function $\Phi(x)$ $(z \geq 1)$ of Eq. (4.1) is tangent to x in $x = 0$ and, therefore, the departure from this repeller requires an elevated number of iterations. This process of slow departure from the repeller is called "Laminar" motion. In the Chaotic region, instead, $\Phi(x)$ is not tangent to x in $x = 1$ and the departure from the repeller is a "fast" one. As result of this slow and fast motion a trajectory will spend more time in the Laminar region than in the Chaotic one.

Let us, now, define in a more rigorous way the intermittent properties of the Manneville map, described above. For this purpose, the following quantities are introduced. The waiting time distribution (wtd) $\psi_{MM}(n)$ in the Laminar region, that is a statistical measure of how many iterations are needed by a trajectory to leave this region of the phase space, and the local departure $\zeta_{MML}(n, x_0)$ and $\zeta_{MMC}(n, x_0)$ of two nearby trajectories (as defined in Sec. 3.1.2) inside the Laminar region and the Chaotic one, respectively[2].

4.1.1   The Wtd Inside the Laminar Region

As stated before, after a trajectory leaves the Laminar region, it remains for few steps in the Chaotic region before being injected back. The process of injection back is uniform, regardless the value of the parameter z. In fact following a trajectory for an extended period of time and recording only the position at which it reenters the Laminar region, gives as result an uniform distribution of such positions. This property is used to adopt a continuous time model with the same characteristics as that of the Manneville map, for which the calculations are easy to carry on. This

---

[1]by trajectory is meant a sequence $\{x_n\}$, $n = 0, 1, 2, \ldots$ of iterates of the map starting from the initial condition x_0. The parameter n can be regarded as a "time".

[2]n is the number of iterations and x_0 a given point inside the Laminar or the Chaotic region.

model, hereby called the Continuous Time Intermittent Model (CTIM), describes the evolution of the position x of an imaginary walker forced to move in the interval $[0:1]$ according to the following rules

$$\frac{dx}{dt} = x^z \qquad\qquad x < 1 \qquad\qquad (4.3)$$

$$x = \text{random} \in]0:1[\qquad x = 1, \qquad\qquad (4.4)$$

where $z \geq 1$ as in the Manneville Map. The motion described by this couple of equations is the following: given an initial condition $x_0 \in]0:1[$, the walker will move using the prescription of Eq. (4.3) until reaching the position $x = 1$ after a time t_0. At this point it is injected back, instantaneously, to a position $x_1 \in]0:1[$ chosen randomly and uniformly, and again it will move towards 1 using Eq. (4.3) and it will reach it after a time t_1, and so on. The properties of a trajectory created in this way are "equivalent" to those of a trajectory of the Manneville Map. In fact Eq. (4.3) is the continuous time approximation in the region x_n close to zero of Eq. (4.1), and, therefore, the motion according to its prescription is a good simulation of the motion of a trajectory in the laminar region of the Manneville map. While the motion inside the Chaotic region is reduced, in the CTIM model, to the process of instantaneous, random and uniform, injection back to the interval $]0:1[$.

The wtd $\psi_{CTIM}(t)$ of the waiting times $t_0, t_1, \ldots$ can be evaluated. The connection between $\psi_{CTIM}(t)$ and $\psi_{MM}(n)$ is the following: for big values of n, it can be regarded as a continuous parameter[3] and the asymptotic properties ($t \to \infty$) of $\psi_{CTIM}(t)$ are the same as those of the wtd $\psi_{MM}(n)$ as $n \to \infty$. This is so, because large values of the waiting time t are obtained in the CTIM model when the injection back happens to be very close to zero, that is, the region where Eq. (4.3) is a good approximation of Eq. (4.1). To evaluate $\psi_{CTIM}(t)$, Eq. (4.3) must be solved, first, with a given initial condition $\tilde{x}$

$$x(t) = \left[\tilde{x}^{1-z} + (1-z)t\right]^{\frac{1}{1-z}}, \qquad\qquad (4.5)$$

---

[3]From now on, the symbol t will be adopted instead of n when considering very large values of n.

and using this solution is possible to evaluate the time t at which the trajectory, starting from the point $\tilde{x}$, reaches 1, thereby obtaining

$$t = \frac{1}{(1 - z)} \times \left(1 - \tilde{x}^{1-z}\right).$$

(4.6)

Finally, using the above result and the fact that the relationship between the pdf relative to the process of injection back $p(\tilde{x})$ and the wtd $\psi_{CTIM}(t)$ is

$$\psi_{CTIM}(t) = p(\tilde{x}) \left| \frac{d\tilde{x}}{dt} \right|,$$

(4.7)

with $p(\tilde{x}) = 1$ in our case, one reaches the conclusion that

$$\psi_{CTIM}(t) = \left[1 + (z - 1) t\right]^{\frac{z}{1-z}}.$$

(4.8)

It is useful to introduce the parameters μ and T, defined as follows

$$\mu = \frac{z}{z - 1} \quad T = \frac{1}{(z - 1)},$$

(4.9)

in term of which $\psi_{CTIM}(t)$ becomes

$$\psi_{CTIM}(t) = \frac{(\mu - 1) T^{\mu-1}}{(T + t)^{\mu}}.$$

(4.10)

The parameters μ and T control two fundamental properties of the wtd. These are the behavior of $\psi_{CTIM}(t)$ for $t \to \infty$

$$\psi_{MM}(t) = \psi_{CTIM}(t) \propto \frac{1}{t^{\mu}} \quad t \to \infty,$$

(4.11)

and the mean waiting time[4],

$$< t > = \int_0^{+\infty} t \, \psi(t) \, dt = \frac{T}{\mu - 2}.$$

(4.12)

---

[4]A generalization of the CTIM consists in using $\dot{x} = kx^z$ with $k \geq 1$. This allows having different mean times, keeping μ unchanged.

In the case of $z = 1$ making properly the limit in all the equations used one gets

$$\psi(t) = e^{-t} \qquad\qquad < t >= 1. \qquad\qquad (4.13)$$

Finally, regarding the Chaotic region it can be shown easily that wtd is, independently from the value of the parameter z, an exponential since Eq.(4.1) can, for the purpose of evaluating the wtd, be approximated linearly.

The following table shows how the change of the parameter of the Manneville map z influences (Eqs. (4.9) and (4.11)) the asymptotic properties of the wtd $\psi_{MM}(t)$.

z	μ	$\psi_{MM}(t)$ as $t \to \infty$	mean time	standard deviation
1	$+\infty$	$exp(-t)$	finite	finite
$]1, 1.5[$	$]+\infty, 3[$	$t^{-\mu}$	finite	finite
$[1.5, 2[$	$[3, 2[$	$t^{-\mu}$	finite	infinite
$[2, +\infty[$	$[2, 1[$	$t^{-\mu}$	infinite	infinite

Table 4.1: The asymptotic property, the mean value and the standard deviation of the wtd $\psi_{MM}(t)$ as a function of the parameter z.

The values of the parameter $z = 1$ ($\mu = +\infty$), $z = 1.5$ ($\mu = 3$) and $z = 2$ ($\mu = 2$) are particular since they signal, respectively, the transition from an exponential wtd to a power law wtd, the transition from finite to infinite fluctuations around the average waiting time and the transition from a finite to an infinite mean waiting time.

Finally, let us discuss, briefly, the statistical characterization of the waiting times in the Chaotic region. it can be shown easily that the wtd in the Chaotic region is, independently from the value of the parameter z, an exponential since Eq.(4.1) can, for the purpose of evaluating the wtd, be approximated linearly.

4.1.2 Departure of Two Nearby Trajectories Inside the Laminar and the Chaotic Region

In this Section, the following quantities will be evaluated:

$$\zeta_{MML}(n, x_0) = \lim_{\Delta(0)\to 0} \frac{\Delta(n)}{\Delta(0)} \qquad\qquad (4.14)$$

$$\zeta_{MMC}(n, x_0) = \lim_{\Delta(0) \to 0} \frac{\Delta(n)}{\Delta(0)}, \tag{4.15}$$

where x_0 and $x_0 + \Delta(0)$ are two starting points chosen inside either the Laminar (Eq. (4.14)) or the Chaotic (Eq. (4.15)) region, and $\Delta(n)$ is the distance, after n iteration of the map. Clearly n cannot be as large as wished because both trajectories must be inside the same region[5]. In order to find $\zeta_{MML}(n, x_0)$, the CTIM approximation is used and the correspondent function $\zeta_{CTIM}(t)$ evaluated. Using Eq. (4.5) for the two initial positions x_0 and $x_0 + \Delta(0)$, one gets

$$x(t) = \left[x_0^{1-z} + (1 - z)t \right]^{\frac{1}{1-z}} \tag{4.16}$$
$$x(t) = \left[(x_0 + \Delta(0))^{1-z} + (1 - z)t \right]^{\frac{1}{1-z}} \tag{4.17}$$

and evaluating the difference between this two expression in the limit $\Delta(0) \to 0$, is possible to obtain

$$\zeta_{CTIM}(t) = \left[1 - (z - 1)x_0^{z-1}t \right]^{\frac{-z}{z-1}} . \tag{4.18}$$

In the case $z = 1$ Eq. (4.18) becomes

$$\zeta_{CTIM}(t) = e^t. \tag{4.19}$$

The function $\zeta_{CTIM}(t)$ can be considerate a good approximation of $\zeta_{MML}(n, x_0)$ when x_0 is close to zero. This power law behavior, a part from the case $z = 1$, of the departure of two nearby trajectory is what characterize the motion of a trajectory inside the Laminar region.

For the behavior of $\zeta_{MMC}(n, x_0)$ instead, using the same argument adopted in the previous section to derive the wtd in the Chaotic region, an exponential is obtained. This is the reason why the interval $]d, 1]$ of the phase space is called the Chaotic region.

---

[5]This exclude those parts of the phase space such that the application of the map result in a jump from the Laminar to the Chaotic region and vice versa: the aim here is to give an idea of what is meant by "slow" motion in the Laminar region and "fast" motion in the Chaotic region.

4.1.3 The Balance Between Order and Randomness

It can be explained, now, why the Manneville map can be considered as a prototype of complexity as balance of order and randomness. This is so because the phase space of the map is divided in two parts: the Laminar region, characterized by a power law departure (see Eq. 4.18) of nearby trajectories and therefore by "predictability", (in the sense that an error on the initial condition is propagated slowly) and the Chaotic region characterized by an exponential departure of nearby trajectories and therefore by "unpredictability". In this sense order and randomness must be thought the motion of trajectory in the phase space is "deterministic" until the trajectory stays in the Laminar part of the phase space, with the excursion in the Chaotic part being a short period of "random" motion.

The relative strength between this two kinds of motion is measured by the wtd $(\psi_{MM}(t))$ in the laminar region, which measures for how much time the motion of a trajectory is "deterministic". This wtd characterizes, therefore, the balance between order and randomness. Eqs. (4.8), (4.9) and (4.10) and the table 4.1 show how this balance is controlled by the parameter z (μ). As z goes from 1 ($\mu = +\infty$) to $+\infty$ ($\mu \to 1$), the systems move from a condition of total randomness to one of total determinism. Finally, for the CTIM the words "determinism" and "randomness" have a literal meaning.

4.2 Invariant Measure, Approach to Equilibrium and the Lyapunov Coefficient

This Section is dedicated to the evaluation of the invariant measure and the Lyapunov coefficient of the Manneville map and how they change, when the parameter z is changed. Moreover, it will be discussed, in the case where an invariant measure exists, how fast the corresponding equilibrium density $\rho_{eq}(x,t)$ is reached, starting from a given initial density $\rho(x,0)$.

4.2.1 Invariant Measure

In order to estimate the invariant measure of the Manneville map is convenient to use the CTIM, discussed in the previous section, and the corresponding GLE (see Sec. 2.4.3)

$$\frac{\partial}{\partial t}\rho(x,t) = -\frac{\partial}{\partial x}(x^z \rho(x,t)) + C(t).\tag{4.20}$$

The first term of the r.h.s. of this equation corresponds to the deterministic motion described by Eq. (4.3) and the second term of the r.h.s., the function $C(t)$, corresponds to the process of random and uniform[6] process of injection back in the Laminar region described by Eq. (4.4). Using the normalization condition

$$\int\limits_0^1 \rho(x,t)dx = 1,\tag{4.21}$$

one finds that $C(t)$ is the the value of the density $\rho(x,t)$ in $x = 1$, namely

$$C(t) = \rho(1,t).\tag{4.22}$$

In order to get the expression for the equilibrium distribution, the l.h.s. of Eq. (4.20) is set to 0 and $t \to +\infty$, so that the following expression is derived

$$\rho_{eq}(x) = \frac{C}{x^{z-1}} \qquad C = \lim_{t\to+\infty} C(t).\tag{4.23}$$

For $z < 2$ $\rho_{eq}(x)$ exists and can be normalized, obtaining

$$\rho_{eq}(x) = \frac{2-z}{x^{z-1}}.\tag{4.24}$$

The fact that there is no equilibrium distribution for $z \geq 2$ is not surprising, since for these values of the parameter the mean value of the wtd (Eq. (4.12)) is infinite (see also table 4.1). In this case any initial distribution will tend to concentrate inside

---

[6]Note how the function C depends only from the time t and not from the position x in the interval $[0, 1]$.

the Laminar region, but no equilibrium is reached. For the real Manneville map, in the case $z \in [1, 2[$, Eq. (4.24) is valid[7] only in the region where the approximation leading to Eq. (4.3) is a good one. Outside this region the $\rho_{eq}(x)$ of the Manneville map can be considered constant.

4.2.2 Lyapunov Coefficient

Sec. 4.1.1 was limited to the study of the departure of two nearby trajectories inside the Laminar or the Chaotic region. Hereby, this restriction is lifted so that the trajectories are free to move in the entire phase space, in order to evaluate the Lyapunov coefficient.

In the case, where an invariant measure exists ($z \in [1, 2[$), the Lyapunov coefficient λ is different from zero [1, 51]. The value of λ can be estimated, using the Pesin theorem (Eq. (3.12)) to write

$$\lambda(z) = \int_0^t \ln(1 + z x^{z-1}) \rho_{eq}(x) dx, \tag{4.25}$$

with Eq. (4.24) for $\rho_{eq}(x)$. Moreover, a fair approximation of the integral in the r.h.s. of the above equation, is the following expression

$$\lambda = (z)(2 - z)\ln 2. \tag{4.26}$$

In the case where no invariant measure exists, $z > 2$, the Lyapunov coefficient is null. Fig. 4.2 shows the results of the numerical evaluation of the Lyapunov coefficient (Eq. (3.11)), the results of the numerical integration of Eq. (4.25) and the function $\lambda(z)$ defined in Eq. (4.26). The last two quantities are practically indistinguishable, while a small discrepancy is observed with numerical evaluation directly from the map. This is due to the fact that the expression of Eq. (4.24) for the equilibrium density is not the exact solution.

The behavior of the Lyapunov coefficient for different values of the parameter z,

---

[7]Numerically, a fair agreement is found till $x \leq 0.2$ for all the value of z for which an equilibrium distribution exists.
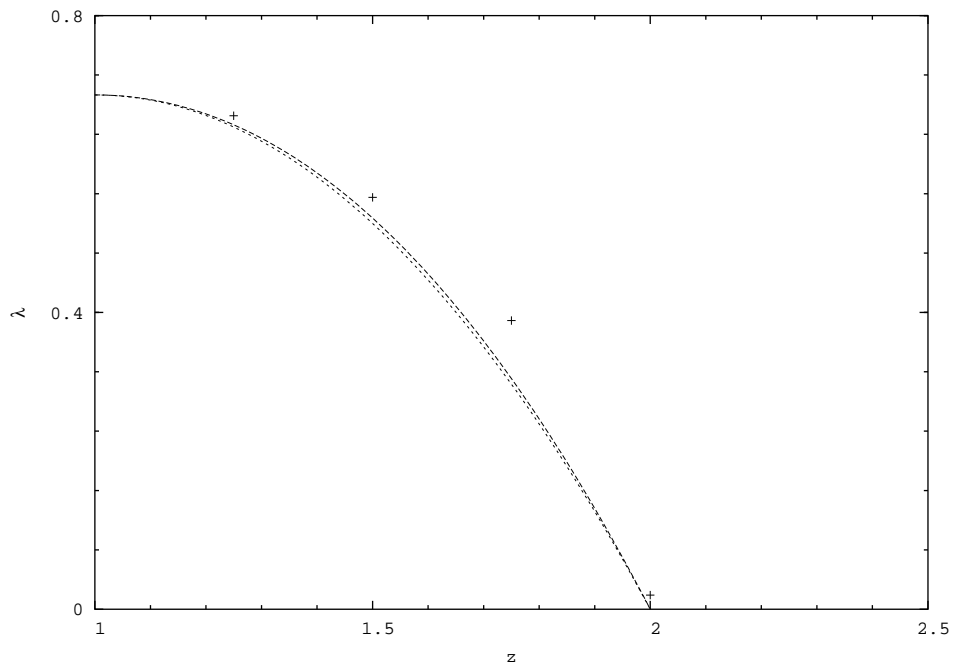
Figure 4.2: The Lyapunov coefficient λ as a function of the parameter z. The dots denotes the results of numerical evaluation, the full line and the dotted line, coinciding in this scale, denote respectively the results of the numerical integration of Eq. (4.25) and the function $\lambda(z)$ of Eq. (4.26).

can be understood in terms of the property of the wtd in the Laminar region. In fact a trajectory, during its motion in the phase space, jumps back and forth between the Laminar region (where the departure of wo nearby trajectory is a power law) and the Chaotic region (where the departure of two nearby trajectory is exponential). If the mean waiting time in the Laminar region is finite, for time much large than this mean value the overall motion of the trajectory can be considered as chaotic, therefore $\lambda \neq$). Moreover as $z \to 2$ the mean time tends to diverge, this is interpreted as a less intense chaotic motion $z \to 2 \mapsto \lambda \to 0$. Finally when the threshold $z = 2$ is reached and passed, with the mean waiting time in the Laminar region being infinite, the overall motion ceases to be chaotic $\lambda = 0$. In this particular condition, it can be proved [51], that the departure of two nearby trajectories become a stretched exponential.

