PARAMETRIC ESTIMATION AND SPECTRAL ANALYSIS OF CHAOTIC TIME SERIES

Artur Lopes and Sílvia Lopes

Instituto de Matemática
Universidade Federal do Rio Grande do Sul
Av. Bento Gonçalves, 9500
Porto Alegre – RS – 91540-000 – Brasil

ABSTRACT

We present an estimation procedure and analyze spectral properties of chaotic stochastic processes as

$$Z_t = X_t + \xi_t = \phi(T^t(\psi)) + \xi_t, \quad \text{for} \quad t \in \mathbf{Z},$$

where T is a deterministic map, ϕ is a given function and ξ_t is the noise process.

The examples considered in this paper generalize the classical harmonic model

$$Z_t = A\cos(\omega_0\, t + \psi) + \xi_t, \quad \text{for} \quad t \in \mathbf{Z}.$$

We also consider large deviation properties of the estimated parameters.

1. INTRODUCTION

Consider the stationary process

$$Z_t = A\cos(\omega_0\, t + \psi) + \xi_t, \quad \text{for} \quad t \in \mathbf{Z}, \tag{1.1}$$

where $A > 0$ and $\omega_0 \in (-\pi, \pi]$ are constants, $\{\xi_t\}_{t\in\mathbf{Z}}$ is a Gaussian white noise with mean zero and variance σ_ξ^2 and ψ is a uniformly distributed random variable in $(-\pi, \pi]$ independent of the noise process.

The process in (1.1) is the classical harmonic model (see Bloomfield (1976)) for time series analysis. Several different procedures to estimate the frequency ω_0 are known. The spectral distribution function of the model (1.1) is

$$dF_Z(\lambda) = \frac{1}{2}(\delta_{\omega_0} + \delta_{-\omega_0}) + \frac{1}{2\pi}, \quad \text{for} \quad \lambda \in [-\pi, \pi].$$

Consider the map $T : (-\pi, \pi) \to (-\pi, \pi)$ given by $T(x) = \omega_0 + x \pmod{2\pi}$ and its iterates

$$T^t = \underbrace{T \circ T \circ \cdots \circ T}_{t \text{ times}}$$

which satisfy $T^t(x) = \omega_0\, t + x \pmod{2\pi}$. We remind the reader that for any given number c, the value $c \pmod{2\pi}$ is the value d where $c = 2\pi n + d$, $0 \leq d < 2\pi$ and $n \in \mathbf{Z}$.

The process (1.1) can be rewritten as

$$Z_t = A \cos(T^t(\psi)) + \xi_t, \quad \text{for} \quad t \in \mathbf{Z}, \tag{1.2}$$

where ψ is a uniformly distributed random variable in $(-\pi, \pi]$.

Our purpose is to analyze stochastic processes of the type (1.2) where the transformation T is a general bijective map from a set $K \subset \mathbf{R}$ (or, more generally, $K \subset \mathbf{R}^n$) to itself.

In a more general setting, given any function $\phi : K \to \mathbf{R}$, consider the stochastic process

$$Z_t = \phi(T^t(x)) + \xi_t, \quad \text{for} \quad t \in \mathbf{Z}, \tag{1.3}$$

where x has a distribution $\mathcal{P}$ absolutely continuous with respect to the Lebesgue measure. Formal definitions will be given in the next section.

The map T will define a dynamical system with chaotic behavior in the examples considered here. We will use techniques from Ergodic Theory (see Cornfeld et al. (1982) and Walters (1981)) and Large Deviations (see Dembo and Zeitouni (1993)) and Ellis (1989)) in order to analyze the process (1.3). We call the models such as (1.3) of *chaotic stochastic processes*. After the general definitions and properties of such processes in Sections 2 and 3, we present Example 1 and Example 2, respectively, in Sections 4 and 5.

The stochastic process (1.1) is a particular case of Example 1 that will be analyzed in Section 4.

We are able to present an estimation procedure to find the parameters (they play the role of the frequency ω_0 in model (1.1)) and also to exhibit explicitly the spectral density function of Example 2 and all the Fourier coefficients of the spectral distribution function of Example 1. A remarkable fact in Example 1 is the appearance of a strong peak of the spectral distribution function in the value corresponding to the rotation number of the map T. The rotation number of a bijective map T is an important invariant previously analyzed in Dynamical System (see Devaney (1989)) and it seems to play also an important role in the spectral analysis properties of certain chaotic time series.

It is well known in the theory of time series analysis that different models require different estimation procedures. We do not know a general procedure that works for all models of the type (1.3). We propose to use here the sample autocovariance functions at low order to estimate the parameters, but each particular model will require a different approach to estimate the involved parameters.

We also carry out a complete analysis of the deviations of the mean estimated values in the case with no noise. In fact, Example 1 and Example 2 (in the case where $\sigma_\xi^2 = 0$) satisfy a Large Deviation principle as will be shown in the sequel.

The large deviation principle for the case $\sigma_\xi^2 \neq 0$ will be presented in a forthcoming paper.

The large deviations properties of the model (see Sections 4 and 5) will assure that the estimation procedure is, in some sense, robust.

We believe that the general techniques presented in Sections 2 and 3 can also be applied to a wide range of different examples of the type (1.3).

Remark: In the literature, different definitions of chaotic systems may be found. According to Devaney (1989), for instance, the transformation $T(x) = \omega_0 + x$ is not chaotic since it does not satisfy the sensitive dependency on *the initial condition property*. However, the temporal evolution $T^t(x)$ of such map, for any $x \in (-\pi, \pi]$, is very erratic and, for abuse of the notation, we also call such systems by chaotic.

As usual, we call $\{\phi(T^t(x))\}_{t\in\mathbf{Z}}$ *the signal process* and $\{\xi_t\}_{t\in\mathbf{Z}}$ *the noise process*. The value σ_ξ determines the strength of the noise. For a given fixed signal $\phi(T^t(x))$, as larger the value σ_ξ is, stronger is the noise with respect to the signal. *The signal to noise ratio* is defined by

$$SNR = 20 \log_{10} \left(\frac{\text{std. signal}}{\text{std. noise}} \right). \tag{1.4}$$

Negative values of the signal to noise ratio mean a stronger noise component than the signal. In the same way as it happens for other kind of time series models, in the present situation, if the noise is much stronger than the signal, that is, if the signal to noise ratio is strongly negative, the estimation procedure works badly.

We present in the end of Sections 4 and 5, a table showing simulations that confirm the good performance of the method for estimation purposes when the signal to noise ratio has reasonable values.

We would like to point out a basic difference between model (1.3) considered here and the previous work of Tong (1990) and others. In Tong (1990), the model is

$$X_{t+1} = \phi(X_t, \xi_t), \quad \text{for} \quad t \in \mathbf{Z}, \tag{1.5}$$

where ϕ and X_t are deterministic and ξ_t is the noise. In this case, for $\phi(x) = x$, for instance, the influence of the noise propagates when time goes on in the following way

$$T(T(x) + \xi_1) + \xi_2. \tag{1.6}$$

In the present situation, the noise propagation is like

$$T(T(x)) + \xi_2. \tag{1.7}$$

The situations in (1.6) and (1.7) are quite different and we are not sure that the results presented here can be applied to processes as in (1.5). When there is no noise, that is, when $\sigma_\xi^2 = 0$, then the expressions (1.6) and (1.7) define, of course, the same process. In this case, the model analyzed here can also be considered a model satisfying the hypothesis of Tong (1990).

We refer the reader to Takens (1994), Kostelich and Yorke (1990), Ding et al. (1993) and Tong (1990) for general properties of time series with chaotic behavior.

2. STATIONARY STOCHASTIC PROCESSES

The general setting of chaotic time series we want to analyze is the following. Consider K a compact subset of $\mathbf{R}^n$ with a given Borel σ-algebra $\mathcal{F}$, an invertible continuous transformation $T : K \to K$, an invariant probability $\mathcal{P}$ on K (that is, $\mathcal{P}(T^{-1}(A)) = \mathcal{P}(A)$, for any set $A \in \mathcal{F}$) and $\phi : K \to \mathbf{R}$ a continuous function. We will analyze the stationary stochastic process $\{Z_t\}_{t\in\mathbf{Z}}$ given by

$$Z_t = X_t + \xi_t = (\phi \circ T)(X_{t-1}) + \xi_t, \quad \text{for } t \in \mathbf{Z}. \tag{2.1}$$

The natural measure on $K^{\mathbf{Z}}$ is the product measure on $K^{\mathbf{Z}}$ and it is invariant for the stationary process $\{X_t\}_{t\in\mathbf{Z}}$ or $\{Z_t\}_{t\in\mathbf{Z}}$. The process $\{\xi_t\}_{t\in\mathbf{Z}}$ is considered to be a Gaussian white noise process (see Brockwell and Davis (1987)) independent of $\{(\phi\circ T)(X_t)\}_{t\in\mathbf{Z}}$, with zero mean and variance σ_ξ^2. One observes that in the model (2.1) the random variables Z_t and Z_{t+1} are generally not independent.

We shall denote the above system by $(K, T, \mathcal{P}, \phi, \mathcal{F}, \sigma_\xi^2)$. Following the terminology in Tong (1990) we may call the system (2.1), when $\sigma_\xi^2 = 0$, the *skeleton* of the system.

For the following definitions we shall not consider the noise process $\{\xi_t\}_{t\in\mathbf{Z}}$ in the model (2.1) and we shall denote the system by $(K, T, \mathcal{P})$. We say that two systems $(K_1, T_1, \mathcal{P}_1)$ and $(K_2, T_2, \mathcal{P}_2)$ (where, for the moment, we do not consider any continuous function ϕ) are *equivalent* in the Ergodic Theory sense if there exists a map $v : K_1 \to K_2$ invertible (that is, there exists $u : K_2 \to K_1$ such that $v \circ u = id$, $\mathcal{P}_1$-almost everywhere and $u \circ v = id$, $\mathcal{P}_2$-almost everywhere) such that

$$\begin{aligned} &(i)\ \ v^*(\mathcal{P}_2) = \mathcal{P}_1, \quad \text{where} \quad v^*(\mathcal{P}_2)(A) = \mathcal{P}_2(v^{-1}(A)), \quad \text{for any set } A \in \mathcal{F}. \\ &(ii)\ \ T_2 \circ v = v \circ T_1, \ \ \mathcal{P}_1 - \text{almost everywhere}. \end{aligned} \tag{2.2}$$

One observes that v plays the role of a change of variables. When v satisfies property (2.2) we say that v is a *conjugacy* between the systems $(K_1, T_1, \mathcal{P}_1)$ and $(K_2, T_2, \mathcal{P}_2)$. We refer the reader to Walters (1981) for precise definitions and general results about equivalence in Ergodic Theory. It is a simple consequence of (ii) in (2.2) that

$$T_2^t \circ v = v \circ T_1^t, \quad \text{for any } t \in \mathbf{Z}.$$

Given a certain measurable function $\phi : K \to \mathbf{R}$ *the autocovariance function at lag h* (see Brockwell and Davis (1987)) *of the process* $\{X_t\}_{t\in\mathbf{Z}}$ as in (2.1) is given by

$$R_{XX}(h) = E(X_t X_{t+h}) - [E(X_t)]^2 = \int \phi(x)\phi(T^h(x))d\mathcal{P}(x) - [\int \phi(x)d\mathcal{P}(x)]^2. \quad (2.3)$$

The autocovariance function $R_{XX}(h)$ in (2.3) measures the covariance between two values of the process $\{X_t\}_{t\in\mathbf{Z}}$ separated by lag h. The *autocorrelation function at lag h of the process* $\{X_t\}_{t\in\mathbf{Z}}$ (see Brockwell and Davis (1987)) is given by

$$\rho_X(h) = \frac{R_{XX}(h)}{R_{XX}(0)}, \quad \text{for} \quad h \in \mathbf{N}, \quad (2.4)$$

where $R_{XX}(0) = E[(X_t - E(X_t))^2] = \text{Var}(X_t)$ is the variance of the process.

