



Instituto
Complutense
de Análisis
Económico

Mathematical framework for pseudo-spectra of linear stochastic difference equations

Marcos Bujosa

Departamento de Fundamentos del Análisis Económico II
Universidad Complutense de Madrid. Campus de Somosaguas
28223 Pozuelo de Alarcón, Spain.

Andrés Bujosa

Departamento de Matemática aplicada a las
Tecnologías de la Información. ETSI Telecomunicación
Universidad Politécnica de Madrid. Avenida Complutense, 30
28040 Madrid, Spain.

Antonio García-Ferrer

Departamento de Análisis Económico: Economía Cuantitativa
Universidad Autónoma de Madrid. Campus de Cantoblanco
28049 Madrid, Spain.

Abstract

Although spectral analysis of stationary stochastic processes has solid mathematical foundations, this is not always so for some non-stationary cases. Here, we establish a rigorous mathematical extension of the classic Fourier spectrum to the case in which there are AR roots in the unit circle, ie, the transfer function of the linear time-invariant filter has poles on the unit circle. To achieve it we: embed the classical problem in a wider framework, extend the Discrete Time Fourier Transform and defined a new Extended Fourier Transform pair *pseudo-covariance function/pseudo-spectrum*. Our approach is a proper extension of the classical spectral analysis, within which the Fourier Transform pair *auto-covariance function/spectrum* is a particular case. Consequently *spectrum* and *pseudo-spectrum* coincide when the first one is defined.

Keywords Spectral analysis, time series, non-stationarity, frequency domain, pseudo-covariance function, linear stochastic difference equations, partial inner product, Extended Fourier Transform.

JL Classification C00, C22

This working paper has been accepted for publication in a future issue of **IEEE Transactions on Signal Processing**. Content may change prior to final publication. Citation information: [DOI:10.1109/TSP.2015.2469640](https://doi.org/10.1109/TSP.2015.2469640)

1053-587X © 2015 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org

See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.



UNIVERSIDAD COMPLUTENSE
MADRID

ISSN: 2341-2356

Working Paper n° 1313

August, 2015

Abstract—Although spectral analysis of stationary stochastic processes has solid mathematical foundations, this is not always so for some non-stationary cases. Here, we establish a rigorous mathematical extension of the classic Fourier spectrum to the case in which there are AR roots in the unit circle, ie, the transfer function of the linear time-invariant filter has poles on the unit circle. To achieve it we: embed the classical problem in a wider framework, extend the Discrete Time Fourier Transform and defined a new Extended Fourier Transform pair *pseudo-covariance function/pseudo-spectrum*. Our approach is a proper extension of the classical spectral analysis, within which the Fourier Transform pair *auto-covariance function/spectrum* is a particular case. Consequently *spectrum* and *pseudo-spectrum* coincide when the first one is defined.

Index Terms—Spectral analysis, time series, non-stationarity, frequency domain, pseudo-covariance function, linear stochastic difference equations, partial inner product, Extended Fourier Transform.

Mathematical Framework for Pseudo-spectra of Linear Stochastic Difference Equations

Marcos Bujosa, Andrés Bujosa, and Antonio García-Ferrer

I. INTRODUCTION

WHEREAS the spectrum describes the frequential content of a stationary signal, the *pseudo-spectrum* describes the frequential content of a non-stationary one. In the time series literature, two approaches to non-stationary stochastic process representations in the frequency domain are found. The first one deals with the time-dependence of the frequential content of the signal. The second, with the frequential content of explosive signals. Whereas in the first case, *pseudo-spectra* are time dependent functions that generalize the traditional Fourier analysis, taking into account possible time variations of spectral characteristics of signals; in the second approach, spectral characteristics are time independent.

The first approach to *pseudo-spectra*, known as time-frequency analysis, has a well established mathematical model. It can be viewed as a time-dependent extension of classical Fourier-based methods for finite-energy signals. Some developments of this approach are time-frequency, time-scale and wavelet analysis, fractional Fourier and linear canonical transforms. These generalizations include cyclostationary signal analysis, multitaper spectral estimation, and evolutionary spectral analysis (see [1]–[11]). These are conjoint time-frequency representations that explicitly consider the time-dependence of the frequential content of the signal. Here we do not follow this time-frequency approach.

The second approach to *pseudo-spectra* deals with models for *signals with infinite energy whose spectral characteristics (like in the wide-sense stationary case) do not depend on time*. Such models naturally arise when the characteristic polynomial of the difference equation has roots on the unit circle and, therefore, the transfer function of the linear time-invariant filter has poles on the unit circle (but note that, since there are poles, the transfer function is not well defined in this case). This constitutes a different paradigm, since models for signals with infinite energy are outside the realm of Hilbert space and, therefore, they can not be treated with the classical

tools. Nevertheless, by abuse of notation and borrowing some operation rules from the stationary case algebra, this approach has brought several statistical methodologies to model non-stationary signals.¹ As a result, most national statistical agencies use these methods to model trends and seasonality (CENSUS Bureau in United States, Eurostat, United Nations, European statistical agencies, European Central Bank, UK, Canada, New Zealand, Japan, etc.). Also, applications in other areas are widespread, i.e. [12]–[16]. Despite the importance and spread of these methods, this approach does not seem to be properly grounded. Given that these methods seem to work, there should be a reason for that.

Here we provide an algebraic model that partly justifies most of the usual practices within this second approach. In particular, it is a common practice to write the pseudo-spectrum with a function that shares identical structure with the spectrum. Our main contribution in this article is a definition of pseudo-spectrum that makes this calculation rigorous, rather than intuitive. To do that, we need to extend some definitions to the non stationary case. Hence, we need a wider framework that includes the Hilbert space $L_2(S, \mathcal{B}, P)$ of scalar random variables with finite variance defined on the probability space (S, \mathcal{B}, P) . The algebraic dual space of $L_2(S, \mathcal{B}, P)$ is not “large enough”². Our strategy is to consider the algebraic dual space of an appropriate subspace of $L_2(S, \mathcal{B}, P)$.

We organize our paper as follows: Section II briefly reviews the standard framework of spectral analysis for the stationary AutoRegressive Moving Average (ARMA) case. Section III describes the state of the art for *pseudo-spectral* theory of Linear Stochastic Difference Equation (LSDE) with AutoRegressive (AR) unit roots. Section IV outlines how our algebraic model for *pseudo-spectra* is developed step by step. There we use the Random Walk (RW) as a simple illustration. Sections V to VII show the technical details. In Section VIII some properties of *pseudo-spectra* are reviewed. In Section IX some examples and applications are presented. Finally, we conclude in Section X.

Notation: Bold symbols denote either sequences of random variables or sequences of functionals of random variables. Uppercase Greek letters denote double infinite sequences of numbers, and $\mathbb{R}^{\mathbb{Z}}$ denotes the set of all those sequences. Consequently, l_1 (the set of absolutely summable sequences) and l_2 (the set of squared summable sequences) are subsets of $\mathbb{R}^{\mathbb{Z}}$. Lowercase Greek letters denote polynomials. The only exceptions are the variance (σ^2) and the standard deviation (σ).

Copyright (c) 2015 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

This work was supported by the Spanish Ministerio de Economía y Competitividad (ECO2012-32854).

M. Bujosa is with the Departamento de Fundamentos del Análisis Económico II. Universidad Complutense de Madrid. Campus de Somosaguas. 28223 Pozuelo de Alarcón, Spain. E-mail: mbujosab@ucm.es.

A. Bujosa is with the Departamento de Matemática aplicada a las Tecnologías de la Información. ETSI Telecomunicación. Universidad Politécnica de Madrid. Avenida Complutense, 30. 28040 Madrid, Spain.

A. García-Ferrer is with the Departamento de Análisis Económico: Economía cuantitativa. Universidad Autónoma de Madrid. Campus de Cantoblanco. 28049 Madrid, Spain.

¹see <https://www.census.gov/srd/www/x13as/> and references therein.

²although it contains the topological dual space of $L_2(S, \mathcal{B}, P)$, a isomorphic “copy” of $L_2(S, \mathcal{B}, P)$.

We assume that the set of polynomials, $\mathbb{R}[X]$, is contained in $\mathbb{R}^{\mathbb{Z}}$ by identifying $a_0 + a_1X + \dots + a_nX^n$ with the sequence $\alpha \equiv \{\alpha_t\}_{t \in \mathbb{Z}}$ where $\alpha_t = a_t$ if $0 \leq t \leq n$, and $\alpha_t = 0$ if $t < 0$ or $t > n$; that is

$$\alpha \equiv \dots 0, 0, \boxed{a_0}, a_1, \dots, a_n, 0, 0, \dots$$

where the coefficient for the zero index ($t = 0$) is boxed. As it is usual, if α is a polynomial and C is an element in an algebra³, $\alpha(C) \equiv a_0 + a_1C + \dots + a_nC^n$, consequently $\alpha(X) = \alpha$. We use this mathematical convention with $\alpha(B)$ where B is the backward shift operator and with $\alpha(X^{-1})$ (where X^{-1} is the sequence in $\mathbb{R}^{\mathbb{Z}}$ that is 1 when $t = -1$ and 0 otherwise)

$$X^{-1} \equiv \dots 0, 1, \boxed{0}, 0, 0, \dots$$

II. STANDARD FRAMEWORK FOR THE STATIONARY CASE

Here, some well known results are stated for further reference along the paper, as well as to show the parallels or differences between the new results, definitions or properties, and those pertaining to the standard framework.

Let (S, \mathcal{B}, P) be a probability space, where S is a nonempty sample space, \mathcal{B} a Borel field of subsets of S , and $P(\cdot)$ a probability measure on \mathcal{B} . To generate the relevant Hilbert space we use zero-mean random variables with finite variance defined on (S, \mathcal{B}, P) , with the inner product $\langle x, y \rangle_{L_2(S, \mathcal{B}, P)} = E[x \cdot y]$, corresponding norm $\|x\| = \sqrt{E[x^2]}$ and metric $\|x - y\|$, where E denotes the expectation operator. If $E[(x - y)^2] = 0$, we say x and y are equivalent. Being equivalent is indeed an equivalence relation, and then, the space $L_2(S, \mathcal{B}, P)$ is the corresponding quotient space, i.e., the collection of these equivalence classes.

Consider a LSDE

$$\sum_{i=0}^p a_i x_{t-i} = \sum_{j=0}^q b_j w_{t-j}, \quad t \in \mathbb{Z}, \quad (1)$$

where $a_0 = b_0 = 1$, and $\{w_t\}_{t \in \mathbb{Z}}$ is a white noise stochastic process with $E(w_t) = 0$ and $\text{Var}(w_t) = \|w_t\|^2 = \sigma^2$ for all t . The characteristic polynomial is $X^p + a_1X^{p-1} + \dots + a_{p-1}X + a_p$, and its reciprocal polynomial $\phi = 1 + a_1X + a_2X^2 + \dots + a_pX^p$ is known as the AR polynomial. In the right sum in (1), $\theta = 1 + b_1X + b_2X^2 + \dots + b_qX^q$ is known as the Moving Average (MA) polynomial.