4.2.3 Approach to Equilibrium

A formal solution of Eq. (4.20), can be obtained using the method of characteristics (for details on the method, the interested reader can look at Ref. [53], which adopts the same method to solve a problem quite similar to that here under study) and reads

$$
\begin{aligned}
\rho(x,t) \;=\; & \int_0^t \frac{C(\tau)}{[(\alpha(t-\tau)-1)x^{z-1}]^\mu} d\tau \\
& + \; \rho(\frac{x}{[(\alpha(t)-1)x^{z-1}]^T}, 0) \times \frac{1}{[(\alpha(t)-1)x^{z-1}]^\mu},
\end{aligned}
\tag{4.27}
$$

where

$$
\alpha(t) \equiv 1 + (1-z)t
\tag{4.28}
$$

and the parameter T and μ are those defined in Eq. (4.9). While for the function $C(t)$ the following expression is obtained

$$
C(t) = \int_0^t \frac{C(\tau)}{(\alpha(t-\tau))^\mu} d\tau + \rho(\frac{1}{(\alpha(t))^T}, 0) \frac{1}{(\alpha(t))^\mu}.
\tag{4.29}
$$

Now, Eqs. (4.27) and (4.29) will be used to monitor the regression to the equilibrium distribution for the CTIM, (for the real Manneville map the qualitative behavior

will be the same). Therefore, only the case where $z \in [1, 2[$ is considered and, for simplicity, the initial condition is chosen to be the uniform distribution in the interval $]0, 1[$. Finally, the following quantity is defined

$$\lambda(t, z) \equiv \int_0^t ln(1 + zx^{z-1})\rho(x, t)dx. \tag{4.30}$$

The quantity $\lambda(t, z)$ is a sort of time dependent Lyapunov coefficient, in the sense that: $t \rightarrow +\infty \Rightarrow \lambda(t, z) \rightarrow \lambda(z)$. Clearly the asymptotic behavior of $\lambda(t, z)$ is a reflection of the asymptotic regression to equilibrium of the density $\rho(x, t)$. Using the simplification $C(t) \approx 2 - z$ when t is big enough (see Eqs. (4.23) and (4.24)) and plugging Eq. (4.27) into Eq. (4.30), one obtains

$$\lambda(t, z) \approx \int_0^1 \frac{2 - z}{x^{z-1}} \left[1 - ((\alpha(t) - 1)x^{z-1})^{-\beta} \right] \times \ln(1 + zx^{z-1}) \, dx. \tag{4.31}$$

Differentiating this equation with respect to time and evaluating the resulting integral with the method of integration by parts, leads to (as $t \rightarrow +\infty$)

$$\lambda(t, z) \approx \lambda(z) + t^{-\frac{1}{z-1}}. \tag{4.32}$$

This asymptotic power law behavior is the same of that of the fraction of trajectories that, still, has to leave the Laminar region for the first time. In fact the initial distribution is the uniform one in the interval $]0, 1[$, therefore there is an infinite number of trajectories all of them, initially, inside the Laminar region. In an infinitesimal time dt a fraction dR of this trajectories will reach $x = 1$ to be injected back in a random and uniform way. A simple relation connects these infinitesimal quantities:

$$\frac{dR}{dt} = \frac{dx}{dt}. \tag{4.33}$$

The interval dx containing the trajectories that reach $x = 1$ in a time dt, can be found using Eq. (4.5). Solving Eq. (4.33) one finds, for the fraction R of trajectories that

still have to leave the Laminar region ($]0, 1[$) once, the following expression:

$$R(t) = \frac{1}{[1 + (z-1)t]^{\frac{1}{z-1}}}, \qquad (4.34)$$

Therefore the regression to equilibrium is driven by the process of first exit from the Laminar region. This is due to the fact that the process of injection back in the laminar region, is a random one, after a trajectory is injected back it has lost all memory of its starting position. This last consideration is valid also in the case of a generic initial density $\rho(x, 0)$ different from the uniform one. What can be different is the power law index characterizing the regression since it strongly depends on the initial condition.

4.3   KS Entropy and Information Complexity of the Manneville Map

Using the theory of Sec. 3.1, Sec. 3.2 and the results of Sec. 4.2, is possible to build the following table for the dependence of the KS entropy and the "quantity of information" of a generic trajectory of the Manneville map:

z	μ	KS entropy	$I(\omega_x^n), \{A\}$
$[1, 2[$	$[+\infty, 2[$	$\approx z(2-z)\ln 2$	$\propto n$
2	2	0	$\propto \frac{n}{\ln n}$
$]2, +\infty[$	$]2, 1[$	0	$\propto n^{\frac{1}{z-1}}$

Table 4.2: The KS entropy and the increase of "quantity of Information" $I(\omega_x^n), \{A\}$ contained in a trajectory starting at x for a given a partition $\{A\}$, as functions of the parameter z (μ).

The KS entropy and "quantity of Information" successfully discriminate the values of the parameter z for which an invariant measure exists, from those for which there is no invariant distribution. The former are those for which a positive KS exists and for which the information contained in a trajectory increase linearly with the length of the trajectory itself. The latter are those characterized by vanishing KS entropy and an increase of the information contained in a trajectory less than linear.

Finally it must be stressed that using the CTIM, one can reach the same results, looking at the function $C(t)$ defined in Eq. (4.29). In fact, this function is the evolution in time of the value of the density at $x = 1$ (Eq. (4.22)), and, in the CTIM, when a trajectory reaches $x = 1$, it is randomly and uniformly injected back. Therefore the function $C(t)$ can be thought as a rate of random events per unit of time and make sense to integrate this quantity to obtain the "number" of random events up to a time t

$$K(t) = \int_0^t C(t')dt'. \tag{4.35}$$

In the case of the existence of the invariant measure the rate $C(t)$ is constant ($C(t) = 2 - z$ from Eqs. (4.23) and (4.24)) and $K(t)$ increase linearly in time, while in the non stationary case $z > 2$ the rate decrease in time ($C(t) \sim t^{-\frac{z-2}{z-1}}$ see [1]) and $K(t)$ has an increase less than linear ($K(t) \sim t^{\frac{1}{z-1}}$) and for $z = 2$, $K(t) = \frac{t}{\ln t}$. In other words the function $K(t)$ behaves exactly has the "quantity of information" $I(\omega_x^n), \{A\}$.

4.4 Dynamic Approach to the SVM Diffusion

In this Section a slightly improved version of the CTIM[8] is used to create a dynamical model for the SVM diffusion an the properties of such a diffusion, studied.

Following the discussion of Sec. 2.3, a dynamical system described by two variables: y and ξ is created. As a result of the coupling with y, the variable ξ fluctuates and its fluctuations are collected by a new variable x, via Eq. (2.15). The variable ξ represents, in the case of the SVM, the state of the velocity of the walker (moving forward or backward) and therefore it is a dichotomous stochastic variable (without any loss of generality one can set the only two possible values of ξ to be $+1$ or -1). The final model that will be adopted, is a "specular" version of the CTIM. This is done, in order to reproduce the dichotomous property of ξ, namely, as in [56], with the extension of the phase space of the CTIM to be the interval $[0, 2]$. In this phase

---

[8]A generalization of the CTIM and not of the Manneville Map is adopted, because the latest is analytically solvable and as shown in the sections previous to this, it has the same property of the Manneville map.

space a trajectory will move according the following prescription: If the trajectory is in the left (right) part $[0, 1[$ $(]1, 2])$ of the interval, then it moves forward (backward) the point $y = 1$, with a dynamic described by Eq. (4.3). After reaching this point the trajectory has equal probability of being, instantaneously, injected back uniformly in the left or right part. Then, to the variable ξ is assigned a value $+1$ (-1) if at the time t the trajectory $x(t)$ is in the left (right) part of the phase space. Finally, in order to create the a diffusion process, one need to consider infinitely many trajectories in the phase space, and since a stationary process is desired, an ensemble of trajectories distributed accordingly the equilibrium distribution of this model, to which we refer as the Generalized CTIM (GCTIM), must be selected. This equilibrium distribution exists only for the value of the parameter z, of Eq. (4.3), inside the interval $[1, 2[$. In fact, as seen in Sec. 4.2.1, Eq. (4.24) defines the equilibrium distribution for the CTIM and for, obvious reason, that one of the GCTIM in the left part of the phase space $([0, 1[)$, while in the right part of the phase space $(]1, 2])$, due to the symmetry of the model, the equilibrium distribution is the mirror image of that one of the left part.

Let us, now, discuss, the property of the diffusion process created in this way. At the beginning all the diffusing trajectories are positioned at the origin $x = 0$ and all the corresponding trajectories in the phase space of the GCTIM are distributed according to the equilibrium distribution. As the diffusion process starts, the diffusing trajectories will split in two parts, the one moving to the right and the one moving to the left, both of them with constant speed and, therefore, these peaks are called ballistic. But, as soon as, this peaks take form they start loosing "population", this is due to the fact that the corresponding trajectories in the phase space of the GCTIM are undergoing the process of random injection back in one of the two Laminar regions. In fact if a trajectory reaches $y = 1$ from the left (right), it has 50% of probability of being injected back in the right (left) Laminar region, the corresponding diffusing trajectory will abandon the left (right) ballistic peak to move backwards to $x = 0$. Clearly after a diffusing trajectory leaves a ballistic peak will never be able to go back. The overall picture of the diffusion process, created in this way, is that one of two ballistic peaks moving one to the left and one to the right, with a decaying intensity

and, as a result of this, an increase of population in between the peaks. Therefore it is necessary, in order to characterize this kind of diffusion, to quantify the decay of the population of the ballistic peaks and the shape of the pdf $p(x, t)$, in between.

- **Population of the ballistic peaks**

  This problem is, first, addressed in the simplified case where every time a trajectory of GCTIM reach $y = 1$, it is injected back in the Laminar region opposite to the one from which is coming. In term of the diffusing trajectory, these implies that no coin tossing is involved to decided the new direction of motion, the new direction is always the opposite of the previous one. This model is also know as the Two State SVM, or the Alternate Sign (AS) case. With this simplification, the population of the left or of right peak $P_{AS}(t)$ (they are the same due to the symmetry of the problem) is just the fraction of the trajectories in the phase space of the GCTIM, who still has to reach $y = 1$ from the left or from the right Laminar region, namely

  $$P_{AS}(t) = \int_0^{\tilde{y}(t)} \rho_{eq}(y) \, dy, \qquad (4.36)$$

  where $\tilde{y}(t)$ is the initial condition in the interval $[0, 1[$ corresponding to an arrival time in $y = 1$ equal to t. Using, Eq. (4.5),

  $$\tilde{y}(t) = [1 - (1 - z)t]^{\frac{1}{1-z}}, \qquad (4.37)$$

  and performing the integration in Eq. (4.36), with the help of Eq. (4.37) and the relationship between z and μ given by Eq. (4.9), one finds

  $$P_{AS}(t) = \left(\frac{T}{T + t} \right)^{\mu - 2}. \qquad (4.38)$$

  This important result can be obtained using the ergodicity of the GCTIM. Instead of performing an integral over the equilibrium distribution $\rho_{eq}(y)$, a single trajectory $y(t)$ in the phase space of the GCTIM and the corresponding

61

fluctuating function $\xi(t)$ are considered. The ergodic condition tell us that the trajectory $y(t)$ will visit in its motion different region of the phase space with a frequency proportional to the invariant natural measure of the region. The function $\xi(t)$ as, stated before, is the velocity of walker of the SVM and in the, AS case, its evolution in time consists of alternating regions of opposite value $+1$ and -1, which duration is dictated by the wtd in the laminar region $\psi(t)$. Thereby, to evaluate the integral of Eq. (4.36), the population of right (left) ballistic peak one must select a window of length t, move it through all the sequence of values of ξ and count the fraction of times this window is contained in one single region of $+1$ (-1). This way of proceeding leads to

$$P_{AS}(t) = \frac{1}{\tau} \int_{t}^{+\infty} (t' - t)\psi(t')dt',$$ (4.39)

where $< \tau >$ is the mean value of ψ. The expression of Eq. (4.39) connects the function $P_{AS}(t)$ with the wtd in the laminar Region $\psi(t)$, and it can be used to derive the peaks' population $P(t)$ in the case of our interest, where the walker chooses its new direction of motion randomly, a case that will be addressed, from now on, as the Random Sign case (RS). The RS case can always be expressed as an AS case, if the wtd $\psi(t)$ is replaced by "effective" wtd $\psi_{eff}(t)$, which Laplace transform $\tilde{\psi}_{eff}(s)$ is connected [36] with the Laplace transform of the original wtd, through

$$\tilde{\psi}_{eff}(s) = \frac{\tilde{\psi}(s)}{2 - \tilde{\psi}(s)}.$$ (4.40)

Therefore Eq. (4.39) can be used also in the RS case, if the "effective" wtd is adopted. Finally, let us discuss, how in the RS case, the peaks' population is connected with the correlation function $\Phi_\xi(t)$ of the variable ξ. The correlation function can be evaluated using the equilibrium distribution or adopting one

single trajectory. In this last case $\Phi_\xi(t)$ reads

$$\Phi_\xi(t) = \lim_{T \to +\infty} \int_0^T \xi(t')\xi(t'+t)\,dt'. \tag{4.41}$$

This means that in order to evaluate the correlation function, one must move a window ot size t along the sequence $\xi(t)$. Such a procedure makes it possible to write the r.h.s. of Eq. (4.41) as follows,

$$\Phi_\xi(t) = p_{++}^{SLR} + p_{--}^{SLR} + p_{++}^{DLR} + p_{--}^{DLR} - p_{+-} - p_{-+}, \tag{4.42}$$

where p_{s_1,s_2} is the probability that, moving a window of size t across the sequence $\xi(t)$, the value of the variable ξ at the initial and the final position, of such a window, are s_1 and s_2 respectively, while the superscript SLR (same laminar region) indicates the fact that between the initial and the final position of the moving window the same symbol is encountered. The superscript DLR, instead, indicates the fact that at least a change of symbol happened in between the initial and final position. Due to the random choice of the new direction of motion, adopted by the walker, in the RS case, $p_{++}^{DLR} = p_{--}^{DLR} = p_{+-} = p_{-+}$ and $p_{++}^{SLR} = p_{--}^{SLR}$, and, using the same arguments that leads to Eq. (4.39), one finds for the correlation function:

$$\Phi_\xi(t) = 2p_{++}^{SLR} = \frac{1}{\tau} \int_t^{+\infty} (t'-t)\psi(t')dt'. \tag{4.43}$$

Therefore, in the case of our interest (the RS case)

$$P(t) = \frac{1}{\tau} \int_t^{+\infty} (t'-t)\psi_{eff}(t')dt' \tag{4.44}$$

$$\Phi_\xi(t) = \frac{1}{\tau} \int_t^{+\infty} (t'-t)\psi(t')dt'. \tag{4.45}$$

and finally, using Eq. (4.40) after, a tedious, calculation

$$P(t) = \Phi_\xi(t) \qquad t \to +\infty. \tag{4.46}$$

- $p(x,t)$ **in between the peaks.**
  What is the pdf, that is created, as time goes on, in between the two ballistic peaks by the trajectories that abandoned the two peaks? Again, it is possible to use the single trajectory approach and a window of size t moving across the sequence $\xi(t)$. Following the prescription of Eq. (2.16), a summation of all the values of ξ founded inside a given window is made. If, inside, a window the walker has decided N times which direction to choose (that is if the corresponding trajectory in the phase space of the GCTIM has escaped the laminar region N times), then sum Σ of ξ inside this window[9] can be written has

$$\Sigma = \sum_{j=1}^{N} s_j \tau_j, \tag{4.47}$$

  where τ_j are the time duration of each laminar phase and s_j can be $+1$ or -1 in a random fashion. Let us recall, that each of the times t_j is a stochastic variable, since the process of injection back in the laminar region is random, and that all of them have the same pdf, the wtd in the Laminar region of Eq. (4.10). Therefore the sum of Eq. (4.47), is the sum of iid stochastic variables and the CLT (Sec. 2.2) or GCLT (Sec. 2.7) can be applied. Asymptotically $(t \to +\infty)$ the wtd of Eq. (4.10) is a power law $t^{-\mu}$, therefore the pdf of the central part (the one in between the two ballistic peaks) of the SVM diffusion becomes, at large times, a Gaussian if $\mu \in]3, +\infty]$ and a Lévy distribution of index $\alpha = \frac{1}{\mu-1}$ if $\mu \in]2, 3[$. Thereby, for the overall pdf p(x,t) in the limit of

---

[9]Eq. (4.47) is valid if the window considered is composed by exactly N laminar regions. Anyway, the argument used to derive the asymptotic from for the pdf inside the ballistic peaks can be applied also in the cases where Eq. (4.47) does not apply.

large times, the following expression can be derived

$$p(x,t) = \Theta(t - \mid x \mid)p_{L_\alpha}(x,t) + \frac{1}{2}\delta(\mid x \mid -t)\Phi_\xi(x,t), \qquad (4.48)$$

where $\Theta(\cdot)$ denotes the Heaviside step function and $p_{L_\alpha}(x,t)$ is a Lévy distribution of index α (the case $\alpha = 2$ being the Gaussian distribution).

4.4.1 Scaling Properties, if Any, of the SVM Diffusion

In this section, the MS analysis is applied to the SVM diffusion process in order to see if a scaling regime, for same value of the index μ, exists. Using Eqs. (3.34) and (3.35) and the asymptotic expression for the pdf $(p(x,t))$, one finds, for the q-th fractional standard deviation $[\sigma_q(t)]^q$, the following asymptotic expression

$$[\sigma_q(t)]^q = A\ t^{q\delta} + B\ t^{q-(\mu-2)}, \qquad (4.49)$$

where A and B are two positive constants and $\delta = \frac{1}{\alpha}$ is the scaling parameter of the the pdf $p_{L_\alpha}(x,t)$. The first term of the r.h.s. of Eq. (4.49) represents the contribution of the central part of the pdf, while the second term is the contribution due to the ballistic peaks. Let us recall the reader that the MS analysis rests on looking for each fractional standard deviation the scaling exponent $\xi(q)$ such that

$$[\sigma_q(t)]^q \propto t^{\xi(q)}. \qquad (4.50)$$

Thus, according to the value of the parameter μ of the wtd in the Laminar region (Eq. (4.10)), the following results apply:

- $\mu = +\infty$

  In this case the wtd ψ is an exponential and also the peaks's population decays as an exponential (Eq. (4.46)). Thus the contribution of the ballistic peaks to the q-th fractional standard deviation becomes irrelevant, with respect the one due to the central part, for any given q. Moreover the asymptotic expression

for the pdf of the central part is a Gaussian distribution. Thus,

$$\xi(q) = 0.5 \times q \qquad \forall q > -1 \tag{4.51}$$

and the diffusion process satisfies the scaling condition with $\delta = 0.5$.

- $\mu \in [3, +\infty[$

  In this case the pdf in the central part tends, asymptotically, to a Gaussian distribution, but the peaks' population decay as a power law of index $\mu - 2$, thus the second term of the r.h.s. of Eq. (4.49) is not negligible, and

$$\xi(q) = 0.5 \times q \qquad -1 < q < 2\mu - 4 \tag{4.52}$$
$$\xi(q) = q - (\mu - 2) \qquad \forall q \geq 2\mu - 4 \tag{4.53}$$

  No scaling condition is, therefore, reached.

- $\mu \in]2, 3[$

  In this case the pdf in the central part tends, asymptotically, to a Lévy distribution and as in the previous case the contribution of the ballistics peaks to the q-th fractional standard deviation cannot be neglected. The expression for the exponents $\xi(q)$ is:

$$\xi(q) = \frac{1}{\mu - 1} \qquad -1 < q < \mu - 1 \tag{4.54}$$
$$\xi(q) = q - (\mu - 2) \qquad q > \mu - 1, \tag{4.55}$$

indicating the absence of the scaling condition. Fig. (4.3) shows the behavior of the $\xi(q)$ in this last case. The agreement with Eqs. (4.54) and (4.55) is perfect.