Proposition 2.1: *If* $(K_1, T_1, \mathcal{P}_1)$ *and* $(K_2, T_2, \mathcal{P}_2)$ *are equivalent as in (2.2) then, for any* ϕ*, the autocovariance functions at lag h of the processes* $X_t = \phi \circ v \circ T_1^t$ *and* $Y_t = \phi \circ T_2^t$ *are the same, that is,*

$$R_{XX}(h) = R_{YY}(h), \quad \textit{for any} \quad h \in \mathbf{Z}.$$

Proof: In fact, given a continuous function $\phi : K_2 \to \mathbf{R}$ and for any $h \in \mathbf{Z}$ then

$$\begin{aligned} R_{YY}(h) &= \int \phi(x)\phi(T_2^h(x))d\mathcal{P}_2(x) - [\int \phi(x)d\mathcal{P}_2(x)]^2 = \\ &= \int \phi(v(y))\phi(T_2^h(v(y)))d\mathcal{P}_1(y) - [\int (\phi\circ v)(y)d\mathcal{P}_1(y)]^2 = \\ &= \int \phi(v(y))\phi(v(T_1^h(y)))d\mathcal{P}_1(y) - [\int (\phi\circ v)(y)d\mathcal{P}_1(y)]^2 = \\ &= \int (\phi\circ v)(y)(\phi\circ v)(T_1^h(y))d\mathcal{P}_1(y) - [\int (\phi\circ v)(y)d\mathcal{P}_1(y)]^2 = \\ &= R_{XX}(h). \end{aligned}$$

The second above equality follows from the fact that $v^*(\mathcal{P}_2) = \mathcal{P}_1$ is equivalent to

$$\int \varphi(x)d\mathcal{P}_2(x) = \int (\varphi\circ v)(y)d\mathcal{P}_1(y),$$

for any continuous function φ.

¿From the Herglotz's theorem (see Brockwell and Davis (1987)) a function $\rho_X(h)$ is non-negative definite if and only if

$$\rho_X(h) = \int_{-\pi}^{\pi} e^{i\lambda h} dF_X(\lambda), \quad \text{for any} \quad h \in \mathbf{Z}, \quad (2.5)$$

where $F_X(\cdot)$ is a right-continuous, non-decreasing, bounded function on $[-\pi, \pi]$ with $F_X(-\pi) = 0$. The function $F_X(\cdot)$ is called *the spectral distribution function* of $\{X_t\}_{t\in\mathbf{Z}}$ and if

$$F_X(\lambda) = \int_{-\pi}^{\lambda} f_X(\omega)\, d\omega, \quad \text{for } -\pi \le \lambda \le \pi, \tag{2.6}$$

then $f_X(\cdot)$ is called *the spectral density function* of the process $\{X_t\}_{t\in\mathbf{Z}}$. When

$$\sum_{h=-\infty}^{\infty} |\rho_X(h)| < \infty,$$

then $\rho_X(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f_X(\lambda)\, d\lambda$, for $h \in \mathbf{Z}$, where $f_X(\cdot)$ is given by

$$f_X(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \rho_X(h). \tag{2.7}$$

Remark: From Proposition 2.1 we conclude that the spectral distribution functions (see (2.5)) of both stochastic processes, $X_t = (\phi \circ v)(T_1^t)$ and $Y_t = \phi(T_2^t)$, are the same. In conclusion, if we are able to analyze the spectrum properties of the system $(K_1, T_1, \mathcal{P}_1, \phi)$ then we are also able to analyze the spectrum properties of any equivalent system $(K_2, T_2, \mathcal{P}_2, \phi \circ v)$.

Example: When the compact subset K is equal to $[-\pi, \pi]$, the transformation T is given by $T(x) = \omega_0 + x \pmod{2\pi}$, with $\omega_0 \in (0, \pi)$, and $\phi(x) = \cos(x)$ (this is the model (1.1)), the spectral distribution function of the process $\{X_t\}_{t\in\mathbf{Z}} = \{(\phi \circ T)(X_{t-1})\}_{t\in\mathbf{Z}}$ as in (2.1) is not a function but a *generalized spectral distribution function* exists and it is given by

$$dF_X(\lambda) = \frac{1}{2}(\delta_{\omega_0} + \delta_{-\omega_0}), \tag{2.8}$$

where δ_{ω_0} is the Dirac delta function concentrated at ω_0.

Remark: Expanding maps (see Lopes (1994) for the definition) always have an exponential decay of autocorrelations, for any ϕ Holder continuous function (see Parry and Pollicott (1990)). Therefore, in this case, the spectral density function always exists and the spectrum is of continuous type. The function T of Example 2 in Section 5 is an expanding map.

3. THE ESTIMATION OF THE PARAMETERS

We shall consider ϕ as a fixed known continuous function, $\mathcal{P}$ is also fixed and T is a unknown transformation indexed by the parameters $(\alpha_1, \alpha_2, \cdots, \alpha_k)$. One of our main purposes in this paper is to estimate the map T, or equivalently, to estimate the parameters $(\alpha_1, \alpha_2, \cdots, \alpha_k)$ from a time series $\{Z_t\}_{t=1}^N$ of size N derived from the stationary stochastic process $\{Z_t\}_{t\in\mathbf{Z}}$ given by (2.1). We also would like to estimate the noise parameter σ_ξ^2.

In the example given before where $T(x) = \omega_0 + x \pmod{2\pi}$ and $\phi(x) = \cos(x)$, one wants to estimate the frequency ω_0.

3.1. Birkhoff's Ergodic Theorem

In the sequel we assume that the system $(K, T, \mathcal{P})$ is *ergodic* (that is, if $T^{-1}(A) = A$ then $\mathcal{P}(A) = 0$ or $\mathcal{P}(A) = 1$, for any $A \in \mathcal{F}$). The Birkhoff's ergodic theorem claims that if $\mathcal{P}$ is ergodic and if $\phi : K \to \mathbf{R}$ is $\mathcal{P}$- integrable then for any y $\mathcal{P}$-almost everywhere

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} \phi(T^j(y)) = \int \phi(x) d\mathcal{P}(x). \tag{3.1}$$

In simple words, the Birkhoff's ergodic theorem says that spatial mean is equal to temporal mean.

Our main tool to estimate the parameters $(\alpha_1, \alpha_2, \cdots, \alpha_k)$ is the ergodic theorem.

Each particular system $(K, T, \mathcal{P}, \phi, \mathcal{F}, \sigma_\xi^2)$ will require a particular method for estimating the parameters $(\alpha_1, \alpha_2, \cdots, \alpha_k)$. It is natural to try to estimate these parameters from the sample autocovariance function at lag h based on the time series $\{Z_t\}_{t=1}^N$, for small values of h. There exist two reasons for possibly small deviations of the estimates $(\hat{\alpha}_1, \hat{\alpha}_2, \cdots, \hat{\alpha}_k)$ from the parameters $(\alpha_1, \alpha_2, \cdots, \alpha_k)$.

(1) There exists a noise process $\{\xi_t\}_{t\in\mathbf{Z}}$ in our system. This generates a small uncertainty in the estimation.

(2) The estimation is based on a finite amount of observations. The value

$$\frac{1}{N} \sum_{j=1}^{N} \phi(T^j(y)) \quad \text{is not equal to} \quad \int \phi(x) d\mathcal{P}(x)$$

but it is very close for sufficiently large N. This also generates a small perturbation on the estimated values $(\hat{\alpha}_1, \hat{\alpha}_2, \cdots, \hat{\alpha}_k)$ (see Section 3.2).

In Lopes (1993) and Lopes and Kedem (1994) the estimation of a certain well known example of system $(K, T, \mathcal{P}, \phi, \mathcal{F}, \sigma_\xi^2)$ is presented. In these works, the compact subset is given by $K = [0, 1]$, the transformation T is such that $T(x) = T_{\omega_0}(x) = \omega_0 + x$, $\phi(x) = \cos(2\pi x)$, where $\omega_0 \in (0, 1)$ is the parameter to be estimated. The spectral distribution function of such system is very well known. It is a Dirac delta function on ω_0 (see (2.8)). Furthermore, in these works the case with p (p ≥ 1) frequencies is analyzed.

Our purpose in this paper is to develop methods to analyze other kind of time series (see Examples 1 and 2) in terms of estimation and spectral analysis.

3.2. Large Deviations

The deviations from $(\hat{\alpha}_1, \hat{\alpha}_2, \cdots, \hat{\alpha}_k)$ to $(\alpha_1, \alpha_2, \cdots, \alpha_k)$ in the case where $\sigma_\xi^2 = 0$, that is, when the model (2.1) has only the signal process, is the content of the Theory of Large Deviations as presented in the book by Ellis (1989). The most important property for the large deviation estimates to be robust is the exponential convergence property (see Definition 3.1 below). This property means that the deviation rate is exponentially decreasing. This is true for Example 2 treated in Section 5 since, in that case, the map T is expanding (see Lopes (1994) for the definition and general properties). The Example 1, presented in Section 4, does not fit in the context of Lopes (1994) since the map T is not an expanding one. It is also true that Example 1 has exponentially decreasing deviation rate and this will be proved in Section 4.2. The case $\sigma_\xi^2 \neq 0$ requires a different analysis and it will be the subject of a forthcoming paper. In this section we consider $\sigma_\xi^2 = 0$.

We shall consider now a general dynamical system $(K, T, \mathcal{P})$ and a continuous function $f : K \to \mathbf{R}$. We assume that $\mathcal{P}$ is an ergodic probability measure.

