This is to certify that the

dissertation entitled

Autoregressive Expansion of Linear Predictor
for Stationary Stochastic Processes

presented by

Jamshid Farshidi

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Statistics

Major professor

Date September 27, 1991

AUTOREGRESSIVE EXPANSION OF LINEAR PREDICTOR FOR STATIONARY STOCHASTIC PROCESSES

by

Jamshid Farshidi

A DISSERTATION

submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Statistics and Probability

1991

ABSTRACT

AUTOREGRESSIVE EXPANSION OF LINEAR PREDICTOR FOR STATIONARY STOCHASTIC PROCESSES

In this thesis *autoregressive expansion* (AR–*expansion*) of 1–step predictor for a univariate stationary stochastic process (SSP) $X = \{X_t;\ t \in \mathbb{Z}\}$ with spectral density f and the optimal factor φ is studied. The main goal is to find necessary and sufficient conditions on f or φ for the existence of a mean–squared convergent AR–expansion. This problem, and it's multivariate case, has been the subject of the study of several authors. N. Wiener and P. Masani [Acta Math; 1958], P. Masani [Acta Math; 1960, and Academic Press; 1966], A.G. Miamee and H. Salehi [Math Mexicana; 1983], M. Pourahmadi [Proc. Am. Math. Soc.; 1984] and others.

The important results in this thesis are: (i) The *uniqueness of* AR–*expansion*, (ii) *necessary and sufficient conditions based on* f *and* φ *guaranteeing the existence* and the *uniqueness of the* AR–*expansion*, (iii) *equivalence of the existence of the* AR–*expansion with the invertibility of* X, and (iv) *sufficiency of the set of conditions* "$(1/f) \in L^p(\lambda)$ *and* $f \in L^q(\lambda)$ *for some* p, $1 \le p < \infty$ *and* $1 < q \le \infty$ *with* $1/p + 1/q = 1$" for the *existence* and the *uniqueness* of AR–expansion.

The important feature of this thesis is to give spectral characterization for an AR–expansion of $\hat{X}_1$, without undue attention to the condition $(1/f)$ in $L^1(\lambda)$ which emerges as a basic restriction in the work of earlier authors. This is accomplished by the consideration of the optimal factor, it's reciprocal, the examination of their Taylor coefficients, and the use of some facts from probability theory, harmonic and functional analysis.

Dedicated to my family

from Mahmoud to Ali

ACKNOWLEDGEMENT

I extend my most sincere gratitude to Professor Habib Salehi for introducing me to this historical problem, and for his encouragement and help during the preparation of this dissertation.

I would like to express my appreciation to Professor James Stapleton for having served on my thesis committee and for his support in performing my teaching duties in the Department. My special thanks go to Professor V. Mandrekar for his valuable comments on the final draft of this dissertation, to Professor Vaclav Fabian for having served on my thesis committee and for mathematical discussions which enhanced my insight into some of the subjects used in this dissertation, and to Professor Shlomo Levental for having served on my thesis committee. Special appreciation to Professor James Hannan for offering valuable advice.

Thanks are also due to Cathy Sparks for her excellent typing of this thesis.

To my wife, I must extend a special appreciation because of her patience with me, and her support of my academic work during my stay at Michigan State University.

TABLE OF CONTENTS

INTRODUCTION

Let $X = \{X_t; t\epsilon\mathbb{Z}\}$ be a univariate *Stationary Stochastic Process* (SSP) with spectral density f and the optimal factor φ. An important problem in the prediction theory of SSP's and in time series analysis is to find conditions on f or φ, which are necessary and sufficient for the existence of a mean–squared convergent *autoregressive expansion* (AR–*expansion*) for the linear least–square predictor $\hat{X}_{t+\tau} = \text{Proj}\,(X_{t+\tau}|\overline{\text{Sp}}\,\{X_s; s \leq t\})$ of a future value $X_{t+\tau}$ based on the observations ... X_{t-1}, X_t; i.e. an expansion of the form

$$\hat{X}_{t+\tau} = \sum_{n=0}^{\infty} a_n(\tau)\, X_{t-n};\ t\epsilon\mathbb{Z},\ \tau \geq 1,$$

for some sequence $\{a_n(\tau); n\epsilon\mathbb{Z}\}$. The problem is fundamental to the entire discussion of SSP's, time series analysis, the problem of rate of convergence of the predictor, and the estimation of the predictor.

This problem, or its equivalent, but simpler form

$$(*) \qquad \hat{X}_1 = \sum_{n=0}^{\infty} a_n\, X_{-n},$$

and its multivariate extension was the subject of study by several authors. N. Wiener and P. Masani [Acta. Math.; 1958] showed the boundedness of f, namely "$0 < c \leq f \leq d < \infty$" is sufficient for the existence and mean–squared convergence of the series in (*). Later P. Masani [Acta. Math.; 1960] weakened these conditions to "$(1/f)\ \epsilon L^1(\lambda)$ and $f\epsilon L^\infty(\lambda)$", where λ is the Lebestue measure on *Borel subsets of* $[-\pi,\pi]$. Since then several attempts have been made to reduce the restrictive condition "$f\epsilon L^\infty(\lambda)$" or to weaken this set of conditions. Recently A.G. Miamee and H. Salehi [Bol. Soc. Math. Mexicana; 1983] gave the necessary and sufficient conditions "$(1/f)\epsilon L^1(\lambda)$, and the convergence of the Fourier series of $1/\varphi$ to $1/\varphi$ in $L^2(f)$" to guarantee an AR–expansion (*) with the condition $\sum_{n=0}^{\infty} |a_n|^2 < \infty$.

Other authors, following the fundamental work of "Helson and Szegö" [Ann. Math. Pura. Appl.; 1960], looked at the problem from a "geometrical" point of view. Among them M. Pourahmadi [Proc. Am. Math. Soc.; 1984] showed that "the positivity of the angle between the past – present and the future of the process X is sufficient to guarantee the existence of a representation as (*)".

The common feature of all these works is the imposition of the condition $(1/f) \epsilon L^1(\lambda)$, and as a result obtaining an expansion of the form (*) with the additional condition $\sum_{n=0}^{\infty} |a_n|^2 < \infty$ as in Miamee–Salehi's. With a fresh look at the problem, in this thesis, the following main results are obtained: (i) the *uniqueness of the expansion* (*), (ii) providing *necessary and sufficient conditions based on* f *and* φ, guaranteeing the *existence* and the *convergence* and the *uniqueness* of the AR–expansion, (iii) establishing the equivalence between the existence of the AR–expansion and the invertibility of X, and (iv) *sufficiency of the set of conditions* "$(1/f) \epsilon L^p(\lambda)$ and $f \epsilon L^q(\lambda)$ for some $1 \leq p < \infty$ and $1 < q \leq \infty$ *with* $1/p + 1/q = 1$" for the *existence* and the *uniqueness* AR–expansion.

The basic tool employed to get the results is the study of the optimal factor φ and it's reciprocal $1/\varphi$, and to look at the Taylor expansions of $1/\varphi$ and $\hat{\psi}$, the $L^2(f)$–analog of $\hat{X}_1$. The results are obtained with the aid of some standard fact in probability theory, harmonic and functional analysis.

