# ITERATION OF MAPPINGS AND FIXED POINTS IN MIXED SPECTRUM ANALYSIS

Silvia Lopes * and Benjamin Kedem **

* Instituto de Matemática Universidade Federal do Rio Grande do Sul Porto Alegre – RS – Brasil and ** Department of Mathematics University of Maryland College Park – MD – USA

Here we will analyze the mixed spectrum model

$$Z_t = \sum_{j=1}^{p} A_j \cos(\omega_j t + \phi_j) + \varepsilon_t = X_t + \varepsilon_t, \quad \text{for} \quad t \in \mathbf{Z},$$

where $p$ is not necessarily known and, for each $j \in \{1, 2, \cdots, p\}$, $A_j$ is an unknown constant, $\omega_j$ is an unknown frequency with value in $(-\pi, \pi]$ and the phase $\phi_j$ is a random variable uniformly distributed in $(-\pi, \pi]$ independent of each other and of the noise component. We assume that the noise component is Gaussian white noise such that $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. Observe that the process $\{Z_t\}_{t \in \mathbf{Z}}$ is not Gaussian. Here we present a recursive method of updating parameters for estimating the frequencies $\omega_j$, $1 \leq j \leq p$. The cosines of the frequencies are obtained as attracting fixed points of a certain map.

## 1. Introduction

Consider the mixed spectrum model

$$Z_t = \sum_{j=1}^{p} A_j \cos(\omega_j t + \phi_j) + \varepsilon_t = X_t + \varepsilon_t, \quad \text{for} \quad t \in T, \tag{1.1}$$

where the set T is either $\mathbf{Z}$ or $\mathbf{R}$ depending on the time parameter being discrete or continuous, $p$ is not necessarily known and, for each $j \in \{1, 2, \cdots, p\}$, $A_j$ is an unknown constant, $\omega_j$ is an unknown frequency with value in $(-\pi, \pi]$ and the phase $\phi_j$ is a random variable uniformly distributed in $(-\pi, \pi]$ independent of each other and of the noise component. We assume here, for simplicity, that the noise component is Gaussian white noise such that $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. The assumption of white noise is not really needed, but it simplifies the exposition. In fact, any continuous spectrum noise will do just as well.

We present a method inspired by the *He and Kedem (HK) Algorithm* (see He and Kedem (1989)) that allows one to obtain, from an iterative procedure, with high order of accuracy, the estimated values of $\omega_j$, $1 \leq j \leq p$. The method is based on *Higher Order Correlation (HOC)* analysis (see Kedem (1990)). The *HOC* analysis is a faster way to estimate the frequencies $\omega_j$, $1 \leq j \leq p$, than the traditional periodogram analysis since the "fast Fourier transform" algorithm requires $O(N \log_2 N)$ computational complexity while in the former we can achieve order of magnitude $O(N)$.

We use successive applications of the *complex filter* (see Definition (3.2)) to obtain all frequencies of the model (1.1) when $T = \mathbf{Z}$ (see Section 4 of this paper).

Let $\{Z_t(\alpha, M)\}_{t\in T}$ be the stochastic process filtered by the *complex filter* (see Definition (3.2)),where $\alpha \in (-1,1)$ and $M \in \mathbf{N} - \{0\}$. Given an initial value $\alpha_0 \in (-1,1)$, the first-order autocorrelation of the complex-filtered process $\{Z_t(\alpha, M)\}_{t\in T}$ is, by definition,

$$\rho_1(\alpha_0) = \frac{\mathcal{R}\{E[Z_t(\alpha_0, M)\overline{Z_{t+1}(\alpha_0, M)}]\}}{E|Z_t(\alpha_0, M)|^2},$$

where here and elsewhere, a bar denotes complex conjugate and $\mathcal{R}\{z\}$ the real part of $z$.

We should write $\rho_1(\alpha_0, M)$ to also show the dependence on M (we use this notation in Sections 3 and 4) but, in order to simplify the notation in this section, we write only $\rho_1(\alpha_0)$. Suppose M is fixed. With $\alpha_1 = \rho_1(\alpha_0) \in (-1,1)$ as the updated filter parameter we calculate again the first autocorrelation of the complex-filtered process and we obtain

$$\alpha_2 \equiv \rho_1(\alpha_1) = \frac{\mathcal{R}\{E[Z_t(\alpha_1, M)\overline{Z_{t+1}(\alpha_1, M)}]\}}{E|Z_t(\alpha_1, M)|^2}.$$

In an analogous way, we define $\alpha_3 = \rho_1(\alpha_2) \in (-1,1)$ from $\alpha_2$, to update the procedure. In general, define

$$\alpha_{k+1} = \rho_1(\alpha_k), \quad \text{for} \quad k \in \mathbf{N}.$$

We consider here the iterative procedure of applying $\rho_1$ successively to the variable $\alpha$. The main point in this paper is to derive useful information on the process (1.1) from the value $\alpha_k$, when $k$ is large, and $\alpha_0$ is chosen at random in (-1,1). Notice that in Section 4 we update just the variable $\alpha$ and not the parameter M. This is the reason why we consider here $\alpha$ as a variable and M a parameter.

Our goal is to show in Section 4 that this iterative procedure of *updating the complex filter parameter* will converge to a value close to the cosine of some frequency. Now if M is large enough the possible values where the iterative procedure converges will give us all frequencies of the model. Taking a large number of different initial values, we are able to locate all frequencies.

Another way is to take just one initial value $\alpha_0$, consider the iterates $\rho_1^k(\alpha_0, M)$ for $k$ large then filter out the value $\alpha_k$ through a bandpass filter. This value $\alpha_k$ will be close to the $\cos(\omega_{l_0})$, for some $l_0 \in \{1, 2, \cdots, p\}$. Now one applies the same above procedure to the resulting time series. Considering again $\alpha_0$ at random, we estimate another value $\cos(\omega_{l_1})$, for $l_1 \in \{1, 2, \cdots, p\} - \{l_0\}$, through the updating procedure described above. In this way we obtain, successively, all frequencies $\omega_j$, for $j \in \{1, 2, \cdots, p\}$.

There is very important information about the iterative updated parameter procedure that can be obtained by analyzing the intersection of the graphs of the function $\rho_1$ with the diagonal line. This will be explored and explained in Section 3.

We applied the method to a simulated model with $p = 2$, $A_1 = A_2 = \sigma_\varepsilon = 1.0$, $\omega_1 = 0.7$ and $\omega_2 = 2.2$ ($\cos(\omega_1) = 0.7684$, $\cos(\omega_2) = -0.5885$) and we find the strong consistent estimates $\hat{\omega}_1 = 0.7044$ and $\hat{\omega}_2 = 2.1965$. We simulate a time series with N = 3000 observations and we considered M = 15. For an initial value as $\alpha_0 = \cos(0.4)$ we obtained

$\alpha_8 = 0.7620 = \cos(0.7044)$ and for $\alpha_0 = \cos(1.6)$ we obtained $\alpha_8 = 0.5857 = \cos(2.1965)$. Therefore, the estimated frequencies when p = 2 are $\hat{\omega}_1 = 0.7044$ and $\hat{\omega}_2 = 2.1965$.

In order to use the data of a time series of length N to obtain information about $\rho_1$, we need the sample autocovariance and variance to be consistent estimators. The strong consistency property of the estimators is proved in the Appendix of this paper.

## 2. General Definitions

For this section we shall give some definitions necessary for the whole understanding of the paper.

**Definition 2.1**: A Borel set $C \subseteq [-1, 1]$ is said to have *full measure* in [-1,1] if and only if

$$\mu(C) = \mu([-1, 1])$$

where $\mu$ is the Lebesgue measure in [-1,1].
Observe that $\mu(C^c) = 0$.

If a property is true in a set of full measure we will say that this *property is true almost surely (a.s.)*.

Suppose we have an updating scheme

$$\alpha_{k+1} = \rho_1(\alpha_k), \quad \text{for} \quad k \in \mathbf{N}, \tag{2.1}$$

applied to the process (1.1). We consider the following definition.

**Definition 2.2**: An updating scheme for the stochastic process (1.1) is *globally convergent* if there exist a set C of full measure in [-1,1] and $l \in \{1, 2, \cdots, p\}$ such that for any $\alpha_0 \in C$,

$$\lim_{k \to \infty} \alpha_k = \cos(\omega_l)$$

where $\omega_l$ is a frequency of the process (1.1).

The above limit expression may depend on $\alpha_0 \in C$.

In the case p = 1, the iterative procedure of *updating the alpha filter parameter* (see definition in He and Kedem (1989)) is *globally convergent*. However, this is not true for p = 2. Therefore, the relevant question is: how do we estimate all frequencies of the process (1.1) when $p \geq 2$? The *alpha filter* is not convenient for our purpose. In Section 4 we show that using the complex filter, and considering $\alpha_k$, with large $k$, we can have as good approximations for the frequencies as we want, by increasing M.

Suppose the updating scheme (2.1) depends on an extra parameter $M \in \mathbf{N} - \{0\}$, that is, $\rho_1(\alpha, M)$ is the first-order autocorrelation of the filtered process $\{Z_t(\alpha, M)\}_{t \in T}$. Then, we consider the following definition.