In general, it may exist points y such that the equality (3.1) does not hold. Given $\epsilon > 0$, consider the set

$$Q_n(\epsilon) = \{y \in K;\ \ |n^{-1} \sum_{j=1}^{n} f(T^j(y)) - \int f(x) d\mathcal{P}(x)| > \epsilon\}.$$

¿From the expression (3.1) it follows that, for any $\epsilon > 0$,

$$\lim_{n \to \infty} \mathcal{P}(Q_n(\epsilon)) = 0. \tag{3.2}$$

If the convergence in (3.2) is very slow, even for large n, we have a certain reasonable large chance of choosing a bad y such that the mean

$$\frac{1}{n} \sum_{j=1}^{n} f(T^j(y))$$

is distant from $\int f(x) d\mathcal{P}(x)$ by more than ϵ. This would be a very bad situation for the estimation purposes presented in Section 4.1.

Definition 3.1: The system $(K, T, \mathcal{P}, f)$ has the *exponential convergence property* if for any $\epsilon > 0$, exists $M > 0$ such that, for any $n > 0$,

$$\mathcal{P}(Q_n(\epsilon)) < e^{-nM}.$$

This property provides a fast decreasing of the probability of choosing a bad y.

For the estimation procedure in Section 4.1 to work properly one should prove that, for any f, the system $(K, T, \mathcal{P}, f)$ satisfies Definition 3.1, where $T = T_{\alpha,\beta}$ (see definition in expression (4.7)). First, we shall prove that this property is true when the transformation T is given by $T(x) = \omega_0 + x$ and then we shall derive, by contraction principle (see Theorem 3.2), the exponential convergence property for $(K, T_{\alpha,\beta}, \mathcal{P}, f)$. This will be the subject of Section 4.2.

Remark: When one needs to estimate

$$E(Z_t Z_{t+1}) = \int \phi(x)\phi(T(x))d\mathcal{P}(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N-1} Z_t Z_{t+1}$$

one should consider large deviations properties for the function $f(x) = \phi(x)\phi(T(x))$ (in the notation of this section).

Definition 3.2: For each $n \in \mathbf{N}$ and $t \in \mathbf{R}$, consider the function

$$c_n(t) = \int e^{t\sum_{j=1}^{n} f(T^j(x))} d\mathcal{P}(x)$$

and the limit

$$c(t) = \lim_{n\to\infty} \frac{1}{n} \log c_n(t).$$

When such limit exists, for all t, we call $c(t)$ *the free energy of* f.

Note that, in Definition 3.2, $c(0) = 0$.

Definition 3.3: Given the free energy $c(t)$, $t \in \mathbf{R}$, of f we define $I(z)$, the Legendre transform of $c(t)$, by

$$I(z) = \sup_{t\in\mathbf{R}}\{t\,z - c(t)\}.$$

We call $I(z)$ *the deviation function of* f.

Remark: When $c(t)$ is differentiable and convex, the deviation function of f is

$$I(z) = t_0\,z - c(t_0), \quad \text{where} \quad c'(t_0) = z.$$

Example: If $c(t)$ is linear with inclination α, then $I(z) = \infty$, for $z \neq \alpha$ and $I(\alpha) = 0$.

Theorem 3.1: *If $c(t)$, the free energy of f, is differentiable on t, then*

$$\lim_{n\to\infty}\frac{1}{n}\log\mathcal{P}(Q_n(\epsilon)) = -\inf_{|z-\int f(x)d\mathcal{P}(x)|>\epsilon} I(z).$$

According to Theorem 3.1, one concludes that if $I(z)$ is such that $I(z) = 0 \iff z = \int f(x)d\mathcal{P}(x)$, otherwise is greater than zero, then the system $(K, T, \mathcal{P}, f)$ has the *exponential convergence property*. In particular, any system with a linear free energy (as presented in the above example) has the *exponential convergence property*. Systems which have a linear free energy present the best possible convergence rate.

We shall prove in Section 4.2 that for the transformation T given by $T(x) = \omega_0 + x$ and for any continuous function f, the free energy is linear.

Given a continuous function f, the deviation function I_f of f can be obtained in the following way (see Ellis (1989))

$$I_f(z) = -\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{n}\log\mathcal{P}\{x\in K; \frac{1}{n}\sum_{j=1}^{n}(f\circ T^j)(x)\in[z-\epsilon, z+\epsilon]\}. \tag{3.4}$$

Now we shall explain the contraction principle for two equivalent systems. Given $(K_1, T_1, \mathcal{P}_1)$ and $(K_2, T_2, \mathcal{P}_2)$, suppose that v is a change of coordinates between two systems in the sense of (2.2). Given a function $f : K_2 \to \mathbf{R}$ one considers its deviation function I_f associated to T_2. Consider the random variable $f\circ v$ defined on K_1. We shall obtain similar properties for the deviation function $I_{f\circ v}(r)$ associated to T_1.

Theorem 3.2 (Contraction Principle for Equivalent Systems): *Let $f : K_2 \to \mathbf{R}$ be a function and let v be a conjugacy between the systems $(K_1, T_1, \mathcal{P}_1)$ and $(K_2, T_2, \mathcal{P}_2)$. Then $I_{f\circ v} = I_f$.*

Proof: Given $z\in\mathbf{R}$, $n\in\mathbf{N}$ and $\epsilon > 0$, then

$$\mathcal{P}_1\{x\in K_1;\ \ \frac{1}{n}\sum_{j=1}^{n}(f\circ v)(T_1^j(x))\in[z-\epsilon, z+\epsilon]\} =$$

$$= \mathcal{P}_1\{x\in K_1;\ \ \frac{1}{n}\sum_{j=1}^{n}(f\circ T_2^j)(v(x))\in[z-\epsilon, z+\epsilon]\} =$$

$$= \int I_A(x)d\mathcal{P}_1(x),$$

where $A = \{x\in K_1;\ \ n^{-1}\sum_{j=1}^{n}(f\circ T_2^j)(v(x))\in[z-\epsilon, z+\epsilon]\}$. Since

$$\int I_A(x)d\mathcal{P}_1(x) = \int (I_A\circ v^{-1})(x)d\mathcal{P}_2(x) = \int I_{v(A)}(x)d\mathcal{P}_2(x) = \mathcal{P}_2(v(A))$$

then, one has

$$\mathcal{P}_1\{x \in K_1; \quad \frac{1}{n}\sum_{j=1}^{n}(f \circ v)T_1^j(x) \in [z-\epsilon, z+\epsilon]\} =$$

$$= \mathcal{P}_2(v(A)) = \mathcal{P}_2\{x \in K_2; \quad \frac{1}{n}\sum_{j=1}^{n}(f \circ T_2^j)(x) \in [z-\epsilon, z+\epsilon]\}.$$

The result follows from expression (3.4) after taking $n^{-1}\log$ and limits on n and ϵ in the above equalities.

Now we shall state the contraction principle (see Orey (1986) and Varadhan (1988)) in a more general form.

Suppose g is a deterministic function $g : \mathbf{R} \to \mathbf{R}$. Suppose also that c is the solution of the equation $g(c) = b$, where b is obtained as

$$b = \int f(y)d\mathcal{P}(y) = \lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}(f \circ T^j)(x),$$

for $\mathcal{P}$-almost every point $x \in K$. Small deviations of the mean $n^{-1}\sum_{j=1}^{n}(f \circ T^j)(x) = \hat{b}_n$ will produce small deviations in the implicit value $\hat{c}_n$ obtained by solving the equation $g(\hat{c}_n) - \hat{b}_n = 0$.

Denote by I the deviation function for f, that is, the deviation function associated to b. We may ask for properties of the deviation function $\tilde{I}$ associated to c, that is,

$$-\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{n}\log\mathcal{P}\{\hat{c}_n \in [z-\epsilon, z+\epsilon]\} = \tilde{I}(z).$$

Assuming that the function g is bijective, the contraction principle (see Orey (1986) and Varadhan (1988)) claims that

$$\tilde{I}(z) = I(r).$$

where $g(z) = r$.

Let us consider now the generalization of the above considerations to a system of equations. When one considers a system $g_1(\alpha, \beta) = k_1 = \int f_1(x)d\mathcal{P}(x)$ and $g_2(\alpha, \beta) = k_2 = \int f_2(x)d\mathcal{P}(x)$, where $g_1, g_2 : \mathbf{R}^2 \to \mathbf{R}$, the analogous property is true (see Orey (1986)). Therefore, the exponential convergence property of I_{f_1} and I_{f_2} implies that $\hat{\alpha}$ and $\hat{\beta}$ have corresponding deviation functions $\tilde{I}_\alpha$ and $\tilde{I}_\beta$ with the exponential convergence property, that is, given $\epsilon > 0$, there exists $M > 0$ such that, for all $n > 0$,

$$\mathcal{P}\{x \in K; \quad |\hat{\alpha} - \alpha| > \epsilon, \quad |\hat{\beta} - \beta| > \epsilon, \quad \text{where} \quad g_1(\hat{\alpha}, \hat{\beta}) = \frac{1}{n}\sum_{j=1}^{n} f_1(T^j(x)) = \hat{k}_1$$

$$\text{and} \quad g_2(\hat{\alpha}, \hat{\beta}) = \frac{1}{n}\sum_{j=1}^{n} f_2(T^j(x)) = \hat{k}_2\} \le e^{-Mn}.$$

In conclusion, if I_{f_1} and I_{f_2} have the exponential convergence property and $\hat{\alpha}$ and $\hat{\beta}$ can be obtained by solving the equations $g_1(\hat{\alpha}, \hat{\beta}) = \hat{k}_1$ and $g_2(\hat{\alpha}, \hat{\beta}) = \hat{k}_2$, then one can use the contraction principle to conclude that $\hat{\alpha}$ and $\hat{\beta}$ satisfy the exponential convergence property.

4. EXAMPLE 1

Consider the two parameters mapping family $\{T_{a,b} : [0,1] \to [0,1]; \quad a, b \in \mathbf{R}\}$ where $T_{a,b}$ is given by

$$T_{a,b}(x) = \begin{cases} a + \dfrac{1-a}{b} x, & \text{if } 0 \le x < b \\ \dfrac{a}{1-b}(x-b), & \text{if } b \le x \le 1, \end{cases} \tag{4.1}$$

with a and b constants. Let α be the derivative of T on $[0, b)$ and β its derivative on $[b, 1]$. Then,

$$\alpha = T'(x) = \frac{1-a}{b}, \quad \text{if} \quad 0 \le x < b \quad \text{and} \quad \beta = T'(x) = \frac{a}{1-b}, \quad \text{if} \quad b \le x \le 1. \tag{4.2}$$

The ergodic properties of the family $\{T_{a,b} : [0,1] \to [0,1]; \quad a, b \in \mathbf{R}\}$ are analyzed in Coelho et al. (1994).