CHAPTER 1

PRELIMINARIES

This chapter is devoted to essential definitions and theorems to be used in this thesis. Other definition and theorem which are considered as general mathematical knowledge will also be used. The chapter is divided into two sections, 1.1 and 1.2, covering materials from *Harmonic* and *Functional analysis*, and *Probability theory*. Some of these theorems are so modified to suit the work.

§1.1. Definitions and Theorems from Harmonic and Functional Analysis

This section contains definitions, theorems and concepts in harmonic and functional analysis which will be used in consequent sections.

1.1.1 Notations. Throughout this thesis:

(i) $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{R}$ stand for the sets of all *natural numbers*, *integers* and *real numbers* respectively. The set of all *complex numbers* is denoted by $\mathbb{C}$.

(ii) $C(0,1)$ denotes the unit circle; i.e. $C(0,1) = \{z\epsilon\mathbb{C}: |z| = 1\}$.

(iii) $D(0,r)$ denotes the open disk with radius r in $\mathbb{C}$; i.e. $D(0,r) = \{z\epsilon\mathbb{C}: |z| < r\}$, and $\overline{D}(0,r)$ denotes the closed disk with radius r in $\mathbb{C}$; i.e. $\overline{D}(0,r) = \{z\epsilon\mathbb{C}: |z| \leq r\}$.

1.1.2 Definition (*Hilbert Spaces*). A complex vector space H with an inner product $< \cdot,\cdot >: H\times H \to \mathbb{C}$ (and norm $\|x\| = < x,x >^{\frac{1}{2}}$ for $x\epsilon H$) which is complete (i.e.

every Cauchy sequence $\{x_n\}$ in H converges in norm to some x in H) is called a *Hilbert space*.

Special types of Hilbert spaces are L^2–spaces, which are used extensively in this thesis. The following definition extends this notion.

1.1.3 Definition (L^p*–spaces*). Let Ω be a non–empty set, $\mathscr{F}$ a σ–algebra of subsets of Ω and μ a positive measure on $\mathscr{F}$.

(i) If $0 < p < \infty$, and X is a complex $\mathscr{F}$–measurable function on Ω, define

$$\|X\|_p^p := \int_\Omega |X|^p d\mu,$$

and let $L^p(\mu)$ consist of all equivalent classes of X (in the sense that $X' \sim X$ if and only if $X' = X$, a.e. $[\mu]$) for which $\|X\|_p < \infty$. We call $\|X\|_p$ the L_p–norm of X.

(ii) If Y is a complex measurable function on Ω, we define $\|X\|_\infty$ to be the essential supremum of $|X|$, and we let $L^\infty(\mu)$ consist of all equivalent classes of X for which $\|X\|_\infty < \infty$.

Note that $H = L^2(\mu)$ with the inner product

$$<X,Y> = EX\overline{Y} = \int_\Omega X\overline{Y}\, d\mu$$

is a Hilbert space.

1.1.4 Notation. Throughout this thesis the symbol λ is used for the *normalized Lebesgue measure* on $\mathscr{B}[-\pi,\pi]$, the Borel sets of $[-\pi,\pi]$, i.e. $d\lambda = \frac{1}{2\pi} d\theta$ and λ_k stands for the *Lebesgue measure* on the *Borel sets* of $\mathbb{R}^k$, $\mathscr{B}(\mathbb{R}^k)$.

If a positive measure μ on $\mathscr{B}([-\pi,\pi])$ or $\mathscr{B}(\mathbb{R}^k)$ has a density f w.r.t. λ or λ_k, we will write $L^p(f)$ or $L_k^p(f)$ for $L^p(\mu)$.

1.1.5 Theorem (*the projection theorem*). If M is a closed subspace of the Hilbert space H, then

(i) there is a unique element $\hat{x} \epsilon M$ such that

$$\|x-\hat{x}\| = \inf\{\|x-y\|; y \epsilon M\},$$

and

(ii) $\hat{x} \epsilon M$ and $\|x-\hat{x}\| = \inf\{x-y\|; y \epsilon M\}$ if and only if $\hat{x} \epsilon M$ and $(x-\hat{x}) \epsilon M^{\perp}$, where $M^{\perp} = \{y \epsilon H: (y,x) = 0, \forall x \epsilon M\}$.

Proof. c.f. [16] Theorems 4.10 and 4.11; or [3], p 51.

1.1.6 Notation. The element $\hat{x}$ in Theorem 1.1.5 is called the *orthogonal projection* of x onto M, and is denoted by

$$\hat{x} = \mathrm{Proj}(x|M).$$

$\mathrm{Proj}(\cdot|M)$ defines a mapping from H onto M which is called a *projection mapping.*

1.1.7 Theorem (*Properties of Projection Mappings*). Let H be a Hilbert space and let $P(\cdot|M)$ denote the projection mapping from H onto a closed subspace M of H. Then

(i) $P(\alpha x + \beta y|M) = \alpha P(x|M) + \beta P(y|M)$

(ii) $P(x_n|M) \to P(x|M)$ if $\|x_n - x\| \to 0$

(iii) $M_1 \subset M_2$ if and only if $P(P(x|M_2)|M_1) = P(x|M_1)$, for all $x \epsilon H$.

(iv) If a family $\{H_t; t \epsilon \mathcal{I}\}$ of subspaces of H has the property that $H_s \subset H_t$ for $s < t$ and $\bigcap_t H_t = \{0\}$, then for any $x \epsilon H$

$$\mathrm{Proj}(x|H_t) \to 0 \text{ as } t \to -\infty.$$

Proof. For proofs of (i) – (iii) see [3] p 52.

(iv) Denote $\mathrm{Proj}(x|H_t)$ by h_t, $t\in\mathbb{I}$. For any sequence $\{t_n\}$ in $\mathbb{I}$ with the property $t_1 > t_2 > \ldots$ and tending to $-\infty$, the elements $h_{t_1} - h_{t_2}, h_{t_2} - h_{t_3}, \ldots$ are mutually orthogonal, and

$$\|h_{t_1}\| \geq \|h_{t_1} - h_{t_{n+1}}\| = \|\sum_{i=1}^{n} (h_{t_i} - h_{t_{i+1}})\| = [\sum_{i=1}^{n} \| h_{t_i} - h_{t_{i+1}}\|^2]^{1/2};\ n\in\mathbb{N}.$$

it follows that the series $\sum_{i=1}^{\infty}(h_{t_i} - h_{t_{i+1}})$ converges in H. Since the ith partial sum of this series is $h_{t_1} - h_{t_i}$, $\lim_{i\to\infty} h_{t_i}$ exists, and since it is evidently contained in each subspace H_{t_i}, for each i, hence $\lim_{i\to\infty} h_{t_i} = 0$. □.

Weak convergence and *weak topology* play a role in some of the theorems in Chapter 2. The following definition and theorem are adapted forms of these notions in Hilbert spaces, which will be used.

1.1.8 Definition (*Weak convergence in Hilbert Spaces*). Let H be a Hilbert space over $\mathbb{C}$.

(i) A sequence $\{x_n\}$ in H is said to be *weakly convergent*, if there is an $x\in H$ with $\lim_{n\to\infty} < x_n,y > = < x,y >$ for every $y\in H$. The point x is called a *weak limit* of $\{x_n\}$, the sequence $\{x_n\}$ is said to converge weakly to x, and we write $x \xrightarrow{w} x$.