**Definition 2.3**: An updating scheme of the form (2.1) is said to be *approximately globally convergent* if for each fixed $M \in \mathbf{N} - \{0\}$ there exists a set $C_M$ of full measure in [-1,1] such that for any $\alpha_0 \in C_M$ there exists the limit

$$\lim_{k \to \infty} \alpha_k = \alpha_M^*.$$

The iterative updated procedure is considered with respect to the filter with parameter M and the value of $\alpha_M^*$ can depend on $\alpha_0$. We shall require in this definition that there exist $p$ of these possible values $\alpha_M^*$ and for each one of them there exists $l \in \{1, 2, \cdots, p\}$ such that

$$\lim_{M \to \infty} \alpha_M^* = \cos(\omega_l).$$

In simple terms, if we take an initial condition $\alpha_0$ at random and iterate the function $\rho_1$ $k$ times, if $k$ is large, then $\rho_1^k(\alpha_0) = \alpha_k = \rho_1(\alpha_{k-1})$ will be very close to the cosine of a frequency by taking M large.

The main purpose of the next section is to define a useful parametric filter family and show in Section 4 how to estimate the frequencies $\omega_j$, $1 \leq j \leq p$, of the process (1.1).

This work is part of the Ph.D. dissertation of the first author under the guidance of the second at the University of Maryland, College Park, 1991.

In a forthcoming paper we will present a method to estimate the amplitudes and the noise variance in the system (1.1).

## 3. Complex Filter

In this section we will consider a parametric family of filters. Consider the stochastic process $\{Z_t\}_{t \in T}$ as in (1.1).

**Definition 3.1**: A *parametric family* $\mathcal{L}_\theta$ *of linear time invariant filters* is defined as the set of filters

$$\{\mathcal{L}_\theta(\cdot) \ \ ; \ \ \theta \in \Theta\},$$

where $\Theta$ is the parameter space, with impulse response function $\{h_n(\theta)\}_{n=-\infty}^{\infty}$ and transfer function $H(\lambda;\theta)$ obtained from the Fourier Transform of the $h_n(\theta)$, that is,

$$H(\lambda;\theta) = \sum_{n=-\infty}^{\infty} \exp(-in\lambda)h_n(\theta).$$

For this to happen we consider the following matching condition

$$\sum_{n=-\infty}^{\infty} |h_n(\theta)|^2 < \infty$$

and that

$$\int_{T'} |H(\lambda;\theta)|^2 \, dF_Z(\lambda) < \infty,$$

where $T' = (-\pi, \pi]$ or $\mathbf{R}$ depending on the process being considered with discrete or continuous time parameter set T.

Let us denote $\{Z_t(\theta)\}_{t \in T}$ the filtered process defined by the convolution

$$Z_t(\theta) \equiv \mathcal{L}_\theta(Z)_t = \sum_{n=-\infty}^{\infty} h_n(\theta) Z_{t-n} = (h_\theta * Z)_t$$

where $*$ denotes convolution.

We shall consider a particular parametric family of linear filters where, from now on, $T = \mathbf{Z}$. Denote $\theta(\alpha) = \cos^{-1}(\alpha)$.

**Definition 3.2**: The *complex filter* applied to the process $\{Z_t\}_{t \in \mathbf{Z}}$ is defined by the transformation

$$Z_t(\alpha, M) = (1 + e^{i\theta(\alpha)}\mathcal{B})^M Z_t, \quad \text{for } \ t \in \mathbf{Z}, \ \ -1 < \alpha < 1 \ \text{ and } \ -\pi < \theta(\alpha) < \pi,$$

where M is a positive integer and $\mathcal{B}$ is the *shift operator* $\mathcal{B}Z_t = Z_{t-1}$. We think of $\theta(\alpha)$ as the *"center of the filter"*.

Clearly,

$$Z_t(\alpha, M) = \sum_{n=0}^{M} \binom{M}{n} e^{i\theta(\alpha)n} Z_{t-n}, \quad \text{for } \ t \in \mathbf{Z}, \ \ -\pi < \theta(\alpha) < \pi \ \text{ and } \ M \in \mathbf{N} - \{0\} \quad (3.1)$$

and the impulse response function is

$$h(n;\alpha,M)=\begin{cases}\binom{M}{n}e^{i\theta(\alpha)n}, & \text{for } 0\le n\le M\\ 0, & \text{otherwise}\end{cases}.$$

The transfer function is

$$H(\lambda;\alpha,M)=(1+e^{i(\theta(\alpha)-\lambda)})^M, \quad \text{for } \ -\pi<\lambda\le\pi$$

and the corresponding square gain function is

$$|H(\lambda;\theta(\alpha),M)|^2=4^M\cos^{2M}\left(\frac{\lambda-\theta(\alpha)}{2}\right), \quad \text{for } -\pi<\lambda,\theta\le\pi \ \text{ and } \ -1<\alpha<1. \quad (3.2)$$

If M is large then (see He and Kedem (1989))

$$\alpha=\cos(\theta(\alpha))\approx\frac{\int_{-\pi}^{\pi}\cos^{2M}(\frac{\lambda-\theta(\alpha)}{2})\cos(\lambda)\,d\lambda}{\int_{-\pi}^{\pi}\cos^{2M}(\frac{\lambda-\theta(\alpha)}{2})\,d\lambda}. \quad (3.3)$$

As we mention before, we will consider only large values of M. Therefore, we can change

$$\int_{-\pi}^{\pi}\cos^{2M}\left(\frac{\lambda-\theta(\alpha)}{2}\right)\cos(\lambda)\,d\lambda \quad \text{by} \quad \alpha\int_{-\pi}^{\pi}\cos^{2M}\left(\frac{\lambda-\theta(\alpha)}{2}\right)\,d\lambda.$$

In fact, already with $M=20$ the approximation is excellent. See Figure 1 for the graph of

$$\frac{\int_{-\pi}^{\pi}\cos^{2M}(\frac{\lambda-\theta}{2})\cos(\lambda)\,d\lambda}{\int_{-\pi}^{\pi}\cos^{2M}(\frac{\lambda-\theta}{2})\,d\lambda} \quad (3.4)$$

as a function of the variable $\theta$ for several values of M (M = 2, 11, 20) to appreciate the closeness of this quotient to $\cos(\theta)$.

Let $\{Z_t\}$ be a time series of length $N+M$ obtained from the process (1.1), when $T=\mathbf{Z}$, and $\{Z_t(\alpha,M)\}$ be the complex-filtered time series version.

Our analysis is based on iterations of the first order autocorrelation function of the filtered time series, that is, on iterations of the quotient between the autocovariance function of the filtered time series at lag 1 and its variance, where $E(Z_t)=0$, for all $t\in\mathbf{Z}$.

The first-order autocorrelation of the complex- filtered process $\{Z_t(\alpha,M)\}_{t\in\mathbf{Z}}$, where $Z_t$ is the process (1.1), is given by

$$\rho_1(\alpha,M)=\frac{\mathcal{R}\{E[Z_t(\alpha,M)\overline{Z_{t+1}(\alpha,M)}]\}}{E|Z_t(\alpha,M)|^2}=$$

$$= \frac{\sum_{j=1}^{p} \frac{A_j^2}{2}[\cos^{2M}(\frac{\omega_j+\theta(\alpha)}{2}) + \cos^{2M}(\frac{\omega_j-\theta(\alpha)}{2})]\cos(\omega_j) + \frac{\sigma_\varepsilon^2}{\pi}\int_{-\pi}^{\pi}\cos^{2M}(\frac{\lambda-\theta(\alpha)}{2})\cos(\lambda)\,d\lambda}{\sum_{j=1}^{p} \frac{A_j^2}{2}[\cos^{2M}(\frac{\omega_j+\theta(\alpha)}{2}) + \cos^{2M}(\frac{\omega_j-\theta(\alpha)}{2})] + \frac{\sigma_\varepsilon^2}{\pi}\int_{-\pi}^{\pi}\cos^{2M}(\frac{\lambda-\theta(\alpha)}{2})\,d\lambda} \tag{3.5}$$

or, from (3.3)

$\rho_1(\alpha, M) =$

$$= \frac{\sum_{j=1}^{p} \frac{A_j^2}{2}[\cos^{2M}(\frac{\omega_j+\theta(\alpha)}{2}) + \cos^{2M}(\frac{\omega_j-\theta(\alpha)}{2})]\cos(\omega_j) + \left[\frac{\sigma_\varepsilon^2}{\pi}\int_{-\pi}^{\pi}\cos^{2M}(\frac{\lambda-\theta(\alpha)}{2})\,d\lambda\right]\alpha}{\sum_{j=1}^{p} \frac{A_j^2}{2}[\cos^{2M}(\frac{\omega_j+\theta(\alpha)}{2}) + \cos^{2M}(\frac{\omega_j-\theta(\alpha)}{2})] + \frac{\sigma_\varepsilon^2}{\pi}\int_{-\pi}^{\pi}\cos^{2M}(\frac{\lambda-\theta(\alpha)}{2})\,d\lambda}. \tag{3.6}$$