In conclusion one can say that a part from the case where the wtd in the Laminar region is exponential, the scaling condition never applies. This is so, because, in this case, the Laminar motion becomes a Chaotic motion and, in this sense, the exponential case is equivalent to the absence of any order. As soon as order comes into play, there is a departure from an exponential wtd, the wtd becomes, in fact, a
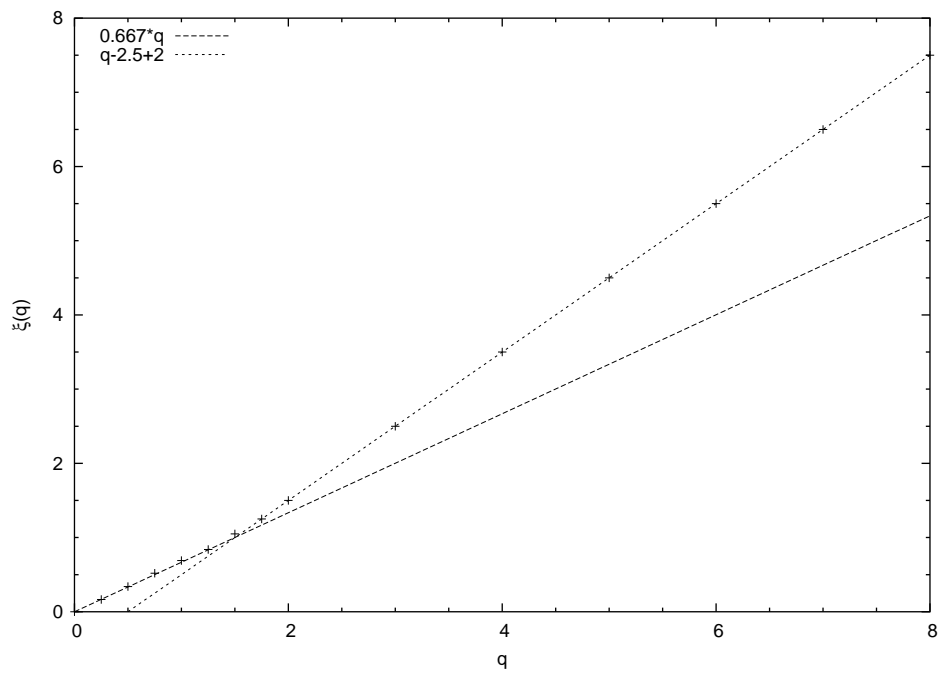
Figure 4.3: The exponent $\xi(q)$ as a function of q. Here $\mu = 2.5$ The change of slope happens at $q = 1.5$ as predicted by Eqs. (4.54) and (4.55).

power law. Even a small amount of order breaks down the scaling condition. This is so because the amount of order is reflected by the way the population in the ballistic peaks decay (Eq. 4.46). The ballistic peaks are responsible for a faster spreading than the spreading taking place inside the central part. Thus, if they do not decay fast enough (in an exponential fashion), they will dominate the high fractional moment, making evident that some form of "order" is present. Finally, moving the balance between order and randomness in favor of the first one, even the diffusion process inside the ballistic peaks undergoes a transition, from ordinary (Gaussian pdf) to anomalous (Lévy pdf). This is the case where $\mu \in]2,3[$.

The lack of scaling above discussed, is very peculiar, since it is stemming from a stationary process (the initial distribution was the equilibrium distribution for GC-TIM), and not from a non stationary one. Therefore, this diffusion process is said to be characterized by "aging", precisely for the lack of a scaling condition. in fact the validity of the condition of Eq. (3.24), implies the impossibility of determining the age of the diffusion process (for how much time the diffusion process has been taking place), whereas in our case, using the peaks' population (Eq. (4.44)), the distance between the two peaks and the parameter b, obtained from the Fourier transform of $p_{L_\alpha}(x,t)$ (Eq. (2.75)), it is possible to evaluate the "age" of the process.

4.4.2   The Conflict Between Trajectories and Densities

Let us now discuss, in some details the breakdown of the equivalence between the trajectory and the density approach, anticipated in Sec. 2.4.5. First of all, one has to write [56] the GLE relative to the GCTIM for the reduced density $\sigma(x,t)$ relative to the diffusing variable x.

$$\frac{\partial}{\partial t}\sigma(x,t) = <\xi^2> \int\limits_0^t \Phi_\xi(t-t^{'}) \frac{\partial^2}{\partial x^2}\sigma(x,t^{'})\, dt^{'}, \qquad (4.56)$$

where $<\xi^2>$ indicates the variance of the fluctuating variable ξ and $\Phi_\xi(t)$ is its correlation function (Eq.(4.41)). If the wtd in the Laminar region $\psi(t)$ is an exponential, then also $\Phi_\xi(t)$ is an exponential and Eq. (4.56) is the "telegrapher's"

equation, whose asymptotic solution is a Gaussian distribution. This solution is also the solution of the trajectory approach, as indicated by the arguments of the previous section and in particular by Eq. (4.51). In the power law case [57] the asymptotic solution still satisfies the scaling condition of Eq. (3.24). The scaling parameter is $\delta = 0.5$ for $\mu \in]3, +\infty[$ (since the asymptotic solution is a Gaussian distribution), while it is $\delta = \frac{4-\mu}{2}$ for $\mu \in]2, 3[$. Both these results are in contrast with the result of the trajectory approach as indicated by Eqs. (4.52), (4.53), (4.54) and (4.55).

An explanation of this, surprising breakdown, can be addressed, comparing the moments (the even moments, since the odd ones are null for the SVM diffusion) of the distributions relative to the case of the density approach and to the case of the trajectory approach. For the density it can be shown [57], that the following expression applies

$$< x^n(t) >_{dens} \propto \int\limits_0^t dt_n \int\limits_0^{t_n} dt_{n-1} \ldots \int\limits_0^{t_2} dt_1 \Phi_\xi(t_2 - t_1) \Phi_\xi(t_4 - t_3) \ldots \Phi_\xi(t_n - t_{n-1}) \quad (4.57)$$

while for the trajectories, using the same argument of Sec. 2.5 one finds

$$< x^n(t) >_{traj} \propto \int\limits_0^t dt_n \int\limits_0^{t_n} dt_{n-1} \ldots \int\limits_0^{t_2} dt_1 < \xi(t_1)\xi(t_2) \ldots \xi(t_n) >, \quad (4.58)$$

where in both Eqs. (4.57) and (4.58) the times are ordinated $t_{j+1} > t_j$, $\Phi_\xi(t_j - t_{j-1})$ is the correlation function of the variable ξ and $< \xi(t_1)\xi(t_2) \ldots \xi(t_n) >$ is the n times correlation function of the variable ξ. The source of disagreement relies on the conditions, under which

$$< \xi(t_1)\xi(t_2) \ldots \xi(t_n) > = \Phi_\xi(t_2 - t_1) \Phi_\xi(t_4 - t_3) \ldots \Phi_\xi(t_n - t_{n-1}) \quad (4.59)$$

is true. For $n = 2$ Eq. (4.59) is trivially true, regardless ot the expression of $\psi(t)$. Let us examine, for example, the case $n = 4$. Using the same arguments involved in the derivation of Eq. (4.42), the condition of Eq. (4.59) can be written in the following

way

$$p(12 \cap 34) = p(12)p(34), \tag{4.60}$$

where the $p(a \cap b)$ denotes the probability that the conditions a and b are both "true" and the condition ij means that the moving a window of length $t_j - t_i$ along the sequence $\xi(t)$, the initial and the final position of such a window are found in the same laminar region (there is no decision made by the walker regarding which new direction of motion to take). Eq. (4.60) can be generalized to the case of the $2n$-th order, as follows

$$p(12 \cap 34 \cap \ldots \cap (2n-1)2n) = p(12)p(34) \ldots p((2n-1)2n). \tag{4.61}$$

Therefore, for the case of the density the events ij are, always, considered as independent, while for the trajectories these events are in principle not independent. This is the "nature" of the discrepancy between trajectories and density.

In the case of an exponential wtd in the Laminar region $(\psi(t))$, Eqs. (4.59), (4.60) and (4.61) are all satisfied [56]. Therefore, the moments of Eq. (4.57) are and Eq. (4.58) are identical for all the values of n, thereby explaining the perfect agreement in the exponential case between trajectory and density (in both case the asymptotic pdf is a Gaussian distribution and the scaling condition is satisfied). In the case of a power law wtd, instead Eqs. (4.59), (4.60) and (4.61) are satisfied only for $n = 2$, and Eq. (4.57) and Eq. (4.58) will induce different moments as described in Eqs. (4.52) and (4.53), or, Eqs. (4.54) and (4.55), depending from the value of the power law index μ.

CHAPTER 5

Complexity and Non-stationarity in Time Series

This chapter is dedicated to the discussion of the important issue of detecting complexity in time series (a sequence $\{\xi_j\}$ $j = 1, 2, \ldots, N$ of seemingly fluctuating numbers), in particular when the time series to be analyzed is not stationary. The kind of non-stationarity examined is that relative to the presence of two kind of biases. A "slow" change in the mean value of the fluctuations ("slow" in the sense that the time scale of the change of the average is much larger then the time scale relative to the fluctuations around the average itself) and a periodic change of the mean value. In Sec. 5.1 and Sec. 5.2 is discussed, with the help of computer generated time series, how the presence of this biases, can lead to misleading conclusion regarding the complexity of the time series analyzed and, consequently, how to eliminate these effects and being able of studying the complexity of the unbiased fluctuations. In the final section (Sec. 5.3) of this chapter, the results of the previous two sections are applied to data of great sociological interest, the teen birth data (the daily number of births, in the whole state of Texas, due to a teenage mother in a span (from 1964 to 1999) of 36 years), in order to discover the real complexity of this sociological process.

5.1 Computer-Generated Data: Effects of a Slow Component

In this Section is addressed the problem of non stationary condition due to a slow change of the average of the fluctuations. The methodology (DE,SM,MS and DAS analysis), described in Sec. 3.4 will show, how the presence of the slow changing bias has the effect of hiding the real complexity of the unbiased fluctuations. The information about the complexity can be recovered, only after a proper detrending of the bias.

Hereby, computer genrated time series are used to illustrate the effects of the bias and how to eliminate the bias, as well. Adopting a discrete-time picture, so that the time series to analyze reads as follows:

$$\xi_j = S_j^T + \zeta_j. \tag{5.1}$$

The function ζ_j represents a stationary stochastic process (a noise). In particular for the numerical simulations of this Section, ζ_j is selected to be a gaussian noise (GN) of zero average and standard deviation $\sigma = 12$ (the theoretical results discussed in the following sections do not depend from the choice of correlated or uncorrelated noise). A diffusion process, generated by this variable alone, yields a scaling condition with scaling parameter $\delta = 0.5$ and no "complexity". The function S_j^T, instead, is the one responsible for the slow change of the average of the fluctuating variable ζ, therefore, from now on, it will be addressed as the slow component. The parameter T indicates the time scale at which changes in the value of the function S_j^T happen. For our computer simulations, three different kind of slow components (see Fig. 5.1) are used, all of them with zero mean: a linear drift with a small slope, SL_j^T, a slowly changing continuous function, SC_j^T, and a step function, SS_j^T. The choice of a standard deviation $\sigma = 12$ for the GN, makes the intensity of the noise component comparable to the intensity of the slow components SC_j^T and SS_j^T, and predominant if compared to that of the SL_j^T component. Moreover the length of all the time series considered will be of 13149: that of the real data (teen birth data) analyzed in Sec. 5.3.
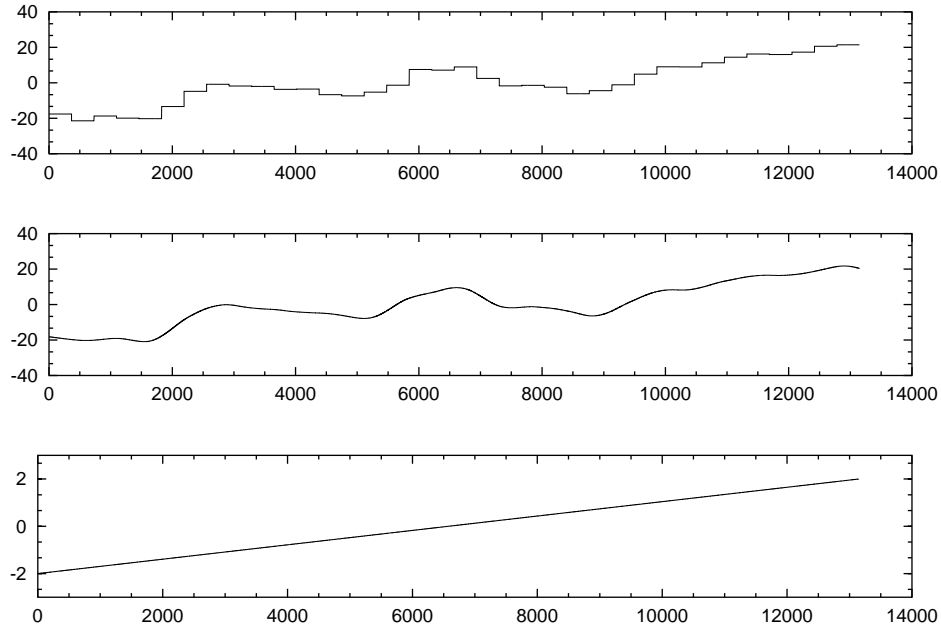
Figure 5.1: The three different slow components adopted. From the top to the bottom frame: step function, continuous smooth function and straight line with a small slope. The initial value of the straight line is -2 and the final, at a time $t = 13149$, is 2.

5.1.1 Diffusion Entropy (DE) and Second Moment (SM) Analysis

The application of DE and the SM analysis to the computer-generated sequences of Eq. (5.1) is shown in Figs. 5.2 and 5.3. In both figures the top, middle and bottom frame refer to SS_j^T, SC_j^T and SL_j^T, respectively. Each frame shows both the results relative to the whole time series and the results obtained when the two components are considered separately.

As far as the slow components SS_j^T and SC_j^T are considered, Fig. 5.2 shows that at short times the DE generated by the time series of Eq. (5.1), is different from both the DE of the slow component and the DE of the GN component. This is explained by the fact that the slow component and GN have comparable intensities. Thus, it is

evident that their joint action generates a faster spreading of the pdf and therefore a faster entropy increase. At long times the joint action of the two components yields the same effects as the slow component alone. This is due to the fact that the slow component generates ballistic diffusion, which is faster than the simple BM diffusion generated by the GN alone.

A different behavior appears with the slow component SL_j^T. In this case, the short-time behavior is dominated by the noise component. This dominance by the noise occurs because the intensity of this component is smaller than the noise intensity. However, even in this case the long-time behavior is dominated, as explained earlier, by the slow component. This is a remarkable property, because in this case a mere visual inspection of the time series is not sufficient to reveal the presence of a bias, which is hidden by very large fluctuations. In any event, the adoption of the DE method yields a large scaling exponent that one might erroneously attribute to highly correlated fluctuations. This is an effect that must be taken into account when analyzing real time series.

The SM analysis reveals properties similar to those emerging from the DE analysis, the only relevant difference is that the convergence to the steady condition of the slow component alone is much faster than in the corresponding case of the DE analysis. It is worthwhile to discuss this result in detail. Let us define $\sigma_{tot}(t)$ as the total standard deviation, at a time t, of the diffusion process relative to the sum of the slow component and the noise. Under the assumption that the individual components of the signal of Eq. (5.1) are mutually independent, one writes

$$\sigma_{tot}^2(t) = \sigma_{slow}^2(t) + \sigma_{noise}^2(t) = \sigma_{slow}^2(t) \left[1 + \frac{\sigma_{noise}^2(t)}{\sigma_{slow}^2(t)} \right]. \tag{5.2}$$

The SM analysis rests on evaluating the increase with time of the logarithm of the standard deviation. With some elementary algebra, it is possible to write Eq. (5.2) as

$$\log\left[\sigma_{tot}(t)\right] = \log\left[\sigma_{slow}(t)\right] + \frac{1}{2} \log\left[1 + \frac{\sigma_{noise}^2(t)}{\sigma_{slow}^2(t)} \right]. \tag{5.3}$$
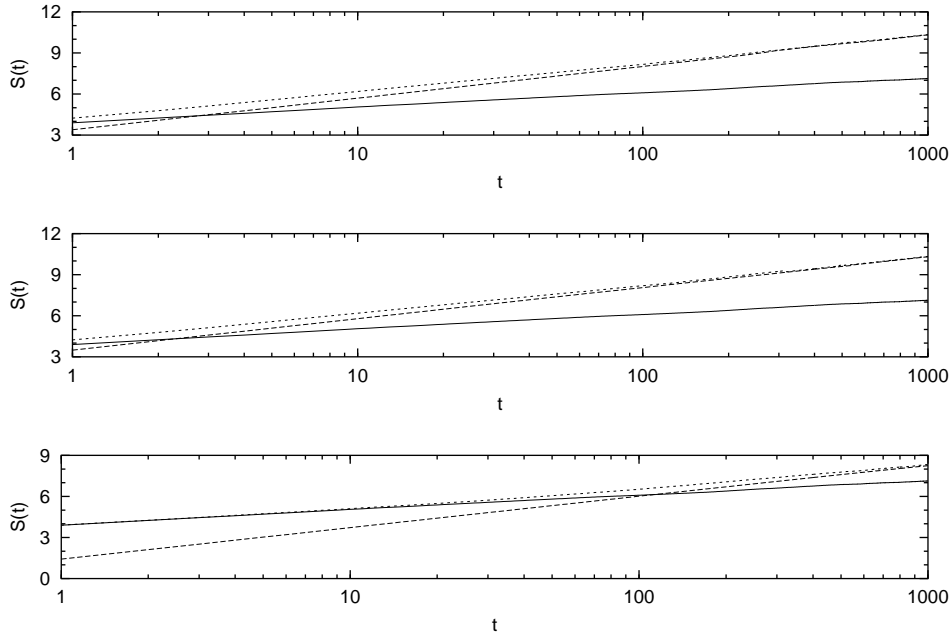
Figure 5.2: The diffusion entropy, $S(t)$, as a function of time t, in a logarithmic scale. Each frame shows three different curves, concerning noise, solid line, slow component, dashed line, and sum of noise and slow component, dotted line. As far as the slow component is concerned, from the top to the bottom frame are shown the results concerning SS_J^T, SC_J^T, and SL_J^T.

When the noise standard deviation is smaller than the slow component standard deviation, namely, $\sigma_{noise}(t) \ll \sigma_{slow}(t)$, using the Taylor series expansion of the logarithm. one finds

$$\log\left[\sigma_{tot}(t)\right] \approx \log\left[\sigma_{slow}(t)\right] + \frac{1}{2}\frac{\sigma_{noise}^2(t)}{\sigma_{slow}^2(t)}. \tag{5.4}$$

Therefore, when $\sigma_{noise}(t) \ll \sigma_{slow}(t)$ the leading contribution of the SM analysis is the logarithm of the slow component standard deviation, and the next expansion term is the square of the ratio of the noise standard deviation to the slow component standard deviation. In the case of diffusion entropy the numerical results indicate
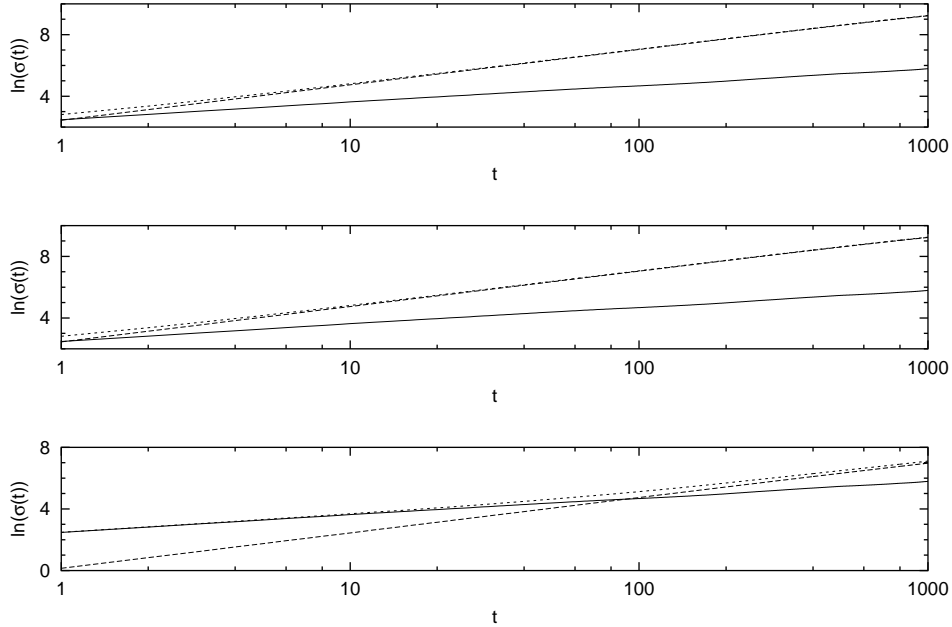
Figure 5.3: The logarithm of the standard deviation, $\ln(\sigma(t))$, as a function of time t, in a logarithmic scale. Each frame contains a set of three curves, the full line referring to the noise alone, the dashed line to the slow component alone and the dotted line to the sum of slow component and noise. From the top to the bottom frame these sets of curves refer to SS_J^T, SC_J^T and SL_J^T, respectively.

that a plausible expression to use is

$$S_{tot}(t) = S_{slow}(t) + O\left[\frac{\sigma_{noise}(t)}{\sigma_{slow}(t)}\right]^{\alpha},\qquad(5.5)$$

with $\alpha < 2$. In fact, the numerical results illustrated in the frames of Fig. 5.2 corresponding to SC_j^T and SS_j^T reveal that the diffusion entropy is more sensitive to the noise component than is the SM analysis. For a deeper understanding of the diffusion created by the superposition of the noise and the slow component is necessary the use of the multiscaling analysis.