(ii) A set $A \subseteq H$ is said to be *weakly sequentially compact* if every sequence $\{x_n\}$ in A contains a subsequence which converges weakly to a point in H.

(iii) Every sequence $\{x_n\}$ in H such that $\{<x_n,y>\}$ is a Cauchy sequence of complex numbers for each $y\in H$ is called a weak *Cauchy sequence*.

1.1.9 Theorem. Let H be a *Hilbert space* over $\mathbb{C}$. Then

(i) A weakly convergent sequence in H has a unique limit.

(ii) A weakly convergent sequence $\{x_n\}$ in H is *bounded.* It's limit x is in the $\overline{sp}\{x_n;\, n\in\mathbb{N}\}$ and $\|x\| \le \underline{\lim}_{n\to\infty} \|x_n\|$.

Proof. c.f. [5] p 68, Lemmas 26 and 27.

1.1.10 Definition (*Fourier Coefficients and Partial Sums*). Let $\varphi \in L^1(\lambda)$. The *nth Fourier coefficient* of φ is defined as

$$\hat{\varphi}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \varphi(t)\, e^{-int}\, dt; \qquad n\in\mathbb{Z},$$

and the *nth order Fourier partial sum* of φ is defined to be

$$S_n(\varphi)(\theta) = \sum_{j=-n}^{n} \hat{\varphi}(j)\, e^{ij\theta}; \qquad -\pi \le \theta \le \pi,\ n\in\mathbb{N}.$$

Note that

$$S_n(\varphi)(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \varphi(t)\, D_n(t-\theta)\, dt = (\varphi * D_n)(\theta)$$

where

$$D_n(\theta) = \sum_{j=-n}^{n} e^{ij\theta} = \begin{cases} \dfrac{\sin[(n+\frac{1}{2})\theta]}{\sin(\theta/2)}; & \theta \ne 0 \\ 2n+1 \quad ; & \theta = 0\,. \end{cases}$$

D_n is called the *Dirichlet kernel.*

In the Hilbert space $H = L^2(\lambda)$, consider the elements e_n, $n\in\mathbb{N}$, defined by $e_n(\theta) = e^{in\theta}$ for $-\pi \le \theta \le \pi$. Then

$$< e_m, e_n > = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(m-n)t}\, dt = \delta_{m-n}$$

i.e. $\{e_n;\, n\in\mathbb{Z}\}$ is an *orthonormal set* in H,

$$\hat{\varphi}(n) = < \varphi, e_n >, \qquad n\in\mathbb{Z}$$

and

$$S_n(\varphi) = \mathrm{Proj}(\varphi | \overline{\mathrm{sp}}\{e_j;\ |j| \le n\}), \qquad n\in\mathbb{N}.$$

1.1.11 Theorem.

(a) The orthonormal set $\{e_n;\ n\in\mathbb{N}\}$ is *complete* in $H = L^2(\lambda)$, in the sense that the only vector in H orthogonal to it is the zero vector.

(b) For every $\varphi\in L^2(\lambda)$ the sequence $\{S_n(\varphi)\}$ converges in $L^2(\lambda)$ to φ.

(c) (Parseval) For every φ and ψ in $L^2(\lambda)$,

$$<\varphi,\psi> = \sum_{n=-\infty}^{\infty} <\varphi,e_n> \overline{<\psi,e_n>},$$

and as a special case

$$\|\varphi\|_2^2 = \sum_{n=-\infty}^{\infty} |<\varphi,e_n>|^2.$$

Proof. c.f. [8] pp 28 and 29.

It was conjectured in 1915 by *Lusin* that if $\varphi\in L^2(\lambda)$ then $S_n(\varphi) \to \varphi$ a.e. $[\lambda]$. In 1966 *Carleson* proved this conjecture, and also showed that for $\varphi \in L^p(\lambda)$ $1 < p < 2$, $S_n(\varphi)(\theta) = o$ (log log log n) a.e. $[\lambda]$, (c.f. [4]). One year later, *Hunt* modified Carleson's technique and proved for $p > 1$, $S_n(\varphi)$ converges a.e. $[\lambda]$.

1.1.12 Theorem (*Carleson*). Let $\varphi\in L^2(\lambda)$. Then $S_n(\varphi) \to \varphi$ a,e,$[\lambda]$. Furthermore, if $\varphi\in L^p(\lambda)$, $1 < p < 2$, then $S_n(\varphi)(\theta) = o$ (log log log n), a.e.$[\lambda]$.

Proof. c.f. [4].

To any function g: $D(0,1) \to \mathbb{C}$, we may associate a family $\{g_r: 0 \le r < 1\}$ of functions defined on T, by

$$g_r(e^{i\theta}) = g(re^{i\theta}); \qquad 0 \le r < 1.$$

1.1.13 Definition (H^p*–spaces*). For g analytic in $D(0,1)$ and $0 < p \leq \infty$, we put

$$\|g\|_p = \sup\{\|g_r\|_p : 0 < r < 1\}.$$

The collection of all g analytic in $D(0,1)$ for which $\|g\|_p < \infty$ is called the H^p–space.

1.1.14 Theorem. If $0 < p < \infty$ and $g \epsilon H^p$, then

(a) The nontangential limits $g^*(e^{i\cdot})$ of g exists a.e.[λ] and $g^*(e^{i\cdot}) \epsilon L^p(\lambda)$.

(b) $\lim_{r \to 1} \|g_r - g^*\|_p = 0$, and

(c) $\|g^*\|_p = \|g\|_p$.

Proof. c.f. [16] p 340 Theorem 17.11.

In this thesis mostly we deal with H^2–spaces. The following theorem concerns H^2–spaces.

1.1.15 Theorem. Suppose g is analytic in $D(0,1)$ and $g(z) = \sum_{n=0}^{\infty} a_n z^n$, $|z| < 1$. Then $g \epsilon H^2$ if and only if $\sum_{n=0}^{\infty} |a_n|^2 < \infty$. Furthermore, when this is the case the boundary function g^* is equal to $\sum_{k=0}^{\infty} a_k e^{ik\theta}$ a.e.[λ].

Proof. c.f. [16] p 341 Theorem 17.12.

1.1.16 Definition (*Inner and outer functions*). An *inner function* is a function $g \epsilon H^\infty$ for which $|g^*(e^{i\cdot})| = 1$ a.e.[λ].

If $g(e^{i\cdot})$ is a positive measurable function on $[-\pi,\pi]$ such that $\log g(e^{i\cdot}) \epsilon L^1(\lambda)$, and if

$$G(z) = c \exp \left\{\frac{1}{2\pi} \int \frac{e^{it} + z}{e^{it} - z} \log g(e^{it}) dt\right\}; \qquad z \epsilon D(0,1),$$

then G is called an *outer function*. Here c is a constant, $|c| = 1$.

The most important properties of outer functions is stated in the following theorem.

1.1.17 Theorem. Suppose G is an outer function related to g as in Definition 1.1.16. Then

(a) $\log|G(\cdot)|$ is the poisson integral of $\log g(e^{i\cdot})$.