From (3.6) $\rho_1(\alpha, M)$ is a weighted average of $\cos(\omega_j)$ and $\alpha$, a crucial observation that helps in recovering all the $\omega_j$. Define by $B_j^M(\alpha)$, for $0 \le j \le p$, the following weights

$$B_j^M(\alpha) = \frac{\frac{A_j^2}{2}[\cos^{2M}(\frac{\omega_j+\theta(\alpha)}{2}) + \cos^{2M}(\frac{\omega_j-\theta(\alpha)}{2})]}{\sum_{l=1}^{p} \frac{A_l^2}{2}[\cos^{2M}(\frac{\omega_l+\theta(\alpha)}{2}) + \cos^{2M}(\frac{\omega_l-\theta(\alpha)}{2})] + \frac{\sigma_\varepsilon^2}{\pi}\int_{-\pi}^{\pi}\cos^{2M}(\frac{\lambda-\theta(\alpha)}{2})\,d\lambda},$$

for $j \in \{1, 2, \cdots, p\}$ and

$$B_0^M(\alpha) = \frac{\frac{\sigma_\varepsilon^2}{\pi}\int_{-\pi}^{\pi}\cos^{2M}(\frac{\lambda-\theta(\alpha)}{2})\,d\lambda}{\sum_{l=1}^{p} \frac{A_l^2}{2}[\cos^{2M}(\frac{\omega_l+\theta(\alpha)}{2}) + \cos^{2M}(\frac{\omega_l-\theta(\alpha)}{2})] + \frac{\sigma_\varepsilon^2}{\pi}\int_{-\pi}^{\pi}\cos^{2M}(\frac{\lambda-\theta(\alpha)}{2})\,d\lambda}.$$

Therefore, $\rho_1(\alpha, M)$ is a weighted average of $\cos(\omega_j)$ and $\alpha$, that is,

$$\rho_1(\alpha, M) = \sum_{j=1}^{p} B_j^M(\alpha)\cos(\omega_j) + \alpha\, B_0^M(\alpha) \tag{3.7}$$

where the weights $B_j^M(\alpha)$, $0 \le j \le p$, are nonnegative, sum up to one and depend on $\alpha$ and M.

Given any $\alpha \in (-1, 1)$, $\rho_1(\alpha, M)$ will be in the convex hull (the intersection of all convex sets that contain the p+1 points) of $\alpha$ and $\cos(\omega_j)$, for $j \in \{1, 2, \cdots, p\}$.

Notice that the weight $B_0^M(\alpha)$, associated to the noise component $\sigma_\varepsilon$, multiplies the variable $\alpha$. The other weights $B_j^M(\alpha)$, $1 \le j \le p$, are associated to the signal component. Then, for $\alpha_1 = \rho_1(\alpha_0, M)$, $\alpha_2 = \rho_1^2(\alpha_0, M)$ is a convex combination of $\cos(\omega_j)$, $1 \le j \le p$, and $\alpha_1$ and so on. In this way, one can see the influence of the weights $B_j^M(\alpha)$. For instance, if one $B_l^M(\alpha)$ is much larger than $B_0^M(\alpha)$ and the others $B_j^M(\alpha)$, $1 \le j \le p$ and $j \ne l$, then there is a strong tendency of converging, by iteration of $\rho_1(\alpha, M)$, to a fixed point close to $\cos(\omega_l)$. Notice the important point that the weights $B_j^M(\alpha)$, $0 \le j \le p$, depend on $\alpha$. The reason for considering $M$ large is that for values of $\alpha$ close to $\cos(\omega_l)$ the relative value of $B_l^M(\alpha)$ is larger than the others $B_j^M(\alpha)$, for $j \in \{1, 2, \cdots, p\} - \{l\}$. Therefore, the weighted average above has more tendency of converging, through the recursive process, to an attracting fixed point very close to $\cos(\omega_l)$ when the value of $\alpha$ is close to $\cos(\omega_l)$. For each $j \in \{1, 2, \cdots, p\}$, $B_j^M(\alpha)$ depends on a multiplicative factor $A_j^2$. These factors have, of course, influence on the relative weights $B_j^M(\alpha)$, $1 \le j \le p$.

We think of $\rho_1(\alpha, M)$ as a function of $\alpha$ with parameter M.

**Remark 3.1**: One can see from the graph of $\rho_1(\alpha) = \rho_1(\alpha, M)$ as in expression (3.6) (see Figure 2) that this mapping is structural stable (Devaney (1989)). Therefore, the properties related to the iterations of the mapping as in (3.6) can be extended to the approximated mapping as in expression (3.5). This is a common procedure in the theory of iterations of mappings and it justifies the use of expression (3.6) instead of (3.5).

**Remark 3.2**: Observe that $\rho_1(\alpha, M)$ in expression (3.5) is defined by mathematical expectations. When we consider a time series of finite length $N + M$, we are assuming that the sample autocorrelation of size N is close to the expression (3.5), that is, $\hat{\rho}_1(\alpha, M)$ is a consistent estimate for $\rho_1(\alpha, M)$. The sample autocorrelation of size N when using the *complex filter* is given by

$$\hat{\rho}_1(\alpha, M) = \frac{\frac{1}{N}\mathcal{R}\{\sum_{j=1}^{N-1}[Z_j(\alpha, M) - \overline{Z(\alpha, M)}][\overline{Z_{j+1}(\alpha, M) - \overline{Z(\alpha, M)}}]\}}{\frac{1}{N}\sum_{j=1}^{N}[Z_j(\alpha, M) - \overline{Z(\alpha, M)}][\overline{Z_j(\alpha, M) - \overline{Z(\alpha, M)}}]}$$

where here the inner bar denotes the mean average value. In order to have $\hat{\rho}_1(\alpha, M)$ as a consistent estimator for $\rho_1(\alpha, M)$ it suffices to show that the sample autocovariance and sample variance are, respectively, consistent estimators for $\mathcal{R}\{E[Z_t(\alpha, M)\overline{Z_{t+1}(\alpha, M)}]\}$ and $E|Z_t(\alpha, M)|^2$. In the Appendix we derive these properties from the Ergodic Theorem.

The following claim is not difficult to show but rather technical and we will not prove it here.

**Claim**: *For each fixed $M \in \mathbf{N} - \{0\}$, $\rho_1(\alpha, M)$, as in expression (3.5), as a function of the variable $\alpha$, is a map from [-1,1] to [-1,1].*

We use the standard notation for iterations of mappings

$$\rho_1^2(\alpha) = \rho_1[\rho_1(\alpha)] = \rho_1 \circ \rho_1(\alpha)$$

and, in general,

$$\rho_1^k(\alpha) = \rho_1[\rho_1^{k-1}(\alpha)].$$

Denote by $\rho_1^k(\alpha)$ *the* $k^{\underline{th}}$ *iterate of the mapping* $\rho_1$ *at the value* $\alpha$. We refer Devaney (1989) for the properties of iterations of mappings that we use here.

We look for a method for finding the frequencies by iterating the mapping $\rho_1(\alpha, M)$, for $\alpha \in (-1, 1)$. Our reasoning is based on the geometric properties of the graph of $\rho_1(\alpha, M)$ and its derivative at the fixed points, for M large but fixed.

**Definition 3.3**: Let f be a smooth mapping from an interval into itself. A *fixed point* for the function f is a value $\alpha^*$ such that $f(\alpha^*) = \alpha^*$.

**Definition 3.4**: A value $\alpha^*$ is called an *attracting fixed point* of a mapping f(x) if $\alpha^*$ is a fixed point and $|f'(\alpha^*)| < 1$. It is called a *repelling fixed point* if $\alpha^*$ is a fixed point and $|f'(\alpha^*)| > 1$.

A graphic way to locate a fixed point of a mapping f is to look at the intersection of its graph with the diagonal line. In the figures presented here we always plot the graph of the function and the diagonal line. This makes easier to visualize the fixed points. We denote

$$\alpha_1 = f(\alpha_0), \quad \alpha_2 = f(\alpha_1) = f^2(\alpha_0)$$

and, in general,

$$\alpha_k = f(\alpha_{k-1}) = f^k(\alpha_0).$$

The set $\{\alpha_0, f(\alpha_0), f^2(\alpha_0), \ldots, f^k(\alpha_0), \ldots\}$ is called *the orbit of the point* $\alpha_0$.

**Property of Attractor Points**: *An attracting fixed point* $\alpha^*$ *has the property that nearby points* $\alpha$ *on both sides of* $\alpha^*$ *are attracted to* $\alpha^*$ *by iterations of f, that is,*

$$\lim_{k\to\infty} f^k(\alpha) = \alpha^* \quad \textit{for} \ \ \alpha \ \ \textit{near} \ \ \alpha^*.$$

Note that this is a local property and it does not mean necessarily that $\alpha^*$ is a global attractor point, that is, almost every point $\alpha_0 \in (-1, 1)$ will converge to $\alpha^*$.

**Property of Repelling Points**: *A repelling fixed point $\alpha^*$ has the property that nearby points $\alpha$, different from $\alpha^*$, on both sides of $\alpha^*$ are repelled from $\alpha^*$ by iterations of $f$.*

In practice, repelling points are not observable but attractor fixed points can be detected by high iterations of the mapping to an initial value $\alpha_0$ chosen at random.