5.1.2 Multiscaling (MS) Analysis
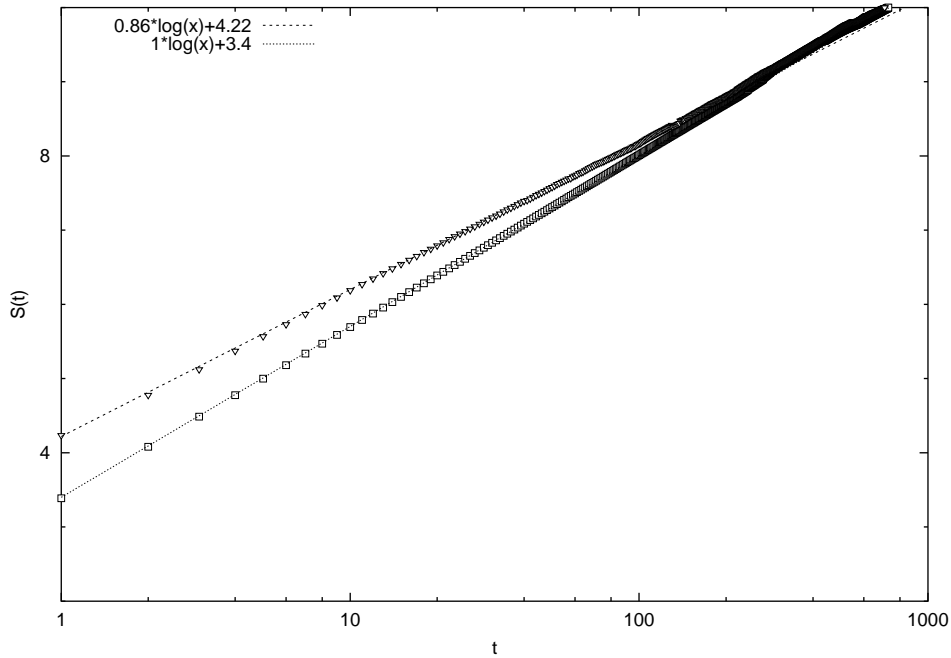
Figure 5.4: The diffusion entropy $S(t)$ as a function of time t, in logarithmic scale. The squares denote the time series corresponding only to the slow component SS_j^T and the triangles denote the slow component SS_j^T plus noise. Two straight lines suggest that in the time range from 10 to 100 the former and the latter curves correspond to $\delta = 0.86$ and $\delta = 1$, respectively.

Fig. 5.4 shows that in the long-time limit, $t > 100$, the DE produced by the time series resulting from the sum of noise and slow component SS_j^T coincides with the DE generated by the slow component alone that increases with a slope equal to 1. This asymptotic property has already been discussed. Our goal here is to shed light on the convergence to the long-time scaling of the whole time series, with the help of multiscaling analysis. To accomplish this goal, one must divide the time range, explored by the DE analysis, into three time regions: times smaller than 10 (early

stage of the diffusion process), times between 10 and 100 (middle stage of the diffusion process) and times from 100 to 1000 (later stage of the diffusion process). Fig. 5.5 shows the results of the multiscaling analysis applied to the two time series in Fig. 5.4, for each of these three time regions.
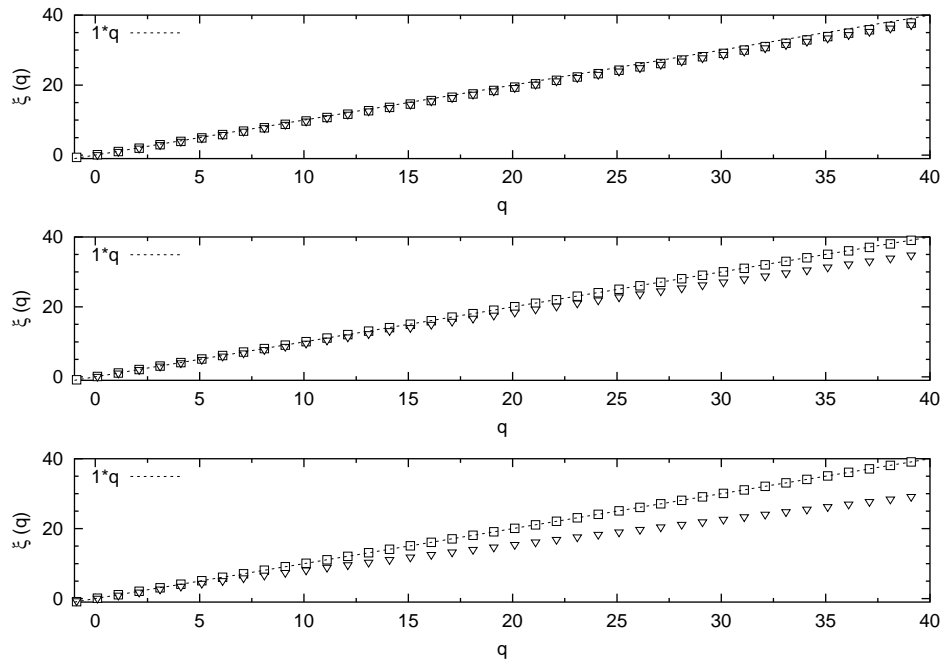
Figure 5.5: The exponent $\xi(q)$ as a function of the parameter q. The squares and the triangles denote the results of the MS analysis applied to the slow component alone and to the the sum of the slow component and noise, respectively. The three frames refer, from the top to the bottom to the short-time region, between 1 and 10, the middle-time region, between 10 and 100, and the large-time scale, between 100 and and 1000.

The three frames of Fig. 5.5, from bottom to top refer to the early, the middle and the later stage of the diffusion process, respectively. The squares denote the numerical results relative to the slow component alone, the triangles denote the numerical results concerning the sum of slow component and noise and the dashed line corresponds to

a straight line of slope 1, as expected for a diffusion process with a ballistic scaling. Moving from the bottom to the top, one can notice that the agreement between triangles and squares tend to increase. At the top, there is an almost complete equivalence throughout the whole q-region explored. Moreover, in the early and middle region the disagreement between triangles and squares tend to increase upon increasing q. This means that the difference between the two cases becomes more and more significant at larger and larger distances. The presence of noise tends to slow down the distribution broadening.

To understand the influence of the GN on the slow component, one must notice that the position of any of the diffusing trajectories can be written at time t as

$$x(t) = x_{slow}(t) + x_{noise}(t), \tag{5.6}$$

where the contributions to $x(t)$ are separated into the slow component and noise, respectively. Now, since the noise has a vanishing mean, there is 50% probability that the absolute value of $x(t)$ is increased by the presence of the noise and 50% that it is decreased. When the absolute value of $x(t)$ is raised to a power of q larger than 1, the half with positive increment contribute to the q-th moment with a greater weight than the other half. So, at any given time, the presence of the noise component makes the q-th moments , with $q > 1$, larger than in the case without noise, namely, the larger q, the larger the discrepancy. At long times, the moments of the two distributions, those with and without noise, must coincide. As earlier pointed out, this occurs, because the noise component has slower diffusion. Consequently, the moments of the distribution with noise undergo a slower increase than the moments of the distribution without noise.

5.1.3 Direct Assessment Scaling (DAS)

Finally, the DAS method is used to shed light on the apparently confusing situation emerging from the use of DE, SM, and MS analysis. As seen in the earlier in this Section, the DE suggests that $\delta = 0.86$ might be a reasonable measure of scaling, thereby suggesting the existence of pronounced cooperation in the system under study.

The MS method, on the contrary, suggests that the ballistic scaling, $\delta = 1$, is a more appropriate representation of the system complexity, at least in the long-time regime. Using the DAS one discovers that neither of these two conditions is a proper representation of the system dynamics. Previously, three distinct time regimes, short, intermediate and long have been studied. Now, just the intermediate time regime, ranging from 10 to 100 is explored. This choice is dictated by the fact that the deviation from a straight line in the middle frame of Fig. 5.5 suggests that the DE prediction of $\delta = 0.86$ is questionable.

The DAS method, namely, the squeezing and enhancing technique, is applied assuming for the scaling parameter δ, the values 1, 0.86 and 0.6. The corresponding results are illustrated in the top, middle and bottom frames of Fig. 5.6, respectively. The adoption of $\delta = 1$ leaves the tail and the peak positions unchanged. However, the peak intensity (in particular the intensity of the central peak) is not invariant. This lack of invariance of the peaks means that $\delta = 1$ is an acceptable scaling parameter for the skeleton of the pdf and its tail, but not for the peak's intensities. The adoption of $\delta = 0.86$, as suggested by the DE analysis, is satisfactory for both peak intensities and position. However, this scaling property is limited to the central part of the histogram. The middle frame of Fig. 5.6 shows that the choice of the scaling parameter indicated by DE has difficulty with the side portions of the histogram. In fact, the peak at $x = -200$ is annihilated by the adoption of the DAS method. Finally, last panel, the scaling parameter $\delta = 0.6$ turns out to be unsatisfactory everywhere.

In summary, the scaling parameter $\delta = 0.86$ emerging from DE analysis depicted in Fig. 5.4 does not reflect the true scaling of the computer-generated process. Fig 5.6 indicates that the scaling $\delta = 0.86$, afforded by DE, refers to the central part of the distribution, whereas the sides of the histogram are more satisfactorily represented by the scaling $\delta = 1$. This anomalous scaling is not the effect of correlated fluctuations, rather it is the consequence of the non-stationary effects produced by the slow component superposed on the GN with our computer-generated data.
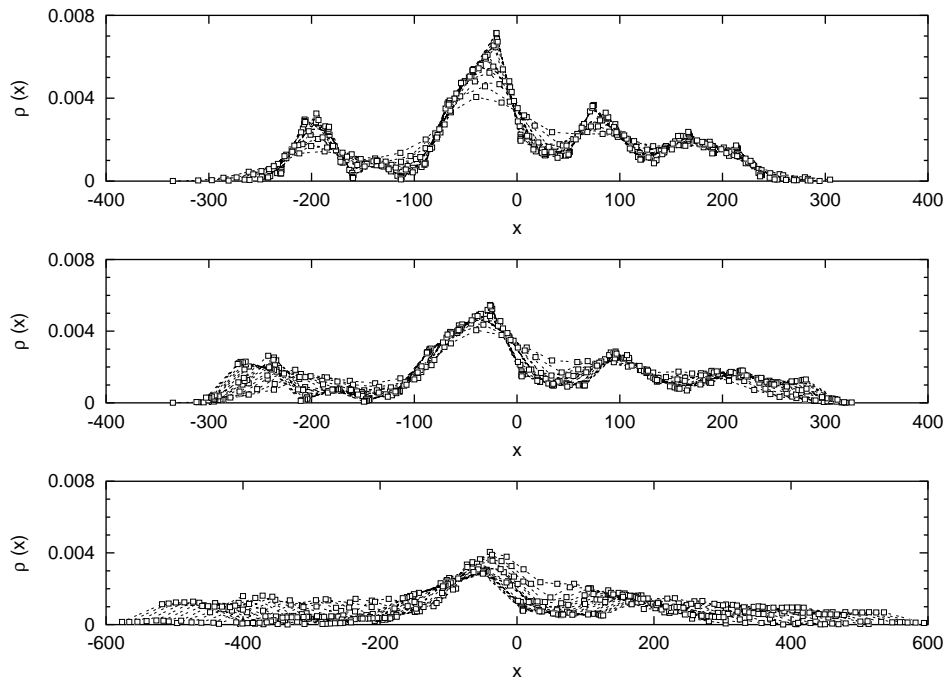
Figure 5.6: The DAS analysis of the time series given by the sum of the slow components SS_j^T plus the GN. The middle-time region, defined in Sec. 5.1.3 is considered. On the axis of the ordinates, $\rho(x)$, is the plot of the histograms produced by adopting different squeezing and enhancing transformations, described in Sec. 3.4 , to assess to what extent the various histograms coincide. The three frames refer, from top to bottom, to DAS analysis with the scaling parameter, $\delta = 1$, 0.86 and 0.6, respectively.

5.1.4 Detrending the Slow Component

The results establish that a slow component can produce misleading effects that have to be removed in order to detect the scaling properties of stationary fluctuations. In order to do this the slow component must be removed. In the following two distinct detrending methods, the "step smoothing" and "wavelet smoothing".

Step Smoothing

The "step smoothing" procedure consists of dividing the time series into non-overlapping patches of length equal to the characteristic time T: the average value inside each patch is evaluated and subtracted from the data. In other words, this procedure consists of approximating the slow component with a step function of the same kind as that of the top frame of Fig.5.1. The details of the procedure are as follows. Start from Eq. (5.1) and create the new variable X_j, the sum of the variable ζ inside the j-th patch, namely,

$$X_j = \sum_{k=jT}^{k<(j+1)T} \xi_k \quad j = 0, 1, 2, .., P, \tag{5.7}$$

where P is the number of patches of length T in the time series. To make explicit the contribution to X_j of the two components (the slow and the random) of the variable ξ_j one can define

$$R_j = \sum_{k=jT}^{k<(j+1)T} \zeta_k \quad j = 0, 1, 2, ..., P \tag{5.8}$$

$$A_j = \sum_{k=jT}^{k<(j+1)T} S_k^T \quad j = 0, 1, 2, ..., P \tag{5.9}$$

and therefore

$$X_j = A_j + R_j. \tag{5.10}$$

If it happens that in any single patch the second term of the right-hand side of Eq. (5.10) is negligible compared to the first term, this equation can be used to evaluate the average of the slow component and consider it as a fair approximation to the slow component inside each patch. Moreover, the noise component has zero mean. Thus, the most probable value for R_j is zero, and the error associated with this prediction is given by the standard deviation σ_{R_j}, therefore $R_j = 0 \pm \sigma_{R_j}$. Then, using the definition given by Eq.(5.8), the following approximation is applied

$$\sigma_{R_j} \approx \sigma_0 T^{\delta_{sm}}, \tag{5.11}$$

where σ_0 is the standard deviation of the variable ζ and the exponent δ_{sm} is a number between 0 and 1. If the scaling condition (Eq. (3.24)) applies, this exponent is the scaling coefficient, that is, $\delta_{sm} = \delta$. If the scaling condition does not apply δ_{sm} can represent the scaling parameter of the standard deviation. In the case where not even the standard deviation rescales, this exponent represents its mean slope in a log-log plot versus the time. Finally, with the help of Eq. (5.11) one can state that $X_j \approx A_j$ when

$$\sigma_0 T^{\delta_{sm}} \ll |a_j| T \Rightarrow \sigma_0 \ll |a_j| T^{1-\delta_{sm}}, \tag{5.12}$$

where a_j is the average amplitude of the slow component in the j-th patch, and, with this equation holding true,

$$a_j \approx \frac{X_j}{T}. \tag{5.13}$$

Fig. 5.7 shows the results of this detrending procedure for the three time series with the three slow components of Fig. 5.1, with characteristic time $T = 365$ and with GN as the noise component. The slow components SS_j^T and SC_j^T (top and middle frame of Fig. 5.7) are fairly well reproduced, while the result for the slow component SL_j^T (bottom frame of Fig. 5.7) is poor. The reason for this behavior is that the first two cases satisfy the condition of Eq. (5.12) while the last one does not. In fact, in the last case the absolute value of the average, $|a_j|$, relative to the j-th patch is considerably smaller than σ_0, the intensity of the GN.
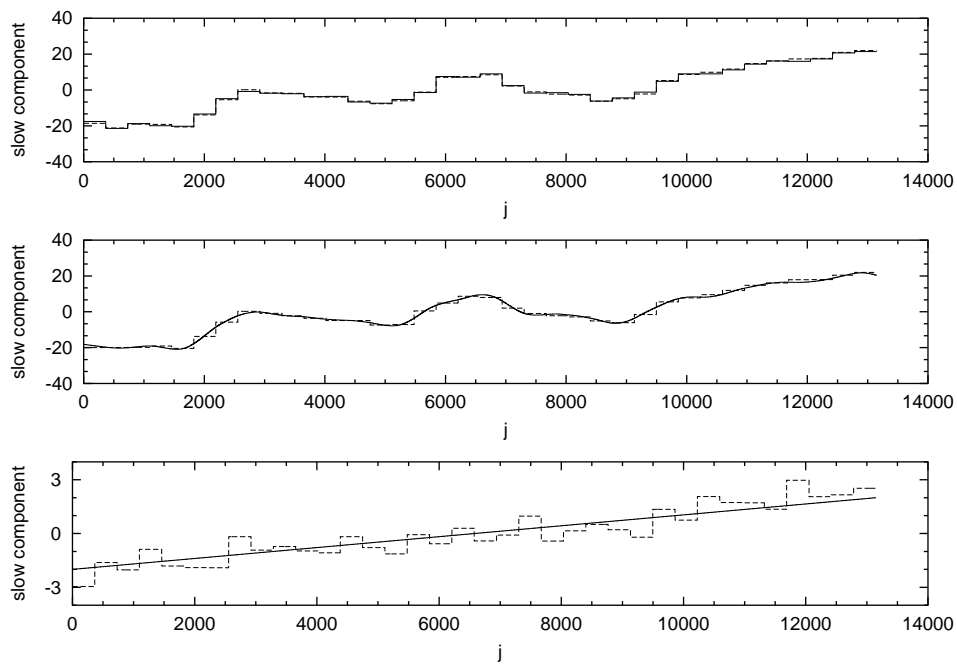
Figure 5.7: The step smoothing technique of Section 5.1.4 at work. The dashed line indicates the results of the step smoothing procedure, and with the full line the slow component to derive. From the top to the bottom frame is represented the case where the slow component are SS_j^T, SC_j^T and ST_j^T, respectively .

Wavelet Smoothing

As suggested by the name, the "wavelet smoothing" is based on the adoption of the wavelet transformation [58, 59]. Therefore, let us discuss briefly the detail of this transformation before explaining the detrending procedure itself. The wavelet transformation is close in spirit to the Fourier transformation, but has a significant advantage. The Fourier transformation decomposes a time series into a superposition of oscillating modes, each of which lasts forever. The wavelet transformation decomposes the time series into "notes" or wavelets [59], localized in time and in frequency. Formally, the wavelet transform $\tilde{f}(s,t)$ of the function $f(u)$ is defined as follows

$$\tilde{f}(s,t) \equiv \int_{-\infty}^{+\infty} du \, |s|^{-p}\psi^*(\frac{u-t}{s})f(u), \qquad (5.14)$$

where the symbol ψ^* denotes the complex conjugate of the wavelet $\psi(u)$. The wavelet $\psi(u)$ is a filter function satisfying the particular condition, known as the "admissibility condition" [58, 59]. The parameters s, t and p, are real numbers. Equation(5.14) shows the advantages of the wavelet transformation in resolving local features of the time series. In fact, the wavelet transformation rests on a convolution of the signal with the wavelet rescaled, through the use of the parameter s, and centered on $u = t$. Therefore the parameter s localizes the frequency domain and the parameter t localizes the time domain.