(b) $\lim_{r \to 1} |G(re^{i\theta})| = g(e^{i\theta})$ a.e.$[\lambda]$

(c) $G \epsilon H^p$ if and only if $g \epsilon L^p(\lambda)$. In this case $\|G\|_p = \|g\|_p$.

Proof. c.f. [16] p 343, Theorem 17.16.

Another useful theorem concerning H^p–spaces, $0 < p \leq \infty$ is the following.

1.1.18 Theorem. If $0 < p < \infty$, $g \epsilon H^p$, and g is not identically 0, then $g^*(e^{i\cdot}) \neq 0$ a.e.$[\lambda]$.

Proof. c.f. [16] p 345, Theorem 17.18.

§1.2. Definitions and Theorems from Probability Theory.

The major definitions and theorems from probability theory to be used in the thesis are as follows:

1.2.1 Definition (*Stochastic Process*). A Stochastic Process (SP) is a family of random variables $X = \{X_t;\ t \epsilon T\}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ for some index set T. The index set T in this thesis is considered to be $\mathbb{Z}$, the set of all integers, in which case the process is sometimes called a *stochastic sequence*.

1.2.2 Definition. The Process $X = \{X_t;\ t\epsilon\mathbb{Z}\}$ is called *stationary stochasic process* (SSP), if

(i) $E|X_t|^2 := \int_\Omega |X_t|^2 dP < \infty$, for all $t\epsilon\mathbb{Z}$,

(ii) $EX_t := \int_\Omega X_t dP = m$, for some constant m and for all $t\epsilon\mathbb{Z}$

(iii) $\gamma_x^*(t,s) := \text{cov}\,(X_t,X_s) = E[(X_t-EX_t)\,\overline{(X_s-EX_s)}]$
$= \gamma_x^*(t+r,s+r)$

for all $r,s,t,\epsilon\mathbb{Z}$.

In condition (iii), the function $\gamma_x^*: \mathbb{Z}\times\mathbb{Z} \to \mathbb{C}$, where $\mathbb{C}$ is the set of complex numbers, is the so called *autocovariance function* of the SSP X.

Remark. Without loss of generality for the ease of writing we may assume the constant m in (ii) is zero. Condition (iii) in Definition 1.2.2 says $\gamma_x^*(t,s) = \gamma_x^*(t-s,0)$ for all $t,s\epsilon\mathbb{Z}$. It is therefore appropriate to consider the *autocovariance function* of a stationary process as a function of just one variable.

1.2.3 Definition (*The Autocovariance Function of a SSP*). For stationary process $X = \{X_t:\ t\epsilon\mathbb{Z}\}$, the function $\gamma_x: \mathbb{Z} \to \mathbb{C}$, defined by

$$\gamma_x(t) := \text{cov}(X_{t+h}, X_h) \qquad \text{for all } t,h\epsilon\mathbb{Z},$$

is called the *autocovariance function* of SSP X.

Autocovariance function γ_x of a SSP X, has the properties: $\gamma_x(0) \geq 0$, $|\gamma_x(t)| \leq \gamma_x(0)$ for all $t\epsilon\mathbb{Z}$, $\overline{\gamma_x(t)} = \gamma_x(-t)$ for all $t\epsilon\mathbb{Z}$, and moreover, it is *non-negative definite* in the sense that for every $n\epsilon\mathbb{N}$, and all $(a_1,\dots,a_n)\epsilon\mathbb{C}^n$ and $(t_1,\dots,t_n)\epsilon\mathbb{Z}^n$,

$$\sum_{i,j=1}^{n} a_i\gamma_x(t_i-t_j)\,\overline{a_j} \geq 0.$$

This property connects the autocovariance functions to the so–called distribution functions, via the following theorem:

1.2.4 Theorem (*Herglotz–Bochner*). A function $\gamma\colon \mathbb{Z} \to \mathbb{C}$ is *non–negative definite* if and only if

$$\gamma(t) = \int_{-\pi}^{\pi} e^{-its}\, dF(s) \qquad \text{for all } t \in \mathbb{Z},$$

where F is a right–continuous, non–decreasing, bounded function on $[-\pi,\pi]$ and $F(-\pi) = 0$.

Proof. c.f. [3] pp 115–116, or [11].

Using Theorem 1.2.4, we have

$$\gamma_x(t) = \int_{-\pi}^{\pi} e^{-its}\, dF_x(x)$$

where F_x is as above.

1.2.5 Definition. For a SSP X, the corresponding function F_x in Theorem 1.2.4 is called the *spectral distribution function* of X. If

$$F_x(t) = \int_{-\pi}^{t} f(s)ds; \qquad -\pi \le \lambda \le \pi,$$

then f is called the *spectral density* of X.

1.2.6 Notation. For the SSP $X = \{X\,;\, t \in \mathbb{Z}\}$ we introduce the following subspaces:

(i) $H_x(t) = \overline{sp}\{X_s\colon s \le t\}$, $t \in \mathbb{Z}$ (*the past of the process*).

(ii) $H_x = \overline{sp}\{X_t\colon t \in \mathbb{Z}\}$ (*the entire space of the process*).

(iii) $H_x(-\infty) = \bigcap_{t \in \mathbb{Z}} H_x(t)$ (*the remote past of the process*).

Here $\overline{sp}A$ denotes the closed linear subspace of $L^2(\Omega, \mathcal{F}, P)$ generated by the elements in the subset A of this space.

The Linear Predictor of a SSP

1.2.7 Definition. Let $X = \{X_t;\ t \in \mathbb{Z}\}$ be an SSP. Let $t \in \mathbb{Z}$ and $\tau \in \mathbb{N}$.

(i) The (best) linear τ–step predictor of $X_{t+\tau}$ based on $H_x(t)$ is an element $\hat{X}(t,\tau)$ of $H_x(t)$ which minimizes the distance between $X_{t+\tau}$ and $H_x(t)$.

(ii) The mean–squared error in prediction in (i):

$$\sigma^2(t,\tau) = E|\hat{X}(t,\tau) - X_{t+\tau}|^2$$

is called the *(mean–squared) τ–step error of the prediction.*

By projection theorem (Theorem 1.1.5)

$$\hat{X}(t,\tau) := \mathrm{Proj}(X_{t+\tau}|H_x(t)), \qquad t \in \mathbb{Z},\ \tau \in \mathbb{N}.$$

An important consequence of *stationarity* of SP's is the existence of a family of *linear operators* $\{U_t;\ t \in \mathbb{Z}\}$ on H_x, which shift the elements and subspaces of H_x. The following theorem proves the existence of such a family, and establishes their basic properties. For a proof c.f. [15], pp 14, 54 (Relations 1.4 and 1.5), or [11].

1.2.8 Theorem (*shift operators*). There exists a unique family of *unitary operators* U_t, $t \in \mathbb{Z}$, on H_x such that for $s,t \in \mathbb{Z}$:

(a) $U_t X_s = X_{t+s}$,

(b) $U_t H_x(s) = H_x(s+t)$,

(c) $U_t \hat{X}(s,\tau) = \hat{X}(t+s,\tau)$.

$\{U_t;\ t \in \mathbb{Z}\}$ is called the *family of shift operators* associated with X.