Using the discrete convolution product “ $*$ ”

$$(f * g)_t = \sum_{m=-\infty}^{\infty} f_m g_{t-m}, \quad t \in \mathbb{Z} \quad (2)$$

we can write (1) as $(\phi * \mathbf{x})_t = (\theta * \mathbf{w})_t$; where $\mathbf{x} \equiv \{x_t\}_{t \in \mathbb{Z}}$ and $\mathbf{w} \equiv \{w_t\}_{t \in \mathbb{Z}}$. Sum limits in (1) are finite since coefficients a_i are zero when $i > p$ or $i < 0$ in AR polynomials, and coefficients b_j are zero when $j > q$ or $j < 0$ in MA polynomials. With $\phi * \mathbf{x}$ and $\theta * \mathbf{w}$ denoting the whole sequences $\{(\phi * \mathbf{x})_t\}_{t \in \mathbb{Z}}$ and $\{(\theta * \mathbf{w})_t\}_{t \in \mathbb{Z}}$ respectively, we can use the following compact notation for (1):

$$\phi * \mathbf{x} = \theta * \mathbf{w}. \quad (3)$$

When ϕ has **no roots on the unit circle** there is a unique stationary solution—a stationary sequence $\mathbf{y} \equiv \{y_t\}_{t \in \mathbb{Z}}$ of random variables in $L_2(S, \mathcal{B}, P)$; see [18]. Using the absolutely summable inverse sequence $\frac{1}{\phi}$, the stationary solution can be written as

$$\mathbf{y} = \frac{1}{\phi} * (\theta * \mathbf{w}). \quad (4)$$

The sequence $\Theta \equiv \frac{1}{\phi} * \theta$ is also absolutely summable and then

$$\mathbf{y} = \frac{1}{\phi} * (\theta * \mathbf{w}) = \left(\frac{1}{\phi} * \theta\right) * \mathbf{w} = \Theta * \mathbf{w}, \quad (5)$$

where the last expression is known as the infinite moving average (Wold) representation of \mathbf{y} . In this special case, with no AR roots on the unit circle, the LSDE is often known as an ARMA(p, q) model, and the stationary solution \mathbf{y} is often known as an ARMA process. It should be noted that, since Θ is a square summable sequence, the convolution product $\Theta * \mathbf{w}$ converges in mean square. Hence, \mathbf{y} is a well-defined (second-order) stationary stochastic process with:

$$\text{cov}(y_i, y_j) = \langle y_i, y_j \rangle_{L_2(S, \mathcal{B}, P)} = \langle y_{i+t}, y_{j+t} \rangle_{L_2(S, \mathcal{B}, P)} \quad (6)$$

for all $i, j, t \in \mathbb{Z}$. The sequence of auto-covariances of \mathbf{y} , known as the *auto-covariance* (generating) *function* is

$$\Gamma_{\mathbf{y}} : \mathbb{Z} \rightarrow \mathbb{R} \\ j \rightarrow \langle y_0, y_j \rangle_{L_2(S, \mathcal{B}, P)} ; \quad (7)$$

so, $\text{cov}(y_i, y_j) = \Gamma_{\mathbf{y}}(j - i)$. Using the infinite moving average representation, it is easy to show that the auto-covariance function is $\Gamma_{\mathbf{y}}(X) = \Theta(X) * \Theta(X^{-1})\sigma^2$ and satisfies

$$\phi(X) * \phi(X^{-1}) * \Gamma_{\mathbf{y}} = \theta(X) * \theta(X^{-1}) \cdot \sigma^2. \quad (8)$$

Its Discrete Time Fourier Transform (DTFT), $\Gamma_{\mathbf{y}}(e^{-i\omega}) = \Theta(e^{-i\omega})\Theta(e^{i\omega})\sigma^2$, is the *spectrum* of \mathbf{y} (see Table I).

III. THE INCOMPLETE GENERALIZATION TO THE NON-STATIONARY CASE

When ϕ has **roots on the unit circle** there is no square summable inverse sequence $\frac{1}{\phi}$ and the stationary solution (4) does not exist, neither does the auto-covariance function nor the spectrum. Nevertheless, in the literature we find references about spectral representation of non-stationary solutions to (3) over the infinite time domain.

Within this approach, used since the late seventies, *pseudo-spectra* are obtained by two alternative ways. First, the *pseudo-spectrum* is described as the limit function of a sequence of spectra of stationary ARMA models, as the modulus of an AR root tends to one (e.g. [19]). But in the limit covariance functions diverge and therefore spectra are not defined. Second, an Autoregressive Integrated Moving Average (ARIMA) process acted upon by a filter $\varphi(B)$ that cancels out the AR roots on the unit circle to make it stationary. Then the spectrum of the stationary filtered process is divided by $\varphi(e^{i\omega t})\varphi(e^{-i\omega t})$; so the spectrum of the filtered process is multiplied by the inverse of the power transfer function of the filter (e.g., [20]). However, a power transfer function is not defined for that filter, since any inverse sequence of ϕ is not absolutely summable when the AR polynomial ϕ has roots on the unit circle. In

³[17, page 229, exercise 5]

TABLE I
STATE OF THE ART. A MODEL THAT FILLS IN THE GAPS **A**, **B** AND **C** IS NEEDED WHEN THE AR POLYNOMIAL HAS ROOTS ON THE UNIT CIRCLE.

STATIONARY CASE (No AR roots on the unit circle)			
Convolution type solution	Covariance generating function of the stationary solution \mathbf{y} satisfies	Discrete Time Fourier Transform	Spectrum
$\mathbf{y} = \frac{1}{\phi} * \theta * \mathbf{w}$ $\phi * \frac{1}{\phi} = 1$ $\left(\frac{1}{\phi} \text{ sq. summable}\right)$	$\phi(X) * \phi(X^{-1}) * \Gamma_{\mathbf{y}} = \theta(X) * \theta(X^{-1}) \cdot \sigma^2$	$\rightarrow \mathcal{F} \rightarrow$	$\frac{\theta(e^{-i\omega})\theta(e^{i\omega})}{\phi(e^{-i\omega})\phi(e^{i\omega})}\sigma^2$
NON-STATIONARY CASE (AR roots on the unit circle)			
Solutions?	Covariance generating function?	Fourier Transform?	Pseudo-Spectrum
Gap A	Gap B	Gap C	$\frac{\theta(e^{-i\omega})\theta(e^{i\omega})}{\phi(e^{-i\omega})\phi(e^{i\omega})}\sigma^2$

many cases *pseudo-spectra* are simply used with no further explanation (e.g., [21]–[23]).

Within this second approach the *pseudo-spectrum* seems to describe “the distribution (over the frequency range) of the energy (per unit time) or variance (possibly infinite) of the process” [24], in a similar way as the spectrum does in the stationary case. Hence, we find that sometimes the term “spectrum” is, in fact, used to denote this type of *pseudo-spectrum* [25]. But there are several major drawbacks in this. This *pseudo-spectrum* cannot be the DTFT of a covariance function, since covariance functions of non-stationary solutions to (3) over the infinite time domain are not defined. Even worse, *pseudo-spectra* are functions outside the Hilbert space $L_2[-\pi, \pi]$, so DTFT is not applicable.

However practitioners and academics use *pseudo-spectra* and operate with their algebraic expressions as if they were spectra and, surprisingly (or not), it seems to work fine! We also find not well defined expressions like $x_t = (\phi^{-1} * \theta) * w_t$, where ϕ has roots on the unit circle in seminal papers as in [21], [26]–[31]. There, $(\phi^{-1} * \theta) * w_t$ is referred to as “the (nonconvergent) infinite moving average representation of x_t ” [26]. Hence, since the energy (the variance) of x is infinite (non-convergent), its physical interpretation is unclear. It should be noted that, even in the stationary case, there is no clear physical interpretation for solutions of infinite duration in the past (... infinitely before “The Big Bang”!). We are used to deal with these mathematical formalisms and, therefore, we don’t usually pay attention on this. In the non-stationary case there is, indeed, a second level of abstraction since the variance of these solutions is also infinite. In spite of that, we often find in the literature that some properties from these models are deduced, and statistical methods developed. These methods are applied to non-stationary finite signals; so there is no infinite energy in practice. However, many properties of these methods are deduced from these mathematical formalisms with no physical interpretation.

IV. A WIDER FRAMEWORK FOR BOTH THE STATIONARY AND THE NON-STATIONARY CASES

In this paper we generalize the spectral theory to the case where $\phi \neq 0$ is an AR polynomial *either with or without roots on the unit circle*. We also show that spectrum and

pseudo-spectrum coincide when the former is defined. This generalization is not straightforward. Several technical steps are needed because, when no restrictions are imposed on the roots of the AR polynomial ϕ , the discussion in Section II is no longer valid. The symbol “ $\frac{1}{\phi}$ ” is used to denote a very particular inverse, the unique absolutely summable sequence such that $\frac{1}{\phi} * \phi = 1$. But that sequence does not exist when ϕ has roots on the unit circle. Nevertheless, the inverse sequences of ϕ are used to denote solutions to $\phi * x = \theta * w$. So, we must consider other (non-summable) inverse sequences. Indeed, there are an infinite number of sequences, Ψ , such that $\Psi * \phi = 1$, that is, there are an infinite number of inverse sequences for each non null degree polynomial. But the corresponding solutions are **non-convergent**. In addition, when the inverse is not unique, convolution products are **no longer associative**⁴: if $\Lambda * \phi = 1$ and $\phi * \Upsilon = 1$, then $(\Lambda * \phi) * \Upsilon \neq \Lambda * (\phi * \Upsilon)$. Here we show how to deal with these issues in order to provide a mathematical model for *pseudo-spectra*. Below, we describe the steps we take in sections V to VII along with the illustration of the RW model.

Step 1 Solutions to (3) are non-convergent when ϕ has roots on the unit circle. One easy example is the RW model $x_t - x_{t-1} = w_t$, where several formal solutions can be found. The backward causal solution is $f_t = \sum_{j=-\infty}^0 w_{t+j}$, and the forward solution is $g_t = -\sum_{j=1}^{\infty} w_{t+j}$. The problem is that those sums are non-convergent. So *we need to provide a new framework where the convergence issues are avoided*. To do so, we embed the standard Hilbert Space in a wider framework. This is done in the first part of Section V.

Step 2 In $\mathbb{R}^{\mathbb{Z}}$, discrete convolution product is not defined for any pair of sequences and, when it is defined, it is not always associative. We need to check under which conditions it is possible to operate as in equation (5), in order to get convolution type solutions in the new and wider framework; that is, solutions in the form⁵ $\mathbf{f} = (\Psi * \theta) * \mathbf{w}^*$ where $\phi * \Psi = 1$. This is done in the second part of Section V.

⁴Although discrete convolution products are associative in l_1 , this is not true when we also consider other inverse sequences outside l_1 .

⁵where \mathbf{w}^* is the embedding of \mathbf{w} . See Section V.

TABLE II

PARALLELISM BETWEEN THE STATIONARY AND NON-STATIONARY CASE. WHEN ϕ HAS NO ROOTS ON THE UNIT CIRCLE, AND THE CO-STATIONARY PAIR OF STATIONARY SOLUTIONS $(\mathbf{y}^*, \mathbf{y}^*)$ IS USED (WHEN $\Psi = \frac{1}{\phi}$), COVARIANCE AND *Pseudo*-COVARIANCE GENERATING FUNCTIONS ARE THE SAME ($\Gamma_{\mathbf{y}} = \Gamma_{\mathbf{y}^*, \mathbf{y}^*}$), AND SO THEY ARE SPECTRUM AND *Pseudo*-SPECTRUM.

STATIONARY CASE (No AR roots on the unit circle)				
Convolution type solution	Covariance generating function of the stationary solution \mathbf{y} satisfies	Discrete Fourier Transform	Time	Spectrum
$\mathbf{y} = \frac{1}{\phi} * \theta * \mathbf{w}$	$\phi(X) * \phi(X^{-1}) * \Gamma_{\mathbf{y}} = \theta(X) * \theta(X^{-1}) \cdot \sigma^2$	$\rightarrow \mathcal{F} \rightarrow$		$\frac{\theta(e^{-i\omega})\theta(e^{i\omega})}{\phi(e^{-i\omega})\phi(e^{i\omega})} \sigma^2$
$\phi * \frac{1}{\phi} = 1$	$(\frac{1}{\phi} \text{ sq. summable})$			
BOTH STATIONARY AND NON-STATIONARY CASES				
Convolution type solutions	<i>Pseudo</i> -covariance generating function of any co-stationary pair of solutions (\mathbf{f}, \mathbf{g}) satisfies	<i>Extended</i> Fourier Transform	Fourier	Spectrum or <i>Pseudo</i> -Spectrum
$\mathbf{f} = \Psi * \theta * \mathbf{w}^*$	$\phi(X) * \phi(X^{-1}) * \Gamma_{\mathbf{f}, \mathbf{g}} = \theta(X) * \theta(X^{-1}) \cdot \sigma^2$	$\rightarrow \overline{\mathcal{F}} \rightarrow$		$\frac{\theta(e^{-i\omega})\theta(e^{i\omega})}{\phi(e^{-i\omega})\phi(e^{i\omega})} \sigma^2$
$\phi * \Psi = 1$				
Proposition VI.1	Proposition VI.1 and Theorem VI.3		Theorem VII.2	

After Section V, Gap A in Table I is filled in (see Table II, column 1). Indeed, for the AR polynomial $\phi = 1 - X$ in our RW illustration, we can consider at least two inverse sequences: $\Lambda = \{c_j\}_{j \in \mathbb{Z}}$, where $c_j = 1$ for $j \leq 0$ and 0 otherwise; and $\Upsilon = \{c_j\}_{j \in \mathbb{Z}}$, where $c_j = -1$ for $j \geq 1$ and 0 otherwise. With the first inverse we can define the backward causal solution sequence of functionals $\mathbf{f} = \Lambda * \mathbf{w}^*$, where each functional $f_t = (\Lambda * \mathbf{w}^*)_t = \sum_{j=-\infty}^0 w_{t+j}^*$ is well defined. Similarly, we can define the forward solution sequence of functionals $\mathbf{g} = \Upsilon * \mathbf{w}^*$, where $g_t = (\Upsilon * \mathbf{w}^*)_t = -\sum_{j=1}^{\infty} w_{t+j}^*$. (see Section VI-A)

At this point new difficulties arise. Since the considered solutions sequences \mathbf{f} are non-convergent in general, its elements f_t are functionals outside $L_2(S, \mathcal{B}, P)$. It follows that neither the covariance nor the auto-covariance generating function are defined for any of these non-convergent solutions \mathbf{f} . Fortunately, it is possible to define *co-stationarity* for some pairs (\mathbf{f}, \mathbf{g}) . Then we define the *pseudo-covariance function* for co-stationary pairs sequences in a similar way as in (7). So similar that, when there are no AR roots on the unit circle and the pair is formed by the stationary solution and itself⁶ $(\mathbf{y}^*, \mathbf{y}^*)$, the covariance and pseudo-covariance functions coincide.