In this case, the discrete version of the wavelet transformation, known as the Maximum Overlap Discrete Wavelet Transform (MODWT) is used. The MODWT gives birth to a decomposition of the signal in terms of "approximations" and "details" relative to different time scales. Consider the integers N and K. The former is the length of the series under study and K is the greatest number satisfying the condition $2^K < N$. In this condition one can consider K different time scales, given by $2^1, 2^2, \ldots, 2^K$. Then, starting from the smallest time scale, the wavelet transformation is used to divide the original time series S into two components

$$S = A_1 + D_1. \qquad (5.15)$$

The component A_1 contains features having a characteristic time scale greater than 2^1, since a local average of the time series over all time scales inferior to 2^1 has been performed. Therefore the component A_1, often referred as "approximation", is a "2^1-smooth" function, in the sense that A_1 can be considered a slow changing function with respect the time scale 2^1. The component D_1, on the other hand, contains all the features of the time series with a characteristic time scale less than 2^1 and therefore is called the "detail". The partitioning procedure described above can then applied to the approximation A_1, splitting it into the functions A_2 and D_2. The latter two functions represent, respectively, the features of the time series with time scales greater than 2^2 and those features with a time scale less than 2^2, but greater than 2^1, that have already been removed. Clearly, at the k-th step of this procedure, with $k < K$, the decomposition reads

$$S = A_k + D_k + D_{k-1} + \ldots + D_1, \tag{5.16}$$

with A_k representing the smooth time series referring to the time scale 2^k and D_j, with $1 < j < k$, the detail of time series with the time scale located in the interval $[2^{k-1}, 2^k]$.

This decomposition of a time series through the wavelet transform, can be used in order to detrend the slow component of Eq. (5.1). In fact, If T is the characteristic time of the slow component, a good approximation is given by the j-th wavelet approximation (the one relative to scale 2^j), where j is such that 2^j is as close as possible to the time T ($T = 365$, so both $j = 8$ and $j = 9$ are good choices). This is so, because the wavelet transformation, as explained above, acts as a filter on the contributions corresponding to time scales smaller than the time scale examined. Therefore the detrending procedure, referred by us as the wavelet smoothing, rests on eliminating the component corresponding to the T-time scale from the data.

The wavelet smoothing is, now, applied to the three time series that have been analyzed, in Fig. 5.7, by means of the step detrending procedure. It is made the choice of using the Daubechies wavelet number 1 and 8, denoted by us as db1 and db8,

respectively[1] and the scale $2^9 = 512$ is adopted. The results of the wavelet smoothing are shown in Fig. 5.8 for the wavelet db1, and in Fig. 5.9 for the wavelet db8. In both cases, the results are very similar to what obtained with the step procedure (the two methods are qualitatively equivalent). This means that the wavelet method very satisfactorily reproduces both the SS_j^T and SC_j^T (top and middle frame) components, but it turns out to be as inaccurate as the step detrending method, when applied to the SL_j^T component (bottom frame). Finally, the slow components obtained with this method partially retain the feature of the particular kind of wavelet adopted (the wavelet db1 is a square wave and this feature is clearly reproduced in the slow components of Fig. 5.8) . This is a consequence of the convolution in the integral of Eq. (5.14).

The Effects of the Detrending Procedure

It is, now, possible to check if after applying the detrending procedures, above described, one can successfully recover the statistical properties of the noise component of Eq. (5.1) (This is a crucial step of our approach to the search of the complexity of a given process). In the case of the computer-generated time series analyzed in this section, one expect, after the detrending procedure, to recover the scaling $\delta = 0.5$ of the diffusion generated by the GN. For brevity, the only results shown are those relative to the DE analysis on the detrended sequences obtained with the step smoothing, Fig. 5.10, and wavelet smoothing, Fig. 5.11 (wavelet db1) and Fig. 5.12 (wavelet db8) procedure, for all the three kind of slow components used so far, SS_j^T, SC_j^T and SL_j^T. In all three cases the detrending procedure works very well, if the saturation taking place at long times is ignored. The saturation effect is a consequence of the detrending procedures adopted and has the, unwanted, property of limiting our ability of monitoring, for an extended period of time, the diffusion process generated by the noise component of Eq. (5.1) alone.

Let us give an explanation of this effects for both the detrending procedures

---

[1]In literature many kind of functions as been used as wavelet $\psi(u)$ in Eq. (5.14), according to the needs of the authors. As far as the wavelet smoothing is involved, there is no particular reason to chose a particular kind of wavelet, since the results are qualitatively the same.
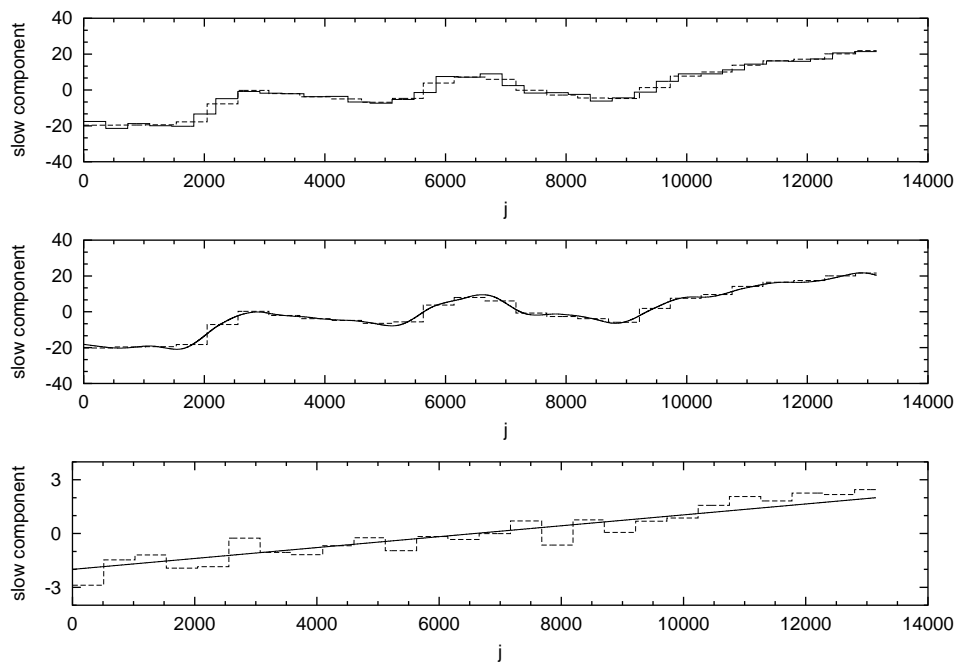
Figure 5.8: The wavelet smoothing technique of Section 5.1.4 with the wavelet db1, at work. The dashed line indicates the results of the wavelet smoothing procedure, and the full line the slow component to derive. From top to bottom, the frames refer to the case where the slow component are SS_j^T, SC_j^T and ST_j^T, respectively.
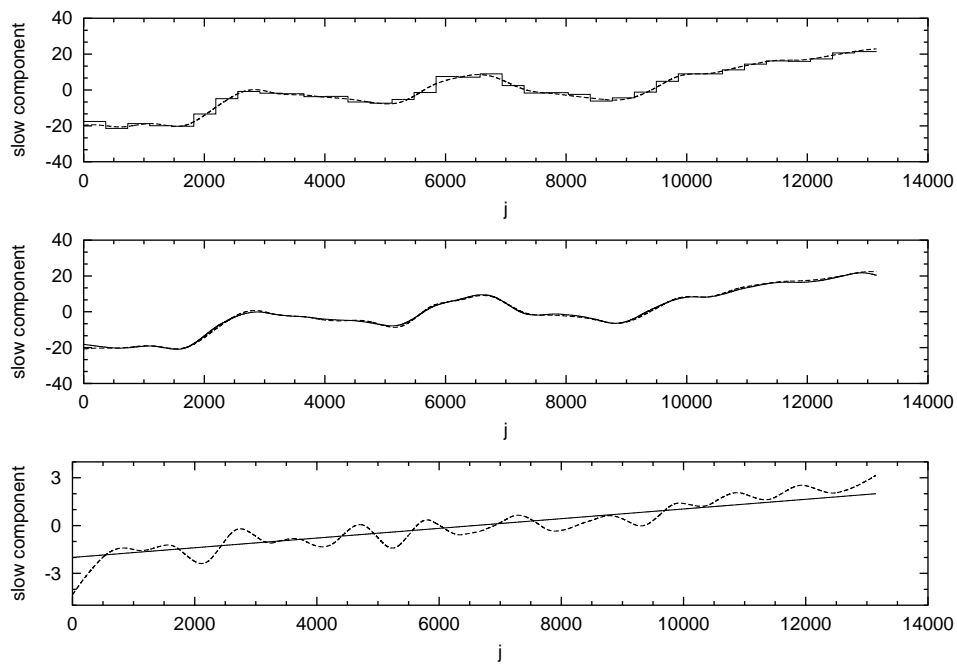
Figure 5.9: The wavelet smoothing technique of Section 5.1.4 with the wavelet db8, at work. The dashed line indicates the results of the wavelet smoothing procedure, and the full line the slow component to derive. From top to bottom, the frames refer to the case where the slow component are SS_j^T, SC_j^T and ST_j^T, respectively.

adopted. Starting with the step smoothing, it is possible to show that with the help of Eq. (5.13), the detrended time series (variable $\tilde{\xi}$) is obtained from the original time series according to the following prescription

$$\tilde{\xi}_k = \xi_k - \sum_j \left(\Pi_k^j \frac{X_j}{T} \right), \tag{5.17}$$

where Π_k^j is the characteristic function of the j-th patch. Using Eqs. (5.1) and (5.10), it is possible to write

$$\tilde{\xi}_k = \zeta_k + S_k^T - \sum_j \left(\Pi_k^j \frac{A_j}{T} \right) - \sum_j \left(\Pi_k^j \frac{R_j}{T} \right). \tag{5.18}$$

The reader must recall, that different diffusing trajectories are created using the method of overlapping windows. According to this method, the position occupied at time t by the m-th trajectory, denoted by $\Gamma_{\tilde{\xi}}(m, t)$, is given by

$$\Gamma_{\tilde{\xi}}(m, t) = \sum_{k=m}^{k=m+t-1} \tilde{\xi}_k. \tag{5.19}$$

The right-hand side of Eq. (5.18) is the sum of four contributions, and, correspondingly, the right-hand side of Eq. (5.19) can be expressed as the sum of four terms. These terms are, with, obvious notations, $\Gamma_\zeta(m, t)$, $\Gamma_{S^T}(m, t)$, $\Gamma_A(m, t)$ and $\Gamma_R(m, t)$. Therefore one can write, for $\Gamma_{\tilde{\xi}}(m, t)$, the following expression

$$\Gamma_{\tilde{\xi}}(m, t) = \Gamma_\zeta(m, t) + \Gamma_{S^T}(m, t) - \Gamma_A(m, t) - \Gamma_R(m, t). \tag{5.20}$$

At this point it is possible to explain the reason for the saturation effect. In fact, when the index m denotes the beginning of a patch and $t = T$, the function $\Gamma_{\tilde{\xi}}(m, t)$, being the sum of $\tilde{\xi}$ within the patch, vanishes (see the definition of R_j and A_j in Eq. (5.8)). For values of m denoting a position different from the first site of a patch, the quantity $\Gamma_{\tilde{\xi}}(m, t)$ can assume non-vanishing values. However, the above mentioned constraint establishes an upper bound on $\Gamma_{\tilde{\xi}}(m, t)$, thereby reducing the spreading of

diffusion process. Fig. 5.10 shows that the reduction of the diffusion process is already significant at times smaller than T. Consequently, the step detrending procedure successfully detrends for only a limited range of times, which can be estimated to be $t < \frac{T}{3}$.
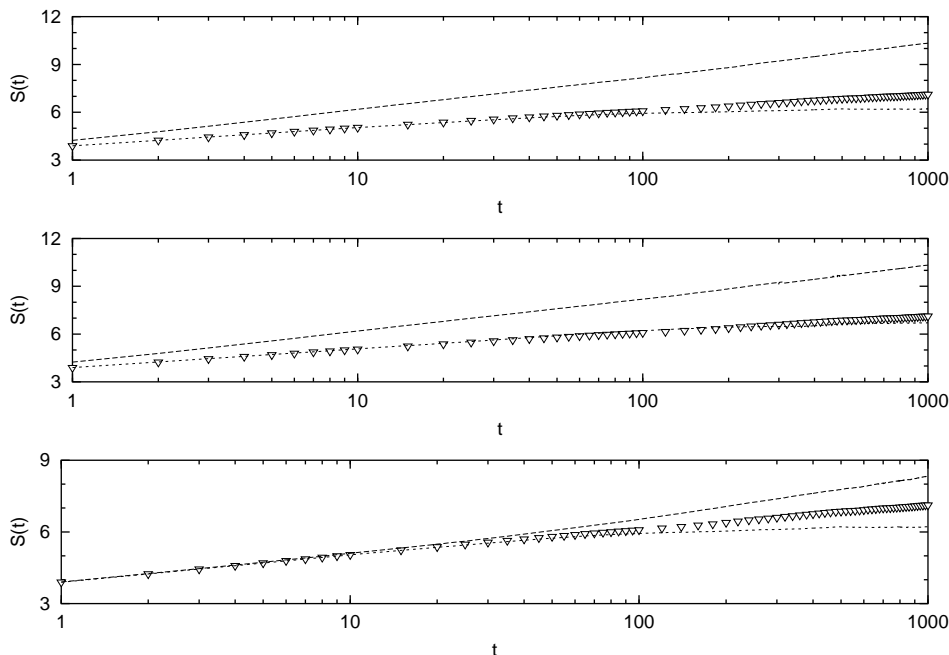

Figure 5.10: The diffusion entropy, $S(t)$, as a function of time t, in a logarithmic time scale. The triangles denote the results of the DE analysis applied to the series produced only by the noise component. The dashed line refers to the DE applied to the sum of noise and slow component. The dotted line denotes the results of the detrending procedure. From top to bottom the three frames refer SS_j^T, SC_j^T and SL_j^T, respectively. The detrending method used is the step smoothing procedure

The wavelet method yields saturation effects for similar reasons. In fact, using the wavelet method the j-th approximation $(2^j \gtrsim T)$ is detrended, and this approximation is obtained through a filtering process averaging all the components with a time scale smaller than 2^j. Therefore, this procedure is similar to the step smoothing, and the
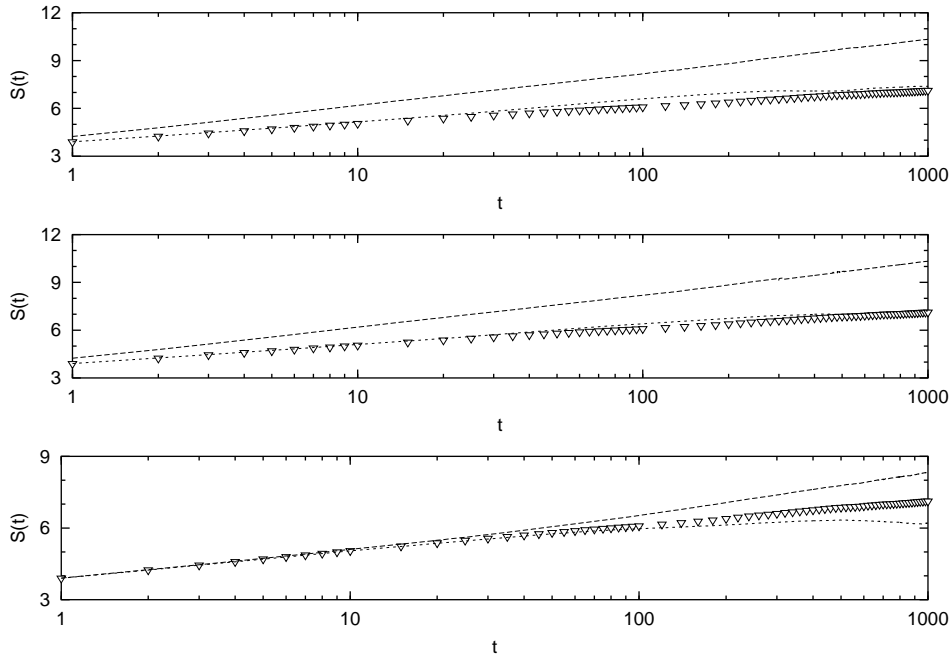
Figure 5.11: The diffusion entropy, $S(t)$, as a function of time t, in a logarithmic time scale. The triangles denote the results of the DE analysis applied to the series produced only by the noise component. The dashed line refers to the DE applied to the sum of noise and slow component. The dotted line denotes the results of the detrending procedure. From top to bottom the three frames refer SS_j^T, SC_j^T and SL_j^T, respectively. The detrending method used rests on the wavelet db1.

Figure 5.12: The diffusion entropy, $S(t)$, as a function of time t, in a logarithmic time scale. The triangles denote the results of the DE analysis applied to the series produced only by the noise component. The dashed line refers to the DE applied to the sum of noise and slow component. The dotted line denotes the results of the detrending procedure. From top to bottom the three frames refer SS_j^T, SC_j^T and SL_j^T, respectively. The detrending method used rests on the wavelet db8.
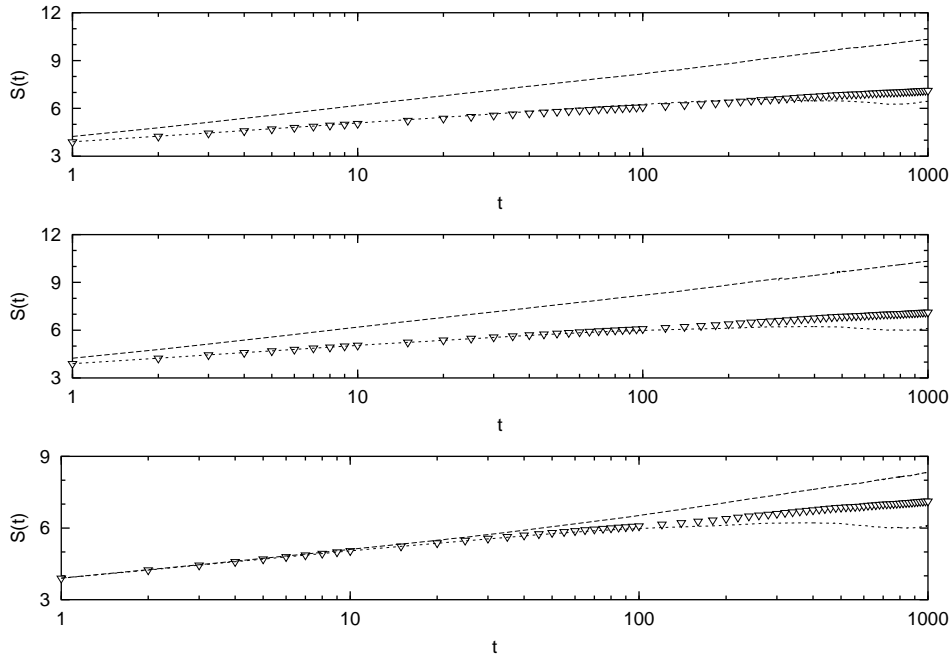
ensuing saturation effects have the same origin.