Property (c) easily implies that $\sigma^2(t,\tau)$ is, in fact, independent of the variable t.

1.2.9 Corollary. For every $s,t\in\mathbb{Z}$ and $\tau\in\mathbb{N}$, $\sigma^2(t,\tau) = \sigma^2(s,\tau)$.

1.2.10 Notation. For SSP $X = \{X_t;\ t\in\mathbb{Z}\}$

(i) Denote the 1–step predictor $\hat{X}(t-1,1)$ by $\hat{X}_t$; i.e.

$$\hat{X}_t := \mathrm{Proj}(X_t \mid H_x(t-1)); \qquad t\in\mathbb{Z}.$$

(ii) Denote the *one–step error of prediction* $\sigma^2(t-1,1)$ by σ^2, for every $t\in\mathbb{Z}$ (c.f. Corollary 1.2.9); i.e.

$$\sigma^2 := E|\hat{X}_t - X_t|^2; \qquad t\in\mathbb{Z},$$

The Wold Decomposition

In this part, we summarize results that are related to the Wold's decomposition which play an important role in the analysis of SSP's.

1.2.11 Definition. A SSP $X = \{X_t;\ t\in\mathbb{Z}\}$ is called

(a) *deterministic*, if $\hat{X}(t_0,\tau_0) = X_{t_0+\tau_0}$ for some $t_0\in\mathbb{Z}$ and $\tau_0\in\mathbb{N} - \{0\}$.

(b) *regular*, if $\lim\limits_{\tau\to+\infty} \hat{X}(t_0,\tau) = 0$ for some $t_0\in\mathbb{Z}$

(c) *white noise* with mean 0 and variance α^2, denoted by $WN(0,\alpha^2)$, if $EX_t = 0$, $t\in\mathbb{Z}$, and

$$\gamma_x(t) = \begin{cases} \alpha^2; & \text{if } t = 0 \\ 0\ ; & \text{if } t \neq 0 \end{cases}.$$

Using the family of the associated shift operators of X, it is easily proved that if X is *deterministic*, then $\hat{X}(t,\tau) = X_{t+\tau}$ for all $t\in\mathbb{Z}$, $\tau\in\mathbb{N}$, and if X is *regular* then $\lim\limits_{\tau\to+\infty} \hat{X}(t,\tau) = 0$ for every $t\in\mathbb{Z}$ (c.f. [11], or [15] p 52).

1.2.12 Theorem. A SSP $X = \{X_t: t\in\mathbb{Z}\}$

(i) is *deterministic* if and only if $H_x = H_x(-\infty)$,

(ii) is *regular* if and only if $H_x(-\infty) = \{0\}$.

Proof. c.f. [11] or [15] p 52.

1.2.13 Theorem (Wold). Every SSP $X = \{X_t; t\in\mathbb{Z}\}$ can be represented as

$$X_t = W_t + V_t; \qquad t\in\mathbb{Z}$$

where $W = \{W_t: t\in\mathbb{Z}\}$ and $V = \{V_t; t\in\mathbb{Z}\}$ are regular and deterministic SSP's, respectively, such that

(i) $H_w(t) \subset H_x(t)$ and $H_v(t) \subset H_x(t)$, $t\in\mathbb{Z}$

and

(ii) V and W are mutually orthogonal; i.e. $EW_tV_s = 0$, $\forall t, s\in\mathbb{Z}$.

(iii) Processes V and W with above properties are unique.

Proof. c.f. [15] Theorem 2.2; or [11].

1.2.14 Theorem (*Moving average representation*). Let $X = \{X_t: t\in\mathbb{Z}\}$ be a regular SSP, then

$$H_x(t) = H_\zeta(t), \quad X_t = \lim_{k\to\infty} \sum_{j=0}^{k} c_j \zeta_{t-j} \text{(in } H_x); \qquad t\in\mathbb{Z}$$

where

$$\zeta_s = (1/\sigma)[X_s - \hat{X}_s]; \qquad s\in\mathbb{Z}$$

and

$$c_j = EX_j \bar{\zeta}_0; \qquad j\in\mathbb{N}.$$

Moreover, $\zeta = \{\zeta_t\} \sim WN(0,1)$, $\sum_{j=0}^{\infty}|c_j|^2 = E|X_0|^2$ (ζ is sometimes called the *innovation process* of X).

Proof. c.f. [15] p 56; or [11].

The following corollary gives the relation between the decomposition in Theorem 1.2.14 of a SSP $X = \{X_t;\ t\epsilon\mathbb{Z}\}$, and the Lebesque decomposition of its spectral distribution function F.

1.2.15 Corollary (*Wold–Cramer concordance*). Let $X = \{X_t;\ t\epsilon\mathbb{Z}\}$ be a non–deterministic SSP with spectral distribution function F_x. Let W and V be SSP's as in Theorem 1.2.13 with spectral distribution functions F_w, and F_v, respectively. Then

(a) $W_t = \sum_{j=0}^{\infty} c_j \zeta_{t-j}$ for $t\epsilon\mathbb{Z}$, where $\{c_j;\ j\epsilon\mathbb{N}\}$ and ζ are as in Theorem 1.2.12; i.e.

$$X_t = \sum_{j=0}^{\infty} c_j \zeta_{t-j} + V_t; \qquad t\epsilon\mathbb{Z}.$$

(b) F_w has the spectral density $f_w(\theta) = \frac{1}{2\pi} |\varphi(e^{i\theta})|^2$ a.e.[$d\theta$], where

$$\varphi(e^{i\theta}) = \sum_{j=0}^{\infty} c_j e^{ij\theta},$$

$$F_x = F_w + F_v$$

is the Lebesque decomposition of F_x.

Proof. Use Theorems 1.2.13, 1.2.14, and [3] p 180 and 183, Theorems 5.7.1 and 5.7.2.

Clearly this theorem implies that every regular SSP has a spectral density, determined completely by Wold decomposition. The next theorem explores the fact that for these processes the coefficients in Wold decompositon can be recovered from the density function.

1.2.16 Theorem.

(a) The SSP $X = \{X_t;\ t\epsilon\mathbb{Z}\}$ is regular if and only if it has an almost everywhere positive spectral density f such that $\int_{-\pi}^{\pi} \log f(\theta)d\theta > -\infty$.

(b) If $X = \{X_t;\ t\in\mathbb{Z}\}$ is a *regular* SSP, then

$$\varphi(z) = \exp\left\{ \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z} \log f(\theta)d\theta \right\}$$

is in H^2. Moreover, $f(\theta) = |\varphi(e^{i\theta})|^2$ a.e.$[d\theta]$,

$$\varphi(z) = \sum_{j=0}^{\infty} c_j z^j, \qquad |z| \leq 1$$

where $\{c_j;\ j\in\mathbb{N}\}$ is the set of coefficients appeared in the moving average representation (Theorem 1.2.14), and

(c) $c_0 = \exp\left\{ \frac{1}{4\pi} \int_{-\pi}^{\pi} \log f(\theta)\, d\theta \right\} = \sigma$

where σ is the one step error.