Step 3 In Section VI we define the co-stationarity of pairs of sequences and its *pseudo-covariance function* (Definitions VI.1 and VI.2).

Two more results are obtained from Section VI: a) as a consequence of Proposition VI.1, the *pseudo-covariance function*, $\Gamma_{\mathbf{f}, \mathbf{g}}$, of any pair (\mathbf{f}, \mathbf{g}) of co-stationary solutions to (3) satisfies

$$\phi(X) * \phi(X^{-1}) * \Gamma_{\mathbf{f}, \mathbf{g}} = \theta(X) * \theta(X^{-1}) \cdot \sigma^2 \quad (9)$$

[note the parallelism with (8).

b) The Existence Theorem VI.3 proves that for any difference equation $\phi * \mathbf{x} = \theta * \mathbf{w}$, the *backward* and the *forward* solutions always form a co-stationary pair.

Hence, after Section VI, Gap B in Table I is also filled in (see Table II, column 2). Following our RW illustration,

the pseudo-covariance function for the *backward* and *forward* solutions pair (\mathbf{f}, \mathbf{g}) is

$$\begin{cases} \Gamma_{\mathbf{f}, \mathbf{g}}(k) = 0 & \text{for } k \geq 0 \\ \Gamma_{\mathbf{f}, \mathbf{g}}(k) = -k\sigma^2 & \text{for } k < 0 \end{cases}; \quad (10)$$

it is easy to check that $(1 - X) * (1 - X^{-1}) * \Gamma_{\mathbf{f}, \mathbf{g}} = \sigma^2$.

Two more difficulties remain. First, since pseudo-covariance functions are not squared summable in general, we need to extend the DTFT outside the Hilbert space. Second, whereas, for a given LSDE, the *pseudo-spectrum* is unique, different pairs of co-stationary solutions have different *pseudo-covariance functions*. For instance, when ϕ has all its roots outside the unit circle, the backward solution with itself form a co-stationary solution pair $(\mathbf{y}^*, \mathbf{y}^*)$, since the backward solution is stationary in this case. Hence, the *covariance function* ($\Gamma_{\mathbf{y}}$) and the *pseudo-covariance function* ($\Gamma_{\mathbf{y}^*, \mathbf{y}^*}$) are the same. But, for the same LSDE, the backward and forward solutions form another co-stationary pair $(\mathbf{y}^*, \mathbf{f})$ with a completely different (non-summable) *pseudo-covariance function* ($\Gamma_{\mathbf{y}^*, \mathbf{f}}$). Both difficulties are solved though the extension of the DTFT.

Step 4 In Section VII we extend the domain outside the Hilbert space l_2 , so it also includes the *pseudo-covariance functions*. In addition, the Extended Fourier Transform $\overline{\mathcal{F}}$ is defined so that for any sequence that satisfies (24), that is, any sequence Ψ such that

$$\phi(X) * \phi(X^{-1}) * \Psi = \theta(X) * \theta(X^{-1}) \cdot \sigma^2, \quad (11)$$

the image is

$$\frac{\theta(e^{-i\omega})\theta(e^{i\omega})}{\phi(e^{-i\omega})\phi(e^{i\omega})} \sigma^2. \quad (12)$$

It follows that, given a difference equation $\phi * \mathbf{x} = \theta * \mathbf{w}$, there is a common image for the *pseudo-covariance functions* of all co-stationary solution pairs. Hence, the *pseudo-spectrum* is unique (Theorem VII.2).

Indeed, with Section VII Gap C in Table I is finally filled in (see Table II, column 3).

It is important to remember that when a stationary solution exists, it forms a co-stationary pair with itself. It immediately follows that **spectra and pseudo-spectra are equal when the**

⁶where \mathbf{y}^* is the embedding of \mathbf{y} . See Section V.

AR polynomial has no roots on the unit circle. Although the construction of the pseudo-spectrum is different, the final structure of the function for spectra and pseudo-spectra is, so to speak, the same. So, **given a difference equation, $\phi * \mathbf{x} = \theta * \mathbf{w}$, we can write the corresponding (pseudo-)spectrum regardless of the roots of the AR polynomial ($\phi \neq 0$).**

Following our illustration, the *pseudo-spectrum* of a RW is:

$$\Gamma(e^{-i\omega}) = \frac{\sigma^2}{(1 - e^{-i\omega})(1 - e^{i\omega})} = \frac{\sigma^2}{2 - 2\cos\omega}. \quad (13)$$

Due to the pole on the frequency zero, the integral of this function is infinite. So it can be viewed as a representation of the infinite variance of the solutions to $x_t - x_{t-1} = w_t$; where an infinite contribution to the variance is concentrated in the zero frequency⁷. Consequently, when a filter whose frequency response has a zero in the zero frequency is applied (for example by taking first differences), then (as a consequence of Proposition VI.1) the output becomes stationary. In the following sections we describe the technical details of this new model for *pseudo-spectra*.

V. A WIDER FRAMEWORK

We, first, need to give meaning to expressions like $\sum_{t \in \mathbb{Z}} a_t w_t$ when $\{a_t\}_{t \in \mathbb{Z}}$ is not square summable (i.e., when the sequence is not in l_2), so as to be able to write $\mathbf{f} = \Psi * \mathbf{w}$ with $\phi * \Psi = \theta$, [in the spirit of (5)] even when $\Psi \notin l_2$. To avoid convergence problems we embed the Hilbert space of the classical framework in a wider space. This wider framework will be specific for each difference equation, since it is defined using the white noise process \mathbf{w} in (3). The only tools we need are the standard scalar product $\langle \cdot, \cdot \rangle$ in the Hilbert space, and the set of finite linear combinations of vectors belonging to a Hilbert basis that includes standardized random variables in \mathbf{w} .

Let $L_2(S, \mathcal{B}, P)$ be the Hilbert space of scalar random variables with finite variance defined on the probability space (S, \mathcal{B}, P) , and consider \mathcal{D} a Hilbert basis (a maximal orthonormal subset) of $L_2(S, \mathcal{B}, P)$ such that $\{\frac{w_t}{\sigma} | t \in \mathbb{Z}\} \subset \mathcal{D}$, where $\frac{w_t}{\sigma}$ are the standardized random variables. The space $L(\mathcal{D})^*$ is the algebraic dual space⁸ of all finite linear combinations of \mathcal{D} . The map we use to embed $L_2(S, \mathcal{B}, P)$ in $L(\mathcal{D})^*$ is given by $f \mapsto f^*$ where $f^*(v) = \langle f, v \rangle$ for all $v \in L(\mathcal{D})$. Since this map is linear and injective,⁹ then for all $\mathbf{y} \in [L_2(S, \mathcal{B}, P)]^{\mathbb{Z}}$,

$$\phi * \mathbf{y} = \theta * \mathbf{w} \quad \text{if and only if} \quad \phi * \mathbf{y}^* = \theta * \mathbf{w}^*; \quad (14)$$

where \mathbf{y}^* and \mathbf{w}^* are the corresponding sequences of embeddings of \mathbf{y} and \mathbf{w} , respectively.

Now, within this framework it is easy to provide a meaning for $\sum_{t \in \mathbb{Z}} a_t w_t^*$ via the following definition:

$$\left[\sum_{t \in \mathbb{Z}} a_t w_t^* \right] (v) \equiv \sum_{w_t^*(v) \neq 0} a_t w_t^*(v), \quad (15)$$

⁷but this physical interpretation is outside the mathematical framework.

⁸although here $L(\mathcal{D}) \subset L_2(S, \mathcal{B}, P) \subset L(\mathcal{D})^*$, this should not be confused with the *Gelfand triple* or the *Rigged Hilbert Space*, since no topology is defined on $L(\mathcal{D})$ (see [32]).

⁹since \mathcal{D} is maximal orthonormal: $f^* = 0$, if and only if $\forall d \in \mathcal{D}, \langle f, d \rangle = 0$, if and only if $f = 0$.

since for all $v \in L(\mathcal{D})$ the set $\{t \in \mathbb{Z} | w_t^*(v) \neq 0\}$ is finite: indeed, if $v \in L(\mathcal{D})$ then $v = \sum_{i=1}^n a_i d_i$, hence $\{t \in \mathbb{Z} | w_t^*(v) \neq 0\} \subset \cup_{i=1}^n \{t \in \mathbb{Z} | w_t^*(d_i) \neq 0\}$, and $\{t \in \mathbb{Z} | w_t^*(d_i) \neq 0\}$ either is empty or it has only one element.

A. Summability in the dual space $L(\mathcal{D})^*$

The goal of this section is to show that some solutions to $\phi * \mathbf{x} = \theta * \mathbf{w}^*$ can be expressed via convolution products just like in (5). To do so, we need a minimal requirement about summability on functionals to get well defined convolution products: we say that a sequence of functionals $\mathbf{f} = \{f_t\}_{t \in \mathbb{Z}}$ of $L(\mathcal{D})^*$ is *summable* if, for all $v \in L(\mathcal{D})$, the subset of indexes $\{t \in \mathbb{Z} | f_t(v) \neq 0\}$ is finite, and its sum is the functional in $L(\mathcal{D})^*$ given by the map $v \mapsto \sum_{f_t(v) \neq 0} f_t(v)$. Note that the embedding of the white noise process, $\mathbf{w}^* \equiv \{w_t^*\}_{t \in \mathbb{Z}}$, is summable.

1) *The convolution product on summable sequences:* despite the fact that *within this framework* the convolution product for any two sequences of functionals is not always defined, and when it is, it is not always associative, we enumerate some useful properties of the convolution product which do hold in some special cases, and which resemble those of the product in (5). These will allow us to carry out, in $L(\mathcal{D})^*$, operations like those involved in (5).

If we let \mathcal{A} denote the set of summable sequences of functionals, and B denotes the backward shift operator, $B(\{x_t\}_{t \in \mathbb{Z}}) = \{x_{t-1}\}_{t \in \mathbb{Z}}$, then:

- \mathcal{A} is a subspace of $[L(\mathcal{D})^*]^{\mathbb{Z}}$ and $B(\mathcal{A}) = \mathcal{A}$. Hence, if $\mathbf{f} \in \mathcal{A}$ and $\theta \in \mathbb{R}[X]$ then $\theta(B)(\mathbf{f}) = \theta * \mathbf{f} \in \mathcal{A}$.
- If $\mathbf{f} \in \mathcal{A}$ and $\Psi \in \mathbb{R}^{\mathbb{Z}}$, then $\Psi * \mathbf{f} \in [L(\mathcal{D})^*]^{\mathbb{Z}}$.
- If $\mathbf{f} \in \mathcal{A}$, $\theta \in \mathbb{R}[X]$ and $\Psi \in \mathbb{R}^{\mathbb{Z}}$, then $\Psi * (\theta * \mathbf{f}) = (\Psi * \theta) * \mathbf{f} = \theta * (\Psi * \mathbf{f})$.

(See the appendix for the corresponding proofs and also for the proofs of all forthcoming propositions, lemmas and theorems).

2) *Solutions to $\phi * \mathbf{x} = \theta * \mathbf{w}^*$ in the form of convolutions:* we are now ready to state the main result of this Section V:

Proposition V.1. *If $\phi \neq 0$ and θ are polynomials and $\Psi \in \mathbb{R}^{\mathbb{Z}}$ such that $\phi * \Psi = 1$ then $(\Psi * \theta) * \mathbf{w}^*$ is a solution of*

$$\phi * \mathbf{x} = \theta * \mathbf{w}^*. \quad (16)$$

Note that in general there is more than one sequence $\Psi \in \mathbb{R}^{\mathbb{Z}}$ such that $\phi * \Psi = 1$. In particular, if ϕ has no roots on the unit circle and $\Psi = \frac{1}{\phi} \in l_1$ then $(\Psi * \theta) * \mathbf{w}^*$ is the embedded version of the solution \mathbf{y} of (5) in $[L(\mathcal{D})^*]^{\mathbb{Z}}$.