Finally, it is noteworthy to point out that the step smoothing is slightly more effective than the wavelet db8 smoothing in the case of the step component SS_j^T, while in the case of the continuous component SC_j^T it is the other way around. In fact, the nature of the two signals SS_j^T and SC_j^T is such that the step smoothing is naturally the best "fit" for SS_j^T and the db8 wavelet smoothing is naturally the best one for SC_j^T. Therefore, in the case of the step smoothing acting on SC_j^T, the detrending process is not so successfull and its effect occur before saturation, becoming thus visible. The same argument applies to the db8 wavelet smoothing acting on SS_j^T, thereby explaining why the db1 wavelet smoothing does not produce excellent results when applied to both SS_j^T and SC_j^T (top and middle frame of Fig. 5.12). Finally, in the case of the SL_j^T all the detrending procedures do not work effectively, given the fact that in this case the noise intensity is much greater than the slow-component intensity.

5.2 Computer-Generated Data: Effects of a Periodic Component

The task of this section is similar to the of the previous one, but, now, a periodic bias is considered. How this bias affects the detection of the complexity of the unbiased fluctuations has been discussed extensively in [3, 60], here the focus will be in how to detrend the periodic (not present [3, 60]) bias. Following the scheme adopted in the previous section, a time series of the kind is considered

$$\xi_j = \Phi_j^T + \zeta_j, \tag{5.21}$$

where Φ_j^T is a periodic function, of period T, denoted as the periodic component, satisfying the relation

$$\sum_{k=j}^{k<(j+1)T} \Phi_k = 0 \quad \forall j. \tag{5.22}$$

and ζ a stationary stochastic process. For the numerical simulation ζ is selected to be a GN, while Φ_j is chosen to be a sine wave (see Fig. 5.13 with a period $T = 365$:

to simulate, for example, an annual periodicity as that found in the teen birth data of Sec. 5.3.

As for the bias studied in the previous section, the presence of the seasonal component has the effect of "masking" the scaling properties of the diffusion process stemming from the variable ζ, by producing additional spreading. However, due to the fact that the external bias is periodic, the additional spreading undergoes regression to the initial condition at times that are an integer multiple of the period T. This yields an alternating sequence of increasing and decreasing spreading phases. For this reason the diffusion effect caused by the periodic component has been called the "accordion" effect.

For a convenient explanation of the "accordion" effect, let us study the variable $\Gamma_\xi(j,t)$. As in the previous section, the quantity t is the length of the window (or diffusion time) and j is the initial position of the sliding window. In the present case $\Gamma_\xi(j,t)$ reads

$$\Gamma_\xi(j,t) = \sum_{k=j}^{k=j+t-1} \left(\Phi_k^T + \zeta_k \right) = \Gamma_{\Phi^T}(j,t) + \Gamma_\zeta(j,t). \tag{5.23}$$

By noticing that

$$t = nT + \tau, \tag{5.24}$$

with n being a given integer depending on t and τ, a real number between 0 and T, and by using (5.22), one obtains

$$\Gamma_\xi(j,t) = \Gamma_{\Phi^T}(j+nT, j+nT+\tau) + \Gamma_\zeta(j,t). \tag{5.25}$$

Now, with t and therefore τ fixed, as the index j runs along the sequence the function $\Gamma_{\Phi^T}(j+nT, j+nT+\tau)$ repeats itself every T steps and inside this time period the function assumes a maximum and a minimum value. The difference between these two extremes is a measure of the spreading due to the seasonal component. It is evident that if $\tau = 0$, that is, if t is a multiple of the period T, the "seasonal" spreading is zero. With τ increasing, the spreading increases, but then it has, eventually, to decrease since for $\tau = T$ the initial condition must be recovered. Using (5.23) or

(5.25) it is evident that if one wants to know the scaling property of the diffusion relative to the variable ζ alone, has to look at the diffusion relative to the variable ξ only at times t that are multiples of the period T, since for these times, the following identity is satisfied

$$\Gamma_\xi(j,t) = \Gamma_\zeta(j,t) \qquad (5.26)$$

throughout the whole sequence. So, in principle, there would be no reason to process the data in any way in order to retrieve the desired information on the variable ζ. One might, in fact, limit him/herself to studying the behavior of the DE or SM at times that are a multiple of T. However, when dealing with real data, the limitation on the number of data points available and the excessive magnitude of T limits the accuracy of the observation of times that are multiples of the time period. For example if the number of data is 13149 and $T = 365$ (the case study in the sequel), then the saturation effect, due to lack of statistics, affects the DE or SM already at times smaller than $2T$. Therefore if one wants insight into the properties of the diffusion process of the noise component, before any saturation takes place, the periodic component must be detrended from the time series.

Let us proceed with the detrending in this case. Consider a time series whose length is NT, N being an integer. In other words, the sequence is assumed to have a length which is a multiple of the time period T. Now, one defines

$$\Sigma_\xi(j) = \sum_{m=0}^{N-1} \xi_{j+mT} = N\Phi_j + \sum_{m=0}^{N-1} \zeta_{j+mT} \qquad (5.27)$$

for all the values of $j \in [0,T]$, or, in a more concise notation,

$$\Sigma_\xi(j) = N\Phi_j + \Sigma_\zeta(j) \quad j \in [0,T]. \qquad (5.28)$$

Eq. (5.28) can be used to evaluate the periodic component Φ_j if $\Sigma_\zeta(j) \ll N\Phi_j$. Using the same argument as that adopted earlier for the detrending of the slow component, it can be establish the condition for operating successfully the detrending procedure

is

$$\sigma_0 N^\delta \ll N\Phi_j^T \Leftrightarrow \sigma_0 \ll \Phi_j^T N^{1-\delta}, \tag{5.29}$$

that is equivalent to

$$\Phi_j^T \approx \frac{\Sigma_\xi(j)}{N}. \tag{5.30}$$

There is a significant difference with respect to the case of the slow changing bias addressed before. In this case the detrended data do not show any saturation effect. In fact

$$\tilde{\xi}_j = \zeta_j - \frac{1}{N}\sum_{m=0}^{N-1}\zeta_{j+mT}, \tag{5.31}$$

does not vanish if the sum of $\tilde{\xi}_j$ is carried out over a patch.

Fig. 5.13 illustrates the results of a test done with a periodic function and a computer-generated sequence having GN. The first frame of Fig. 5.13 shows the periodic component used and the one resulting from the detrending procedure described above, the agreement is good. The second frame depicts the diffusion entropy applied to the sequence with both noise and periodic component, to the sequence with only the noise component and to the sequence derived from the detrending procedure. The agreement between the DE applied to the original noise and the DE applied to the detrended sequence is impressive, and, as expected, there is no indication of a saturation effect.

5.3   The Teen Birth Data

This section is dedicated to show the approach illustrated in the earlier sections in action on the teen birth phenomenon. The daily number of births to teenager mothers in the whole state of Texas from 1964 to 1999 (the teen birth data) is analyzed . These data are illustrated in Fig. 5.14. A closer look to the data reveals immediately an annual periodicity. It has been pointed out [61] that this kind of periodicity is related to the natural annual cycle of a woman's fertility. Therefore the power spectrum of the teen birth data is evaluated. The results are illustrated in Fig.5.15. In addition
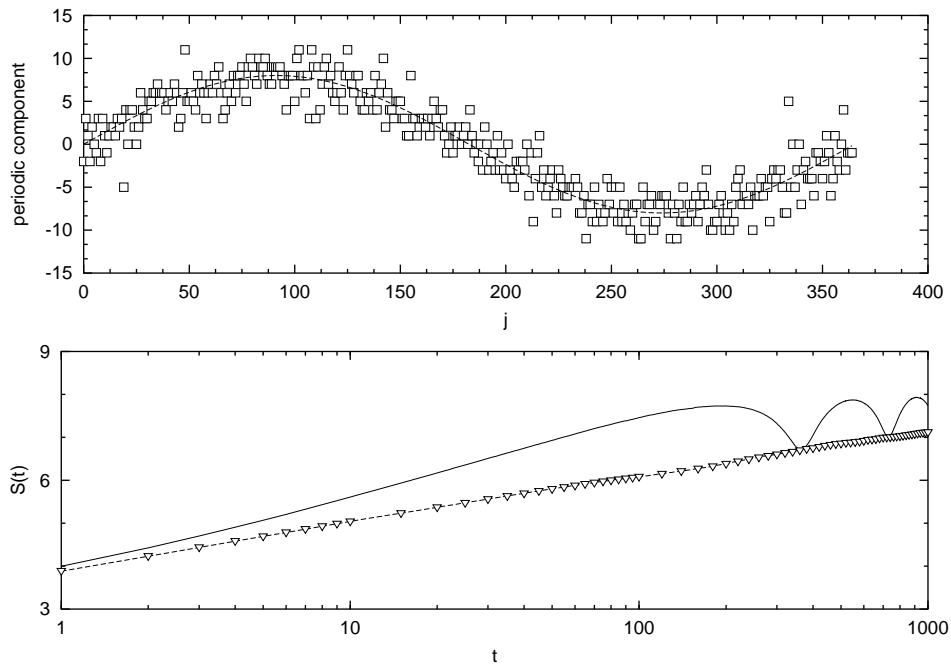
Figure 5.13: Top frame: The periodic component as a function of the linear time. The full lines denotes the real component and the squares the results of a procedure based on the use of Eq.(5.30). Bottom frame: diffusion entropy $S(t)$ as a function of time t in a logarithmic time scale. The triangles denotes the result of the DE analysis applied to the noise component alone. The full line illustrates the result of the DE analysis applied to the time series stemming from the sum of noise and periodic component. The dashed line refers to the the results of the DE applied to the signal resulting from the detrending procedure.
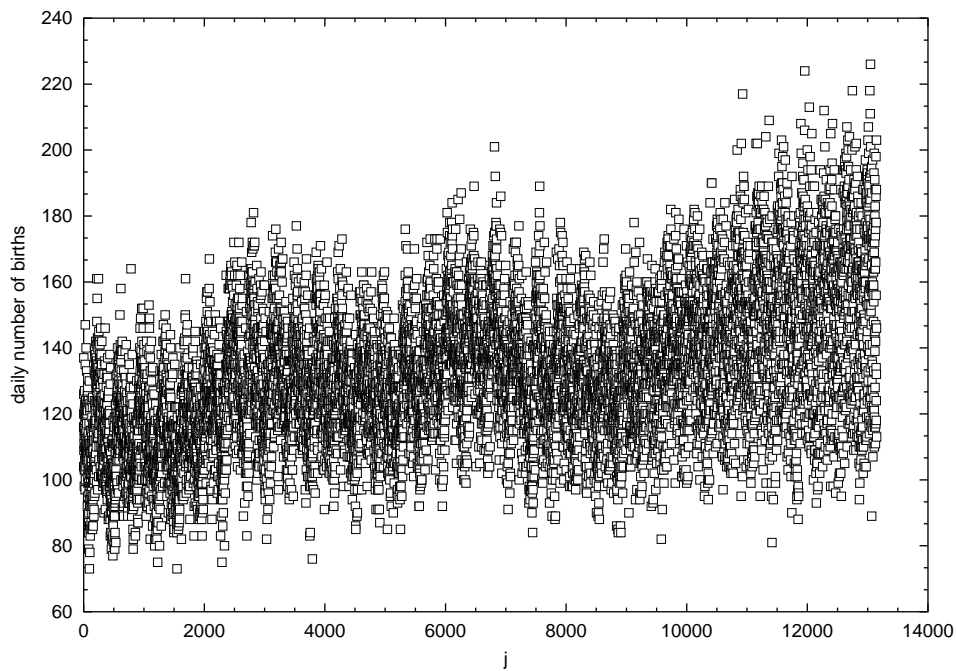
Figure 5.14: The daily number of births from a teenager mother in the whole state of Texas, from the 1st January 1964 ($j = 1$) to the 31 st December 1999 ($j = 13149$).
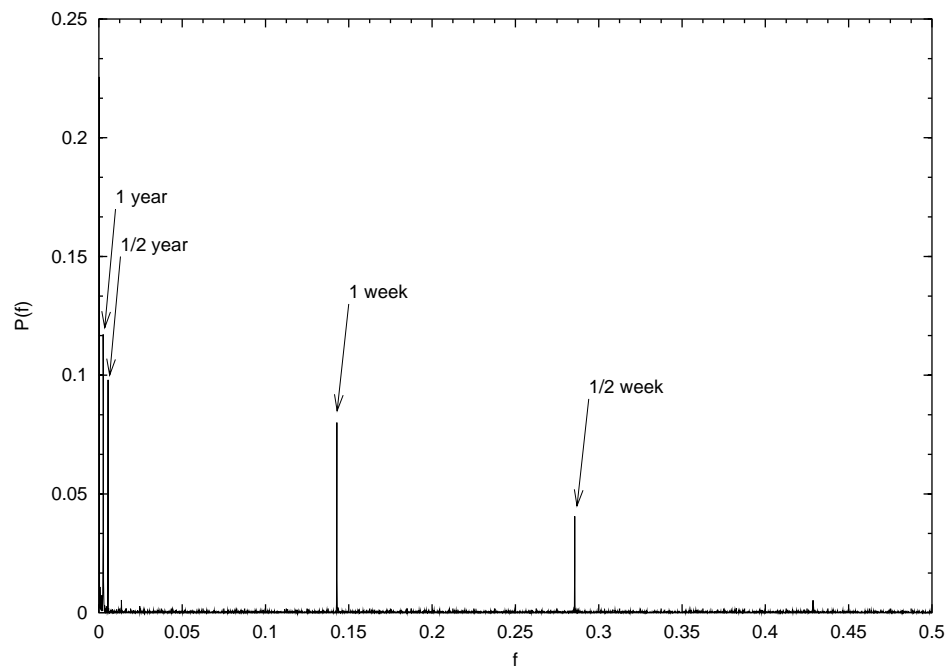
Figure 5.15: The power $P(f)$ as a function of the frequency f for the teen birth data of Fig. 5.14.

to the expected annual periodicity, there are also exist a six month, a one week and a half-week component. The weekly periodicity can be related to the practice of delaying or anticipating the birth in order to have less births on the weekend. It is important to point out that in addition to seasonal cycles one has to consider another important feature of the teen birth data. This is the fact that the annual average of births changes, sensibly, during the 36 year span considered. For example, starting from the year 1989, $j > 9000$ in Fig. 5.14, there is a steady increase. On the basis of these remarks, a reasonable attempt at describing the teen birth data is based on the assumption that the numbers of births per day $\{\xi_j\}$ are given by the following expression:

$$\xi_j = \Phi_j^{year} + \Phi_j^{week} + S_j + S + \zeta_j. \tag{5.32}$$

where S is the mean value of the data, Φ_j^{year} and Φ_j^{week} are respectively an annual and a weekly periodic component satisfying the condition of zero mean, S_j is a slow changin bias and ζ_j is an uncorrelated random variable with zero mean and fixed variance σ_ζ. The assumption that all the components of the right-hand side of Eq. (5.32) are independent of one another is made and, therefore, following Sec. 5.1, the function S_j is called the "slow component".

The main goal of the following sections, is to verify the validity of (5.32) as a model of the teen birth phenomenon and to detect the correlation properties of the fluctuation ζ.

5.3.1   Processing the Teen Birth Data: Detrending the Slow Component

In the previous sections of this chapter have been studied cases of time series where to a stationary noise is added, either a slow component (Eq. (5.1) of Sec. 5.1), either a periodic one (Eq. (5.21) of Sec. 5.2) separately , and it has been discussed how to remove these influences. In the teen birth data, however, presents a more complicated situation, since both biases are superposed on the random fluctuations in the number of births, and in addition there are two periodic components rather than one. Therefore some care must be used in applying the detrending procedures discussed in Sec. 5.1 and Sec. 5.2. Let us consider, first, the case of a time series

were, in addition to a slow component, a periodic is present. In this case the two detrending procedures do not commute: detrending the periodic component prior to the detrending of the slow component S_j, would lead to a distorted periodic bias, while the opposite is possible if the function S_j is constant for an interval of time equal, at least, to the time period of the periodic component. With two or more periodic components in the data, the function S_j has to be assumed constant for a time interval equal to the smallest common multiple of the time periods of the periodic components. The teen birth data show a yearly and a weekly periodicity added to the slow component S_j and considering that a year is equivalent to "almost" 52 weeks, one can take the length of a year as the smallest common multiple of the two periods of these two periodic components. Therefore the scale 2^9 is adopted for the wavelet detrending procedure (here, it is made use of the Daubechies wavelet number 8 (db8)) and the characteristic time $T = 365$ for the step detrending procedure. In this last case, using the fact that both Φ_j^{year} and Φ_j^{week} have a zero mean, one can write for the sum of the variable ξ inside the j-th patch of length 365 (366 in a the leap year)

$$X_j = \sum_{k=bj}^{k=bj+365} \xi_k \approx \sum_{k=bj}^{k=bj+365} \zeta_k + \sum_{k=bj}^{k=bj+365} S_k, \qquad (5.33)$$

where b_j is the index of the time series relative to the first January of the j-th (starting from the year 1964) year. Eq. (5.33) allows an approximation of the slow component S_j with a step function of step length 365 (366 in a the leap year). Fig. 5.16 shows both the slow components obtained with the wavelet smoothing and the step smoothing. It is easy see that these are the SS_j^T and SC_j^T of Fig . 5.1.

Before proceeding further, an important question to address is whether the component S_j can really be considered slow with respect the time scale $T = 365$, when the step smoothing procedure is adopted, or the time scale $T = 2^9 = 512$, when the wavelet smoothing procedure is adopted. S_j has to be considered fairly smooth because S_j represent a sort of daily average number of the births. This numbers depend on the daily total teen population, which, in turns, depends on the number of births occurring from 13 to 19 years earlier, and so on. Moreover one has to take
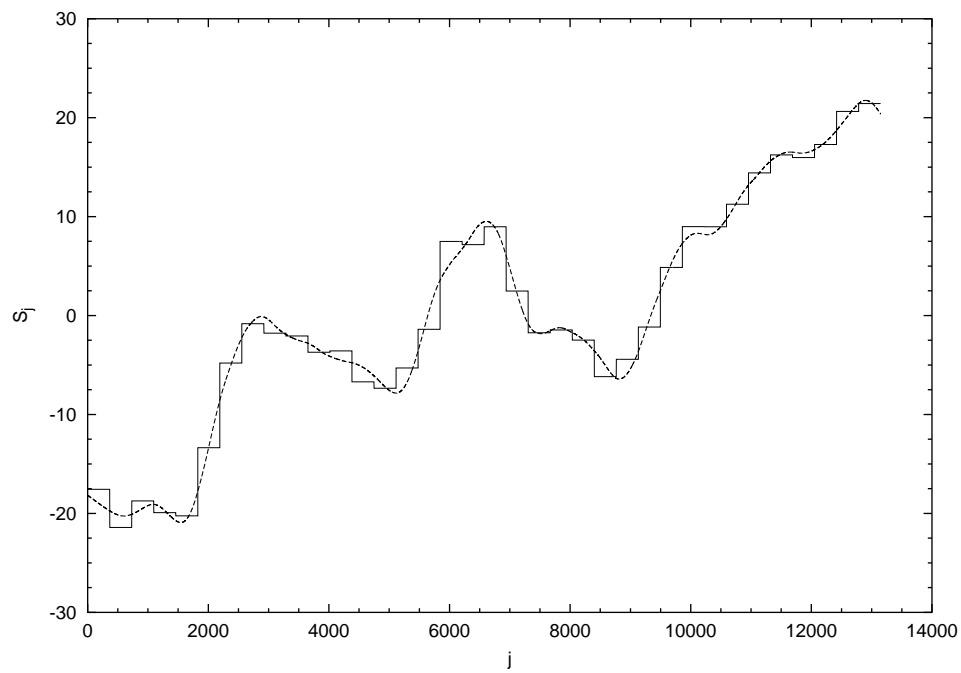
Figure 5.16: The slow component S_j as a function of time j. The full and dashed lines denote the result produced by the step smoothing of Section 5.1.4 and by the wavelet smoothing of Section 5.1.4, with wavelet db8, respectively.

into account social factors such as immigration or change of policy. therefore, it is plausible to assume the time scale relative to the interplay of all these factors to be of the order of one year.