We know that in the case where p = 1, by using the *alpha filter*, we can determine the unique frequency $\omega_1$. This is not true when p = 2.

The main point here is to show that for the *complex filter* the cosines of the frequencies $\omega_j$ are arbitrarily close to the attracting fixed points of the mapping $\rho_1$ when $M$ is large.

## 4. Fixed Bandwidth

Our interest is to estimate the frequencies $\omega_j$. With this purpose in mind we shall consider the parameter $M$ as being large but fixed. The frequencies $\omega_j$ are obtained by increasing the iterations of the mapping $\rho_1(\alpha_0, M)$ with $M$ fixed.

He and Kedem (1989) discuss the complex filter when there is no noise, that is when $\sigma_\varepsilon = 0$, if we take an $\alpha_0 \in (-1, 1)$ then

$$\lim_{M \to \infty} \rho_1(\alpha_0, M) = \cos(\omega_l)$$

where $\omega_l$ is the closest frequency to $\cos^{-1}(\alpha_0)$, that is,

$$|\alpha_0 - \cos(\omega_l)| < |\alpha_0 - \cos(\omega_j)|, \quad \text{for} \;\; 1 \le j \le p, \;\; j \ne l.$$

In this case, there is no need to iterate the mapping $\rho_1(\alpha, M)$.

The main result in this section is Theorem 4.1 that claims, in the presence of noise ($\sigma_\varepsilon \ne 0$), if we consider an $\alpha_0 \in (-1, 1)$ chosen at random and iterate $\rho_1^k(\alpha_0, M)$ then

$$\lim_{k \to \infty} \rho_1^k(\alpha_0, M)$$

will exist and it will be close to $\cos(\omega_l)$, where $\omega_l$ is such that

$$|\alpha_0 - \cos(\omega_l)| < |\alpha_0 - \cos(\omega_j)|, \quad \text{for} \;\; 1 \le j \le p, \;\; j \ne l.$$

Here we consider M large but fixed.

In a real situation we do not know *a priori* where the frequencies are. In any case, using the method described here, if one iterates $\rho_1(\alpha) = \rho_1(\alpha, M)$ starting from any initial value $\alpha_0$ (a.s. with respect to the Lebesgue measure on [-1,1]) then $\rho_1^k(\alpha_0, M)$ converges to a fixed point. Denote this fixed point by $\alpha^*_M$. Now, as M goes to infinity, the sequence $\{\alpha^*_M\}_{M>0}$ will converge to $\cos(\omega_j)$ (no iterations are used here when $M \to \infty$), where $\omega_j$ is the closest frequency to $\cos^{-1}(\alpha_0)$. In this way if we consider a sufficient large number of initial values chosen at random in the interval (-1,1) and iterate each one of them by $\rho_1(\alpha, M)$, we shall find very good approximations to all frequencies (if M is large enough). Therefore, we are also able to estimate the number of frequencies.

An alternative way to locate all frequencies is the following: for an initial value $\alpha_0$, consider the iterated function $\rho_1^k(\alpha_0, M)$ for $k$ large. In this way we will locate one frequency. Now we apply a very narrow bandpass filter to isolate this located frequency and we obtain a new time series with $p-1$ frequencies. The next step is to apply the same iterative procedure as above and then to locate another frequency. Now filter out this located frequency. Therefore, by using the same procedure again and again we locate all frequencies.

We show in Theorem 4.1 below that the method is *approximately globally convergent*, that is, there is no way that the iteration of $\rho_1(\alpha, M)$, beginning at an initial value $\alpha_0$ almost surely, will converge to something else that is not the cosine of an approximated frequency (if M is large enough). Before the proof of this theorem we show several pictures that will help to understand, in an intuitive way, why the method is *approximately globally convergent*.

A good indication of the above fact can be observed in the graph of $\rho_1(\alpha, M)$ in [-1,1] in Figures 2 and 3. These figures show, for the complex filter with M = 15 (in the case when p = 2) with $A_1 = A_2 = \sigma_\varepsilon = 1.0$, $\omega_1 = 0.7$ and $\omega_2 = 2.2$ ($\cos(\omega_1) = 0.7684$ and $\cos(\omega_2) = -0.5885$) the graph of, respectively, $\rho_1(\alpha, M)$ and $\rho_1^5(\alpha, M)$. We also plot the graph of the constant functions $\cos(\omega_1)$ and $\cos(\omega_2)$ in order to see how precise the method is when M is large. In all figures the graphs of the constant functions $\cos(\omega_1)$ and $\cos(\omega_2)$ are plotted by dotted lines.

Figures 4 and 5 show (in the case when p = 3) the graph of, respectively, $\rho_1(\alpha, M)$ and $\rho_1^{30}(\alpha, M)$ for the complex filter with M = 40, $A_1 = A_2 = A_3 = \sigma_\varepsilon = 1.0$ and frequencies $\omega_1 = 0.5$, $\omega_2 = 1.7$ and $\omega_3 = 2.4$ ($\cos(\omega_1) = 0.8775$, $\cos(\omega_2) = -0.1288$ and $\cos(\omega_3) = -0.7373$). From the graph of $\rho_1(\alpha, M)$ in Figures 2 and 4 one can see that the only attracting fixed points are very close to the cosine of the true frequencies. There exist other fixed points but they are repelling ones. From the considerations made just after the Definition 3.4 for repelling fixed points, we know that they will not attract iterations of an initial value $\alpha_0$ by $\rho_1(\alpha, M)$ (a.s.).

Note that there exist other fixed points for the mapping $\rho_1$ different from $\cos(\omega_j)$, $j \in \{1, 2, \cdots, p\}$, but they are all repelling fixed points.

**Remark 4.1**: In a compact set, the number of zeros of a real analytic function f(x) - x is finite (see Rudin (1987)). Then, the set of fixed points and, more specifically, the set of repelling fixed points is finite. Therefore, it has Lebesgue measure zero. The set $C_M$ in Definition 2.1 is the interval [-1,1] without the repelling fixed points of $\rho_1(\alpha, M)$.

If we increase M the attracting fixed points $\alpha^*_M$ will be closer and closer to the true frequencies $\omega_j$. This can be seen in Figure 6 where we consider the graph of $\rho_1(\alpha, M)$ (in the case when p = 2) in an interval very close to the cosine of the frequency $\omega_1 = 0.7$, where $\cos(\omega_1) = 0.7684$. Notice that by increasing M, the fixed point $\alpha_M^*$ will be as close as one wants to $\cos(\omega_1)$. The intersection of the diagonal line and the graph of $\rho_1(\alpha, M)$ show where the fixed point is located. Figure 7 shows the same situation when we look at the graph of $\rho_1(\alpha, M)$ in a small interval very close to the cosine of the other frequency $\omega_2 = 2.2$, where $\cos(\omega_2) = -0.5885$.

**Remark 4.2**: We would like to mention here that one must not confuse M and $k$. First, we fix M and consider an $\alpha_0$ chosen at random, and then we consider $\rho_1^k(\alpha_0, M)$ for a large $k$. The sequence $\{\rho_1^k(\alpha_0, M)\}_{k\geq 1}$ converges to a fixed point very close to the cosine of one of the frequencies. We consider $\rho_1^k(\alpha_0, M)$, for a certain large M, as a good approximation for the cosine of the frequency to be detected. This is a different approach from the one in Kedem and Lopes (1991), where one shrinks the bandwidth (by increasing $M$) at each iteration of $\alpha_{k+1} = \rho_1(\alpha_k, M_k)$.

The method presented above is reminiscent of Newton's Method for locating the roots of a polynomial equation. If the initial value is very close to a certain root, the iterative procedure of Newton's Method will converge to this root. If the initial value is close to another root then the iterative procedure will converge to that other root.

We now give a rigorous proof of why this method works well. The important point here is the weighted average property of $\rho_1(\alpha, M)$ (see (3.7)).

**Theorem 4.1**: *The family of complex filters is approximately globally convergent.*

**Proof**:

First we give the proof of the following claim.