VI. CO-STATIONARITY

By embedding the problem in $L(\mathcal{D})^*$ we have avoided convergence issues. But, is there any relation between the new solutions and stationarity? Our next step is to find an expression similar to (6) in $L(\mathcal{D})^*$. In this section we define *co-stationarity* for pairs of sequences, and we search for pairs of co-stationary solutions to (16). To do so, we shall use a so called [33] *partial inner product* in $L(\mathcal{D})^*$: two functionals $f, g \in L(\mathcal{D})^*$ are said *compatible* if $\sum_{d \in \mathcal{D}} |f(d)g(d)| < \infty$, and in this case $\langle f, g \rangle_{L(\mathcal{D})^*} = \sum_{d \in \mathcal{D}} f(d)g(d)$ is their *inner product*. This is a *partial inner product* since it is only defined

for *compatible* functional pairs. Clearly, if $f, g \in L_2(S, \mathcal{B}, P)$ then $f = \sum_{d \in \mathcal{D}} a_d d$ and $g = \sum_{d \in \mathcal{D}} b_d d$, and therefore

$$\begin{aligned} \langle f, g \rangle_{L_2(S, \mathcal{B}, P)} &= \sum_{d \in \mathcal{D}} a_d b_d = \sum_{d \in \mathcal{D}} \langle f, d \rangle \langle g, d \rangle \\ &= \sum_{d \in \mathcal{D}} f^*(d) g^*(d) = \langle f^*, g^* \rangle_{L(\mathcal{D})^*}. \end{aligned} \quad (17)$$

It follows that the embedding of $L_2(S, \mathcal{B}, P)$ in $L(\mathcal{D})^*$ is also an isometry, and $\langle w_i^*, w_j^* \rangle_{L(\mathcal{D})^*} = \delta_{ij}$ (the Kronecker delta). From now on, we shall simply write $\langle f, g \rangle$ instead of $\langle f, g \rangle_{L(\mathcal{D})^*}$.

Now we can extend the notion of stationarity to pairs of sequences of functionals in $[L(\mathcal{D})^*]^{\mathbb{Z}}$ and define the pseudo-covariance function for co-stationary pairs:

Definition VI.1 (Co-stationarity). *The sequences $\mathbf{f}, \mathbf{g} \in [L(\mathcal{D})^*]^{\mathbb{Z}}$ are co-stationary if*

- 1) f_i and g_j are compatible for all $i, j \in \mathbb{Z}$, and
- 2) $\langle f_i, g_j \rangle = \langle f_{i+t}, g_{j+t} \rangle$ for all $i, j, t \in \mathbb{Z}$.

We then say that (\mathbf{f}, \mathbf{g}) is a co-stationary pair.

Definition VI.2. *If \mathbf{f} and \mathbf{g} are co-stationary, we define their pseudo-covariance function as:*

$$\Gamma_{\mathbf{f}, \mathbf{g}}: \begin{array}{ccc} \mathbb{Z} & \longrightarrow & \mathbb{R} \\ j & \longrightarrow & \langle f_0, g_j \rangle \end{array}, \quad (18)$$

and therefore $\langle f_i, g_j \rangle = \Gamma_{\mathbf{f}, \mathbf{g}}(j - i)$.

Moreover, a usual stochastic process $\mathbf{y} \in [L_2(S, \mathcal{B}, P)]^{\mathbb{Z}}$ is (second-order) stationary if and only if $(\mathbf{y}^*, \mathbf{y}^*)$ is a co-stationary pair (where \mathbf{y}^* is the embedding of \mathbf{y} in $L(\mathcal{D})^*$). Therefore, we can also refer to $\Gamma_{\mathbf{y}^*, \mathbf{y}^*} = \Gamma_{\mathbf{y}}$ as the usual *auto-covariance function* of \mathbf{y} [see (7)].

To complete these extensions we give one more statement pertaining linear filters in the new framework, which we use in Section VII to find the domain of the Extended Fourier Transform.

Proposition VI.1. *Given two polynomials $\theta(X), \phi(X)$, if \mathbf{f} and \mathbf{g} are co-stationary, then $\theta * \mathbf{f}$ and $\phi * \mathbf{g}$ are so, and*

$$\Gamma_{\theta * \mathbf{f}, \phi * \mathbf{g}} = \theta(X^{-1}) * \phi(X) * \Gamma_{\mathbf{f}, \mathbf{g}}. \quad (19)$$

A. Co-stationary solution pairs

Our results in the previous section clearly show that if the AR polynomial $\phi \neq 0$ has no roots on the unit circle, there is a co-stationary pair of solutions $(\mathbf{y}^*, \mathbf{y}^*)$. But we need co-stationary pairs of solutions to any LSDE. Indeed, we are almost ready to show (Theorem VI.3) that there is at least one co-stationary pair of solutions for any LSDE: the pair consisting of what could be interpreted as the backward and the forward solutions to (16). First we need to remember that for any $\phi(X) = \phi_k X^k + \phi_{k+1} X^{k+1} + \dots + \phi_n X^n$ with $\phi_k \neq 0 \neq \phi_n$, we can find two sequences Λ and Υ such that

$$\phi * \Lambda = 1 = \phi * \Upsilon. \quad (20)$$

Those sequences are given by the following formulas:

$$\Lambda_t = \begin{cases} 0 & \text{if } t < -k \\ \frac{1}{\phi_k} & \text{if } t = -k \\ \frac{-1}{\phi_k} \sum_{i=1}^{n-k} \Lambda_{t-i} \phi_{k+i} & \text{if } t > -k \end{cases} \quad (\text{Backward inverse}) \quad (21)$$

which is zero for all but finitely many negative indices j (a formal Laurent series); while the sequence

$$\Upsilon_t = \begin{cases} 0 & \text{if } t > -n \\ \frac{1}{\phi_n} & \text{if } t = -n \\ \frac{-1}{\phi_n} \sum_{i=1}^{n-k} \Upsilon_{t+i} \phi_{n-i} & \text{if } t < -n \end{cases} \quad (\text{Forward inverse}) \quad (22)$$

is zero for all but finitely many positive indices j . From Proposition VI.1 we know that $(\Lambda * \theta) * \mathbf{w}^*$ and $(\Upsilon * \theta) * \mathbf{w}^*$ are solutions to (16), which we name the *backward* and the *forward* solutions, respectively.

Lemma VI.2. *If $\Psi \in \mathbb{R}^{\mathbb{Z}}$ is zero for all but finitely many negative indices j , and $\Omega \in \mathbb{R}^{\mathbb{Z}}$ is zero for all but finitely many positive indices j , then $\Psi * \mathbf{w}^*$ and $\Omega * \mathbf{w}^*$ are co-stationary.*

It follows then that:

Theorem VI.3. *For any polynomials $\phi \neq 0$ and θ at least one pair of co-stationary solutions of*

$$\phi * \mathbf{x} = \theta * \mathbf{w}^* \quad (23)$$

exists in $[L(\mathcal{D})^]^{\mathbb{Z}}$: the pair consisting of the backward and the forward solutions.*

VII. THE UNIQUENESS OF THE PSEUDO-SPECTRUM

The pseudo-spectrum associated to $\phi * \mathbf{x} = \theta * \mathbf{w}^*$ is defined in the literature as

$$\sigma^2 \frac{\theta(e^{-i\omega}) \theta(e^{i\omega})}{\phi(e^{-i\omega}) \phi(e^{i\omega})}.$$

Although *pseudo-spectrum* is unique for each LSDE, we can find more than one *pseudo-covariance function*. For example, backward and forward solutions form a co-stationary pair, however, when ϕ has all roots outside the unit circle, the backward solution with itself is also so (since backward solution is stationary in this case). It is easy to see that *pseudo-covariance function* for the backward-forward solution pair is zero for all but finitely many positive indices, whereas *auto-covariance function* for the backward solution is symmetric.

Therefore, our target in this section is twofold: we want firstly to extend the DTFT outside the Hilbert space; and secondly, to do so in such a way that for a given LSDE, it links the same *pseudo-spectrum* to all *pseudo-covariance functions* of co-stationary solutions. Fortunately those *pseudo-covariance functions* have something in common. Indeed, consider, \mathbf{g} and \mathbf{h} , two co-stationary solutions to $\phi * \mathbf{x} = \theta * \mathbf{w}^*$; since $\Gamma_{\phi * \mathbf{g}, \phi * \mathbf{h}} = \Gamma_{\theta * \mathbf{w}^*, \theta * \mathbf{w}^*}$, then, by Proposition VI.1

$$\phi(X^{-1}) * \phi(X) * \Gamma_{\mathbf{g}, \mathbf{h}} = \theta(X^{-1}) * \theta(X) \cdot \sigma^2. \quad (24)$$

[note the similarity between (24) and (8)].

It follows that for any m

$$X^m * \phi(X^{-1}) * \phi(X) * \Gamma_{g,h} = X^m * \theta(X^{-1}) * \theta(X) \cdot \sigma^2, \quad (25)$$

but $X^m * \phi(X^{-1}) * \phi(X)$ and $X^m * \theta(X^{-1}) * \theta(X)$ are polynomials, provided that we take a large enough m . Thus, *pseudo-covariance functions* are sequences such that when multiplied by a particular polynomial, we get another polynomial.

A. The Extended Fourier Transform, $\overline{\mathcal{F}}$

Our task is now to extend the DTFT in such a way that, for any given pair of polynomials $\psi \neq 0$ and θ , it assigns the same image to all Ψ which verify $\psi * \Psi = \theta$. To retain some of the properties of the DTFT, we should require that if $\phi * \Psi = \theta$ then $\overline{\mathcal{F}}(\phi)\overline{\mathcal{F}}(\Psi) = \overline{\mathcal{F}}(\theta)$ and therefore $\overline{\mathcal{F}}(\Psi) = \frac{\overline{\mathcal{F}}(\theta)}{\overline{\mathcal{F}}(\phi)} = \frac{\mathcal{F}(\theta)}{\mathcal{F}(\phi)}$. Hence, to get a meaningful extension we need to include the set $\left\{ \frac{\mathcal{F}(\theta)}{\mathcal{F}(\phi)} \mid \text{where } \phi \neq 0 \text{ and } \theta \text{ are polynomials} \right\}$ in the transform co-domain and to prove that if $\phi' * \Psi = \theta'$ then $\frac{\mathcal{F}(\theta)}{\mathcal{F}(\phi)} = \frac{\mathcal{F}(\theta')}{\mathcal{F}(\phi')}$. First, we establish the domain and co-domain of the Extended Fourier Transform, and then we define the extension.

a) *Extending the domain of the DTFT*: consider, in addition to the Hilbert space l_2 , the following subspace¹⁰ \mathcal{S} of $\mathbb{R}^{\mathbb{Z}}$:

$$\mathcal{S} \equiv \left\{ \Psi \in \mathbb{R}^{\mathbb{Z}} \mid \begin{array}{l} \text{there exists a polynomial } \phi \neq 0 \\ \text{such that } \phi * \Psi \text{ is also a polynomial} \end{array} \right\}. \quad (26)$$

With the aid of the following statement:

Lemma VII.1. *If $\phi * \Psi = \theta$ and $\phi' * \Psi = \theta'$ with $\phi, \phi' \in \mathbb{R}[X] - \{0\}$ and $\theta, \theta' \in \mathbb{R}[X]$ then $\theta * \phi' = \theta' * \phi$,*

we can show that the map with domain \mathcal{S} and co-domain \mathcal{Q} (see below) given by

$$\Psi \mapsto \frac{\mathcal{F}(\theta)}{\mathcal{F}(\phi)}, \quad (27)$$

where Ψ is such that $\phi * \Psi = \theta$, ($\phi \neq 0$) and where \mathcal{F} is the DTFT¹¹, is well defined; since it does not depend on the election of the pair (ϕ, θ) (see the appendix for the proof). Hence, the extension's domain will be $\mathcal{S} + l_2$.

b) *Extending the co-domain of the DTFT*: the usual Hilbert space $L_2[-\pi, \pi]$ (the DTFT co-domain) is a quotient set where two functions belong to the same *almost everywhere (a.e.) equivalence class* if they differ only in a set of measure zero. If we consider the same equivalence relationship over the set of all complex functions in $\mathbb{C}^{[-\pi, \pi]}$, the *a.e. equivalence class* $[f]$ of a function f has a multiplicative inverse if and only if the set of zeros of f has zero measure. As the DTFT of a non-zero polynomial only has a finite number of zeros, the DTFT has a multiplicative inverse in $\mathbb{C}^{[-\pi, \pi]}/(a.e.)$. Moreover,

¹⁰Note that $(\phi * \Psi \in \mathbb{R}[X] \text{ and } \psi * \Omega \in \mathbb{R}[X]) \Rightarrow (\phi * \psi) * (a\Psi + b\Omega) \in \mathbb{R}[X]$, and $a, b \in \mathbb{R}$. Note also that, from (25), *pseudo-covariance functions* are vectors in this subspace.