In Example 1 we want to analyze the estimation of the parameters a and b and the spectral analysis of the process $\{X_t\}_{t \in \mathbf{Z}}$ defined in (4.3) below.

Notice that when $b = 1-a$, the transformation $T_{a,b}$ of Example 1 is $T(x) = a + x$ (mod 1), which corresponds to the model (1.1) analyzed by Lopes and Kedem (1994). Therefore, the presented analysis of Example 1 is a generalization of that work when $p = 1$, that is, the case with only one frequency.

First we examine the system with no noise. The case with noise can be analyzed in a simple way afterwards.

4.1. Estimation

By using the notation introduced in Section 2, for a given transformation $T_{a,b}$ and $\phi(x) = x$ one considers the signal process $\{X_t\}_{t \in \mathbf{Z}}$ given by

$$X_t = T_{a,b}(X_{t-1}), \quad \text{for} \quad t \in \mathbf{Z}. \tag{4.3}$$

To estimate the unknown constants a and b is the same as to estimate α and β, since one has the following identities

$$\alpha = \frac{1-a}{b} \quad \text{and} \quad \beta = \frac{a}{1-b} \Longleftrightarrow a = \frac{\beta(\alpha - 1)}{\alpha - \beta} \quad \text{and} \quad b = \frac{1-\beta}{\alpha - \beta}. \tag{4.4}$$

Therefore, for the sake of simplicity in our analysis we shall estimate the parameters α and β.

The invariant measure $\mathcal{P}_{\alpha,\beta} = \mathcal{P}$ (see Coelho et al. (1994)) for the process $\{X_t\}_{t\in\mathbf{Z}}$, in terms of α and β, is given by the density

$$\varphi_{\alpha,\beta}(x) = \varphi(x) = \frac{1}{c}\,\frac{1}{x + \frac{\beta}{\alpha}(1-x)} = \frac{1}{c}\,\frac{1}{(\alpha-\beta)x + \beta}, \tag{4.5}$$

where

$$c = \frac{1}{\beta-\alpha}\log\left(\frac{\beta}{\alpha}\right) = \frac{1}{\frac{\beta}{\alpha}-1}\log\left(\frac{\beta}{\alpha}\right). \tag{4.6}$$

For a set $A \subset [0,1]\times[0,1]$, with Lebesgue measure equal to 1, for all $(\alpha,\beta)\in A$, the map $T_{\alpha,\beta}$ is ergodic for $\mathcal{P}_{\alpha,\beta} = \mathcal{P}$. We will assume $(\alpha,\beta)\in A$ in the sequel.

In other words, in this case $\mathcal{P}$ given by

$$\mathcal{P}(A) = \int_A \varphi(x)dx, \quad \text{for all } A \in \mathcal{F},$$

where now $\mathcal{F}$ is the Borel σ-algebra in $[0,1]$, defines an invariant ergodic probability measure for T.

¿From the expressions (4.1) and (4.4) the transformation $T_{\alpha,\beta}$ is given by

$$T_{\alpha,\beta}(x) = \begin{cases} \dfrac{\beta(\alpha-1)}{\alpha-\beta} + \alpha x, & \text{if } 0 \le x < \dfrac{1-\beta}{\alpha-\beta} \\ \beta\left(x - \dfrac{1-\beta}{\alpha-\beta}\right), & \text{if } \dfrac{1-\beta}{\alpha-\beta} \le x \le 1. \end{cases} \tag{4.7}$$

The list of integrals below are useful to understand the estimation and the spectral analysis that we shall present in the sequel.

1. $\displaystyle\int_0^y \varphi(x)dx = \frac{\log(\frac{(\alpha-\beta)y+\beta}{\beta})}{\log(\frac{\alpha}{\beta})}.$
2. $\displaystyle E(Z_t) = E(X_t) = \int_0^1 x\varphi(x)dx = \frac{1}{\log(\frac{\alpha}{\beta})} - \frac{\beta}{\alpha-\beta}.$
3. $\displaystyle E(Z_t^2) = E(X_t^2) + \sigma_\xi^2 = \int_0^1 x^2\varphi(x)dx + \sigma_\xi^2 = \left(\frac{\beta}{\alpha-\beta}\right)^2 + \frac{\alpha-3\beta}{2(\alpha-\beta)\log(\frac{\alpha}{\beta})} + \sigma_\xi^2.$
4. $\displaystyle E(Z_tZ_{t+1}) = E(X_tX_{t+1}) = \int_0^1 xT(x)\varphi(x)dx = \left(\frac{\beta}{\alpha-\beta}\right)^2 + \frac{1+\alpha\beta-4\beta}{2(\alpha-\beta)\log(\frac{\alpha}{\beta})}.$

(4.8)

Some of these integrals are obtain after long calculations.

For estimation purposes, in the case where $\phi(x) = x$, one needs the integrals 2. and 4. in the expression (4.8). Suppose we are able to estimate from a time series (that is, by the ergodic theorem), respectively, the integrals 2. and 4. by $\hat{k}_1$ and $\hat{k}_2$. That is,

$$\frac{1}{N}\sum_{t=1}^{N} Z_t = \hat{k}_1 \approx E(Z_t).$$
$$\frac{1}{N}\sum_{t=1}^{N-1} Z_t Z_{t+1} = \hat{k}_2 \approx E(Z_t Z_{t+1}).$$

Then, the estimates of α and β are obtained as the solutions of the following equations

$$\begin{aligned}
\int_0^1 x\varphi(x)dx = k_1 &\Longleftrightarrow g_1(\alpha,\beta) = \frac{1}{\log(\frac{\alpha}{\beta})} - \frac{1}{\frac{\alpha}{\beta}-1} = k_1\\
\int_0^1 xT(x)\varphi(x)dx = k_2 &\Longleftrightarrow g_2(\alpha,\beta) = \left(\frac{\beta}{\alpha-\beta}\right)^2 + \frac{1+\alpha\beta-4\beta}{2(\alpha-\beta)\log(\frac{\alpha}{\beta})} = k_2.
\end{aligned} \tag{4.9}$$

¿From the second equivalence in (4.9) above one can see that

$$\begin{aligned}
\left(\frac{\beta}{\alpha-\beta}\right)^2 + \frac{\frac{1}{\beta}+\alpha-4}{2\left(\frac{\alpha-\beta}{\beta}\right)\log(\frac{\alpha}{\beta})} = k_2 \Longleftrightarrow \left(\frac{\beta}{\alpha-\beta}\right)^2 - \frac{2}{\left(\frac{\alpha}{\beta}-1\right)\log(\frac{\alpha}{\beta})} +\\
+ \frac{\frac{1}{\beta}+\alpha}{2\left(\frac{\alpha}{\beta}-1\right)\log(\frac{\alpha}{\beta})} = k_2.
\end{aligned} \tag{4.10}$$

If we consider $\alpha = \Delta\beta$, then the first two terms and also the denominator of the third term in the last equality in (4.10) depend only on Δ. Hence, from (4.10) one has

$$\frac{1}{\beta} + \alpha = 2\left(\frac{\alpha}{\beta}-1\right)\log\left(\frac{\alpha}{\beta}\right)\left[k_2 - \left(\frac{\alpha}{\beta}-1\right)^{-2} + \frac{2}{\left(\frac{\alpha}{\beta}-1\right)\log(\frac{\alpha}{\beta})}\right] = k_3 = k_3(\Delta).$$

The estimates of the parameters α and β, given by (4.9), are alternatively given as the solutions of the two equations below.

$$\begin{aligned}
\frac{1}{\log(\frac{\alpha}{\beta})} - \frac{1}{\frac{\alpha}{\beta}-1} &= k_1\\
\frac{1}{\beta} + \alpha &= k_3.
\end{aligned} \tag{4.11}$$

For numerical analysis reasons, given $\hat{k}_1$, we shall find first the value $\Delta = \frac{\alpha}{\beta}$ in the equation $g_1(\alpha, \beta) = \hat{k}_1$ of (4.9). For this purpose, one can use Newton's method in $g_1(\Delta) \in \mathbf{R}$, for $\Delta \in \mathbf{R}$. After that, we find the value of $k_3 = k_3(\Delta)$, that depends only on $\Delta = \frac{\alpha}{\beta}$ (and also on k_2). Hence, we solve the equations

$$\frac{\alpha}{\beta} = \Delta$$
$$\frac{1}{\beta} + \alpha = k \ . \tag{4.12}$$

in α and β (this will require to find the roots of a polynomial of degree 2). Therefore, one can have the estimates $(\hat{\alpha}, \hat{\beta})$, or equivalently, the estimates $(\hat{a}, \hat{b})$, by getting the solutions of equations in (4.11), where

$$\hat{k}_1 = \frac{1}{N} \sum_{t=1}^{N} Z_t \approx E(Z_t) \quad \text{and} \quad \hat{k}_2 = \frac{1}{N} \sum_{t=1}^{N-1} Z_t Z_{t+1} \approx E(Z_t Z_{t+1})$$

with $\{Z_t\}_{t=1}^N$ a time series derived from the process $\{Z_t\}_{t\in\mathbf{Z}}$. After we find $\hat{\alpha}$ and $\hat{\beta}$, the value σ_ξ^2 can be easily estimated from the integral 3. in (4.8) and from the value $N^{-1} \sum_{t=1}^N Z_t^2 \approx E(Z_t^2)$ obtained from a time series derived from the process $\{Z_t\}_{t\in\mathbf{Z}}$.

Remark: Notice that from the expression (4.12) one obtain two pairs of solutions. One pair is the value (α, β). The other pair is $(\tilde{\alpha}, \tilde{\beta})$ such that

$$T_{\alpha,\beta}^{-1} = T_{\tilde{\alpha},\tilde{\beta}}.$$

The stationary processes as in (1.3) generated, respectively, by $T_{\alpha,\beta}$ and $T_{\tilde{\alpha},\tilde{\beta}}$ have the same spectral distribution. This indeterminacy is analogous to the one observed in the harmonic model (1.1) where ω_0 and $-\omega_0$ determine the same spectral measure $\frac{1}{2}(\delta_{\omega_0} + \delta_{-\omega_0})$.

One can ask about the deviations of the time series estimates $\hat{k}_1$ and $\hat{k}_2$ to the values k_1 and k_2. The large deviations of k_1 and k_2 are determined, respectively, by deviation functions I_1 and I_2 (see Definition 3.3) of the kind

$$I_i(z) = 0 \quad \text{for} \quad z = k_i \quad \text{and} \quad I_i(z) = \infty \quad \text{for } z \neq k_i, \quad i \in \{1, 2\}.$$

3

This will be shown in Section 4.2. The parameters α and β are obtained from $g_1(\alpha, \beta) = k_1$ and $g_2(\alpha, \beta) = k_3$ by solving the two equations in (4.11). The deviation functions of α and β can be obtained by means of a contraction principle (see the end of Section 3.2). Thus, the considered system has the exponential convergence property for the large deviation rate. The above results are presented in Section 4.2.