Proof. Proofs can be seen respectively in [15] Theorem 5.1 p 64, Theorem 5.2 p 65 and also pp 57 and 58, and relation (5.12) p 66; see also [11]. □

The outer function φ appeared in part (b) of Theorem 1.2.16 plays an important role in the prediction theory of regular SSP's. It has a *maximal* property in the sense that among all $\psi\in H^2$ satisfying the boundary condition $|\psi(e^{i\theta})|^2 = f(\theta)$, $\varphi(0)$ is positive, $\varphi(0) \geq |\psi(0)|$, and that such a φ is *unique*. (c.f. [15] p 60 Theorem 4.2; see also [11]).

1.2.17 Definition. If $X = \{X_t;\ t\in\mathbb{Z}\}$ is a regular SSP with density function f, the outer function

$$\varphi(z) = \exp\left\{ \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z} \log f(\theta)d\theta \right\}$$

is called the *optimal factor* of X.

The second formulation of φ stated in the paragraph preceding Definition 1.2.17 extends to the multivariate case; see [11].

Spectral Representation of SSP's

1.2.18 Theorem. If $X = \{X_t;\ t\in\mathbb{Z}\}$ is a SSP with $EX_0 = 0$ and spectral distribution function F_x, there exists a right continuous orthogonal–increment process $\{Z_x(\theta);\ -\pi \le \theta \le \pi\}$ such that

(i) $$X_t = \int_{-\pi}^{\pi} e^{-it\theta}\, dZ_x(\theta) \qquad \text{a.e.}[P],\ t\in\mathbb{Z}$$

(ii) $$E|Z(\theta) - Z(-\pi)|^2 = F_x(\theta); \qquad -\pi \le \theta \le \pi$$

(representation (i) is called the *spectral representation* of X).

Proof. c.f. [3] Theorem 4.8.2 p 140.

1.2.19 Theorem. Let $X = \{X_t;\ t\in\mathbb{Z}\}$ be a *regular* SSP with optimal factor φ. Then

(a) for $t\in\mathbb{Z}$ and $\tau\in\mathbb{N} - \{0\}$,

$$\hat{X}(t,\tau) = \int_{-\pi}^{\pi} \hat{\psi}(e^{i\theta},t,\tau)\, Z_x(d\theta)$$

where

$$\hat{\psi}(e^{i\theta},t,\tau) = e^{-i\theta(t+\tau)}\Big[1 - \frac{\sum_{n=0}^{\tau-1} c_n e^{in\theta}}{\varphi(e^{i\theta})}\Big] = [e^{-i(t+\tau)\theta}\varphi(e^{i\theta})]_+/\varphi(e^{i\theta}); \text{ a.e. } [d\theta].$$

Here for $\psi(e^{i\theta}) = \sum_{k=-\infty}^{\infty} b_k e^{ik\theta}$ in $L^2(\lambda)$, $(\psi)_+(e^{i\theta}) = \sum_{k=0}^{\infty} b_k e^{ik\theta}$, (c.f. [11]).

(b) If $\zeta = \{\zeta_t\} \sim WN(0,1)$ is the white noise (innovation process) corresponsing to X, then

$$\zeta_t = \int_{-\pi}^{\pi} \frac{e^{-i\theta t}}{\varphi(e^{i\theta t})} Z_x(d\theta) \qquad \text{a.e. } [P];\ t\in\mathbb{Z},$$

where φ is the optimal factor of X.

Proof. c.f. [11] or [15] Theorem 5.3 p 68, and use Theorem 1.2.14 and Theorem 1.2.16 part (c).

1.2.20 Theorem. Let $X = \{X_t;\ t \in \mathbb{Z}\}$ be a SSP with spectral distribution function F. Let μ_F be the positive measure induced by F on $\mathscr{B}([-\pi,\pi])$. Then the linear transformation

$$L\colon L^2(\mu_F) \longrightarrow H_x$$

defined by

$$L(h(e^{i\cdot})) = \int_{-\pi}^{\pi} h(e^{-i\theta})\, Z_x\,(d\theta)$$

is the unique isomorphism between $L_2(\mu_F)$ and H_x with the property that

$$L(e^{-it\cdot}) = X_t; \qquad t \in \mathbb{Z}.$$

Proof. c.f. [3] Theorem 4.8.1 p 139.

1.2.21 Notation. If μ_F has a density f with respect to λ, we denote $L^2(\mu_F)$ by $L^2(f)$. (see Notation 1.1.4).

A geometric way of looking at the problem of autoregressive expansion is the approach "the angle between the past – present and the future", followed by some autors in this area. The following definition and theorem concern this concept to be used in this thesis. For detailed discussion see [7] and [9].

1.2.22 Definition (*Angle between "the past – present" and "the future"*). For SSP X with spectral density f, let

$$\rho(f) = \sup\{|(x,y)|:\ x \in H_x(0) \text{ and } y \in H_x^+(1) \text{ and } \|x\| \le 1,\ \|y\| \le 1\}$$

where $H_x^+(1) = \overline{sp}\{X_t\colon t \ge 1\}$. We say the "the past and present" and "the future" of X are at *positive angle* if $\rho(f) < 1$.

The definition naturally can be extended for any two closed subspaces of a Hilbert space (c.f. [9], p 107, Definition 2.1). The important contribution of "angle approach" to the discussion of autoregressive expansion for linear predictor of SSP's is the following theorem which is a consequence of a theorem of Helson and Szegõ.

1.2.23 Theorem. Let $X = \{X_t;\ t \epsilon \mathbb{Z}\}$ be a SSP with spectral density f, such that $\log f \epsilon L^1(\lambda)$.

(a) Then $\rho(f) < 1$, if and only if $L^2(f) \subset L^1(\lambda)$ and the Fourier series of any $\varphi \epsilon L^2(f)$ converges to φ in the $L^2(f)$ norm; and

(b) $\rho(f) < 1$ implies $1/f \epsilon L^1(\lambda)$.

Proof. For part (a) c.f. [9], p 131. Also see [13] p 318 Theorem 2.1. For part (b) c.f. [9]; p 110 Corollary 2.9.

CHAPTER 2

AUTOREGRESSIVE EXPANSION

In this chapter, first, *autoregressive* (AR) *expansion of the predictor* for a stationary stochastic process (SSP) is introduced (c.f. Definition 2.2.1). The *existence* and *uniqueness* of this expansion for *regular* SSP's, and the *invertibility* of processes admitting such expansions are the subject of study of the rest of this chapter. The necessary and sufficient conditions to achieve such a representation are given in terms of the spectral *density function* f and the *optimal factor* φ of regular SSP's; however, all sufficient condition(s) are based on the density function.

Section 1 consists of the definition, and a summary of the earlier study of several authors on AR–expansion. Section 2 is devoted to a new approach to the problem, using the optimal factor as the main tool. The role of the optimal factor is completely clarified in Section 3. In this section, moreover, the *uniqueness* of AR–expansion of the linear predictor for a regular–SSP is proved. In Sections 4 and 5 two types of convergence of the AR–expansion are explored. Especially, in Section 5 the strong convergence of the AR–expansion, based on a condition on the density function and the optimal factor, is achieved. Consequently in Section 5, also an important extension of earlier results (1958 and 1960) of N. Wiener and P. Masani is derived. The invertibility of regular–SSP's and the equivalence with the existence (and the uniqueness) of AR–expansion of one–step predictor is discussed in Section 6. The Chapter is concluded with a brief discussion of sufficiency of the condition $1/f \epsilon L^1(\lambda)$, in Section 7.