¹¹For a sequence $\Phi \equiv \{\phi_t\}_{t \in \mathbb{Z}}$ in l_2 , the DTFT is defined as $[\mathcal{F}(\Phi)](\omega) = \sum_{t=-\infty}^{\infty} \phi_t e^{-2\pi i \omega t}$, belonging to $L_2[-\pi, \pi]$.

since for any $\phi \neq 0$ and θ polynomials, $\frac{\mathcal{F}(\theta)}{\mathcal{F}(\phi)} \in \mathbb{C}^{[-\pi, \pi]}/(a.e.)$, we shall choose \mathcal{Q} as

$$\mathcal{Q} \equiv \left\{ \frac{\mathcal{F}(\theta)}{\mathcal{F}(\phi)} \in \mathbb{C}^{[-\pi, \pi]}/(a.e.) \mid \phi \neq 0, \theta \text{ are polynomials} \right\}, \quad (28)$$

and then the extension's co-domain will be $\mathcal{Q} + L_2[-\pi, \pi]$.

c) *The Extended Fourier Transform, $\overline{\mathcal{F}}$* : both \mathcal{Q} and $L_2[-\pi, \pi]$ are subspaces of $\mathbb{C}^{[-\pi, \pi]}/(a.e.)$ so we define the Extended Fourier Transform as:

$$\begin{aligned} \overline{\mathcal{F}}: \mathcal{S} + l_2 &\longrightarrow \mathcal{Q} + L_2[-\pi, \pi] \\ \Psi + \Omega &\longrightarrow \frac{\mathcal{F}(\theta)}{\mathcal{F}(\phi)} + \mathcal{F}(\Omega) \end{aligned}, \quad (29)$$

where $\phi * \Psi = \theta$, ($\phi \neq 0$). This extension is well defined since given $\Psi, \Psi' \in \mathcal{S}$ and $\Omega, \Omega' \in l_2$ such that $\Psi + \Omega = \Psi' + \Omega'$ then $\frac{\mathcal{F}(\theta)}{\mathcal{F}(\phi)} + \mathcal{F}(\Omega) = \frac{\mathcal{F}(\theta')}{\mathcal{F}(\phi')} + \mathcal{F}(\Omega')$ (see the appendix).

B. The pseudo-spectrum

We are now ready to state our main result:

Theorem VII.2. *Given a pair of polynomials $\psi \neq 0$ and θ , for any pair of co-stationary solutions to $\phi * \mathbf{x} = \theta * \mathbf{w}^*$ the Extended Fourier Transform of its pseudo-covariance function is*

$$\sigma^2 \frac{\theta(e^{-i\omega}) \cdot \theta(e^{i\omega})}{\phi(e^{-i\omega}) \cdot \phi(e^{i\omega})}. \quad (30)$$

This common image for the *Extended Fourier Transforms* of any pair of co-stationary solutions to $\phi * \mathbf{x} = \theta * \mathbf{w}^*$ is the *pseudo-spectrum*.

VIII. WHICH OF LOYNES' DESIRABLE PROPERTIES DOES THIS PSEUDO-SPECTRUM SATISFY?

In a seminal paper, Loynes proposed a list of eight desirable properties regarding the *pseudo-spectrum* [2]. He deals with representations of continuous-time processes which explicitly consider the time-dependence of the frequency content of the signal. But we deal with discrete time processes where the frequency content has no time dependence. Since, as far as we know, our paper is the first formal approach to this type of pseudo-spectrum, we could not find a more appropriate list of desirable properties within our context.

Let us see which of Loynes' properties satisfies our pseudo-spectrum:

- A1: *The pseudo-spectrum is a real function of time and of "frequency", completely determined by the covariance function.* Loynes says this is the minimum that could be assumed about a spectrum. By Theorem VII.2, our *pseudo-spectrum* is a time independent real function completely determined by any of the *pseudo-covariance functions*.
- A2: *The pseudo-spectrum describes the distribution of energy over frequency.* This physical interpretation holds, but here each root (or each complex conjugate pair of roots) on the unit circle produces a pole in the *pseudo-spectrum* with an infinite contribution to the variance (energy).
- A3: *The pseudo-spectrum transforms reasonably, and preferably simply, when the process $\{x_t\}_{t \in \mathbb{Z}}$ is transformed*

linearly. In particular, a knowledge of the spectrum of x determines the spectrum of the transformed process. Loynes says that it would seem that one of properties A2 and A3 is essential if the name spectrum is to be justifiable; A2 describes what it is, and A3 how it can be used. Clearly property A3 is stated by Proposition VI.1.

- A4: *The relationship between pseudo-spectrum and the covariance function is one to one.* Loynes says that this is not altogether essential, but one would not wish to lose too much information in passing from covariance to spectrum. He speculates that this is probably the simplest way of ensuring that the second part of A3 is satisfied. Although our *pseudo-spectrum* does not satisfy A4 (the Extended Fourier Transform is not invertible), Proposition VI.1 holds; and the *pseudo-spectrum* is unique for each LSDE (Theorem VII.2).
- A5: *The pseudo-spectrum reduces to the ordinary spectrum, or some simple transformation if x is in fact stationary.* Loynes says this is probably essential. In our case, since for any (second-order) stationary stochastic process y the pair (y^*, y^*) is co-stationary, the pseudo-covariance function $\Gamma_{y^*, y^*} = \Gamma_y \in l_2$ is the *auto-covariance function* of y (see Section VI), and therefore its Extended Fourier Transform $\bar{\mathcal{F}}(\Gamma_y)$ and its DTFT $\mathcal{F}(\Gamma_y)$ are the same.
- A6: *If the process is composed of a succession of stationary parts, say $\{x_{1t}\}_{t \leq 0}$ and $\{x_{2t}\}_{t > 0}$, then the spectrum is also composed of the corresponding succession of (stationary) spectra.* This property is conceived in the context of time-variable parameter models which is not our context (see [2]).
- A7: *The pseudo-spectrum is estimable in principle, probably from (infinite) length of record.* In our case, a parametric estimation is possible following the Box-Jenkins modeling approach to identify and estimate the AR, ϕ , and MA, θ , polynomials.
- A8: *The pseudo-spectrum is the Fourier transform, or some related transform, of some apparently meaningful quantity.* Loynes says that such a property would be welcome, but it does not seem important. Our *pseudo-spectrum* is the Extended Fourier Transform of the *pseudo-covariance function* of any co-stationary pair of solutions of LSDE. In addition, it is the quotient of two positive defined functions and, hence, it is also positive defined (another desirable feature).

Hence, our *pseudo-spectrum* satisfies six out of eighth properties of Loynes' list; particularly all those qualified as "essential".

IX. EXAMPLES AND APPLICATIONS

A. The simplest example: AR(1) models

Consider the set of models

$$x_t - ax_{t-1} = w_t \quad (31)$$

where $a \in [0, 1]$. When a is zero, x is the white noise process w ; when $a = 1$, it corresponds to the Random Walk model we have used as illustration in Section IV. For $0 < a < 1$, there are no roots on the unit circle. Following (12), since

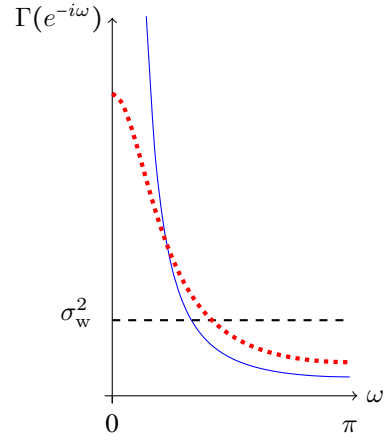


Fig. 1. Pseudo-spectrum for the white noise, $a = 0$ (dashed); the AR(1) model with $a = 0.5$ (dotted); and for the Random Walk model, $a = 1$ (continuous).

$\phi = 1 - aX$ and $\theta = 1$, these models have a *pseudo-spectrum* whose formula is

$$f_x(\omega) = \frac{\sigma_w^2}{(1 - ae^{-i\omega})(1 - ae^{i\omega})} = \frac{\sigma_w^2}{1 - 2a \cos(\omega) + a^2}. \quad (32)$$

Fig. 1 shows the *pseudo-spectrum* for models in three particular cases.

It should be noted that, when $0 \leq a < 1$, *pseudo-spectrum* and spectrum of the stationary solution to (31) is the same function. But, when $a = 1$ only the *pseudo-spectrum* is defined. Hence, *pseudo-spectrum* generalize the spectrum to the non-stationary case.

B. Sum of non-stationary signals

Suppose that $Z = S + N$, where S , and N are unobservable signal and noise components that follow the models,

$$\phi_S * S = \theta_S * b \quad (33)$$

$$\phi_N * N = \theta_N * c, \quad (34)$$

where each of the pairs of polynomials $\{\phi_S, \theta_S\}$, and $\{\phi_N, \theta_N\}$ have their zeros lying on or outside the unit circle and have no common zeros, and b , and c are mutually independent white noise processes. The non-stationary signal extraction problem in this framework has been studied since the sixties (see [21], [26], [28], [34]–[38]). As Bell pointed out in [35]: *typically, the solution for the stationary case has been borrowed and used in the nonstationary case.* But this has been done without a proper definition of *pseudo-spectra* and *pseudo-covariance functions*. In this section we justify a result frequently used in the literature: the “spectrum” of Z is the sum of the “spectra” of S and N even in the non-stationary case.

Consider ϕ , the *least common multiple* of ϕ_S and ϕ_N ; and let φ_S and φ_N be the polynomials such that $\phi = \varphi_S * \phi_S = \varphi_N * \phi_N$. We can multiply (33) and (34) by φ_S and φ_N respectively to get

$$\phi * S = \varphi_S * \theta_S * b \quad (35)$$

$$\phi * N = \varphi_N * \theta_N * c. \quad (36)$$

Solutions to (33) and (34) are also solutions for (35) and (36) respectively. Adding the last two equations we get

$$\begin{aligned}\phi * (\mathbf{S} + \mathbf{N}) &= \phi * \mathbf{Z} = \vartheta_S * \mathbf{b} + \vartheta_N * \mathbf{c} \\ &= \theta * \mathbf{a},\end{aligned}\quad (37)$$

where we write each product $\varphi_k * \theta_k$ as ϑ_k ; and where the last equality follows from the fact that \mathbf{b} and \mathbf{c} are independent white noise processes, so the right hand side is a stationary process with a finite moving average representation $\theta * \mathbf{a}$.