The conclusion is that, for any ϕ, with very high probability the mean value of the time series, $N^{-1} \sum_{j=1}^N \phi(T^j(y))$, will be very close to $\int \phi(x) d\mathcal{P}(x)$.

In the simulations, where the sample size is $N = 5,000$ whenever σ_ξ^2 is equal to zero and $N = 2,000$ otherwise, we obtained the following table.

Table 1: Parameters of Example 1 and their respective estimates.

α	β	σ_ξ	snr	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}_\xi$
0.63049	3.31683	0.000	∞	0.63383	3.26947	0.00031
0.99566	1.03169	0.100	9.208	1.00230	0.98326	0.10487
1.21035	0.44972	0.100	9.138	1.21481	0.44430	0.08482
1.21035	0.44972	0.295	-0.258	1.21481	0.44298	0.08601
1.19998	0.80002	0.000	∞	1.19983	0.79988	0.00000
2.32675	0.19141	0.100	8.796	2.40095	0.19690	0.09674
2.32675	0.19141	0.430	-3.874	6.47962	0.03473	0.30758

In the simulation procedure we found the solution Δ of $g_1(\Delta) = \hat{k}_1$ very easily by using the software Mathematica (see Wolfram (1991)). Notice that the function

$$g_1(\Delta) = \frac{1}{\log(\Delta)} - \frac{1}{\Delta - 1}$$

is bijective and, therefore, the solution of $g_1(\Delta) = \hat{k}_1$ is unique. In Table 1 above, the values of $\hat{\alpha}$ and $\hat{\beta}$ were obtained up to the indeterminacy mentioned in the last remark.

4.2. Large Deviations

In this section we shall analyze the large deviations associated to the mean

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} (f \circ T^j)(x)$$

where f is a continuous function on $[0,1]$ and $T = T_{\alpha,\beta}$ is the map defined by (4.7).

In Coelho et al. (1994) is shown that the function

$$v(y) = \int_0^y \varphi(x)dx = \frac{\log\left(\frac{(\alpha-\beta)y+\beta}{\beta}\right)}{\log(\frac{\alpha}{\beta})}$$

is a change of coordinates in the sense of (2.2) between the systems $([0,1], T_{\alpha,\beta}, dx)$ and $([0,1], T_{\omega_0}, dx)$, where $T_{\omega_0}(x) = \omega_0 + x$ with $\omega_0 = \frac{\log\alpha}{\log(\frac{\alpha}{\beta})}$. The value ω_0 is called the *rotation number of* $T_{\alpha,\beta}$ (see Devaney (1989) for definitions). We shall analyze first the large deviation properties of $T_{\omega_0}(x) = \omega_0 + x$ and then, after that, we shall derive by a contraction principle argument (see Theorem 3.2) the exponential decreasing property for the system $([0,1], T_{\alpha,\beta}, dx)$ by using the change of coordinates v.

Consider a rotation $T(x) = \omega_0 + x$, where ω_0 is an irrational number. It is well known that, in this case, the Lebesgue measure dx is ergodic for T (see Devaney (1989)) where $\mathcal{P}(A)$ is the length of A, for any interval A. We shall analyze first the deviation properties for the transformation $T(x) = \omega_0 + x$ and a continuous function $f : [0,1] \to \mathbf{R}$.

We shall present now several results for a continuous function $f : [-1,1] \to \mathbf{R}$ such that $f(-1) = f(1)$. These results can be applied to the case when $f(x) = A\cos(x)$ considered by Lopes and Kedem (1994).

The result is also true for any continuous function f, by using an approximation argument in $L^1(dx)$.

Proposition 4.1: *For any* $y \in S_1 = [0,1]$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} f(T^j(y)) = \int f(z)dz.$$

Proof: From the ergodic theorem, for almost every $x \in S_1$, the above limit is true since ω_0 is irrational and the Lebesgue measure dx is ergodic for T.

As $f(-1) = f(1)$ and f is continuous, it is also uniformly continuous. Thus, given $\epsilon > 0$ there exists $\delta > 0$ such that if $|x - y| < \delta$ then $|f(x) - f(y)| < \epsilon$, for all $x, y \in [0,1]$.

Fix a certain $y \in S_1$. By the ergodic theorem, there exists $x \in (y - \delta, y + \delta)$ such that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} f(T^j(x)) = \int f(z)dz.$$

It is easy to see that

$$|T^j(x) - T^j(y)| = |x - y| < \delta, \quad \text{for all} \quad j \in \mathbf{N}.$$

Therefore,

$$|N^{-1} \sum_{j=1}^{N} f(T^j(x)) - N^{-1} \sum_{j=1}^{N} f(T^j(y))| < \epsilon, \quad \text{for all} \quad N \in \mathbf{N}.$$

As $\lim_{N\to\infty} N^{-1} \sum_{j=1}^{N} f(T^j(x)) = \int f(z)dz$ and by using a limit sup and limit inf argument and by taking $\epsilon \to 0$, then

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} f(T^j(y)) = \int f(z)dz.$$

So, the proposition holds.

Corollary 4.2: *For any open interval* $[a, b]$ *and any* $x \in S_1$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} I_{[a,b]}(T^j(x)) = b - a,$$

where $I_{[a,b]}$ is the indicator function of the interval $[a,b]$.

The corollary is an easy consequence of a step function approximation by continuous functions in L_1 norm.

Proposition 4.3: *Given $\epsilon > 0$, there exists $M > 0$ such that, for all $x \in S_1$ and all $N > M$,*

$$\frac{1}{N} \sum_{j=1}^{N} f(T^j(x)) \in \left[\int f(z)dz - \epsilon, \int f(z)dz + \epsilon\right].$$

Proof: Fix $y \in (0,1)$. Given $\frac{\epsilon}{4} > 0$, let $\delta > 0$ be such that

$$|x-y| < \delta \Rightarrow |f(x) - f(y)| < \frac{\epsilon}{4}.$$

¿From corollary 4.2, for any $x \in S_1$ there exists $m(x) \in \mathbf{N}$ such that $T^{m(x)}(x) \in (y-\delta, y+\delta)$.

Claim: *There exists $M_1 > 0$ such that $m(x) < M_1$, for all $x \in S_1$.*

We will prove the claim later. Suppose the claim is true. Then,

$$|N^{-1}\sum_{j=1}^{N} f(T^j(x)) - \int f(z)dz| \leq |N^{-1}\sum_{j=1}^{m(x)} f(T^j(x)) + N^{-1}\sum_{j=m(x)+1}^{N} f(T^j(x)) -$$

$$- \int f(z)dz| \leq N^{-1}M_1K + |N^{-1}\sum_{j=m(x)+1}^{N} f(T^j(x)) - \int f(z)dz|, \qquad (4.13)$$

where $K = \sup\{|f(x)|; \;\; x \in S_1\}$. As $|T^{m(x)}(x) - y| < \delta$ then,

$$|N^{-1}\sum_{j=m(x)+1}^{N} f(T^j(x)) - N^{-1}\sum_{j=1}^{N} f(T^j(y))| \leq |N^{-1}\sum_{j=m(x)+1}^{N+m(x)} f(T^j(x)) -$$

$$-N^{-1}\sum_{j=N}^{N+m(x)} f(T^j(x)) - N^{-1}\sum_{j=1}^{N} f(T^j(y))| \leq |N^{-1}\sum_{j=m(x)+1}^{N+m(x)} f(T^j(x)) -$$

$$-N^{-1}\sum_{j=1}^{N} f(T^j(y))| + N^{-1}M_1K = |N^{-1}\sum_{j=1}^{N} f(T^j(T^{m(x)}(x)))-$$

$$-N^{-1}\sum_{j=1}^{N} f(T^j(y))| + N^{-1}M_1K \leq \frac{\epsilon}{4} + N^{-1}M_1K. \tag{4.14}$$

¿From expressions (4.13) and (4.14) one has

$$|N^{-1}\sum_{j=1}^{N} f(T^j(x)) - \int f(z)dz| \leq N^{-1}M_1K + |(N^{-1}\sum_{j=m(x)+1}^{N} f(T^j(x))-$$

$$-N^{-1}\sum_{j=1}^{N} f(T^j(y))) + \left(N^{-1}\sum_{j=1}^{N} f(T^j(y)) - \int f(z)dz\right)| \leq 2N^{-1}M_1K+$$

$$+\frac{\epsilon}{4} + |N^{-1}\sum_{j=1}^{N} f(T^j(y)) - \int f(z)dz|. \tag{4.15}$$

Since

$$\lim_{N\to\infty} \frac{1}{N}\sum_{j=1}^{N} f(T^j(y)) = \int f(z)dz,$$

given $\frac{\epsilon}{4} > 0$, there exists $M_2 > 0$ such that, for all $N > M_2$, one has

$$|N^{-1}\sum_{j=1}^{N} f(T^j(y)) - \int f(z)dz| < \frac{\epsilon}{4}.$$

Consider now $M \in \mathbf{N}$ such that

$$M > \sup\{M_2, \frac{8M_1K}{\epsilon}\}.$$

Then, for any $N > M$, $\frac{2M_1K}{N} < \frac{\epsilon}{4}$. Therefore, from (4.15) and for any $x \in S_1$, we have

$$|N^{-1}\sum_{j=1}^{N} f(T^j(x)) - \int f(z)dz| < \epsilon.$$

And the proposition holds.

Now we shall prove the Claim. For each $x \in S_1$, let $m(x)$ be such that $f^{m(x)} \in (y-\delta, y+\delta)$. There exists a small neighborhood $A(x)$ of x such that for any $z \in A(x)$,

$f^{m(x)}(z) \in (y-\delta, y+\delta)$. It is easy to see that $\cup_{x\in S_1} A(x) = S_1$. As S_1 is a compact set, there exist $x_1, \cdots, x_k$ such that $A(x_1) \cup A(x_2) \cup \cdots \cup A(x_k) \supset S_1$. Denote by M_1 the supremum

$$M_1 = \sup_{1\le j\le k} \{m(x_j)\}.$$

Then, for all $x \in S_1$, there exists $M_1 > m(x)$ such that

$$f^{m(x)}(x) \in (y-\delta, y+\delta).$$

And the Claim is proved.

Remark: We shall consider the function $f(x) = xT(x)$ to estimate large deviations of the autocovariance at lag 1 of the process $\{X_t\}_{t\in\mathbf{Z}} = \{T(X_{t-1})\}_{t\in\mathbf{Z}}$.

Now we show that the free energy $c(t)$ is linear.