§2.1. Definition of AR–expansion, and Historical Notes

Let $X = \{X_t;\ t\epsilon\mathbb{Z}\}$ be a Stationary Stochastic Process (SSP). The τ–step ahead linear predictor $\hat{X}(t,\tau)$, as defined earlier (Definition 1.2.7) and by projection theorem (Theorem 1.1.5), is

$$\hat{X}(t,\tau) := \mathrm{Proj}(X_{t+\tau} | H_x(t)); \qquad t\epsilon\mathbb{Z},\ \tau\epsilon\mathbb{N} - \{0\}.$$

2.1.1 Definition. Let $X = \{X_t;\ t\epsilon\mathbb{Z}\}$ be an SSP. Let $\tau\epsilon\mathbb{N} - \{0\}$. If there exists a sequence $\{b_k(\tau)\}$ of complex numbers such that

$$(2.1.1) \qquad \hat{X}(t,\tau) = \sum_{k=0}^{\infty} b_k(\tau)\, X_{t-k} \ \ (\text{in } H_x),$$

we say that τ–step predictors have *autoregressive expansions with the coefficients* $\{b_k(\tau)\}$.

Using the family of shift operators of X (c.f. Theorem 1.2.8), it is necessary and sufficient to have (2.1.1) holds for $t = 0$; i.e.

$$(2.1.2) \qquad \hat{X}(0,\tau) = \sum_{k=0}^{\infty} b_k(\tau)\, X_{-k}; \qquad \tau\epsilon\mathbb{N}.$$

This is because (2.1.1) implies (2.1.2), and if (2.1.2) holds, then

$$\begin{aligned}\hat{X}(t,\tau) = U_t(\hat{X}(0,\tau)) &= U_t\left(\sum_{k=0}^{\infty} b_k(\tau)\, X_{-k}\right)\\ &= \sum_{k=0}^{\infty} b_k(\tau)\, U_t(X_{-k})\\ &= \sum_{k=0}^{\infty} b_k(\tau)\, X_{t-k}.\end{aligned}$$

In this chapter we are mostly concerned with the special case of $t = 0$, $\tau = 1$; i.e. $\hat{X}_1 = \hat{X}(0,1)$ (c.f. Notation 1.2.10), the 1–step ahead linear predictor of X_1 based on $H_x(0)$. With a minor change of notation, taking $b_k(1) = b_k$ for $k\epsilon\mathbb{N}$, (2.1.1) in this case is restated as follows:

$$(2.1.3) \qquad \hat{X}_1 = \sum_{k=0}^{\infty} b_k\, X_{-k}.$$

An important problem in time series analysis and prediction theory of SSP's is to find condition(s), preferably on the spectral density function of the regular part of the process (c.f. Definitions 1.2.5 and 1.2.11 also Theorems 1.2.12 and 1.2.16) required to achieve the existence and the mean–squared convergence of (2.1.1) (or its equivalent form (2.1.3)), and its generalization to the multivariate case. This problem has been the subject of study of several authors. Among all, the following works, demonstrating the diversity of approaches to the problem, are notable.

N. Wiener and *P. Masani*, in their paper [17; 1958] proved that: "*the boundedness of the density function* f; *i.e.* $0 < c \leq f \leq d < \infty$, *is sufficient for the existence of an AR–expansion* for $\hat{X}_1$". Later *P. Masani* [10; 1960] weakened these conditions to: "$(1/f) \in L^1(\lambda)$ and $f \in L^\infty(\lambda)$". Since then, several attempts have been made to relax or weaken the very restrictive condition "$f \in L^\infty(\lambda)$". *A.G. Miamee* and *H. Salehi* [12; 1983] gave the following necessary and sufficient condition to maintain an AR–expansion as in (2.1.3) with the condition $\sum_{k=0}^{\infty} |b_k|^2 < \infty$ for $\hat{X}_1$: (i) $(1/f) \in L^1(\lambda)$, (ii) *The convergence of the Fourier series of* $(1/\varphi)$ *the reciprocal of the optimal factor* (c.f. Definition 1.2.17) *to* $(1/\varphi)$ *in the space* $L^2(f)$. The second condition on their set of conditions, although weaker than Masani's "$f \in L^\infty(\lambda)$", is not easily verifiable. In addition their condition $\sum_{k=0}^{\infty} |b_k|^2 < \infty$ restrict the applicability of their work. *M. Pourahmadi* [13; 1984], with an eye on the famous theorem of *Helson* and *Szegö* on the *positively of the angle between past – present and future* of a stationary process (c.f. Theorem 1.2.23), examined the problem and proved: "*the positivity of the angle between the past–present, and the future of the process is sufficient to guarantee the existence of an AR–expansion.*"

In the forementioned works the basic imposition $(1/f) \in L^1(\lambda)$ emerges as a common restriction on the spectral density f. This requirement translates into the squared summability of the coefficients in the resulting AR–expansion. In the following sections, by a *fresh look at the problem*, this restrictive condition is removed, and especially the results of Wiener–Masani, Masani, Miamee–Salehi and Pourahmadi will be derived as a consequence of an essential theorem (see Corollaries 2.5.2 through 2.5.5). The stepping stone of the entire discussion is the analysis of the optimal factor φ of f, discussed below.

§2.2. The Role of the Optimal Factor

Let $X = \{X_t; t \in \mathbb{Z}\}$ be a regular SSP with density function f and optimal factor φ:

$$\varphi(z) = \exp\{\frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{e^{i\theta}+z}{e^{i\theta}-z}\log f(\theta)d\theta\}$$

$$= \sum_{j=0}^{\infty} c_j z^j; \qquad |z| < 1,$$

$$f(\theta) = |\varphi(\theta)|^2; \qquad -\pi \leq \theta \leq \pi,\ a\cdot e[\lambda]$$

(c.f. Theorem 1.2.16 and Definition 1.2.17). For the sake of clarity, the assumptions on f that have been used so far are stated here:

$$(1)\ f \in L^1(\lambda) \quad \text{and} \quad (2) \log f \in L^1(\lambda).$$

The 1–step predictor $\hat{X}_1$ has the spectral representation

(2.2.1) $$\hat{X}_1 = \int_{-\pi}^{\pi} e^{-i\theta}[1 - c_0/\varphi(e^{i\theta})]\, Z_x(d\theta),$$

where $c_0^2 = \sigma^2 = E|\hat{X}_1 - X_1|^2 > 0$ (c.f. Theorem 1.2.19). Since φ is in H^2 and has no zero in $D(0,1)$, $1/\varphi$ is analytic (holomorphic) in $D(0,1)$. Let