Now, let's consider \mathcal{D} a Hilbert basis of $L_2(S, \mathcal{B}, P)$ such that $\{\frac{b_t}{\sigma} | t \in \mathbb{Z}\} \subset \mathcal{D}$ and $\{\frac{c_t}{\sigma} | t \in \mathbb{Z}\} \subset \mathcal{D}$, where $\frac{b_t}{\sigma}$, $\frac{c_t}{\sigma}$ are the standardized random variables of the two mutually independent white noise processes \mathbf{b} and \mathbf{c} . Using the corresponding sequences of embeddings \mathbf{b}^* and \mathbf{c}^* , we can consider the backward-forward co-stationary solution pairs $(\mathbf{S}_\bullet, \mathbf{S}^\flat)$, and $(\mathbf{N}_\bullet, \mathbf{N}^\flat)$ to

$$\phi_S * \mathbf{S} = \vartheta_S * \mathbf{b}^* \quad (38)$$

$$\phi_N * \mathbf{N} = \vartheta_N * \mathbf{c}^*. \quad (39)$$

By Proposition VI.1, the corresponding pseudo-covariance functions verify

$$\phi(x^{-1}) * \phi(x) * \Gamma_{\mathbf{S}_\bullet, \mathbf{S}^\flat} = \vartheta_S(x^{-1}) * \vartheta_S(x) \cdot \sigma_b^2 \quad (40)$$

$$\phi(x^{-1}) * \phi(x) * \Gamma_{\mathbf{N}_\bullet, \mathbf{N}^\flat} = \vartheta_N(x^{-1}) * \vartheta_N(x) \cdot \sigma_c^2. \quad (41)$$

Adding these equations we get

$$\begin{aligned}\phi(x^{-1}) * \phi(x) * (\Gamma_{\mathbf{S}_\bullet, \mathbf{S}^\flat} + \Gamma_{\mathbf{N}_\bullet, \mathbf{N}^\flat}) \\ = \vartheta_S(x^{-1}) * \vartheta_S(x) \cdot \sigma_b^2 + \vartheta_N(x^{-1}) * \vartheta_N(x) \cdot \sigma_c^2.\end{aligned}\quad (42)$$

If we choose the solution pair $\mathbf{Z}_\bullet = \mathbf{S}_\bullet + \mathbf{N}_\bullet$ and $\mathbf{Z}^\flat = \mathbf{S}^\flat + \mathbf{N}^\flat$ to the embedding of (37), then

$$\phi(x^{-1}) * \phi(x) * \Gamma_{\mathbf{Z}_\bullet, \mathbf{Z}^\flat} = \theta(x^{-1}) * \theta(x) \cdot \sigma_a^2. \quad (43)$$

Since \mathbf{b} and \mathbf{c} are mutually independent white noise processes, the backward-forward co-stationary pair $(\mathbf{Z}_\bullet, \mathbf{Z}^\flat)$ has a particular property: the *pseudo-covariance* generating function of the ‘‘addition’’ pair $(\mathbf{Z}_\bullet, \mathbf{Z}^\flat)$ is the addition of the *pseudo-covariance* generating functions of $(\mathbf{S}_\bullet, \mathbf{S}^\flat)$, and $(\mathbf{N}_\bullet, \mathbf{N}^\flat)$:

$$\Gamma_{\mathbf{S}_\bullet + \mathbf{N}_\bullet, \mathbf{S}^\flat + \mathbf{N}^\flat} = \Gamma_{\mathbf{S}_\bullet, \mathbf{S}^\flat} + \Gamma_{\mathbf{N}_\bullet, \mathbf{N}^\flat}; \quad (44)$$

(see the appendix for the proof) so, the left hand sides of (42) and (43) are equal (cp. [28, Equation 1.4]).

From (44), it follows a more general result. Since the image of the Extended Fourier Transform of the *pseudo-covariance function* of any pair of co-stationary solutions of a linear stochastic difference equation is common, then the *pseudo-spectrum* associated to (37) is the sum of the *pseudo-spectra* associated to (33) and (34)

$$\Gamma_Z(e^{-i\omega}) = \Gamma_S(e^{-i\omega}) + \Gamma_N(e^{-i\omega}); \quad (45)$$

or

$$\begin{aligned}\frac{\theta(e^{-i\omega}) \cdot \theta(e^{i\omega})}{\phi(e^{-i\omega}) \cdot \phi(e^{i\omega})} \sigma_a^2 = \\ \sigma_b^2 \frac{\theta_S(e^{-i\omega}) \cdot \theta_S(e^{i\omega})}{\phi_S(e^{-i\omega}) \cdot \phi_S(e^{i\omega})} + \sigma_c^2 \frac{\theta_N(e^{-i\omega}) \cdot \theta_N(e^{i\omega})}{\phi_N(e^{-i\omega}) \cdot \phi_N(e^{i\omega})}.\end{aligned}\quad (46)$$

This decomposition of a *pseudo-spectrum* as a sum of *pseudo-spectra* is widely used in the literature of unobserved component models, and Seasonal Adjustment of Economic Time Series (see references in Section I). The above is a formal justification for this decomposition, that it is used in the next example.

C. Dynamic Harmonic Regression

The Dynamic Harmonic Regression (DHR) model [20] is based on a spectral approach, under the hypothesis that the observed time series z is periodic or quasi-periodic and can be decomposed into several components whose variances are concentrated around certain frequencies: e.g. at a fundamental frequency and its associated sub-harmonics. This is an appropriate hypothesis if the observed time series has well defined spectral peaks, which implies that its variance is concentrated around narrow frequency bands. By ‘quasi-periodic’ we mean that the amplitude and the phase of the periodicity may vary over time. The DHR model is the sum of several Unobserved Components:

$$\sum_{j=0}^R \mathbf{s}^j + \mathbf{e} \quad (47)$$

where the irregular component, \mathbf{e} , is normally distributed with zero mean and variance σ_e^2 ; and each DHR component \mathbf{s}^j has the form

$$\mathbf{s}_t^j = a_t^j \cos(\omega_j t) + b_t^j \sin(\omega_j t). \quad (48)$$

Oscillations of each DHR component, \mathbf{s}^j , are modulated by the stochastic processes $\{a_t^j\}_{t \in \mathbb{Z}}$ and $\{b_t^j\}_{t \in \mathbb{Z}}$. Both stochastic processes, \mathbf{a}^j and \mathbf{b}^j , are solutions to the same AR(1) or AR(2) difference equations with at least one root on the unit circle. The frequency ω_j is associated to the j th component. Usually $j = 0$ corresponds to the zero frequency term, that is, the trend; and the other components ($j = 1, \dots, R$) correspond to the seasonal frequency and its harmonics. Hence, the complete DHR model is

$$\sum_{j=0}^R \{a_t^j \cos(\omega_j t) + b_t^j \sin(\omega_j t)\} + e_t \quad (49)$$

This model can be considered a straightforward extension of the classical harmonic regression model, in which the gain and phase of the harmonic components can vary randomly due to the stochastic processes \mathbf{a}^j and \mathbf{b}^j .

In [39] it is shown that each \mathbf{s}^j has an alternative representation as a solution to

$$\phi^j * \mathbf{s}^j = \theta^j * \mathbf{w}^j, \quad j = 0, \dots, R; \quad (50)$$

where ϕ^j has roots on the unit circle, and $\mathbf{w}^j \sim N(0, \sigma_j^2)$. Hence, the *pseudo-spectrum* of the DHR model is the sum of the *pseudo-spectra* of its components:

$$f_{dhr}(\omega, \sigma^2) = \sum_{j=1}^R \sigma_j^2 \frac{\theta^j(e^{-i\omega}) \theta^j(e^{i\omega})}{\phi^j(e^{-i\omega}) \phi^j(e^{i\omega})} + \sigma_e^2; \quad (51)$$

where the variances in vector $\sigma^2 = [\sigma_0^2, \dots, \sigma_R^2, \sigma_e^2]$ are the unknown hyper-parameters of the model.

The model is fitted in the frequency domain by seeking the vector σ^2 that minimizes the euclidean distance¹²

$$\min_{\sigma^2 \in \mathbb{R}^{R+1}} \|f_z(\omega) - f_{dhr}(\omega, \sigma^2)\|, \quad (52)$$

where $f_z(\omega)$ is the spectrum of the observed time series. This strategy has an intuitive appeal but, since DHR components are non-stationary, the corresponding *pseudo-spectra* have poles; and therefore the norm is not defined.

In order to find the Ordinary Least Squares (OLS) solution, it is needed to eliminate the unit modulus AR roots of $f_{dhr}(\omega, \sigma^2)$. Fortunately, it is possible to exploit the algebraic structure of the *pseudo-spectra* model. If (52) is multiplied by the function $\Psi(\omega) = \varphi(e^{-i\omega})\varphi(e^{i\omega})$, where φ is the minimum order polynomial with all unit modulus AR roots of the complete DHR model, then we can try the alternative minimization problem:

$$\min_{\sigma^2 \in \mathbb{R}^{R+2}} \|\Psi(\omega) \cdot f_y(\omega) - \Psi(\omega) \cdot f_{dhr}(\omega, \sigma^2)\|. \quad (53)$$

Since $\Psi(\omega) \cdot f_y(\omega)$ and $\Psi(\omega) \cdot f_{dhr}(\omega, \sigma^2)$ are functions in $L_2[-\pi, \pi]$, it follows that (53) can be solved by OLS (see [39]).

X. CONCLUSIONS

If the spectrum is defined within the algebraic framework provided in this paper, spectrum and pseudo-spectrum are the same functions. But even when the spectrum is not-defined, the pseudo-spectrum is well defined for any LSDE.

Contrary to the case of the spectrum, and since the pseudo-spectrum is the Extended Fourier Transform of the pseudo-covariance function of *any pair* of co-stationary solutions of $\phi * \mathbf{x} = \theta * \mathbf{w}^*$, the pseudo-spectrum is not associated to any particular solution, neither to any particular pair of co-stationary solutions. The Extended Fourier Transform is not invertible as an operator and therefore the pseudo-spectrum is associated to the difference equation itself.

The convolution type solutions to (23) that we use in the paper ($\mathbf{f} = (\Psi * \theta) * \mathbf{w}^*$ where $\phi * \Psi = 1$) closely resemble the ones used in the literature (see the references given in Section III). However, we had to define them in a different framework, the space $[L(\mathcal{D})^*]^{\mathbb{Z}}$, so as to avoid the convergence issues of convolution expressions. The embedding of $L_2(S, \mathcal{B}, P)$ in $L(\mathcal{D})^*$ is an isomorphic isometry, and therefore our algebraic model constitutes a generalization proper of the spectral theory to the case in which the AR polynomial ϕ has roots on the unit circle.

APPENDIX

If $\mathbf{f} \in \mathcal{A}$, $\theta \in \mathbb{R}[X]$ and $\Psi \in \mathbb{R}^{\mathbb{Z}}$ then $\Psi * (\theta * \mathbf{f}) = (\Psi * \theta) * \mathbf{f}$.

Proof: Let $\theta = \theta_0 + \dots + \theta_n X^n$. We only need to check that for any $i \in \mathbb{Z}$ and any $v \in L(\mathcal{D})$, $[\Psi * (\theta * \mathbf{f})]_i(v) =$

$[(\Psi * \theta) * \mathbf{f}]_i(v)$. Since $v \in L(\mathcal{D})$ we can choose m, m' such that $f_i(v) \neq 0 \Rightarrow m \leq i \leq m'$. Then

$$\begin{aligned} & [\Psi * (\theta * \mathbf{f})]_i(v) \\ &= \sum_{r+s=i} \Psi_r (\theta * \mathbf{f})_s(v) = \sum_{r+s=i} \Psi_r \sum_{p+q=s} \theta_p \mathbf{f}_q(v) \\ &= \sum_{r+s=i} \Psi_r \sum_{q=m}^{m'} \theta_{s-q} \mathbf{f}_q(v) \quad (\text{making } t = s - q) \\ &= \sum_{r+(t+q)=i} \Psi_r \sum_{q=m}^{m'} \theta_t \mathbf{f}_q(v) = \sum_{t=0}^n \Psi_{i-t-q} \sum_{q=m}^{m'} \theta_t \mathbf{f}_q(v) \\ &= \sum_{t=0}^n \sum_{q=m}^{m'} \Psi_{i-t-q} \theta_t \mathbf{f}_q(v) = \sum_{q=m}^{m'} \sum_{t=0}^n \Psi_{i-t-q} \theta_t \mathbf{f}_q(v) \\ &= \sum_{q=m}^{m'} \left(\sum_{t=0}^n \Psi_{i-t-q} \theta_t \right) \mathbf{f}_q(v) = \sum_{q=m}^{m'} (\Psi * \theta)_{i-q} \mathbf{f}_q(v) \\ &= \sum_{r+s=i} (\Psi * \theta)_r \mathbf{f}_s(v) = [(\Psi * \theta) * \mathbf{f}]_i(v). \end{aligned}$$

If $\mathbf{f} \in \mathcal{A}$, $\Psi \in \mathbb{R}^{\mathbb{Z}}$ and $\phi \in \mathbb{R}[X]$ then $\phi * (\Psi * \mathbf{f}) = (\phi * \Psi) * \mathbf{f}$.

Proof: Let $\phi = \phi_0 + \dots + \phi_n X^n$, as before; we again only need to check that for any $i \in \mathbb{Z}$ and any $v \in L(\mathcal{D})$, $[\phi * (\Psi * \mathbf{f})]_i(v) = [(\phi * \Psi) * \mathbf{f}]_i(v)$. Since $v \in L(\mathcal{D})$ we chose m, m' such that $f_i(v) \neq 0 \Rightarrow m \leq i \leq m'$. Then

$$\begin{aligned} & [\phi * (\Psi * \mathbf{f})]_i(v) = \sum_{k=0}^n \phi_k (\Psi * \mathbf{f})_{i-k}(v) \\ &= \sum_{k=0}^n \phi_k \sum_{r+s=i-k} \Psi_r \mathbf{f}_s(v) = \sum_{k=0}^n \phi_k \sum_{s=m}^{m'} \Psi_{i-k-s} \mathbf{f}_s(v) \\ &= \sum_{s=m}^{m'} \left(\sum_{k=0}^n \phi_k \Psi_{i-k-s} \right) \mathbf{f}_s(v) = \sum_{s=m}^{m'} (\phi * \Psi)_{i-s} \mathbf{f}_s(v) \\ &= \sum_{s=-\infty}^{\infty} (\phi * \Psi)_{i-s} \mathbf{f}_s(v) = [(\phi * \Psi) * \mathbf{f}]_i(v). \end{aligned}$$

Proof of Proposition VI.1: $\phi * (\Psi * (\theta * \mathbf{w}^*)) = (\phi * \Psi) * (\theta * \mathbf{w}^*) = 1 * (\theta * \mathbf{w}^*)$.