Theorem 4.4: *The free energy $c(t)$ is linear and, therefore, the deviation function I satisfies $I(z) = \infty$ for $z \neq \int f(x)d\mathcal{P}(x)$, otherwise it is zero.*

Proof: One needs to show that

$$c(t) = t\int f(x)d\mathcal{P}(x).$$

One observes that

$$|n^{-1}\log\int e^{t\sum_{j=1}^{n} f(T^j(x))}d\mathcal{P}(x) - t\int f(x)d\mathcal{P}(x)| = |n^{-1}\log(\int(e^{t\sum_{j=1}^{n} f(T^j(x))} -$$
$$- e^{nt\int f(x)d\mathcal{P}(x)})d\mathcal{P}(x))| = |n^{-1}\log\left[\int e^{nt\left(n^{-1}\sum_{j=1}^{n} f(T^j(x)) - \int f(x)d\mathcal{P}(x)\right)}d\mathcal{P}(x)\right]|.$$

¿From Proposition 4.3, given $\epsilon > 0$, there exists $M > 0$ such that, for any $x \in S_1$ and all $n > M$,

$$|n^{-1}\sum_{j=1}^{n} f(T^j(x)) - \int f(x)d\mathcal{P}(x)| < \epsilon.$$

Therefore, for all $n > M$,

$$|n^{-1}\log\int e^{t\sum_{j=1}^{n} f(T^j(x))}d\mathcal{P}(x) - t\int f(x)d\mathcal{P}(x)| \le |n^{-1}\log$$
$$\left[\int e^{nt\left(n^{-1}\sum_{j=1}^{n} f(T^j(x)) - \int f(x)d\mathcal{P}(x)\right)}d\mathcal{P}(x)\right] | \le |n^{-1}\log\int e^{nt(\pm\,\epsilon)}d\mathcal{P}(x)| =$$
$$= n^{-1}\log e^{\pm\,\epsilon nt} = \pm\,\epsilon t.$$

As t is fixed, by taking $\epsilon \to 0$ one concludes that $c(t) = t\int f(x)d\mathcal{P}(x)$.

Since for $T(x) = \omega_0 + x$ and for any continuous function f the deviation function I_f has the exponential convergence property and since $v(y) = \int_0^y \varphi_{\alpha,\beta}(x)dx$ defines an equivalence between the systems (K, T, dx) and $(K, T_{\alpha,\beta}, \varphi_{\alpha,\beta})$ then, from Theorem 3.2, one concludes that, for a given continuous function g, the deviation function I_g (associated to the system $(K, T_{\alpha,\beta}, \varphi_{\alpha,\beta}, g)$ has also the exponential convergence property. This follows from the fact that v is a continuous function and by considering, in Theorem 3.2, $g = f \circ v$, with $f = g \circ v^{-1}$.

4.3. Spectral Analysis

For a given $T = T_{\alpha,\beta}$ and the corresponding invariant density $\varphi = \varphi_{\alpha,\beta}$ we consider the signal process $\{X_t\}_{t\in\mathbf{Z}} = \{(\phi \circ T_{\alpha,\beta})(X_{t-1})\}_{t\in\mathbf{Z}}$.

¿From the expressions (4.5) and (4.6) one observes that the density function $\varphi_{\alpha,\beta}(x)$ depends only on the quotient $\Delta = \frac{\alpha}{\beta}$. Consider now the transformation T^h, for any $h \in \mathbf{Z}$, where $T = T_{\alpha,\beta}$ is given by the expression (4.7). ¿From Coelho et al. (1994) it is known that

$$T^h(x) = T_{\alpha_h,\beta_h}(x) \quad \text{where} \quad \alpha_h = \frac{b_h}{1-a_h} \quad \text{and} \quad \beta_h = \frac{a_h}{1-b_h}$$

with $a_h = T^h(0)$ and $b_h = T^{-h}(0)$. ¿From Coelho et al. (1994) it is also known that

$$\frac{\alpha_h}{\beta_h} = \frac{\alpha}{\beta}, \quad \text{for any} \quad h \in \mathbf{N},$$

and hence

$$\varphi_{\alpha_h,\beta_h} = \varphi_{\alpha,\beta}, \quad \text{for any} \quad h \in \mathbf{N}.$$

The conclusion is that, for any continuous function ϕ and $h \in \mathbf{N}$,

$$E(X_tX_{t+h}) = \int \phi(x)\phi(T^h(x))\varphi(x)dx = \int \phi(x)\phi(T_{\alpha_h,\beta_h}(x))\varphi_{\alpha,\beta}(x)dx =$$
$$= \int \phi(x)\phi(T_{\alpha_h,\beta_h}(x))\varphi_{\alpha_h,\beta_h}(x)dx. \tag{4.16}$$

As we know $\int \phi(x)\phi(T_{\alpha,\beta}(x))\varphi_{\alpha,\beta}(x)dx$ (see integral 4. in (4.8)), for any α and β, one can calculate $\int \phi(x)\phi(T_{\alpha_h,\beta_h}(x))\varphi_{\alpha_h,\beta_h}(x)dx$, for any $h \in \mathbf{N}$.

Notice that $E(X_tX_{t+h}) = E(X_tX_{t-h})$, for all $h \in \mathbf{N}$.

Therefore, we are able to obtain the exact values of $R_{XX}(h)$, for all $h \in \mathbf{Z}$, from the positive and negative orbit of zero by T (since α_h and β_h depend only on a_h and b_h).

We now consider $\phi(x) = x$. It is known that, for fixed α and β, there exists Δ such that $\alpha_h = \Delta\beta_h$, for all $h \in \mathbf{Z}$. ¿From integral 4. in (4.8), a simple calculation shows that there exist $c_1(\Delta)$ and $c_2(\Delta)$ such that

$$\int_0^1 xT^h(x)\varphi(x)dx = c_1(\Delta) + c_2(\Delta)\left(\frac{1}{\beta_h} + \alpha_h\right).$$

As α_h and $\frac{1}{\beta_h}$ wander around the interval $[0,1]$, then the above integral does not converge to zero as $h \to \infty$. Therefore, the spectral density function is not a function, but there exists a spectral distribution function also called a *generalized spectral density function*.

First one observes that the process $\{X_t\}_{t\in\mathbf{Z}} = \{T_{\alpha,\beta}(X_{t-1})\}_{t\in\mathbf{Z}}$ has mathematical expectation given by the integral 2. in expression (4.8), that is,

$$E(X_t) = \frac{1}{\log\left(\frac{\alpha}{\beta}\right)} - \frac{\beta}{\alpha-\beta}, \quad \text{for all} \quad t \in \mathbf{Z}.$$

We want to derive the spectral distribution function of the process $\{Z_t\}_{t\in\mathbf{Z}}$. We first consider the autocorrelation $\rho_X(h)$ at lag h of the process $\{X_t\}_{t\in\mathbf{Z}} = \{T_{\alpha,\beta}(X_{t-1})\}_{t\in\mathbf{Z}}$ and then use the Herglotz's theorem (see (2.5)) for the process $\{X_t\}_{t\in\mathbf{Z}}$.

Remark: The Fourier coefficients of the spectral distribution function in the case where $T(x) = \omega_0 + x$ are given by $\rho_X(h) = \cos(h\omega_0) = \cos(T^h(0))$, for $h \in \mathbf{Z}$, that is, they are determined by the iterates T^h of zero. The next theorem claims a similar property for the transformation $T_{\alpha,\beta}$ and $\phi(x) = x$.

Theorem 4.5: *The spectral distribution function of the process*

$$Z_t = T^t_{\alpha,\beta}(\cdot) + \xi_t = T_{\alpha,\beta}(X_{t-1}) + \xi_t, \quad \textit{for } t \in \mathbf{Z},$$

is given by

$$dF_Z(\lambda) = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty} e^{-i\lambda h}\rho_X(h) + \frac{1}{2\pi}, \quad \textit{for } \lambda \in [-\pi,\pi] \tag{4.17}$$

where $\rho_X(h)$ is given by $\frac{R_{XX}(h)}{R_{XX}(0)}$ (see the expression (2.4)) with

$$R_{XX}(h) = \frac{1+\alpha_h\beta_h}{2(\alpha_h-\beta_h)\log(\frac{\alpha_h}{\beta_h})} - \frac{1}{\left[\log(\frac{\alpha_h}{\beta_h})\right]^2} \tag{4.18}$$

and

$$R_{XX}(0) = \frac{\alpha+\beta}{2(\alpha-\beta)\log(\frac{\alpha}{\beta})} - \frac{1}{\left[\log(\frac{\alpha}{\beta})\right]^2} \tag{4.19}$$

where α and β are given by the expression (4.2) and

$$\alpha_h = \frac{1-a_h}{b_h}, \quad \beta_h = \frac{a_h}{1-b_h}, \quad a_h = T^h(0) \quad \textit{and} \quad b_h = T^{-h}(0).$$

Proof: From the expression (4.16) and integrals 4. and 3. in (4.8) we have

$$R_{XX}(h) = \frac{1+\alpha_h\beta_h}{2(\alpha_h-\beta_h)\log(\frac{\alpha_h}{\beta_h})} - \frac{1}{\left[\log(\frac{\alpha_h}{\beta_h})\right]^2}$$

and

$$R_{XX}(0) = \frac{\alpha+\beta}{2(\alpha-\beta)\log(\frac{\alpha}{\beta})} - \frac{1}{\left[\log(\frac{\alpha}{\beta})\right]^2},$$

where $X_t = T_{\alpha,\beta}(X_{t-1})$, which give the expressions (4.18) and (4.19). By adding the noise process $\{\xi_t\}_{t\in\mathbf{Z}}$, independent of $\{X_t\}_{t\in\mathbf{Z}}$, we obtain the expression (4.17).

Now we consider $\phi(x) = \cos(2\pi x)$. One wants to calculate the spectral distribution of the process

$$Z_t = X_t + \xi_t = \cos(2\pi T_{\alpha,\beta}(X_{t-1})) + \xi_t, \quad \text{for} \quad t \in \mathbf{Z}.$$

For this purpose we need the following integral:

$$E(X_tX_{t+1}) = \int_0^1 \cos(2\pi x)\cos(2\pi T(x))\varphi(x)dx = \frac{1}{2\log\left(\frac{\alpha}{\beta}\right)} \times k \tag{4.20}$$

where

$$\begin{aligned} k = {} & \cos(2d\beta)[ci(d(\alpha+1)) + ci(d\alpha(\beta+1)) - ci(d\beta(\alpha+1)) - ci(d(\beta+1))] + \\ & + \sin(2d\beta)[si(d(\alpha+1)) + si(d\alpha(\beta+1)) - si(d\beta(\alpha+1)) - si(d(\beta+1))] + \\ & + ci(d(\alpha-1)) + ci(d\alpha(\beta-1)) - ci(d(\beta-1)) - ci(d\beta(\alpha-1)), \end{aligned}$$

with $d = \frac{2\pi}{\alpha-\beta}$, $ci(x)$ is the cosine integral and $si(x)$ is the sine integral (see Gradshteyn and Ryzhik (1965), page 928). The integral (4.20) comes after a long calculation.