**Claim 1**: *The relative masses of the weights $B_j^M(\alpha)$ (see expression (3.7)) are in such way that if the initial value $\alpha_0$ is closer to $\cos(\omega_l)$ than the other $\cos(\omega_j)$, $1 \leq j \leq p$, $j \neq l$, then*

$$\lim_{M\to\infty} \frac{B_j^M(\alpha_0)}{B_l^M(\alpha_0)} = 0, \quad \textit{for} \ \ 0 \leq j \leq p \ \ \textit{and} \ \ j \neq l. \tag{4.1}$$

**Proof of Claim 1**:

Since $\cos(\frac{x}{2})$ is monotone decreasing in $(0, \pi)$, then

$$\left|\cos(\frac{\omega_l - \theta(\alpha_0)}{2})\right| > \left|\cos(\frac{\omega_j \pm \theta(\alpha_0)}{2})\right|, \quad \text{for} \ \ 1 \leq j \leq p \ \ \text{and} \ \ j \neq l \tag{4.2}$$

where $\theta(\alpha_0) = \cos^{-1}(\alpha_0)$ because $\theta(\alpha_0)$ is closer to $\omega_l$ than $\omega_j$, for $j \neq l$.
Notice that for a fixed M, each $B_j^M(\alpha_0)$, $1 \leq j \leq p$, is related to

$$\cos^{2M}\left(\frac{\omega_j + \theta(\alpha_0)}{2}\right) + \cos^{2M}\left(\frac{\omega_j - \theta(\alpha_0)}{2}\right). \tag{4.3}$$

Observe that the values in (4.2) are raised to the power 2M in the expression (4.3). Therefore, if M is large then the quotient

$$\left| \frac{B_j^M(\alpha)}{B_l^M(\alpha)} \right|, \quad \text{for } 1 \le j \le p \text{ and } j \ne l,$$

will be close to zero if $\alpha_0$ is close to $\cos(\omega_l)$ since large values of $M$ amplify the difference between the weights when one is close to the cosine of one frequency. For each $j \in \{1, 2, \cdots, p\}$, the influence of the multiplicative term (4.3) is of higher order than the multiplicative term $A_j^2$ when M is large. In this case, the successive iteration of $\rho_1(\alpha_0, M)$ (the weighted average applied to an initial value $\alpha_0$) will have a strong bias in the direction of $\cos(\omega_l)$ (the closest one to $\alpha_0$). For $j = 0$, the above claim follows from the fact that

$$\lim_{M \to \infty} \int_{-\pi}^{\pi} \cos^{2M} \left( \frac{\lambda - \theta(\alpha_0)}{2} \right) d\lambda = 0.$$

This is the end of the proof of Claim 1.

Now we want to prove that for any small interval $I = (\alpha_a, \alpha_b) \subseteq (-1, 1)$ containing $\cos(\omega_l)$, for some $l \in \{1, 2, \cdots, p\}$, there exists $M_0$ large enough such that, for all $M > M_0$, there exists $\alpha_M^* \in I$ satisfying

$$\rho_1(\alpha_M^*, M) = \alpha_M^*,$$

that is, $\rho_1(\alpha, M)$ has a fixed point $\alpha_M^* \in I$.

Let $\alpha_a$ and $\alpha_b$ be any two values in (-1,1) such that there exists $l \in \{1, 2, \cdots, p\}$ where

$$\alpha_a \le \cos(\omega_l) \le \alpha_b$$

and

$$|\alpha_a - \cos(\omega_l)|, \quad |\alpha_b - \cos(\omega_l)| < \min_{\substack{1 \le j \le p \\ j \ne l}} \{|\alpha_a - \cos(\omega_j)|, \quad |\alpha_b - \cos(\omega_j)|\}$$

such that

$$\cos\left(\frac{\omega_l \pm \theta(\alpha)}{2}\right) \ge \left|\cos\left(\frac{\omega_j \pm \theta(\alpha)}{2}\right)\right|, \quad \text{for } 1 \le j \le p \quad \text{and} \quad j \ne l,$$

with $\theta(\alpha) = \cos^{-1}(\alpha)$.

We will show that for M large enough

$$\alpha_a < \rho_1(\alpha_a, M) < \cos(\omega_l) < \rho_1(\alpha_b, M) < \alpha_b.$$

Since $\rho_1(\alpha, M)$ is the convex combination (3.7) and the weights $B_j^M(\alpha)$, $j \in \{0, 1, \cdots, p\}$, sum up to one for any $\alpha \in (-1, 1)$, we have

$$
\begin{aligned}
\rho_1(\alpha, M) &= \sum_{j=1}^{p} B_j^M(\alpha)\cos(\omega_j) + \alpha\, B_0^M(\alpha) \\
&= B_l^M(\alpha)\cos(\omega_l) + \sum_{\substack{j=1\\ j\neq l}}^{p} B_j^M(\alpha)\cos(\omega_j) + \alpha\, B_0^M(\alpha) \\
&= (1 - B_0^M(\alpha) - \sum_{\substack{j=1\\ j\neq l}}^{p} B_j^M(\alpha))\cos(\omega_l) + \alpha\, B_0^M(\alpha) + \sum_{\substack{j=1\\ j\neq l}}^{p} B_j^M(\alpha)\cos(\omega_j) \\
&= \cos(\omega_l) + B_0^M(\alpha)(\alpha - \cos(\omega_l)) - \cos(\omega_l)\sum_{\substack{j=1\\ j\neq l}}^{p} B_j^M(\alpha) + \sum_{\substack{j=1\\ j\neq l}}^{p} B_j^M(\alpha)\cos(\omega_j).
\end{aligned} \tag{4.4}
$$

As $\omega_l$ is the frequency closest to the cosine of the initial value $\alpha_b$ with $\cos(\omega_l) < \alpha_b$ then for any fixed $j \in \{1, 2, \cdots, p\}$, $j \neq l$, it holds (4.1), that is,

$$
\limsup_{M\to\infty} \frac{B_j^M(\alpha_b)}{B_l^M(\alpha_b)} = 0, \quad \text{for} \quad j \neq l.
$$

This fact was shown in Claim 1.

Hence, from (4.1), taking limit sup in both sides of expression (4.4), for $\alpha = \alpha_b$, we have

$$
\limsup_{M\to\infty} \frac{\rho_1(\alpha_b, M)}{B_l^M(\alpha_b)} = \limsup_{M\to\infty} \frac{\cos(\omega_l) + B_0^M(\alpha_b)[\alpha_b - \cos(\omega_l)]}{B_l^M(\alpha_b)}.
$$

However, $0 < B_0^M(\alpha_b) < 1$ (one notices here the strict inequality) and $\alpha_b - \cos(\omega_l) > 0$. Hence,

$$
\limsup_{M\to\infty} \frac{\rho_1(\alpha_b, M)}{B_l^M(\alpha_b)} < \limsup_{M\to\infty} \frac{\cos(\omega_l) + 1\,(\alpha_b - \cos(\omega_l))}{B_l^M(\alpha_b)} = \limsup_{M\to\infty} \frac{\alpha_b}{B_l^M(\alpha_b)}.
$$

The weights $B_j^M(\alpha)$, for $j \in \{0, 1, \cdots, p\}$, sum up to one for any $\alpha \in (-1, 1)$. Hence, from (4.1) one has

$$
\lim_{M\to\infty} B_j^M(\alpha) = 0, \quad \text{for} \quad j \neq l, \quad \text{and} \quad \lim_{M\to\infty} B_l^M(\alpha) = 1.
$$

Therefore, since

$$\limsup_{M\to\infty} \rho_1(\alpha_b, M) = \limsup_{M\to\infty} \frac{\rho_1(\alpha_b, M)}{B_l^M(\alpha_b)} < \alpha_b,$$

there exists $M \in \mathbf{N} - \{0\}$ such that

$$\rho_1(\alpha_b, M) < \alpha_b.$$

Note the strict inequality.

By similar argument, we will show that $\cos(\omega_l) \leq \rho_1(\alpha_b, M)$. Again, from equality (4.4), we have

$$\begin{aligned}\frac{\rho_1(\alpha_b, M)}{B_l^M(\alpha_b)} &= \frac{\cos(\omega_l)}{B_l^M(\alpha_b)} + \frac{B_0^M(\alpha_b)(\alpha_b - \cos(\omega_l))}{B_l^M(\alpha_b)} \\ &\quad - \cos(\omega_l) \sum_{\substack{j=1 \\ j\neq l}}^{p} \frac{B_j^M(\alpha_b)}{B_l^M(\alpha_b)} + \sum_{\substack{j=1 \\ j\neq l}}^{p} \frac{B_j^M(\alpha_b)}{B_l^M(\alpha_b)} \cos(\omega_j).\end{aligned}$$

Since $0 < B_0^M(\alpha_b) < 1$ and $\alpha_b - \cos(\omega_l) > 0$, from expression (4.1) we have

$$\cos(\omega_l) = \frac{\cos(\omega_l)}{\lim_{M\to\infty} B_l^M(\alpha_b)} = \limsup_{M\to\infty} \frac{\cos(\omega_l)}{B_l^M(\alpha_b)} \leq \limsup_{M\to\infty} \frac{\rho_1(\alpha_b, M)}{B_l^M(\alpha_b)} = \limsup_{M\to\infty} \rho_1(\alpha_b, M).$$

Therefore, there exists $M \in \mathbf{N} - \{0\}$ such that

$$\cos(\omega_l) \leq \rho_1(\alpha_b, M).$$

We conclude that

$$\cos(\omega_l) \leq \rho_1(\alpha_b, M) < \alpha_b.$$

Similarly, since $\omega_l$ is the frequency closest to the cosine of the initial value $\alpha_a$, one can show that

$$\alpha_a < \rho_1(\alpha_a, M) \leq cos(\omega_l).$$

Recall that $I = (\alpha_a, \alpha_b) \subset (-1, 1)$. We have shown that $\rho_1(I, M) \subset I$. Therefore, by Brouwer Fixed Point Theorem (see Proposition 2.11 in Devaney (1989)), there exists a real value $\alpha_M^* \in (-1, 1)$ such that

$$\rho_1(\alpha_M^*, M) = \alpha_M^* \quad \text{and} \quad \alpha_M^* \in I.$$

Now we want to show that the fixed point $\alpha_M^*$ is unique and attracting.