Proof of Proposition VI.1: By definition:

$$\begin{aligned} & \left\langle \sum_{k=0}^n \theta_k f_{i+t-k}, \sum_{l=0}^m \phi_l g_{j+t-l} \right\rangle \\ &= \sum_{k=0}^n \theta_k \sum_{l=0}^m \phi_l \langle f_{i+t-k}, g_{j+t-l} \rangle = \sum_{k=0}^n \theta_k \sum_{l=0}^m \phi_l \langle f_{i-k}, g_{j-l} \rangle \\ &= \left\langle \sum_{k=0}^n \theta_k f_{i-k}, \sum_{l=0}^m \phi_l g_{j-l} \right\rangle; \end{aligned}$$

hence $\theta * \mathbf{f}$ and $\phi * \mathbf{g}$ are co-stationary, and then

$$\begin{aligned} & \Gamma_{\theta * \mathbf{f}, \phi * \mathbf{g}}(j) = \sum_{k=0}^n \theta_k \sum_{l=0}^m \phi_l \langle f_{-k}, g_{j-l} \rangle \\ &= \sum_{k=0}^n \theta_k \sum_{l=0}^m \phi_l \Gamma_{\mathbf{f}, \mathbf{g}}(j - l + k) = \sum_{k=0}^n \theta_k [\phi(X) * \Gamma_{\mathbf{f}, \mathbf{g}}](j + k) \\ &= \sum_{k=0}^n [\theta(X^{-1})]_{-k} [\phi(X) * \Gamma_{\mathbf{f}, \mathbf{g}}](j + k) \\ &= [\theta(X^{-1}) * \phi(X) * \Gamma_{\mathbf{f}, \mathbf{g}}](j). \end{aligned}$$

¹²The algorithm proposed in [20] seeks the vector $\mathbf{NVR} = [1, \text{NVR}_0, \dots, \text{NVR}_R]$, where $\text{NVR}_j = \sigma_j^2 / \bar{\sigma}^2$, using the residual variance $\bar{\sigma}^2$ from a fitted AR model of the observed series z .

Proof of Lemma VI.2: We know there are integers m and m' such as $k < m \Rightarrow \Psi_k = 0$ and $k > m' \Rightarrow \Omega_k = 0$. Then:

$$\begin{aligned} & \langle [\Psi * \mathbf{w}^*]_{i+t}, [\Omega * \mathbf{w}^*]_{j+t} \rangle \\ &= \sum_{d \in \mathcal{D}} [\Psi * \mathbf{w}^*]_{i+t}(d) [\Omega * \mathbf{w}^*]_{j+t}(d) \\ &= \sum_{h \in \mathbb{Z}} [\Psi * \mathbf{w}^*]_{i+t}(w_h/\sigma) [\Omega * \mathbf{w}^*]_{j+t}(w_h/\sigma) \\ &= \sigma^2 \sum_{h \in \mathbb{Z}} \Psi_{i+t-h} \Omega_{j+t-h} = \sigma^2 \sum_{h' \in \mathbb{Z}} \Psi_{i-h'} \Omega_{j-h'} \\ &= \sigma^2 \sum_{h'=j-m'}^{h'=i-m} \Psi_{i-h'} \Omega_{j-h'} \end{aligned}$$

Proof of Theorem VI.3: Let Λ and Υ be the backward and forward inverses of ϕ defined above in the text; from Proposition V.1 we know that $(\Lambda * \theta) * \mathbf{w}^*$ and $(\Upsilon * \theta) * \mathbf{w}^*$ are solutions to (23). On the other hand, from Lemma VI.2 we know that $\Lambda * \mathbf{w}^*$ and $\Upsilon * \mathbf{w}^*$ are co-stationary, it follows that $\theta * \Lambda * \mathbf{w}^*$ and $\theta * \Upsilon * \mathbf{w}^*$ are also co-stationary (Proposition VI.1). Further, from the properties in Section V-A1 we get $\theta * (\Lambda * \mathbf{w}^*) = (\theta * \Lambda) * \mathbf{w}^* = (\Lambda * \theta) * \mathbf{w}^*$, which it is the backward solution. A similar argument shows that $\theta * (\Upsilon * \mathbf{w}^*) = (\Upsilon * \theta) * \mathbf{w}^*$ gives the forward solution. ■

Proof of Lemma VII.1: Multiplying $\phi * \Psi = \theta$ by θ' ; and $\phi' * \Psi = \theta'$ by θ we get

$$\begin{aligned} \theta' * \phi * \Psi &= \theta' * \theta \\ \theta * \phi' * \Psi &= \theta * \theta', \end{aligned}$$

thus $(\theta' * \phi - \theta * \phi') * \Psi = 0$. Thus, there are only two possibilities:

- 1) If $(\theta' * \phi - \theta * \phi') = 0$, from which $\theta * \phi' = \theta' * \phi$.
- 2) If $(\theta' * \phi - \theta * \phi') \neq 0$, since $0 = \phi * \left[(\theta' * \phi - \theta * \phi') * \Psi \right] = (\theta' * \phi - \theta * \phi') * [\phi * \Psi] = (\theta' * \phi - \theta * \phi') * \theta$; we conclude $\theta = 0$. And for the same reason $\theta' = 0$. Consequently we also get $\theta * \phi' = \theta' * \phi$. ■

Proof of “Map (27) is well defined”: By Lemma VII.1, the fraction $\frac{\mathcal{F}(\theta)}{\mathcal{F}(\phi)}$ is uniquely determined by the condition $\phi * \Psi = \theta$: indeed, if we also had $\phi' * \Psi = \theta'$, then

$$\begin{aligned} \theta * \phi' = \theta' * \phi &\Rightarrow \mathcal{F}(\theta * \phi') = \mathcal{F}(\theta' * \phi) \\ &\Rightarrow \mathcal{F}(\theta) \cdot \mathcal{F}(\phi') = \mathcal{F}(\theta') \cdot \mathcal{F}(\phi) \\ &\Rightarrow \frac{\mathcal{F}(\theta)}{\mathcal{F}(\phi)} = \frac{\mathcal{F}(\theta')}{\mathcal{F}(\phi')}. \end{aligned}$$

Proof of “The Extended Fourier Transform is well defined”: Let us assume that $\Psi + \Omega = \Psi' + \Omega'$, where $\psi * \Psi = \beta$, $\psi' * \Psi' = \beta'$ (with $\psi, \psi' \in \mathbb{R}[X] - \{0\}$ and $\beta, \beta' \in \mathbb{R}[X]$) and $\Omega, \Omega' \in l_2$. Then, since $\Psi - \Psi' = \Omega' - \Omega \in \mathcal{S} \cap l_2$, there exists $\phi \in \mathbb{R}[X] - \{0\}$ and $\theta \in \mathbb{R}[X]$ such that

$$\phi * (\Psi - \Psi') = \theta = \phi * (\Omega' - \Omega).$$

Thus we get on the one hand, $\mathcal{F}(\theta) = \mathcal{F}(\phi) \cdot \mathcal{F}(\Omega' - \Omega)$ and consequently $\mathcal{F}(\Omega') - \mathcal{F}(\Omega) = \frac{\mathcal{F}(\theta)}{\mathcal{F}(\phi)}$. And on the other hand, since $\psi * \Psi = \beta$ and $\psi' * \Psi' = \beta'$, it follows that

$$\begin{aligned} \psi' * \psi * \Psi &= \psi' * \beta \\ \psi * \psi' * \Psi' &= \psi * \beta'; \end{aligned}$$

and therefore $\psi' * \psi * (\Psi - \Psi') = \psi' * \beta - \psi * \beta'$. Now, using Lemma VII.1

$$\psi' * \psi * \theta = (\psi' * \beta - \psi * \beta') * \phi;$$

hence $\mathcal{F}(\psi') \cdot \mathcal{F}(\psi) \cdot \mathcal{F}(\theta) = (\mathcal{F}(\psi') \cdot \mathcal{F}(\beta) - \mathcal{F}(\psi) \cdot \mathcal{F}(\beta')) \cdot \mathcal{F}(\phi)$, and then

$$\frac{\mathcal{F}(\theta)}{\mathcal{F}(\phi)} = \frac{\mathcal{F}(\psi') \cdot \mathcal{F}(\beta) - \mathcal{F}(\psi) \cdot \mathcal{F}(\beta')}{\mathcal{F}(\psi') \cdot \mathcal{F}(\psi)} = \frac{\mathcal{F}(\beta)}{\mathcal{F}(\psi)} - \frac{\mathcal{F}(\beta')}{\mathcal{F}(\psi')}.$$

Thus $\mathcal{F}(\Omega') - \mathcal{F}(\Omega) = \frac{\mathcal{F}(\theta)}{\mathcal{F}(\phi)} = \frac{\mathcal{F}(\beta)}{\mathcal{F}(\psi)} - \frac{\mathcal{F}(\beta')}{\mathcal{F}(\psi')}$, and therefore

$$\frac{\mathcal{F}(\beta)}{\mathcal{F}(\psi)} + \mathcal{F}(\Omega) = \frac{\mathcal{F}(\beta')}{\mathcal{F}(\psi')} + \mathcal{F}(\Omega').$$

Proof of Theorem VII.2: From (25) we know $\Gamma_{g,h}$ is in \mathcal{S} . Therefore, since $\mathcal{F}(\Phi * \Omega) = \mathcal{F}(\Phi)\mathcal{F}(\Omega)$, the pseudo-spectrum, $\mathcal{F}(\Gamma_{g,h})$, is

$$\begin{aligned} \sigma^2 \frac{\mathcal{F}(X^m * \theta(X^{-1}) * \theta(X))}{\mathcal{F}(X^m * \phi(X^{-1}) * \phi(X))} &= \sigma^2 \frac{\mathcal{F}(\theta(X^{-1})) \cdot \mathcal{F}(\theta(X))}{\mathcal{F}(\phi(X^{-1})) \cdot \mathcal{F}(\phi(X))} \\ &= \sigma^2 \frac{\theta(e^{-i\omega}) \cdot \theta(e^{i\omega})}{\phi(e^{-i\omega}) \cdot \phi(e^{i\omega})}. \end{aligned}$$