In order to calculate the spectral distribution function, one should obtain the Fourier coefficients of such distribution by substituting in (4.20) the values of α and β by α_h and β_h (see expression (4.16)).

Theorem 4.6: *The spectral distribution function of the process*

$$Z_t = T^t_{\alpha,\beta}(\cdot) + \xi_t = \cos(2\pi\, T_{\alpha,\beta}(X_{t-1})) + \xi_t, \quad \textit{for } t \in \mathbf{Z},$$

is given by

$$dF_Z(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-i\lambda h} \rho_X(h) + \frac{1}{2\pi}, \quad \textit{for } \lambda \in [-\pi, \pi], \tag{4.21}$$

where $\rho_X(h)$ is given by $\frac{R_{XX}(h)}{R_{XX}(0)}$ (see the expression (2.4)) with

$$R_{XX}(h) = \frac{1}{2\log(\frac{\alpha_h}{\beta_h})} \times k_h - \frac{1}{[\log(\frac{\alpha_h}{\beta_h})]^2} \times l_h$$

where

$$\begin{aligned} k_h = &\cos(2d_h\beta_h)[ci(d_h(\alpha_h+1)) + ci(d_h\alpha_h(\beta_h+1)) - ci(d_h\beta_h(\alpha_h+1)) - ci(d_h(\beta_h+1))] \\ &+ \sin(2d_h\beta_h)[si(d_h(\alpha_h+1)) + si(d_h\alpha_h(\beta_h+1)) - si(d_h\beta_h(\alpha_h+1)) - si(d_h(\beta_h+1))] \\ &+ ci(d_h(\alpha_h-1)) + ci(d_h\alpha_h(\beta_h-1)) - ci(d_h(\beta_h-1)) - ci(d_h\beta_h(\alpha_h-1)), \end{aligned}$$

and

$$l_h = \{\cos(d_h\beta_h)[ci(d_h\alpha_h) - ci(d_h\beta_h)] + \sin(d_h\beta_h)[si(d_h\alpha_h) - si(d_h\beta_h)]\}^2$$

with

$$d_h = \frac{2\pi}{\alpha_h - \beta_h}, \quad \alpha_h = \frac{1-a_h}{b_h}, \quad \beta_h = \frac{a_h}{1-b_h},$$

$a_h = T^h(0)$ and $b_h = T^{-h}(0)$. The variance $Var(X_t)$ is given by

$$\begin{aligned} R_{XX}(0) = &\frac{1}{2\log(\frac{\alpha}{\beta})}\{\cos(2d\beta)[ci(2d\alpha) - ci(2d\beta)] + \sin(2d\beta)[si(2d\alpha) - si(2d\beta)]\} + \frac{1}{2} - \\ &- \frac{1}{[\log(\frac{\alpha}{\beta})]^2} \times l, \end{aligned}$$

where

$$l = \{\cos(d\beta)[ci(d\alpha) - ci(d\beta)] + \sin(d\beta)[si(d\alpha) - si(d\beta)]\}^2$$

with

$$d = \frac{2\pi}{\alpha - \beta}, \quad \alpha = \frac{1-a}{b} \quad \textit{and} \quad \beta = \frac{a}{1-b}.$$

In Figure 1 we plot the graph of the Fourier series $\frac{1}{2\pi}\sum_{h=-100}^{100} e^{-i\lambda h}\rho_X(h)$ when $\alpha = 2.41809$ and $\beta = 0.22052$. Therefore, we are considering here an approximation of the generalized spectral density function $f_X(\lambda)$ up to an order of 100.

Figure 1: The generalized spectral density function $f_X(\lambda)$ for Example 1 as in (4.21) when $\sigma_\xi^2 = 0$, $\alpha = 2.41809$ and $\beta = 0.22052$.

Remark: The rotation number of $T_{\alpha,\beta}$ is

$$\theta_1 = \frac{\log(\alpha)}{\log(\frac{\alpha}{\beta})}$$

and the rotation number of $T_{\tilde{\alpha},\tilde{\beta}} = T_{\alpha,\beta}^{-1}$ is

$$\theta_2 = \frac{\log(\beta)}{\log(\frac{\beta}{\alpha})}.$$

One observes that $\theta_1 + \theta_2 = 1$. We denote by ζ the smallest value between θ_1 and θ_2. Therefore, $\zeta \leq 0.5$. We call ζ *the rotation number of the stochastic process*.

It is extremely interesting the fact that, for any α and β, the spectral measure is not a Dirac delta function concentrated on the rotation number of $T_{\alpha,\beta}$ (we checked the coefficients $\rho_X(h)$) but it has a very strong peak on the value $2\pi\zeta$ where ζ is the rotation number of the process. In other words, the spectral distribution is very close to

$$\frac{1}{2}(\delta_{2\pi\,\zeta} + \delta_{-2\pi\,\zeta}) = \frac{1}{2}(\delta_{2\pi\,\theta_1} + \delta_{-2\pi\,\theta_1}),$$

where $\theta_1 \leq 0.5 \leq \theta_2$ were defined above.

In conclusion if one applies the Fourier transform to the data it will appear a strong peak in the rotation number.

This property requires, in the future, a deeper analysis in order to understand the spectral distribution function given by (4.21). Notice in Figure 1 the strong peak in the value $2\pi\,\zeta = 2.31671$, where ζ is the rotation number of the process when $\alpha = 2.41809$ and $\beta = 0.22052$ (corresponding to the values $a = 0.1423$ and $b = 0.3547$).

We remind the reader that if $a = 1 - b$ then the rotation number of $T_{\alpha,\beta}$ is equal to a and, in fact, in this case, the spectral distribution function is a Dirac delta function $\frac{1}{2}(\delta_{\pi a} + \delta_{-\pi a})$, when $\phi(x) = \cos(2\pi\, x)$.

Notice that for $T_{\alpha,\beta}(x) = a + x \pmod 1$, the inverse map $T^{-1}_{\alpha,\beta} = T_{\tilde{\alpha},\tilde{\beta}}$ is such that $T_{\tilde{\alpha},\tilde{\beta}}(x) = x - a \pmod 1$. In this case, $\zeta = \pi|a|$.

5. EXAMPLE 2

Sakai and Tokumaru (1980) introduce the following model of chaotic time series. For a given constant $a \in (0, 1)$ consider the transformation $T_a : [0, 1] \to [0, 1]$ given by

$$T_a(x) = \begin{cases} \dfrac{x}{a}, & \text{if } 0 \leq x < a \\ \dfrac{1-x}{1-a}, & \text{if } a \leq x \leq 1. \end{cases} \tag{5.1}$$

The Lebesgue measure dx is invariant and ergodic for the transformation T_a (see Li and Yorke (1975)). In the notation of Section 2, $\mathcal{P}(A)$ is the length of A, for any interval A.

We now consider the stochastic process

$$Z_t = X_t + \xi_t = T_a(X_{t-1}) + \xi_t, \quad \text{for} \quad t \in \mathbf{Z}, \tag{5.2}$$

where $\phi(x) = x$.

The autocovariance function at lag h of the process $\{X_t\}_{t\in\mathbf{Z}}$ in (5.2) (see Sakai and Tokumaru (1980)) is given by

$$R_{XX}(h) = \int_0^1 xT^h(x)dx - [E(X_t)]^2 = \frac{1}{12}(2a-1)^h, \quad \text{for} \quad h > 0, \tag{5.3}$$

where $E(X_t) = \frac{1}{2}$ and $R_{XX}(0) = \mathrm{Var}(X_t) = \frac{1}{12}$.

One can use the above integral to estimate a from the system $(K, T, \mathcal{P}, \phi, \mathcal{F}, \sigma_\xi^2)$ where $\phi(x) = x$. This will be done in Section 5.1. After that, we shall analyze the spectral properties of the process in (5.2).

5.1. Estimation

Let us consider now the case with noise. We want to estimate the parameters $a \in (0,1)$ and σ_ξ^2. From the ergodic theorem, one observes that

$$\begin{aligned}\lim_{N\to\infty} N^{-1}\sum_{t=1}^{N-1} Z_t Z_{t+1} &= \lim_{N\to\infty} N^{-1}\sum_{t=1}^{N-1} X_t X_{t+1} + \lim_{N\to\infty} N^{-1}\sum_{t=1}^{N-1} X_t \xi_{t+1} + \\ &+ \lim_{N\to\infty} N^{-1}\sum_{t=1}^{N-1} X_{t+1}\xi_t + \lim_{N\to\infty} N^{-1}\sum_{t=1}^{N-1} \xi_t\xi_{t+1} = \\ &= \int_0^1 x T_a(x)dx = \frac{a+1}{6} \qquad (5.4)\end{aligned}$$

since $\{X_t\}_{t\in\mathbf{Z}}$ and $\{\xi_t\}_{t\in\mathbf{Z}}$ are independent processes and since

$$E(\xi_t) = 0 \quad \text{and} \quad E(\xi_t\xi_{t+h}) = \begin{cases} \sigma_\xi^2, & \text{if } h = 0 \\ 0, & \text{if } h \neq 0. \end{cases}$$

for all $t \in \mathbf{Z}$. The last equality in expression (5.4) comes from (5.3) when $h = 1$ and from the fact that $E(X_t) = \frac{1}{2}$. ¿From the ergodic theorem and the independence, one has

$$\lim_{N\to\infty} N^{-1}\sum_{t=1}^{N} Z_t^2 = \lim_{N\to\infty} N^{-1}\sum_{t=1}^{N} X_t^2 + \lim_{N\to\infty} N^{-1}\sum_{t=1}^{N} \xi_t^2 = \int_0^1 x^2\, dx + \sigma_\xi^2 = \frac{1}{3} + \sigma_\xi^2. \quad (5.5)$$

Therefore, by using a time series $\{Z_t\}_{t=1}^N$ of size N derived from the stochastic process $\{Z_t\}_{t\in\mathbf{Z}}$ given by (5.2), the estimators $\hat{a}$ and $\hat{\sigma}_\xi^2$ of a and σ_ξ^2 can be obtained implicitly from expressions (5.4) and (5.5) and, thus, are given by

$$\begin{aligned}\hat{a} &= 6\left(N^{-1}\sum_{t=1}^{N-1} Z_t Z_{t+1}\right) - 1 \approx 6\int_0^1 x T_a(x)dx - 1 \\ \hat{\sigma}_\xi^2 &= N^{-1}\sum_{t=1}^{N} Z_t^2 - \frac{1}{3} \approx \int_0^1 x^2\, dx - \frac{1}{3}.\end{aligned}$$

In the simulations, where the sample size is $N = 5{,}000$ whenever σ_ξ^2 is equal to zero and $N = 2{,}000$ otherwise, we obtained the following table.