To strengthen these argument a study the annual moving average is made. The annual moving average is defined by

$$\Gamma(j, 365) \equiv \frac{1}{365} \sum_{k=j}^{k=j+365-1} \xi_k \approx \frac{1}{365} \sum_{k=j}^{k=j+365-1} S_k + \frac{1}{365} \sum_{k=j}^{k=j+365-1} \zeta_k, \qquad (5.34)$$

where j goes from 1 to $(13149 - 365 + 1)$. Using the fact that ζ is a noise whose average intensity is null, one can write, following the arguments exposed in Sec. 5.1.4,

$$\Gamma(j, 365) = \frac{1}{365} \sum_{k=j}^{k=j+365-1} S_k. \qquad (5.35)$$

with the condition that

$$\sigma_j \times 365^{\delta_{sm}-1} \ll \frac{1}{365} \sum_{k=j}^{k=j+365-1} S_k. \qquad (5.36)$$

The expression Eq. (5.35) for the annual moving average $\Gamma(j, 365)$ affords interesting, though indirect, information on the behavior of function S_j. The top frame of Fig. 5.17 shows the function $\Gamma(j, 365)$. It is evident that $\Gamma(j, 365)$ is not dominated by the noise (it is a smooth function). Therefore the approximation leading to Eq. 5.35 is a good one. In the bottom frame of Fig. 5.17 the annual moving average of the data is compared with the corresponding quantities provided by the slow components determined with both the step and the wavelet smoothing. Both methods yield results in good agreement with the the annual moving average applied to the real data. Therefore, it is plausible to think that this slow component is a real property of the data that ought to be detrended in order to determine the genuine complexity of the time series.
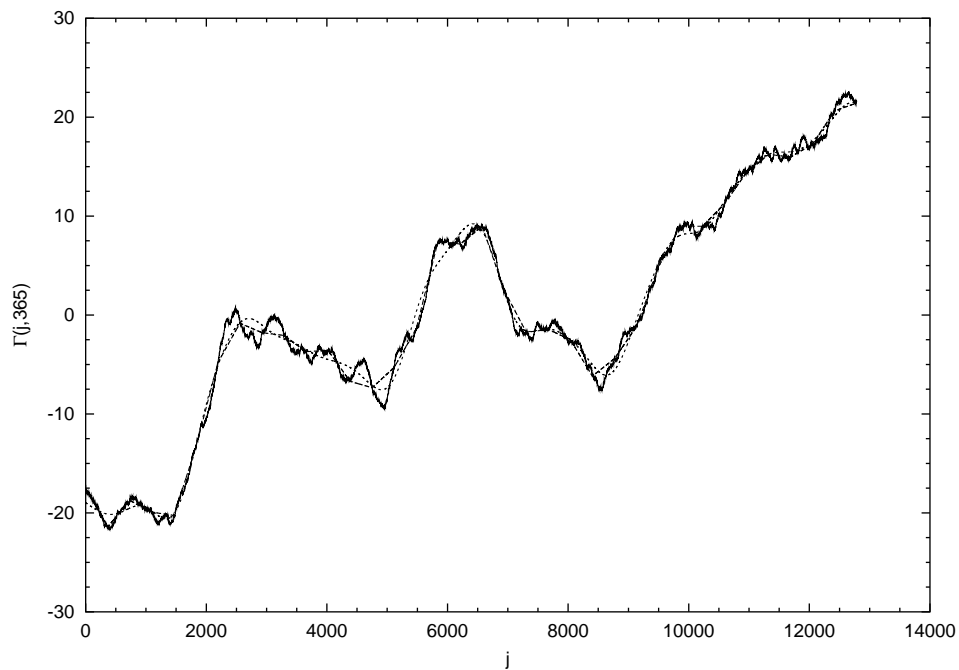
Figure 5.17: The annual moving average of Eq.(5.34). The full line indicate the result of the numerical analysis applied to the data, as they are. The dashed line denotes the the annual moving average applied to the slow component determined by the step smoothing method. The dotted line denotes the the annual moving average applied to the slow component determined by the wavelet smoothing.

5.3.2 Processing the Teen Birth Data: Detrending the Periodic Components

After detrending the slow component one can proceed to detrending the two periodic components. As in the case of the arguments yielding Eqs. (5.33) and (5.34), the fact that one year is almost a multiple of one week simplifies our efforts. In fact when Eq. (5.27) is applied to detrend the year periodicity, virtually no weekly contribution enters into play. Therefore the yearly is detrended first and the weekly component after. The top frame of Fig. 5.18 shows the evaluated yearly periodicity. It can noticed a sudden drop in the number of births in the correspondence of the 4th of July ($= 186$), the 1st of September ($j = 245$) and the 24th and 25th of December ($j = 359$ and $j = 360$). Moreover there is no appreciable difference between the results of the two different recipes, used to detrend the slow component, in both the top and the bottom frame. The bottom frame shows the power spectrum after detrending the yearly component, with only the weekly periodicity remaining.

Now, the weekly periodicity can be detrended. Fig. 5.19 illustrates the corresponding results. The top frame shows the evaluated weekly periodicity, with the number 1 representing Monday, 2 Tuesday, and so on. As expected, on the weekend there is a significant drop in the number of births. The bottom frame illustrates the power spectrum of the detrended signal. It is surprising that the power spectrum still shows a sign of a week component. The reason for this unexpected effect is that the result is based on the implicit assumption that the weekly component is the same throughout all the years: an incorrect assumption. Therefore it is made an attempt to evaluate the weekly component year by year, to see if, in so doing no sign of weekly periodicity on the power spectrum is left. Fig (5.20) shows that the conjecture of a weekly periodicity changing during the years is correct. In fact the bottom frame shows no signs of weekly periodicity. Note that the ordinate scale is smaller than the scale of the previous figures by a factor of 10. The top frame shows the intensity of the weekly component through the years. This intensity increases and this phenomenon becomes especially significant in the last ten years.
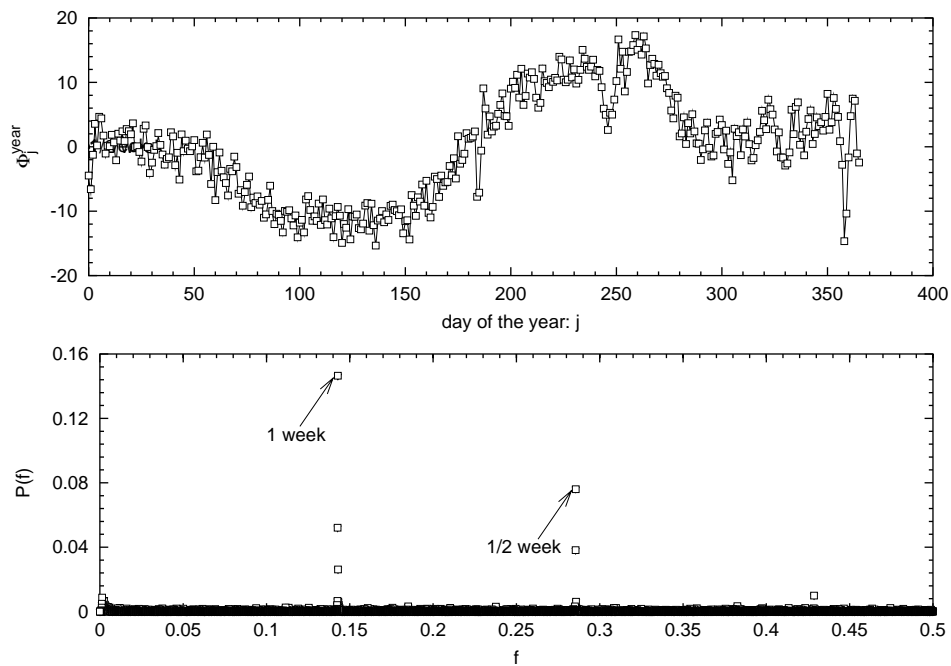
Figure 5.18: Top frame: the annual periodicity Φ_j^{year} of the teen birth data. The squares indicate the result of the wavelet detrending method. The full line denotes the result of the step smoothing method. Bottom frame: the power spectrum of the signal after detrending the yearly periodicity has been detrended. The squares and the triangles indicate the results of the wavelet detrending method and of the step smoothing method, respectively.
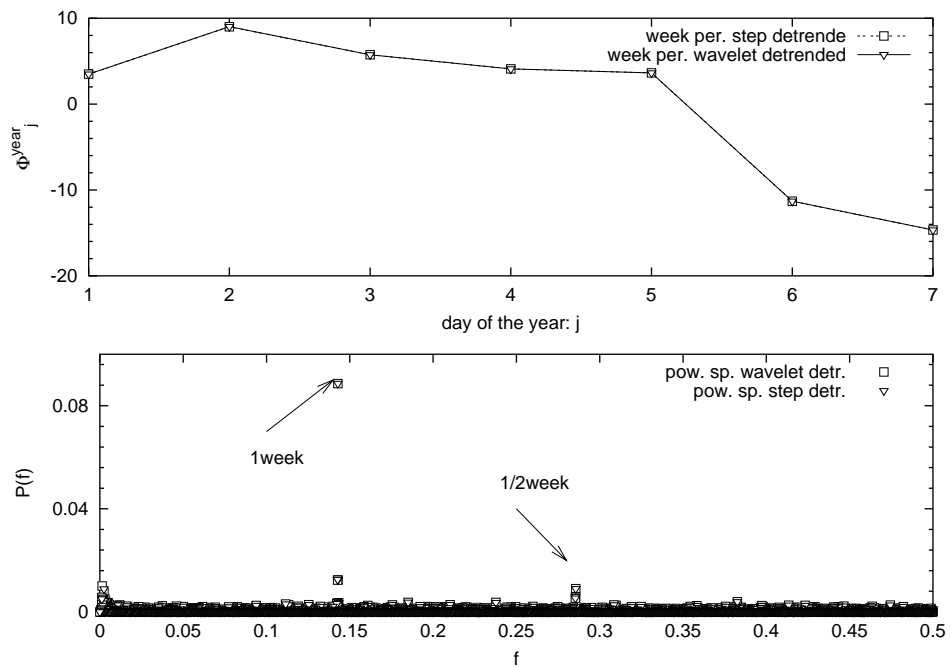
Figure 5.19: The week periodicity (top frame) Φ_j^{week} as a function of the days of the week, from Monday, 1 to Sunday 7. The power spectrum (bottom frame), after detrending the week periodicity. In both frames the squares and the triangle denote the result obtained using the step smoothing procedure and the wavelet method, respectively.
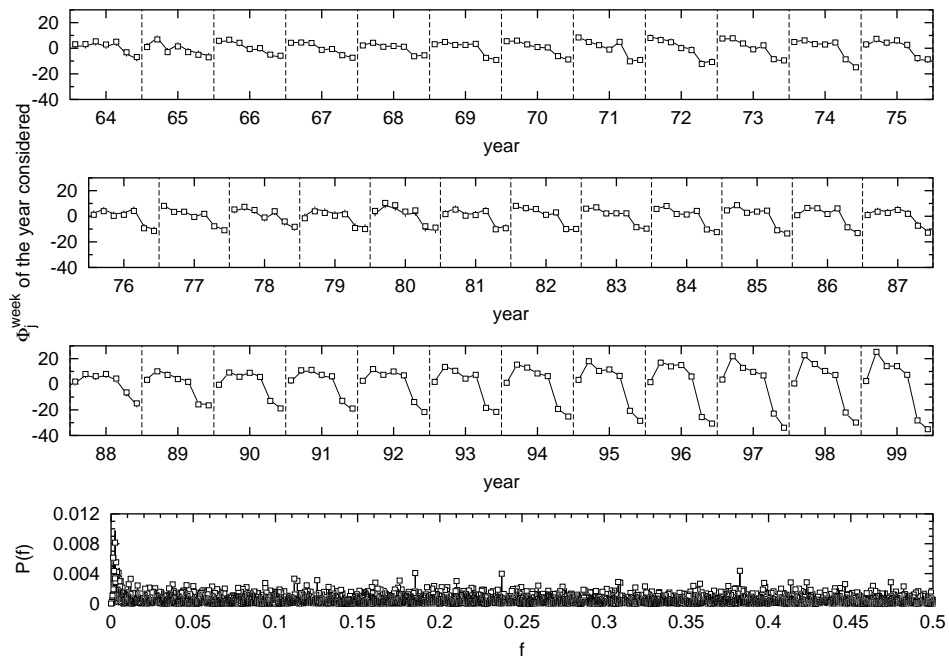
Figure 5.20: Effect of detrending weekly periodicity year by year. The three top frames illustrate how the weekly periodicity changes over the years, from 1964 to 1999. For each year seven values are reported corresponding to the seven days of the week, from Monday to Sunday. The squares and the full line denote the results stemming from the step detrending method and from the wavelet decomposition, respectively. The bottom frame is the power spectrum of the detrended data.

5.3.3 Processing the Teen Birth Data: The Fluctuations

Finally, the diffusion entropy analysis (DEA) is applied to the data with all possible forms of bias, seasonal and demographic, detrended. Fig.5.21 shows that the first ten days are characterized by $\delta_{de} = 0.58$ and that immediately before the saturation regime, caused by the detrending procedure, in the time region between 10 and 80, an even larger value of scaling index, $\delta_{de} = 0.67$, emerges. Do these parameters correspond to a proper realization of the scaling definition of Eq. (3.24)? To answer this question the DAS and the MS are applied to both time regions. Figs 5.22, 5.23 and 5.24 show the DAS in the time region of the first ten days, with $\delta_{de} = 0.5$, $\delta_{de} = 0.57$ and $\delta_{de} = 0.67$, respectively. The same time region is analyzed in Fig 5.25 by means of the MS method. These figures indicate clearly that $\delta_{de} = 0.57$ can be considered a genuine scaling parameter. The top frame of Fig. (5.26) shows the results of the DAS in the region where $\delta_{de} = 0.67$. The result of the DAS applied to the time series stemming from the step detrending procedure is virtually equivalent to that of the wavelet procedure. For simplicity in Fig(5.26) only the case of the wavelet detrending procedure is reported. The result is that δ_{de} does not correspond to a satisfactory realization of Eq. (3.24).

On the basis of these results, one would be tempted to conclude that the scaling $\delta = 0.57$ is genuine and $\delta = 0.67$ is not. However, to be as rigorous as possible, one has to discuss, first, the intriguing issue of the difference between *real* and *genuine* scaling. By real is meant a scaling corresponding to a realization of Eq. (3.24). By genuine is meant a scaling reflecting the cooperative properties of the process under study. It cannot be ruled out the possibility that the scaling $\delta = 0.67$ is real, but not genuine. In fact, since the length of the time series of the teen birth data is 13149, several artificial sequences are generated, with the same length, with the algorithmic prescription that Ref. [62] proposes to build up fractional Brownian motion, with $H = 0.67$. This corresponds to a real scaling with $\delta = 0.67$. Then the DAS is applied to the same time region where the real data yield $\delta_{de} = 0.67$. The results are reported in the bottom frame of Fig(5.26) and should be compared to top frame, illustrating the analysis of the real data. The results of the artificial sequences as "good" or as

"bad" as those of the real data. On the basis of this, one cannot dismiss the possibility that $\delta = 0.67$ is essentially real scaling.

Would the scaling also be genuine? Here there are two different eventualities. The first is that the genuine fluctuations remaining after detrending are a generalization of the dynamic model proposed years ago in Ref. [63]. This model, the Copying Mistake Map (CMM), assumes that the time series is generated by a composite of two mechanisms. The first, adopted with larger probability, is a prescription generating uncorrelated fluctuations, and the second, applying with a much smaller probability, is a prescription generating correlated fluctuations, and consequently a diffusion process faster than the uncorrelated component. In the large time scale regime the second component dominates the diffusion process, thereby producing a crossover from normal to anomalous scaling. If one replaces the random component with fluctuations characterized by anomalous scaling, weaker than the second component, a crossover from a scaling larger than the ordinary to an even larger scaling is expected. This might be the model behind the results illustrated in Fig. 5.21. In this case the scaling $\delta = 0.67$ would be genuine as well as real. However, one cannot rule out the possibility that the effect is not genuine. This effect might be due to the presence of a residual contribution of the slow component that would generate in the long-time regime a diffusion faster than that produced by the correlation stemming from the genuine complexity of the process under study, in the same way as the generalized CMM model would do. However, in this case the effect might be real or not, but it would not be genuine.

With all these caveats in mind, the following conclusion can be reached. The scaling $\delta = 0.57$ is both real and genuine, while a definite conclusion about the scaling $\delta = 0.67$ can not be reached.
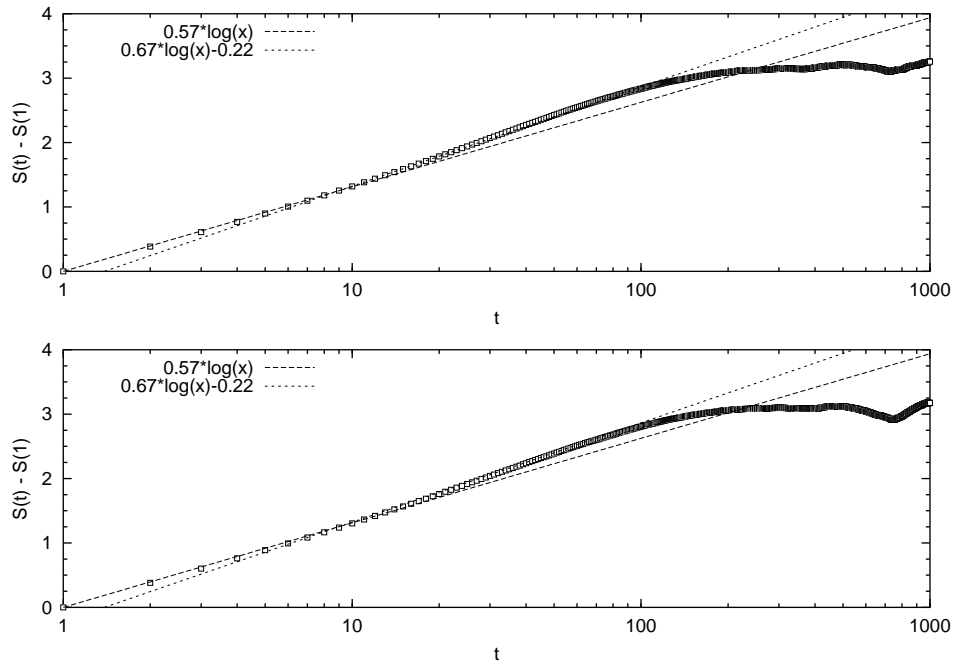
Figure 5.21: The DE of the detrended data. The squares denotes the results of the
DE analysis on the data detrended with the step smoothing (top frame), and wavelet
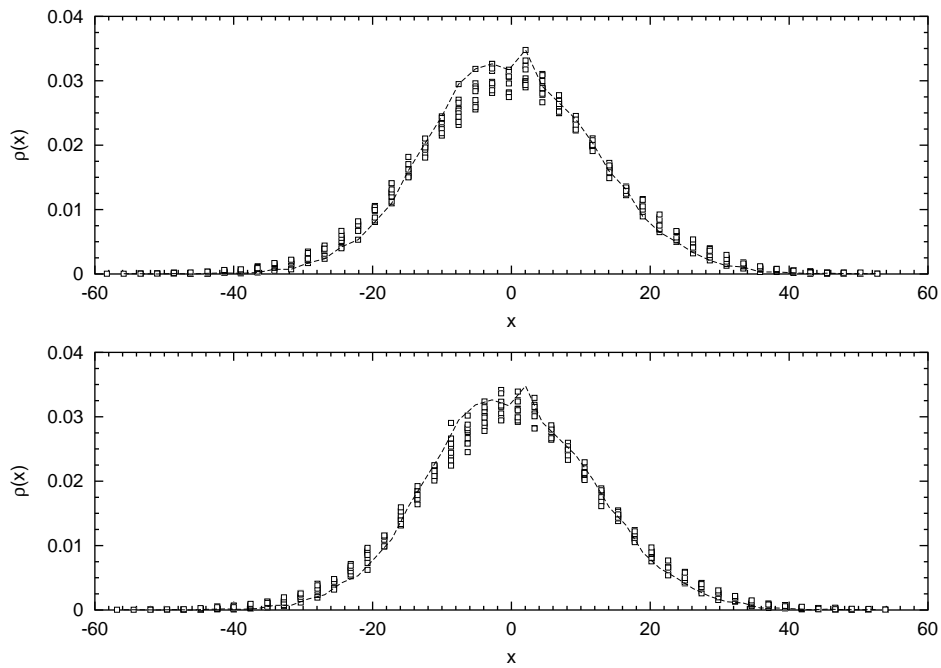smoothing (bottom frame).