The mapping $\rho_1(\cdot\,, M)$ is defined on the variable $\theta(\alpha) \in (-\pi, \pi)$, for any $\alpha \in (-1, 1)$. The extension of the mapping $\rho_1(\cdot\,, M)$ to a neighborhood $V$ of $(-1, 1)$ in the Complex Plane is now considered. This extension will be necessary for using a strong form of Schwarz Lemma (see Hervé (1963)) and it will be made clear later. Let us consider the complex analytic mapping

$$\rho_1(\cdot\,, M) \colon V \to \mathbf{C}$$

such that $\alpha \to \rho_1(\alpha, M)$ is given by

$$\rho_1(\alpha, M) = \sum_{j=1}^{p} B_j^M(\alpha) \cos(\omega_j) + \alpha\, B_0^M(\alpha)$$

for $\alpha \in V \subset \mathbf{C}$, where the weights $B_j^M(\alpha)$, for $0 \leq j \leq p$, are defined by the expressions given before (3.7).

Let a function f be defined by

$$f \colon (-\pi, \pi) \to \mathbf{C}$$

such that $t \to f(t) = \cos(\omega_l) + re^{it}$ with $r \in \mathbf{R}$, $r > 0$.

We want to show that any circle of center $\cos(\omega_l)$ and small radius $r$ is contracted by the transformation $\rho_1(\cdot\,, M)$, if M is large enough. That is, we want to show that, for a fixed small value $r$, there exists $M \in \mathbf{N} - \{0\}$ such that

$$\left| \frac{\rho_1(\alpha, M) - \cos(\omega_l)}{\alpha - \cos(\omega_l)} \right| < 1,$$

uniformly for all $\alpha$ of the form $f(t) = \cos(\omega_l) + re^{it}$, $t \in (-\pi, \pi)$. Observe that

$$|\alpha - \cos(\omega_l)| = |f(t) - \cos(\omega_l)| = |re^{it}| = r.$$

**Claim 2**:

$$\limsup_{M\to\infty}\left|\frac{B_j^M(\alpha_b)}{B_l^M(\alpha_b)}\right| = 0 \quad \textit{for } \ j \neq l \ \ \textit{and any} \ \ \alpha \in V \ \ \textit{close enough to} \ \ \omega_l. \tag{4.5}$$

**Proof of Claim 2**:

The proof of this claim is similar to the case when the mapping $\rho_1(\cdot\,, M)$ was considered defined only in the interval (-1,1). Then, we suppose Claim 2 is proved.

Since $\sum_{j=0}^{p} B_j^M(\alpha) = 1$, from the above claim one has

$$\lim_{M\to\infty} |B_j^M(\alpha)| = 0, \quad \text{for all } \ j \neq l, \quad \text{and} \quad \lim_{M\to\infty} |B_l^M(\alpha)| = 1$$

for all $\alpha \in (-1, 1)$. Considering the analytic function $\rho_1(\cdot\,, M)$ at $\alpha = \cos(\omega_l) + re^{it}$, for $t \in (-\pi, \pi)$, and applying Claim 2 we get

$$\limsup_{M\to\infty} |\rho_1(\alpha, M) - \cos(\omega_l)| = \frac{\limsup_{M\to\infty} |\rho_1(\alpha, M) - \cos(\omega_l)|}{\limsup_{M\to\infty} |B_l^M(\alpha)|}$$

$$= \limsup_{M\to\infty}\left|\frac{\rho_1(\alpha, M) - \cos(\omega_l)}{B_l^M(\alpha)}\right|$$

$$= \limsup_{M\to\infty}\left|\frac{B_0^M(\alpha)(\alpha - \cos(\omega_l)) - \cos(\omega_l)\sum_{\substack{j=1\\ j\neq l}}^{p} B_j^M(\alpha) + \sum_{\substack{j=1\\ j\neq l}}^{p} B_j^M(\alpha)\cos(\omega_j)}{B_l^M(\alpha)}\right|$$

$$\leq \limsup_{M\to\infty} \frac{|B_0^M(\alpha)|}{|B_l^M(\alpha)|}|\alpha - \cos(\omega_l)| + 0 < \limsup_{M\to\infty} \frac{1}{|B_l^M(\alpha)|}|re^{it}| = r.$$

Therefore, there exists $M \in \mathbf{N} - \{0\}$ such that

$$|\rho_1(\alpha, M) - \cos(\omega_l)| < r = |\alpha - \cos(\omega_l)|,$$

for all $\alpha$ of the form $f(t) = \cos(\omega_l) + re^{it}$, for $t \in (-\pi, \pi)$. Note the strict inequality.

We conclude that the circle with center in $\cos(\omega_l)$ and small radius $r > 0$ is contracted by $\rho_1(\cdot\,, M)$.

Denote by $U$ the ball of center $\cos(\omega_l)$ and radius $r$ in $\mathbf{C}$. Since an analytic mapping is an open mapping, the set $U$ is mapped by $\rho_1(\cdot\,, M)$ inside $U$.

Now we recall a strong version of Schwarz Lemma (see Hervé (1963), page 83).

**Theorem**: *Suppose $U$ is a simply connected open subset of* $\mathbf{C}$ *not equal to* $\mathbf{C}$ *itself. Suppose $F: U \to U$ is complex analytic and the closure of $F(U)$ is contained in $U$. Then $F$ has a fixed point $z_0 \in U$ and*

- $|F'(z_0)| < 1$ *and* $F^n(z) \to z_0$, *for all* $z \in U$.

Therefore, there exists a unique fixed point for $\rho_1(\cdot\,, M)$ in the set $U = B(\cos(\omega_l), r)$ in $\mathbf{C}$ and this fixed point is an attractor for the set $U$. We have shown before the existence of a real fixed point $\alpha_M^*$. From the above we conclude that $\alpha_M^*$ is the unique attracting fixed point in $U$. Then, this point $\alpha_M^*$ attracts all the real values in a small neighborhood of $\cos(\omega_l)$.

■

**Remark 4.3**: The expressions (3.5) and (3.6) are very close if M is large. This follows from (3.3). In Figures 8 and 9 we plot the graph of (3.5) and (3.6) for p = 2, M = 20 where $\omega_1 = 0.7$, $\omega_2 = 2.2$, $A_1 = A_2 = \sigma_\varepsilon = 1.0$. One can see that, if M is large then the graph given by the expressions (3.5) and (3.6) are almost the same. Theorem 4.1 shows, among other things, that the mapping given by the expression (3.6) is structural stable and, therefore, our reasoning using (3.6) instead of (3.5) is justifiable.

**Conclusion**:

First consider, for simplification of the argument, the case p = 2. We consider a large number of equally spaced initial values $\alpha_0$ (for instance, 10) in (-1,1). For each $\alpha_0$ we take $\rho_1^8(\alpha_0, M)$ as a good approximation for the cosine of the frequencies. Some of these values will be very close to $\cos(\omega_1)$ and some of them will be close to $\cos(\omega_2)$. If M is large, the values obtained by the above procedure will be so close to $\cos(\omega_1)$ or $\cos(\omega_2)$ as one wants. We will choose among these ten values $\rho_1^8(\alpha_0, M)$ two of them that are distant apart. We will denote these two values $\hat{\omega}_1$ and $\hat{\omega}_2$ the estimated frequencies. In general, M = 20 is good enough when p = 2 and $\omega_1$ and $\omega_2$ are distant apart. Two examples are provided in Table 4.1 and Table 4.2.

In the general case, when we have $p \geq 3$ frequencies, we will consider a large number of equally spaced initial values $\alpha_0 \in (-1, 1)$ and with the same iterative procedure we will get approximated values of $\omega_j$, $1 \leq j \leq p$, when M is large. We will choose p among these values $\rho_1^8(\alpha_0, M)$ that are distant apart and we will denote $\hat{\omega}_j$, $1 \leq j \leq p$, the estimated frequencies.

**Remark 4.4**: A problem that can appear in the method is when two frequencies are close. In some cases, we will need to take M = 200 requiring more computational time due to the calculations of the binomial coefficients (see expression (3.1)). The method still works but the convergence can be very slow for some initial values $\alpha_0$. Other filters, with narrower band like $AR(2)$, for instance, will do better.

It is convenient in the numerical implementation of the method to compute the binomial coefficients (see expression (3.1)) in the beginning of the code and store them in order to decrease the computational complexity (remember that M is fixed) .

## Appendix: Ergodicity of the Stochastic Process

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be the probability space where $\Omega$ is the sample space, $\mathcal{F}$ is the $\sigma$-algebra of Borel sets and $\mathbf{P}$ is a probability function on $\Omega$. Consider T a transformation defined from $\Omega$ to itself, so that T is measurable and also measurably invertible.

**Definition 1**: We say that $\mathbf{P}$ is an *invariant measure for $T$* or *$T$ is measure-preserving* if $\mathbf{P}(T^{-1}(A)) = \mathbf{P}(A)$, for any Borel set $A \in \mathcal{F}$.

**Definition 2**: We say that $\mathbf{P}$ is *ergodic for $T$*, if for any Borel set A such that $T^{-1}(A) = A$, we have that $\mathbf{P}(A) = 0$ or $\mathbf{P}(A) = 1$.