Proof of Equation (44): Let $\{\frac{b_t}{\sigma} | t \in \mathbb{Z}\} = \mathcal{B}$, $\{\frac{c_t}{\sigma} | t \in \mathbb{Z}\} = \mathcal{C}$, then we have $\mathcal{B} \cap \mathcal{C} = \emptyset$ and $\mathcal{B} \cup \mathcal{C} \subset \mathcal{D}$. Using the backward and forward recursive formulas (21) and (22); if denote the backward solution \mathcal{S}_\bullet to (38) as $\Psi_S * \mathbf{b}^*$, where $\Psi_S = \Lambda_S * \theta_S$, and the forward solution \mathcal{S}_\circ to (38) as $\Omega_S * \mathbf{b}^*$, where $\Omega_S = \Upsilon_S * \theta_S$; and if we follow the same notation convention for the backward and forward solution pair $(\mathcal{N}_\bullet, \mathcal{N}^\circ)$ to (39), then $\Gamma_{\mathcal{S}_\bullet + \mathcal{N}_\bullet, \mathcal{S}^\circ + \mathcal{N}^\circ}$ is

$$\begin{aligned} & \langle [\Psi_N * \mathbf{b}^* + \Psi_S * \mathbf{c}^*]_{i+t}, [\Omega_N * \mathbf{b}^* + \Omega_S * \mathbf{c}^*]_{j+t} \rangle = \\ & \sum_{d \in \mathcal{D}} [\Psi_N * \mathbf{b}^* + \Psi_S * \mathbf{c}^*]_{i+t}(d) \cdot [\Omega_N * \mathbf{b}^* + \Omega_S * \mathbf{c}^*]_{j+t}(d) = \\ & \sum_{d \in \mathcal{B}} [\Psi_N * \mathbf{b}^* + \Psi_S * \mathbf{c}^*]_{i+t}(d) \cdot [\Omega_N * \mathbf{b}^* + \Omega_S * \mathbf{c}^*]_{j+t}(d) + \\ & \sum_{d \in \mathcal{C}} [\Psi_N * \mathbf{b}^* + \Psi_S * \mathbf{c}^*]_{i+t}(d) \cdot [\Omega_N * \mathbf{b}^* + \Omega_S * \mathbf{c}^*]_{j+t}(d) + \\ & \sum_{d \in \mathcal{D} - (\mathcal{B} \cup \mathcal{C})} [\Psi_N * \mathbf{b}^* + \Psi_S * \mathbf{c}^*]_{i+t}(d) \cdot [\Omega_N * \mathbf{b}^* + \Omega_S * \mathbf{c}^*]_{j+t}(d) = \\ & \sum_{d \in \mathcal{B}} [\Psi_N * \mathbf{b}^*]_{i+t}(d) \cdot [\Omega_N * \mathbf{b}^*]_{j+t}(d) + \\ & \sum_{d \in \mathcal{C}} [\Psi_S * \mathbf{c}^*]_{i+t}(d) \cdot [\Omega_S * \mathbf{c}^*]_{j+t}(d) + 0 = \\ & \sum_{d \in \mathcal{D}} [\Psi_N * \mathbf{b}^*]_{i+t}(d) \cdot [\Omega_N * \mathbf{b}^*]_{j+t}(d) + \\ & \sum_{d \in \mathcal{D}} [\Psi_S * \mathbf{c}^*]_{i+t}(d) \cdot [\Omega_S * \mathbf{c}^*]_{j+t}(d) = \\ & \langle [\Psi_N * \mathbf{b}^*]_{i+t}, [\Omega_N * \mathbf{b}^*]_{j+t} \rangle + \langle [\Psi_S * \mathbf{c}^*]_{i+t}, [\Omega_S * \mathbf{c}^*]_{j+t} \rangle \end{aligned}$$

So $\Gamma_{\mathcal{S}_\bullet + \mathcal{N}_\bullet, \mathcal{S}^\circ + \mathcal{N}^\circ}$ is equal to $\Gamma_{\mathcal{S}_\bullet, \mathcal{S}^\circ} + \Gamma_{\mathcal{N}_\bullet, \mathcal{N}^\circ}$. ■

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers and the associate editor Dr. A. Napolitano for their comments. Their suggestions helped improve this article significantly. We are also indebted to Dr. R. Banerjee, Dr. J. Crespo, Dr. A. García-Hiernaux and Dr. M. Jerez.

REFERENCES

- [1] M. B. Priestley, "Evolutionary spectra and non-stationary processes," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 27, no. 2, pp. 204–237, 1965.
- [2] R. M. Loynes, "On the concept of the spectrum for non-stationary processes," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 30, no. 1, pp. 1–30, 1968. [Online]. Available: <http://www.jstor.org/stable/2984457>
- [3] D. Tjøstheim, "Spectral generating operators for non-stationary processes," *Advances in Applied Probability*, vol. 8, no. 4, pp. 831–846, 1976.
- [4] W. Martin, "Line tracking in nonstationary processes," *Signal Processing*, vol. 3, no. 2, pp. 147–155, 1981.
- [5] W. Martin and P. Flandrin, "Wigner-Ville spectral analysis of non-stationary processes," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 33, no. 6, pp. 1461–1470, December 1985.
- [6] C. Detka and A. El-Jaroudi, "The transitory evolutionary spectrum," in *Acoustics, Speech, and Signal Processing, 1994. ICASSP-94., 1994 IEEE International Conference on*, vol. 4. IEEE, 1994, pp. IV–289.
- [7] R. Dahlhaus, "Fitting time series models to nonstationary processes," *The Annals of Statistics*, vol. 25, no. 1, pp. 1–37, 1997.
- [8] G. Matz, F. Hlawatsch, and W. Kozek, "Generalized evolutionary spectral analysis and the weyl spectrum of nonstationary random processes," *Signal Processing, IEEE Transactions on*, vol. 45, no. 6, pp. 1520–1534, 1997.
- [9] G. Matz and F. Hlawatsch, "Nonstationary spectral analysis based on time-frequency operator symbols and underspread approximations," *Information Theory, IEEE Transactions on*, vol. 52, no. 3, pp. 1067–1086, 2006.
- [10] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal processing*, vol. 86, no. 4, pp. 639–697, 2006.
- [11] P. Flandrin, M. Amin, S. McLaughlin, and B. Torresani, "Time-frequency analysis and applications [from the guest editors]," *Signal Processing Magazine, IEEE*, vol. 30, no. 6, pp. 19–150, 2013.
- [12] W. Tych, D. J. Pedregal, P. C. Young, and J. Davies, "An unobserved component model for multi-rate forecasting of telephone call demand: the design of a forecasting support system," *International Journal of Forecasting*, vol. 18, no. 4, pp. 673–695, Oct. 2002. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0169207002000717>
- [13] T. Vercauteren, P. Aggarwal, X. Wang, and T.-H. Li, "Hierarchical forecasting of web server workload using sequential monte carlo training," *Signal Processing, IEEE Transactions on*, vol. 55, no. 4, pp. 1286–1297, April 2007.
- [14] S. Becker, C. Halsall, W. Tych, R. Kallenborn, Y. Su, and H. Hung, "Long-term trends in atmospheric concentrations of α - and γ -hch in the arctic provide insight into the effects of legislation and climatic fluctuations on contaminant levels," *Atmospheric Environment*, vol. 42, no. 35, pp. 8225–8233, Nov. 2008. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1352231008006857>
- [15] T. Vogt, E. Hoehn, P. Schneider, A. Freund, M. Schirmer, and O. A. Cirpka, "Fluctuations of electrical conductivity as a natural tracer for bank filtration in a losing stream," *Advances in Water Resources*, vol. 33, no. 11, pp. 1296–1308, Nov. 2010. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0309170810000394>
- [16] M. Bujosa, A. García-Ferrer, and A. de Juan, "Predicting recessions with factor linear dynamic harmonic regressions," *Journal of Forecasting*, vol. 32, no. 6, pp. 481–499, 2013.
- [17] T. W. Hungerford, *Algebra*. Hold, Rinehart and Winston, inc, 1974.
- [18] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*, ser. Springer series in Statistics. New York: Springer-Verlag, 1987.
- [19] J. Haywood and G. Tunnicliffe Wilson, "An improved state space representation for cyclical time series," *Biometrika*, vol. 87, no. 3, pp. 724–726, 2000. [Online]. Available: <http://biomet.oxfordjournals.org/cgi/content/abstract/87/3/724>
- [20] P. C. Young, D. Pedregal, and W. Tych, "Dynamic harmonic regression," *Journal of Forecasting*, vol. 18, pp. 369–394, November 1999.
- [21] J. P. Burman, "Seasonal adjustment by signal extraction," *Journal of the Royal Statistical Society. Series A*, vol. 143, no. 3, pp. 321–337, 1980.
- [22] W. R. Bell and S. C. Hillmer, "Issues involved with seasonal adjustment of economic time series," *Journal of Business and Economic Statistics*, vol. 2, pp. 291–320, 1984.
- [23] J. Haywood and G. Tunnicliffe Wilson, "Fitting time series models by minimizing multistep-ahead errors: a frequency domain approach," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 59, no. 1, pp. 237–254, 1997.
- [24] A. Maravall and C. Planas, "Estimation error and the specification of unobserved component models," *Journal of Econometrics*, vol. 92, pp. 325–353, 1999.
- [25] A. Maravall, "Unobserved components in econometric time series," in *The Handbook of Applied Econometrics*, ser. Blackwell Handbooks in Economics, H. H. Pesaran and M. Wickens, Eds. Oxford, UK: Basil Blackwell, 1995, ch. 1, pp. 12–72.
- [26] D. A. Pierce, "Signal extraction error in nonstationary time series," *The Annals of Statistics*, vol. 7, no. 6, pp. 1303–1320, November 1979. [Online]. Available: <http://www.jstor.org/stable/2958546>
- [27] G. E. P. Box, S. Hillmer, and G. C. Tiao, "Analysis and modeling of seasonal time series," in *Seasonal Analysis of Economic Time Series*, ser. NBER Chapters. National Bureau of Economic Research, Inc, August 1979, pp. 309–346. [Online]. Available: <http://ideas.repec.org/h/nbr/nberch/3904.html>
- [28] S. C. Hillmer and G. C. Tiao, "An arima-model-based approach to seasonal adjustment," *Journal of the American Statistical Association*, vol. 77, no. 377, pp. 63–70, Mar 1982. [Online]. Available: <http://www.jstor.org/stable/2287770>
- [29] T. C. Mills, "Signal extraction and two illustrations of the quantity theory," *The American Economic Review*, vol. 72, no. 5, pp. 1162–1168, December 1982.
- [30] A. C. Harvey and P. H. J. Todd, "Forecasting economic time series with structural and box-jenkins models: A case study," *Journal of Business and Economic Statistics*, vol. 1, no. 4, pp. 299–307, 1983.
- [31] C. Chen and G. C. Tiao, "Random level-shift time series models, ARIMA approximations, and level-shift detection," *Journal of Business and Economic Statistics*, vol. 8, no. 1, pp. 83–97, 1990.
- [32] I. M. Gelfand and N. J. Vilenkin, *Some Applications of Harmonic Analysis. Rigged Hilbert Spaces*, ser. Generalized Functions. New York: Academic Press, 1964, vol. 4.
- [33] J.-P. Antoine and A. Grossmann, "The partial inner product spaces. i. general properties," *Journal of Functional Analysis*, vol. 23, pp. 369–378, 1976.
- [34] G. C. Tiao and S. C. Hillmer, "Some consideration of decomposition of a time series," *Biometrika*, vol. 65, no. 3, pp. 497–502, Dec 1978.
- [35] W. Bell, "Signal extraction for nonstationary time series," *The Annals of Statistics*, vol. 12, no. 2, pp. 646–664, June 1984.
- [36] E. J. Hannan, "Measurement of a wandering signal amid noise," *Journal of Applied Probability*, vol. 4, no. 1, pp. 90–102, Apr 1967. [Online]. Available: <http://www.jstor.org/stable/3212301>
- [37] E. Sobel, "Prediction of noise-distorted, multivariate, non-stationary signal," *Journal of Applied P*, vol. 4, no. 2, pp. 330–342, Aug 1967. [Online]. Available: <http://www.jstor.org/stable/3212027>
- [38] W. P. Cleveland and G. C. Tiao, "Decomposition of seasonal time series: a model for the x-11 program," *Journal of the American Statistical Association*, vol. 71, no. 355, pp. 581–587, Sep 1976. [Online]. Available: <http://www.jstor.org/stable/2285586>
- [39] M. Bujosa, A. García-Ferrer, and P. C. Young, "Linear dynamic harmonic regression," *Comput. Stat. Data Anal.*, vol. 52, no. 2, pp. 999–1024, October 2007. [Online]. Available: <http://dx.doi.org/10.1016/j.csda.2007.07.008>

Marcos Bujosa was born in Madrid, Spain, on March 25, 1969. He received the B.S. degree in Economics from Universidad Autónoma de Madrid, Spain, in 1996 and the Ph.D. degree in Economics at the same university in 2001.

He is an Associate Professor of Econometrics at Universidad Complutense de Madrid, Spain. His research has been concerned with modelling in the frequency domain and forecasting seasonal economic time series.

Andrés Bujosa was born in Leeds, UK, on March 26, 1960. He received the B.S. degree in Mathematics from Universidad Complutense de Madrid, Spain, in 1988 and the Ph.D. degree in Mathematics at Universidad Politécnica de Madrid, Spain, in 1993.

He is an Associate Professor of Applied Mathematics at Universidad Politécnica de Madrid, Spain. His research interests include Artificial Intelligence and Computacional Logic.

Antonio García-Ferrer was born in La Roda (Albacete), Spain, on December 13, 1950. He received the B.S. degree in Economics from Universidad Autónoma de Madrid, Spain, in 1973 and the Ph.D. degree in Economics at U.C. Berkeley in 1978.

From 1978 to 1980, he was Assistant Professor at the Universidad de Alcalá de Henares, Spain. From 1984 to 1985, he was Fulbright Visiting Professor at the Booth GSB of the University of Chicago. He is currently Full Professor of Econometrics at Universidad Autónoma de Madrid, Spain. His research has been concerned with modeling and forecasting seasonal time series, turning point predictions, and leading indicators.

Dr. García-Ferrer was President of the International Institute of Forecasters from 2008 to 2012.