Table 2: Parameters of Example 2 and their respective estimates.

a	σ_ξ	snr	$\hat{a}$	$\hat{\sigma}_\xi$
0.273001011	0.100	9.208	0.27249	0.11406
0.273001011	0.295	-0.188	0.25564	0.29990
0.273001011	3.000	-20.334	2.33035	3.08462
0.273001011	0.000	∞	0.27551	0.01262
0.400010101	0.100	9.208	0.38870	0.08430
0.400010101	0.000	∞	0.39978	0.00852
0.400010101	0.000	∞	0.40707	0.02615
0.500010111	0.000	∞	0.49147	0.01856
0.783000101	0.000	∞	0.77875	0.06136

5.2. Large Deviations

The map T is an expanding one and the function $\phi(x) = x$ is Holder continuous, therefore, from the differentiability of the free energy (see Lopes (1994)), the exponential convergence property is true. The conclusion is that, with very high probability the samples autocovariances at lags 1 and 0, $N^{-1}\sum_{t=1}^{N-1} Z_t Z_{t+1} - \left(N^{-1}\sum_{t=1}^{N} Z_t\right)^2$ and $N^{-1}\sum_{t=1}^{N} Z_t^2 - \left(N^{-1}\sum_{t=1}^{N} Z_t\right)^2$, estimate with high accuracy, respectively, the autocovariance of lag 1 and the variance of the process.

Finally, by the contraction principle (see the end of Section 3.2) the estimates $\hat{a}$ and $\hat{\sigma}_\xi^2$ also satisfy the exponential convergence property.

5.3. Spectral Analysis

The main obstacle to proceed in the spectral analysis of Example 2 is that the map T_a is not invertible. Therefore, the autocovariance function $R_{XX}(h)$ of the process $\{X_t\}_{t\in\mathbf{Z}}$, given by expression (5.2), for negative lag h does not have a precise meaning. For the estimation of the parameters there is no problem, since we just need the positive lag h. In fact, $h = 0$ and $h = 1$ were enough.

We propose to analyze the natural extension F of T_a, instead of T_a itself.

The natural extension is a canonical way of embedding a non-invertible dynamical system in an invertible one. We refer the reader to Pollicott (1986) and Adler (1991) for general considerations about the natural extension map.

In Example 2, the natural extension of T_a is the map $F : [0,1] \times [0,1] \to [0,1] \times [0,1]$ such that

$$F(x, y) = (T(x), G(x, y)), \quad \text{for any} \quad (x, y) \in [0,1] \times [0,1],$$

where

$$G(x,y) = \begin{cases} ya, & \text{if } 0 \le x < a \\ (a-1)y+1, & \text{if } a \le x \le 1. \end{cases}$$

The map F is invertible and it is easy to see that the Lebesgue measure $dxdy$ is invariant and ergodic for F.

As a particular example, we mention that the Baker map is the natural extension of the tent map (with inclination 2).

Therefore, we shall consider the dynamical system $(K, F, \mathcal{P})$ where $K = [0,1] \times [0,1]$ and $\mathcal{P}$ is the Lebesgue measure $dxdy$ on $[0,1]\times[0,1]$. Instead of $\phi(x) = x$, one can consider $\phi(x,y) = \Pi(x,y) = x$ for any $(x,y) \in [0,1] \times [0,1]$ as a random variable. In the setting of Section 2, we shall analyze in this section the system $(K, F, \mathcal{P}, \Pi, \mathcal{F}, \sigma_\xi^2)$. Now, if $h \ge 0$ then

$$\int_0^1 xT^h(x)dx = \int_0^1 \int_0^1 x\Pi(F^h(x,y))dxdy = \int_0^1 \int_0^1 \Pi(x,y)\Pi(F^h(x,y))dxdy$$

and we obtain, from the expression (5.3), $R_{XX}(h)$ for positive h when $X_t = \Pi \circ F^t$. As the map F is invertible, it makes sense to estimate, for $h > 0$, the integral

$$\int_0^1 \int_0^1 \Pi(x,y)\Pi(F^{-h}(x,y))dxdy.$$

Denote by Inv the function such that $Inv(x,y) = (y,x)$. ¿From an easy calculation one can derive that $F^{-1} = Inv \circ F \circ Inv$. Then, $F^{-h} = Inv \circ F^h \circ Inv$. Now, from a change of variables, one obtain the following

$$\int_0^1 \int_0^1 \Pi(x,y)\Pi(F^{-h}(x,y))dxdy = \int_0^1 \int_0^1 \Pi(x,y)\Pi(F^h(x,y))dxdy = \int_0^1 xT^h(x)dx.$$

After these results one can obtain the spectral density function associated to the stochastic process $\{X_t\}_{t\in\mathbf{Z}}$. The last term in the above equalities has already been calculated (see (5.3)).

Theorem 5.1: *The spectral density function of the stochastic process*

$$Z_t = X_t + \xi_t = (\Pi \circ F)(X_{t-1}) + \xi_t, \quad \textit{for } t \in \mathbf{Z},$$

is given by

$$f_Z(\lambda) = \frac{2a(1-a)}{\pi[1 - 2(2a-1)\cos(\lambda) + (2a-1)^2]} + \frac{1}{2\pi}, \quad \textit{for } \lambda \in [-\pi, \pi]. \tag{5.6}$$

Proof: Since $R_{XX}(h)$ is given by the expression (5.3) and goes to zero exponentially when $h \to +\infty$, the spectral density function (see (2.7)) does exist and it is given by

$$
\begin{aligned}
f_X(\lambda) &= \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-i\lambda h} \rho_X(h) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-i\lambda h} (2a-1)^{|h|} = \\
&= \frac{1}{2\pi} \left[\sum_{h \geq 0} ((2a-1)e^{-i\lambda})^h + \sum_{h=-\infty}^{-1} ((2a-1)e^{i\lambda})^{-h} \right] = \\
&= \frac{1}{2\pi} \left[\frac{1}{1-(2a-1)e^{-i\lambda}} + \frac{(2a-1)e^{i\lambda}}{1-(2a-1)e^{i\lambda}} \right] = \\
&= \frac{2a(1-a)}{\pi[1-2(2a-1)\cos(\lambda)+(2a-1)^2]},
\end{aligned}
$$

for all $\lambda \in [-\pi, \pi]$, since $|(2a-1)e^{\pm i\lambda}| < 1$ when $a \in (0,1)$. The spectral density function of the process $\{Z_t\}_{t \in \mathbf{Z}}$ follows from this.

The spectrum of the signal process $\{X_t\}_{t \in \mathbf{Z}}$ is continuous and its graph is shown in Figure 2 (a), (b) and (c). Notice that if a is small then there exists a peak on π and if a is large the peak is on zero.

Figure 2: The spectral density function $f_X(\lambda)$ for Example 2 as in (5.6) when $\sigma_\xi^2 = 0$ and

(a) $a = 0.15240$**;** **(b)** $a = 0.36570$**;** **(c)** $a = 0.93459$**.**

(a)

(b)

(c)

REFERENCES

[1] Adler, R.L., "Geodesic Flows, Interval Maps and Symbolic Dynamics". In *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces*, T. Bedford et al. (eds.), Oxford University Press, Oxford, pp. 93-123, 1991.

[2] Bloomfield, P., *Fourier Analysis of Time Series: An Introduction*. John Wiley, New York, 1976.

[3] Brockwell, P.J. and R.A. Davis, *Time Series: Theory and Methods*. Springer-Verlag, New York, 1987.

[4] Coelho, Z., A. Lopes and L.F. Rocha, "Absolutely continuous invariant measures for a class of affine interval exchange maps". To appear in *Proceedings of the American Mathematical Society*, 1994.

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

[6] Dembo, A. and O. Zeitouni, *Large Deviations Techniques and Applications*. Jones and Bartlett, Boston, 1993.

[7] Devaney, R.L., *An Introduction to Chaotic Dynamical System*. Addison-Wesley, Redwood City, 1989.

[8] Ding, M.Z., C. Grebogi, E. Ott, T. Sauer and J.A. Yorke, "Estimating Correlation Dimension from a Chaotic Time Series: when does a plateau onset ocurr?". *Physica D*, Vol. 69, No. 3, 4, pp. 404-424, 1993.

[9] Ellis, R., *Entropy, Large Deviations and Statistical Mechanics*. Springer-Verlag, New York, 1989.

[10] Gradshteyn, I.S. and I.M. Ryzhik, *Table of Integrals, Series and Products*. Academic Press, New York, 1965.

[11] Kostelich, E. and J.A. Yorke, "Noise Reduction: Finding the Simplest Dynamical System Consistent with Data". *Physica D*, Vol. 41, pp. 183-196, 1990.

[12] Li, T.-Y. and J.A. Yorke, "Period three implies chaos". *American Mathematical Monthly*, Vol. 82, pp. 985-992, 1975.

[13] Lopes, A., "Entropy, Pressure and Large Deviation". In *Cellular Automata, Dynamical Systems and Neural Networks*, E. Goles and S. Martinez (eds.), Kluwer, Massachusetts, pp. 79-146, 1994.

[14] Lopes, S., "Amplitude Estimation in Multiple Frequency Spectrum". *Communications in Statistics, Theory and Methods*, Vol. 22, No. 10, pp. 2955-2967, 1993.

[15] Lopes, S. and B. Kedem, "Iteration of Mappings and Fixed Points in Mixed Spectrum Analysis". *Stochastic Models*, Vol. 10, No. 2, pp. 309-333, 1994.

[16] Orey, S., "Large Deviations in Ergodic Theory". In *Seminar on Stochastic Processes*, K.L. Chung et al. (eds.), Birkhäuser, Boston, pp. 195-249, 1986.

[17] Parry, W. and M. Pollicott, "Zeta Functions and the Periodic Orbit Structure of Hyperbolic Dynamics". *Asterisque*, Vol. 187-188, 1990.

[18] Pollicott, M., "Distribution of Closed Geodesics on the Modular Surface and Quadratic Irrationals". *Bulletin de la Société Mathématique de France*, Vol. 114, pp. 431-446, 1986.

[19] Sakai, H. and H. Tokumaru, "Autocorrelations of a certain chaos". *IEEE Trans. Acoust., Speech and Signal Processing*, Vol. ASSP-28, No. 5, pp. 588-590, 1980.

[20] Takens, F., "Analysis of non-linear time series, a survey". Pre-print, April, 1994.

[21] Tong, H., *Non-linear Time Series: A Dynamical System Approach.* Clarendon, Oxford, 1990.

[22] Varadhan, S.R.S., "Large Deviations and Applications". *Lecture Notes in Mathematics*