Figure 5.22: The DAS at work in the time region from 1 to 10. The squares denote the pdf rescaled with $\delta = 0.5$ for the data detrended with the step smoothing (top frame) and with the wavelet smoothing (bottom frame), the dashed line correspond to the pdf at time $t = 1$. The overlap between squares and dashed line is not so good
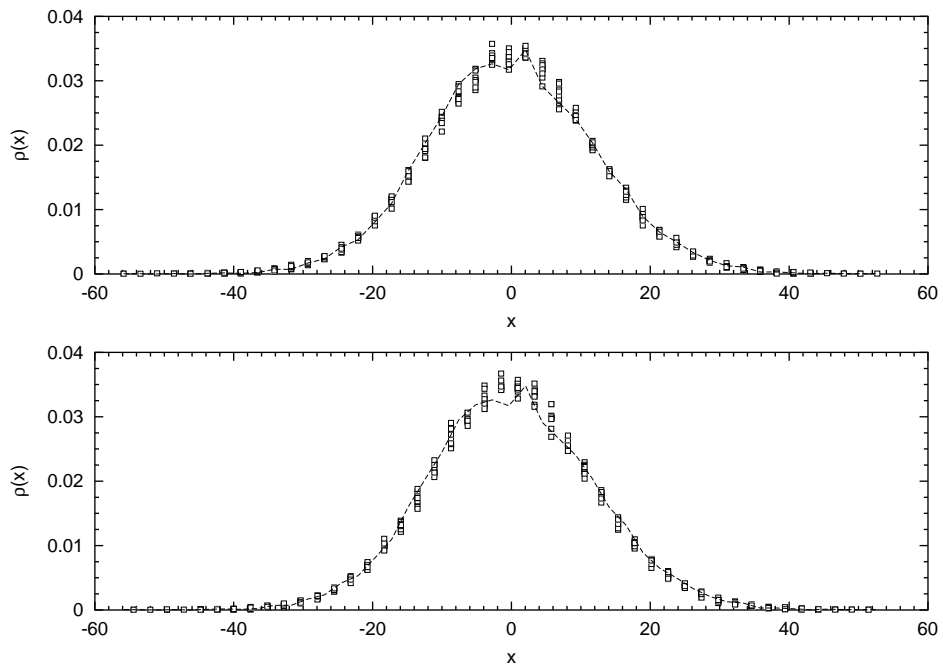
Figure 5.23: The DAS at work in the time region from 1 to 10. The squares denote the pdf rescaled with $\delta = 0.57$ for the data detrended with the step smoothing (top frame) and with the wavelet smoothing (bottom frame), the dashed line correspond to the pdf at time $t = 1$. The overlap between squares and dashed line is good
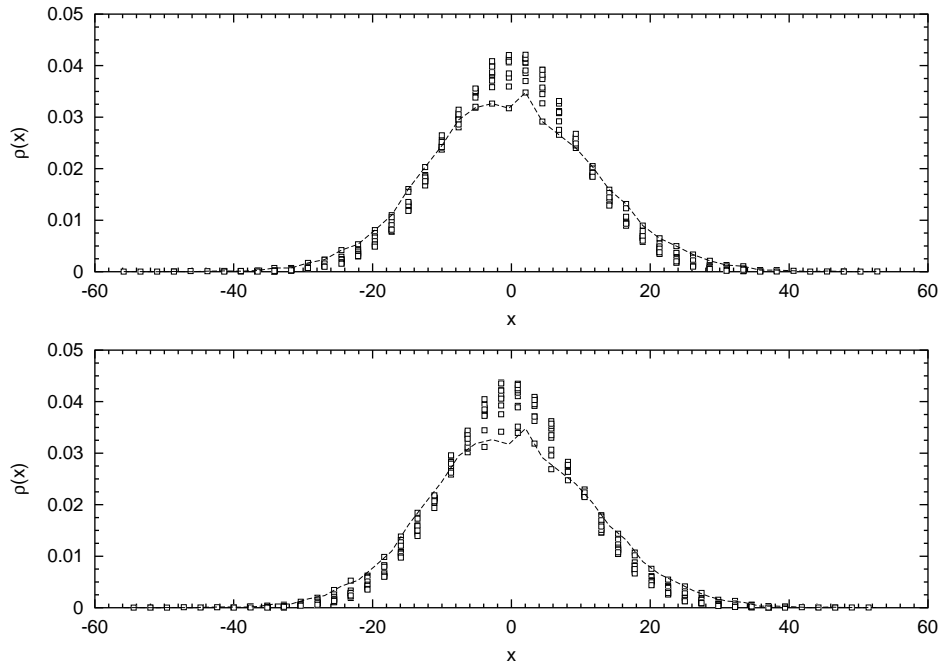
Figure 5.24: The DAS at work in the time region from 1 to 10. The squares denote
the pdf rescaled with $\delta = 0.67$ for the data detrended with the step smoothing (top
frame) and with the wavelet smoothing (bottom frame), the dashed line correspond
to the pdf at time $t = 1$. The overlap between squares and dashed line is bad in
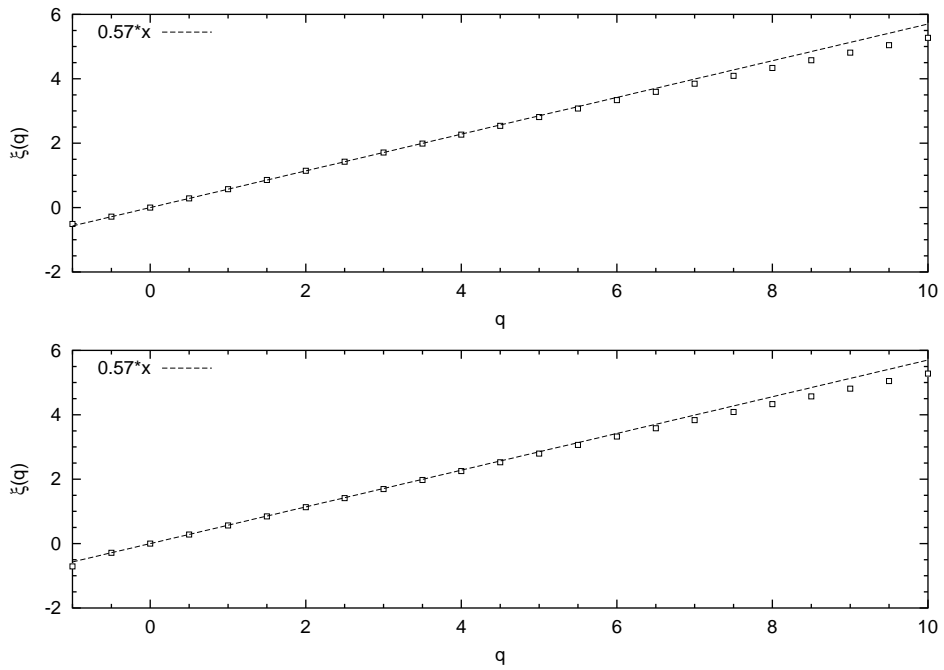particular in the central part of the pdf

Figure 5.25: The MS at work in the time region from 1 to 10. The squares denote the results for the data detrended with the step smoothing (top frame) and with the wavelet smoothing (bottom frame), the dashed line correspond to a straight line of slope of 0.57.
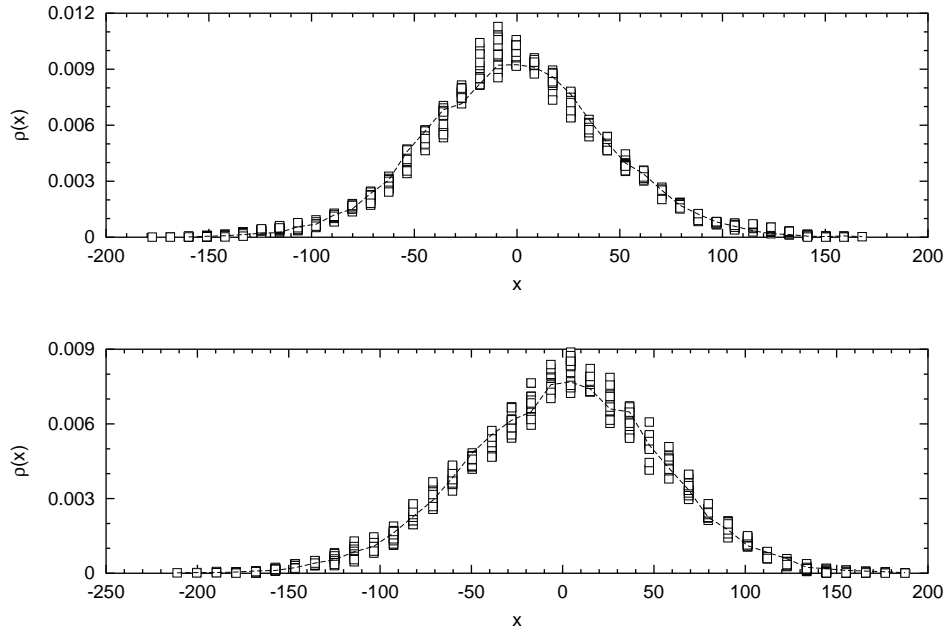
Figure 5.26: The DAS at work in the time region from 10 to 80. The squares denote the pdf rescaled with $\delta = 0.67$ for the data detrended with the wavelet smoothing (top frame) and for an artificial data corresponding to a FBM with $H = 0.67$t (bottom frame) the dashed line correspond to the pdf at time $t = 10$.

CHAPTER 6

Concluding Remarks

This thesis yields two important results. The first one is the scaling breakdown of the SVM diffusion process, created using the GCTIM intermittent dynamical model. This departure from the scaling condition is a signature of the interplay between order and randomness and as a "dramatic" consequence also for the equivalence between the density and trajectory approach for the description of dynamical systems.

The second result of this thesis has to do with the search of complexity in the case of a real process of sociological interest, the teen birth phenomenon. I found that the existence of either periodic or non-periodic biases can create the impression of a high complexity that actually does not exist. I have developed a new procedure to detrend biases that is competitive with the wavelet analysis and finally, measured the genuine complexity of the system. What about this resulting complexity and the complexity of living system introduced in the first part of this thesis? The analysis carried out by us seems to yield to the conclusion that this complexity is a form of FBM. Apparently, this discover seems to lead to a different model of complexity with respect to the earlier one. However I can not rule out the possibility that the model described in Chapter 4 is somewhat related to the dynamic derivation of the FBM, since this is still an open question.

BIBLIOGRAPHY

[1] M. Ignaccolo, P. Grigolini and A. Rosa, Phys. Rev. E **64**, 026210 (2001)

[2] P. Allegrini, J. Bellazzini, G. Bramanti, M. Ignaccolo, P. Grigolini, and J. Yang, Phys. Rev. E 66, 015101 (2002).

[3] P. Allegrini, V. Benci, P. Grigolini, P. Hamilton, M. Ignaccolo, G. Menconi, L. Palatella, G. Raffaelli, N. Scafetta, M. Virgilio, J. Yang, Chaos, Solitons & Fractals, **15**, 517 (2003)

[4] M.Ignaccolo, P. Allegrini, P. Grigolini, P. Hamilton and B. J. West, *Scaling in Non-stationary time series*, submitted to Physica A

[5] M.Ignaccolo, P. Allegrini, P. Grigolini, P. Hamilton and B. J. West, *Search for the complexity of the Teen Birth Phenomenon*, submitted to Physica A

[6] H. E. Hurst, Trans. Am. Soc. Civ. Eng. **116**, 770-808 (1951).

[7] R. Brown, Phil. Mag. **4**, 161 (1828).

[8] A. Fick, Ann. Phys. (Liepzig) **170**, 50 (1855).

[9] A. Einstein, Ann. der Physik **17**, 549 (1905).

[10] L. E. Reichl, *A Modern Course in Statistical Physics*, Wiley, New York (2nd edition) 1998.

[11] R. Metzler and J. Klafter, Phys. Rep. **339**, 1 (2000)

[12] H. Mori, Prog. Theo. Phys. **33**, 423 (1965).

[13] H. Mori, Prog. Theo. Phys. **34**, 399 (1965).

[14] R. Zwanzig, J. Chem. Phys. **33**, 1338 (1960).

[15] R. Zwanzig, *Lectures in Theoretical Physics (Boulder)*, p. 135 Vol. III Wiley, New York 1961.

[16] H. E. Hurst, R. P. Black and Y. M. Simaika, *Long-Term Storage:An Experimental Study*, Constable, London 1965.

[17] B. B. Mandelbrot and J. W. Van Ness, SIAM Rev. **10**, 422 (1968).

[18] P. Lévy, *Théory de l'Addition des Variables Aléatoires*, Gauthier-Villars, Paris 1937.

[19] B. V. Gnedenko, A.N. Kolmogorov, *Limit Distributions for Sum of Independent Random Variables*, Addison Wesley, Reading 1954.

[20] I. S. Gradstein, I. M. Ryzhik, *Table of integrals, series and products*, Academic Press, San Diego 1994, (5th edition) p. 943.

[21] H. Scher and E. W. Montroll, Phys. Rev. B **12**, 2455 (1975).

[22] M. F. Shlesinger, J. Klafter and Y. M. Wong, J. Stat. Phys. **27**, 499 (1982).

[23] G. H. Weiss and R. J. Rubin, Adv. Chem. Phys. **52**, 363 (1983)].

[24] S. Havlin and D. Ben Avraham, Adv. Phys. **36**, 695 (1987).

[25] J. W. Haus and K. W. Kehr, Phys. Rep. **150**, 263 (1987).

[26] J.-P. Bouchaud and A. Georges, Phys. Rep. **195**, 127 (1990).

[27] A. Blumen, J. Klafter, and G. Zumofen in *Optical Spectroscopy of Glasses*, edited by I. Zschokke, Reidel, Dordrecht, 1986, p. 199.

[28] M. F. Shlesinger and J. Klafter, Phys. Rev. Lett. **54**, 2551 (1985).

[29] M. F. Shlesinger and J. Klafter, in *Transport and Relaxation in Random Materials* edited by J. Klafter, R. J. Rubin, and M. F. Shlesinger, World Scientific, Singapore, 1986.

[30] J. Klafter, A. Blumen and M. F. Shlesinger, Phys. Rev. A **35**, 3081 (1987).

[31] G. H. Weiss, *Aspects and Applications of the Random Walk*, North-Holland, Amsterdam, 1994.

[32] P. Grigolini, L. Palatella and G. Raffaelli, Fractals **9**, 439 (2001).

[33] P. Allegrini, P. Grigolini, P. Hamilton, L. Palatella and G. Raffaelli, Phys. Rev. E **65**, 041926 (2002).

[34] J. Yang and P. Grigolini work in progress.

[35] P. Ch. Ivanon, L. A. Nunes Amaral, A. L. Goldberg, S. Havlin, M. G. Rosenblum, Z. R. Struzik and H. E. Stanley, Nature (London) **399**, 461 (1999).

[36] G. Zumofen and J. Klafter, Phys. Rev. E **47**, 851 (1993).

[37] M. F. Shlesinger, B. J. West and J. Klafter, Phys. Rev. Lett. **58**, 1100 (1987).

[38] M. F. Shlesinger and J. Klafter, J. Phys. Chem. **93**, 7023 (1989).

[39] P. Allegrini, M. Barbi, P. Grigolini and B. J. West Phys. Rev. E **52**, 5281 (1995).

[40] P. Allegrini, P. Grigolini and B. J. West, Phys. Lett. A **211**, 217 (1996).

[41] P. Allegrini, M. Buiatti, P. Grigolini and B. J. West, Phys. Rev. E **57**, 4558 (1998).

[42] C. Shannon and W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana 1948.

[43] C. Beck and F. Schlögl, *Thermodynamics of Chaotic Systems : An Introduction*, Cambridge Nonlinear Science, Cambridge 1993.

[44] J.R. Dorfman, *An Introduction to Chaos in Nonequilibrium Statistical Mechanics*, Cambridge University Press, Cambridge 1999.

[45] I. P. Cornfeld, S. V. Fomin, Ya. G. Sinai, *Ergodic Theory*, Springer, New York 1982.

[46] Ya. Pesin, Russ. Math. Surveys **32**, 55 (1977).

[47] D. Ruelle, Bol. Soc. Bras. Mat. **9**, 83 (1978).

[48] V. Benci, C. Bonanno,S. Galatolo, G. Menconi, F. Ponchio, http://arXiv.org/abs/math.DS/0107067 (2001).

[49] G. J. Chaitin, *Information, randomness and incompleteness. Papers on algorithmic information theory*, World Scientific, Singapore 1987.

[50] S. Galatolo, arXiv E-print no. math.DS/0102187 (2001).

[51] P. Gaspard and X. J. Wang, Proc. Natl. Acad. Sci. U.S.A. **85**, 4591 (1988).

[52] P. Manneville, J. Phys. (Paris) **41**, 1235-1243 (1980).

[53] N. Goldenfeld, *Lectures on Critical Phenomena and the Renormalization Group*, Addison-Wesley, New York, 1992.

[54] T. Geisel and S. Thomae, Phys. Rev. Lett. **52**, 1936 (1984).

[55] T. Geisel, J. Nierwetberg, and A. Zacharl, Phys. Rev. Lett. **58**, 616 (1985).

[56] P. Allegrini, P. Grigolini, L. Palatella and A. Rosa, submitted to Phys. Rev. E.

[57] M. Bologna, P. Grigolini, B. J. West, "*Strange Kinetics: conflict between density and trajectory description*", in press on Chem. Phys.

[58] D.B. Percival and A.T. Walden, *Wavelet Methods for Time Series Analysis*, Cambridge University Press, Cambridge 2000.

[59] G. Kaiser, *A Friendly Guide to Wavelets*, Birkhäuser, Boston 1994.

[60] P. Allegrini, P. Grigolini, P. Hamilton, L. Palatella, G. Raffaelli and M. Virgilio, in *Emergent Nature*, Ed. M.M. Novak, World Scientific, pp. 173, 2002.

[61] N. Scafetta, P. Grigolini, P. Hamilton, B. J. West, submitted to Phys. Rev. E.

[62] J. Feder, *fractals* Plenum Press, New York 1988.

[63] P. Allegrini, M. Barbi,P. Grigolini and B.J. West, Phys. Rev. E **52**, 5281 (1995).

[64] P. Allegrini, M. Buiatti, P. Grigolini and B. J. West, Phys. Rev. E 57, 4558 (1998).