A very important result is the Birkhoff Ergodic Theorem (see Skorokhod (1989)). We next state this theorem.

**Birkhoff Ergodic Theorem**: Suppose V is an integrable random variable on $\Omega$, $\mathbf{P}$ is a probability invariant measure on $\Omega$ and T is a measurable transformation on $\Omega$. Let $\mathcal{G}$ be the smallest $\sigma$-algebra of sets in $\mathcal{F}$ with respect to which all random variables $W$ with $W(T^t(\omega)) = W(\omega)$ for $\mathbf{P}$-almost all $\omega$ and for $t > 0$ are measurable. Then,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} V(T^t(\omega)) = E(V/\mathcal{G})(\omega) \qquad \mathbf{P}\text{ \_ } a.s..$$

When $\mathbf{P}$ is ergodic (that is, $\mathcal{G}$ is trivial) then $E(V/\mathcal{G})$ reduces to $E(V) =$ constant and the above result essentially says that for the typical trajectory with respect to $\mathbf{P}$, time average of V converge to spatial average of V.

In terms of stochastic processes, we are considering in the above setting the stationary process $X_t(\omega) = V(T^t(\omega))$, $\omega \in \Omega$ and $t \in \mathbf{Z}$. This is the standard way to transfer results from transformations with invariant measures to stationary processes (we refer to Lamperti (1977), chapter 5 for further details). Basically, one has to consider on the space $\Omega^{\mathbf{N}}$, the product measure generated by $\mathbf{P}$ on $\Omega$ and the above defined stochastic process $X_t$. We remark here that $\mathbf{P}$ will be a product measure in the case of independent and identical distributed coordinates.

**Remark 1**: Suppose that $\int V(\omega)\,\mathbf{P}(d\omega) = 0$. Then, in this case, if the probability is ergodic, the autocovariance at lag $k$

$$\int V(\omega)V(T^k(\omega))\,\mathbf{P}(d\omega)$$

can be obtained as the almost-sure limit of the mean

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} V(T^t(\omega))V(T^{t+k}(\omega)), \quad \text{for} \quad k \geq 0.$$

In this way, we can say that the sample autocovariance (the case k=1) and variance (the case k=0) are consistent estimators.

In our case we will need to consider $\Omega = (-\pi, \pi]$ and for any $\omega \in \Omega$, we have

$$T(\omega) = \omega + \omega_1 \pmod{2\pi},$$

where $\omega_1$ is a fixed number in the interval $(-\pi, \pi]$. Now $\mathbf{P}$ will be the normalized Lebesgue measure on $(-\pi, \pi]$, and this probability $\mathbf{P}$ is clearly invariant for T.

It is well known (see Cornfeld, Fomin and Sinai (1984), page 64) that when $\frac{w_1}{2\pi}$ is an irrational number, then $\mathbf{P}$ is *ergodic for* $T$.

**Remark 2**: The Ergodic Theorem in the case when $\frac{w_1}{2\pi}$ is irrational, is true in a stronger form than the one provided by Birkhoff Ergodic Theorem. In fact, if $V$ is continuous, the statement about time averages is true, not only $\mathbf{P}$- almost surely, but in fact for *all* $\omega \in (-\pi, \pi]$. The analogous statement for numbers $\omega_1$, such that $\frac{\omega_1}{2\pi}$ is rational is false.

Now let us concentrate on the specific case we want to understand here. We will denote elements in our space $\Omega$ by $\phi$, in order to have a coherent notation with the one we used previously. We will need here to consider the random variable $V(\omega) = V(\phi) = A\cos(\phi)$. Notice that

$$\int V(\omega)\, P(d\omega) = A \int \cos(\phi)\, \mathbf{P}(d\phi) = 0.$$

Therefore, the assumption of Remark 1 is satisfied.

Note that for any $n \in \mathbf{N}$ and $\phi \in (-\pi, \pi]$, we have that $T^n(\phi) = \phi + n\, w_1 \pmod{2\pi}$.

If $\frac{\omega_1}{2\pi}$ is irrational then we can apply the Ergodic Theorem for the random variable $V(\phi)V(T^k(\phi))$, because $\mathbf{P}$ is ergodic (see Remark 1). In this way, we have a consistent estimator for the autocovariance.

Therefore, from the Ergodic Theorem it follows that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} V(T^t(\phi))V(T^{t+k}(\phi)) = \lim_{N\to\infty} \frac{1}{N} A^2 \sum_{t=0}^{N-1} \cos(\omega_1\, t + \phi)\cos(\omega_1\,(t+k) + \phi)$$
$$= A^2 \int \cos(\phi)\cos(\omega_1\, k + \phi)\, \mathbf{P}(d\phi) = A^2 \int \cos(\phi)\cos(T^k(\phi))\, \mathbf{P}(d\phi), \quad \text{for} \quad k \geq 0.$$

Therefore, for any $\phi \in (-\pi, \pi]$, we have that the sample means of the autocovariance give an almost sure consistent estimator for the autocovariance of the process $\{X_t\}_{t \in \mathbf{Z}}$. The analogous statement for the variance is also true.
Recall that we first consider the stochastic process

$$X_t(\phi) = V(T^t(\phi)) = V(\phi + t\,\omega_1) = A\cos(\omega_1\, t + \phi).$$

Now we will add Gaussian white noise process $\varepsilon_t$ to $X_t$.

Therefore, we want to analyze autocovariance and variance for the process $Z_t$ given by

$$Z_t = X_t + \varepsilon_t = V(T^t(\cdot)) + \varepsilon_t.$$

For the autocovariance we have to consider the following sum

$$Z_t(\phi) Z_{t+1}(\phi) = X_t(\phi) X_{t+1}(\phi) + X_t(\phi)\,\varepsilon_{t+1} + X_{t+1}(\phi)\,\varepsilon_t + \varepsilon_t\,\varepsilon_{t+1}.$$

In the case we want to take samples for the autocovariance the above equality is given as follows

$$\begin{aligned}\frac{1}{N}\sum_{t=0}^{N-1} Z_t(\phi) Z_{t+1}(\phi) &= \frac{1}{N}\sum_{t=0}^{N-1} V(T^t(\phi)) V(T^{t+1}(\phi)) \\ &+ \frac{1}{N}\sum_{t=0}^{N-1} V(T^t(\phi))\,\varepsilon_{t+1} + \frac{1}{N}\sum_{t=0}^{N-1} V(T^{t+1}(\phi))\,\varepsilon_t \\ &+ \frac{1}{N}\sum_{t=0}^{N-1} \varepsilon_t\,\varepsilon_{t+1}.\end{aligned}$$

The sample means corresponding to the first term on the right hand side of the above equality were analyzed by previous considerations using the Ergodic Theorem (that means, when we have only the process $X_t$).

The sample means corresponding to the second and third terms on the right hand side of the above equality converge to zero since from the uncorrelatedness of the variables $\varepsilon_t$ and $X_t = V(T(\phi))$ we have

$$\lim_{N\to\infty} \frac{1}{N}\sum_{t=0}^{N-1} V(T^t(\phi))\,\varepsilon_{t+k} = 0, \quad \text{for} \quad k \geq 0.$$

Finally, the sample means in the fourth term converge to zero, from the hypothesis of uncorrelatedness of the variables $\varepsilon_t$ in the definition of white Gaussian noise. More precisely,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} \varepsilon_t \, \varepsilon_{t+k} \; = \; 0.$$

Therefore, we conclude that for the process

$$Z_t = A\cos(\omega_1 \, t + \phi) + \varepsilon_t,$$

where $\frac{\omega_1}{2\pi}$ is irrational and $\{\varepsilon_t\}_{t\in\mathbf{Z}}$ is Gaussian white noise, the sample autocovariance is an almost sure consistent estimator for the autocovariance of the process $\{Z_t\}_{t\in\mathbf{Z}}$. We mention here that, for simplicity, the noise component is assumed to be Gaussian white noise but the reasoning holds more generally for any ergodic colored noise.

When one wants the variance, one just has to consider the case *k=0* in the above considerations.

Therefore, we conclude that for the process

$$Z_t = A\cos(\omega_1 \, t + \phi) + \varepsilon_t,$$

where $\frac{\omega_1}{2\pi}$ is irrational and $\{\varepsilon_t\}_{t\in\mathbf{Z}}$ is Gaussian white noise, the sample variance is an almost sure consistent estimator for the variance of the process $\{Z_t\}_{t\in\mathbf{Z}}$.

Then, we can also take the sample variance to estimate the variance of the process $\{Z_t\}_{t\in\mathbf{Z}}$. Therefore, $\hat{\rho}_1(\alpha)$ is an almost sure consistent estimator for $\rho_1(\alpha)$ as mentioned in Remark 3.2 of Section 3.

It follows that the sample means of the variance and autocovariance converge, respectively, to

$$E[Z_t^2] = \frac{A^2}{2} + \sigma_\varepsilon^2$$

and

$$E[Z_t Z_{t+1}] = \frac{A^2}{2} \cos(\omega_1).$$

Now we will briefly explain how to extend the above results to the process

$$Z_t = \sum_{j=1}^{p} A_j \cos(\omega_j \, t + \phi_j) + \varepsilon_t$$

where $A_j$, $\omega_j$, $\phi_j$ and the $\varepsilon_t$ were previously defined (see expression (1.1)).

In this case we consider $\Omega$ as the $p$-torus $(-\pi, \pi]^p$, and let the transformation T be defined as

$$T(\phi_1, \phi_2, \cdots, \phi_p) = (\omega_1 + \phi_1 \,(\text{mod } 2\pi), \omega_2 + \phi_2 \,(\text{mod } 2\pi), \cdots, \omega_p + \phi_p \,(\text{mod } 2\pi))$$

for any p-uple $\omega = (\phi_1, \phi_2, \cdots, \phi_p)$.

The measure $\mathbf{P}$, in this case, will be the normalized product measure on the torus when we consider the Lebesgue measure in the interval $(-\pi, \pi]$.

We refer to Cornfeld, Fomin and Sinai (1984), page 64, for a careful analysis of the above mentioned situation.

First we want to analyze the signal component in $\{Z_t\}_{t\in\mathbf{Z}}$, that is, we want to analyze the process

$$X_t = \sum_{j=1}^{p} A_j \cos(\omega_j\, t + \phi_j), \quad \text{where } \ A_j \ \text{ are unknown constants.}$$

In the case when all $\frac{\omega_j}{2\pi}$ are irrational and rationally independent (that is, $\sum_{j=1}^{p} s_j \frac{\omega_j}{2\pi} = q$ where $s_j$ and $q$ are integers, is possible only when $s_1 = s_2 = \cdots = s_p = 0$), for any $j \in \{1, 2, \cdots, p\}$, the above probability $\mathbf{P}$ in the torus is ergodic for the map T defined above (see page 64 in Cornfeld, Fomin and Sinai), and results similar to the ones in Remark 1 can also be applied to the random variable

$$V(\phi_1, \phi_2, \cdots, \phi_p) = A_1 \cos(\phi_1) + A_2 \cos(\phi_2) + \cdots + A_p \cos(\phi_p).$$

Notice that we can assume, without loss of generality, that the frequencies are irrational and also rationally independent, because the set of such frequencies has probability one among the possible values of frequencies.

Therefore, it follows that the samples of the autocovariance and variance are almost sure consistent estimators also in the case when we have p irrational frequencies.

Now if we introduce an additive white noise to the above defined stochastic process $\{X_t\}_{t\in\mathbf{Z}}$, we will have the model that we called $\{Z_t\}_{t\in\mathbf{Z}}$ with p frequencies and additive noise component.

With the same reasoning as before, when $p = 1$, we can transfer results from $\{X_t\}_{t\in\mathbf{Z}}$ to $\{Z_t\}_{t\in\mathbf{Z}}$. This means that we just have to use the fact that the noise is white and Gaussian with mean zero and variance $\sigma_\varepsilon^2$, and also that $\mathbf{P}$ is ergodic for T and V is uniformly bounded. In this case we can also conclude that the empirical autocovariance and variance are consistent estimators for the autocovariance and variance of the process $\{Z_t\}_{t\in\mathbf{Z}}$.

We recall here that the sum of any two independent ergodic stochastic process is also an ergodic process and any linear transformation of an ergodic stochastic process gives rise to an ergodic process. So, if $\{X_t(\theta)\}$ and $\{\varepsilon_t(\theta)\}$ are uncorrelated ergodic stochastic process then so is the process $\{Z_t(\theta)\}$. That is, the sample autocovariance and variance are strongly consistent estimators.

## References


1. Billingsley, P., *Probability and Measure*, John Wiley, New York, 1986.

2. Brillinger, D. R., *Times Series Data Analysis and Theory*, Holt, Rinehart and Winston, New York, 1975.

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

4. Conway, J. B., *Functions of one Complex Variable*, Springer-Verlag, New York, 1978.

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

6. Devaney, R., *An Introduction to Chaotic Dynamical System*, Benjamin, $2^{\underline{nd}}$ edition, New York, 1989.

7. He, S. and B. Kedem, "Higher Order Crossings Spectral Analysis of an Almost Periodic Random Sequence in Noise", *IEEE Transactions on Information Theory*, 35, pp. 360-370, 1989.

8. Hervé, M., *Several Complex Variables: Local Theory*, Oxford University Press, Bombay, 1963.

9. Kedem, B., "Contraction Mappings in Mixed Spectrum Estimation", Presented at Institute for Mathematics and Its Applications, University of Minnesota, Minneapolis, 1990.

10. Kedem, B. and S. Lopes, "Fixed points in mixed spectrum analysis", To appear in *Proceedings of NATO ASI*, July 14-27, Il Ciocco, Italy, 1991.

11. Koopmans, L. H., *The Spectral Analysis of Time Series*, Academic Press, New York, 1974.

12. Lamperti, J., *Stochastic Processes - A Survey of the Mathematical Theory*, Springer-Verlag, New York, 1977.

13. Lopes, S., *Spectral Analysis in Frequency Modulated Models*, Ph.D. Dissertation, University of Maryland, College Park, 1991.

14. Rudin, W., *Real and Complex Analysis*, McGraw-Hill, New York, 1987.

15. Skorokhod, A. V., *Asymptotic Methods in the Theory of Stochastic Differential Equations*, Vol. 78, AMS, Providence, RI, 1989.

16. Yaglom, A. M., *Correlation Theory of Stationary and Related Random Functions I*, Springer-Verlag, New York, 1987.

**Figure 1**: Graph of the expression (3.3) as a function of $\theta \in [-\pi, \pi]$ with values $M = 2, 11, 20$. The dotted line is the function $y = \cos(\theta)$, for $\theta \in [-\pi, \pi]$.

**Figure 2**: Fixed points in $\rho_1(\alpha, 15)$ from the complex filter for $p = 2$, $A_1 = A_2 = \sigma_\varepsilon = 1.0$, $\omega_1 = 0.7$ and $\omega_2 = 2.2$ ($\cos(\omega_1) = 0.7684$, $\cos(\omega_2) = -0.5885$).

**Figure 3**: Fixed points in $\rho_1^5(\alpha, 15)$ from the complex filter for $p = 2$, $A_1 = A_2 = \sigma_\varepsilon = 1.0$, $\omega_1 = 0.7$ and $\omega_2 = 2.2$ ($\cos(\omega_1) = 0.7684$, $\cos(\omega_2) = -0.5885$).

**Figure 4**: Fixed points in $\rho_1(\alpha, 40)$ from the complex filter for $p = 3$, $A_1 = A_2 = A_3 = \sigma_\varepsilon = 1.0$, $\omega_1 = 0.5$, $\omega_2 = 1.7$ and $\omega_3 = 2.4$ ($\cos(\omega_1) = 0.8775$, $\cos(\omega_2) = -0.1288$ and $\cos(\omega_3) = -0.7373$).

**Figure 5**: Fixed points in $\rho_1^{30}(\alpha, 40)$ from the complex filter for $p = 3$, $A_1 = A_2 = A_3 = \sigma_\varepsilon = 1.0$, $\omega_1 = 0.5$, $\omega_2 = 1.7$ and $\omega_3 = 2.4$ ($\cos(\omega_1) = 0.8775$, $\cos(\omega_2) = -0.1288$ and $\cos(\omega_3) = -0.7373$).

**Figure 6**: $\rho_1(\alpha, M)$ from the complex filter for $p = 2$ in a neighborhood of the cosine of the frequency $\omega_1 = 0.7$ ($\cos(\omega_1) = 0.7684$). The graph of the constant function $y = \cos(\omega_1)$ and the diagonal line are also plotted.
(a) $M = 8$; (b) $M = 11$; (c) $M = 15$.

**(a)**

**(b)**

**(c)**

**Figure 7**: $\rho_1(\alpha, M)$ from the complex filter for $p = 2$ in a neighborhood of the cosine of the frequency $\omega_2 = 2.2$ ($\cos(\omega_2) = -0.5885$).
(a) $M = 8$; (b) $M = 11$; (c) $M = 15$.

**(a)**

**(b)**

**(c)**

**Figure 8**: $\rho_1(\alpha, 20)$ as in expression (3.5) for $p = 2$, $\omega_1 = 0.7$, $\omega_2 = 2.2$, $A_1 = A_2 = \sigma_\varepsilon = 1.0$.

**Figure 9**: $\rho_1(\alpha, 20)$ as in expression (3.6) for $p = 2$, $\omega_1 = 0.7$, $\omega_2 = 2.2$, $A_1 = A_2 = \sigma_\varepsilon = 1.0$.

**Table 4.1**: Estimation of the frequency $\omega_j$, $j = 1, 2$, from the Complex Filter. $p = 2$, $\omega_1 = 2.2$, $\omega_2 = 0.8$, $N = 3,000$ and $SNR = 20\log_{10}\left(\frac{\text{std. signal}}{\text{std. noise}}\right) dB$. Number of iterations $= 8$.

**Table 4.2**: Estimation of the frequency $\omega_j$, $j = 1, 2$, from the Complex Filter. $p = 2$, $\omega_1 = 2.5$, $\omega_2 = 0.5$, $N = 3,000$ and $SNR = 20\log_{10}\left(\frac{\text{std. signal}}{\text{std. noise}}\right) dB$. Number of iterations $= 8$.