(2.2.2) $$1/\varphi(z) = \sum_{n=0}^{\infty} d_n z^n; \qquad |z| < 1$$

be the Taylor expansion of $1/\varphi$ in $D(0,1)$. Define

(2.2.3) $$\hat{\psi}(z) = z^{-1}[1 - c_0/\varphi(z)]; \qquad |z| < 1.$$

Clearly $\hat{\psi}$ is analytic in $D(0,1)$ since $c_0 d_0 = 1$ and consequently the constant term of the Taylor series of the function $1 - c_0/\varphi$ equals 0. Let

$$\hat{\psi}(z) = \sum_{n=0}^{\infty} a_n z^n; \qquad |z| < 1$$

be its Taylor expansion. The relations between the coefficients a_n, d_n, and c_n (which are the coefficients in the moving average representation, c.f. Theorem 1.2.14) sheds light on the entire discussion. In fact, since

$$\hat{\psi}(z) = \sum_{n=0}^{\infty} a_n z^n = z^{-1}[1 - c_0/\varphi(z)]$$

$$= z^{-1}[1 - c_0 \sum_{n=0}^{\infty} d_n z^n]$$

$(c_0 d_0 = 1)$

$$= -\sum_{n=1}^{\infty}(c_0 d_n) z^{n-1}$$

$$= \sum_{n=0}^{\infty} (-c_0 d_{n+1}) z^n; \qquad |z| < 1,$$

we have

(2.2.4)
$$a_n = -c_0 d_{n+1}; \qquad n \in \mathbb{N},$$

and by multiplying both sides of (2.2.3) by $\varphi(z)$, we get

$$z^{-1}[\varphi(z) - c_0] = (\sum_{n=0}^{\infty} a_n z^n)(\sum_{n=0}^{\infty} c_n z^n); \quad |z| < 1.$$

This is equivalent to

$$\sum_{k=1}^{\infty} c_k z^{k-1} = \sum_{k=0}^{\infty} [\sum_{n=0}^{k} (c_{k-n} \cdot a_n)] z^k; \qquad |z| < 1.$$

Notice that, since $\sum_{n=0}^{\infty} c_n z^n$ is absolutely convergent in $D(0,1)$, the so called Cauchy product theorem is applicable. The last equality implies

(2.2.5)
$$c_{k+1} = \sum_{n=0}^{k} c_{k-n} \cdot a_n; \qquad k \in \mathbb{N}.$$

Equations in (2.2.5) form a triangular system of equations in unknowns $\{a_n;\ n \in \mathbb{N}\}$, which has a unique solution. The unique solution, by a little effort is found as follows:

$$a_0 = c_1/c_0; \ a_n = 1/c_0 [c_{n-1} - \sum_{k=0}^{n-1} a_n c_{n-k}];$$

As a result, using (2.2.4)

(2.2.6)
$$\hat{\psi}(z) = \sum_{n=0}^{\infty} (-c_0 d_{n+1}) z^n$$

where $c_0 = \sigma_0 > 0$ and d_n, $n \in \mathbb{N}$, are the coefficients of Taylor expansion of $1/\varphi$. The relation (2.2.5) between a_n's and c_n's and expansion (2.2.6) will play important roles in the analysis of AR–expansion of $\hat{X}_1$ in subsequent sections. It should be noted that the Taylor expansions of φ and $1/\varphi$, and the relations between the coefficients c_n's and d_n's in multivariate case are discussed by P. Masani in his 1960 and 1966 paper (c.f. [10] p 146, and [11] p 375).

§2.3. Uniqueness of Autoregressive Expansion

In this section the *uniqueness* of AR–expansion for regular SSP's, which admit an AR–expansion, is proved.

2.3.1 Theorem (*uniqueness*). Let $X = \{X_t;\ t \in \mathbb{Z}\}$ be a regular SSP with the optimal factor φ. Let $1/\varphi(z) = \sum_{n=0}^{\infty} d_n z^n$ for $|z| < 1$, and $c_0 = \sigma$. If for some sequence $\{b_n\}$,

$$\hat{X}_1 = \sum_{n=0}^{\infty} b_n X_{-n}, \tag{2.3.1}$$

then $b_n = -c_0\, d_{n+1}$ for $n \in \mathbb{N}$.

Proof. By the the moving average representation (c.f. Theorem 1.2.14) for $t \in \mathbb{Z}$

$$X_t = \sum_{j=0}^{\infty} c_j\, \zeta_{t-j}, \quad \{\zeta_j\} \sim WN(0,1),\ c_0 = \sigma_0.$$

Therefore, since $H_x(0) = H_\zeta(0)$ and $\zeta_1 \perp H_\zeta(0)$, using Theorem 1.1.7 (i) we have

$$\hat{X}_1 = \mathrm{Proj}(X_1 | H_x(0)) = \sum_{j=0}^{\infty} c_{j+1}\, \zeta_{-j},$$

and

$$\begin{aligned} \hat{X}_1 - \sum_{n=0}^{k} b_n X_{-n} &= \sum_{n=0}^{\infty} c_{n+1}\, \zeta_{-n} - \sum_{n=0}^{k} b_n \sum_{s=0}^{\infty} c_s\, \zeta_{-n-s} \\ &= \sum_{n=0}^{\infty} c_{n+1}\, \zeta_{-n} - \sum_{s=0}^{\infty} \sum_{n=0}^{k} b_n\, c_s\, \zeta_{-n-s} \\ &= \sum_{r=0}^{\infty} c_{r+1}\, \zeta_{-r} - \sum_{r=0}^{\infty} \Big[\sum_{n=0}^{r \wedge k} b_n\, c_{r-n}\Big] \zeta_{-r} \end{aligned} \tag{2.3.2}$$

(in the second sum, take $n + s = r$, then $0 \le r < \infty$ and $0 \le n \le r \wedge k$).

$$= \sum_{r=0}^{\infty} [c_{r+1} - \sum_{n=0}^{r \wedge k} b_n c_{r-n}] \zeta_{-r}.$$

Now, suppose (2.3.1) holds for some sequence $\{b_n\}$. Then, for every $k \in \mathbb{N}$,

$$\hat{X}_1 - \sum_{n=0}^{k} b_n X_{-n} \in H_x(-k-1).$$

According to the last relation the coefficient of ζ_{-r} for $r \in \{0,1,\ldots,k\}$ should vanish; i.e.

$$c_{r+1} - \sum_{n=0}^{r} b_n c_{r-n} = 0; \qquad r \in \{0,1,\ldots,k\},$$

for every $k \in \mathbb{N}$. This is equivalent to

(2.3.3) $$c_{r+1} - \sum_{n=0}^{r} b_n c_{r-n} = 0; \qquad r \in \mathbb{N}.$$

The system of equations (2.3.3) for b_n's is the same as (2.2.5) for a_n's, $n \in \mathbb{N}$. Since the latter system has the unique solution $a_n = -c_0 d_{n+1}$, $n \in \mathbb{N}$, we have $b_n = -c_0 d_{n+1}$ for $n \in \mathbb{N}$. □

§2.4. Weak Convergence of AR–expansion

In search for the convergence of the series in AR–expansion of predictor for a regular SSP we start off by the study of weak convergence of this series. The following theorem shows the very special nature of this expansion.

2.4.1 Theorem. Let $X = \{X_t;\ t \in \mathbb{Z}\}$ be a *regular* SSP with optimal factor φ. Let $1/\varphi(z) = \sum_{n=0}^{\infty} d_n z^n$, $|z| < 1$. Set $a_n = -c_0 d_{n+1}$ for $n \in \mathbb{N}$, where $c_0 = \sigma$. Then $\sum_{n=0}^{k} a_n X_{-n} \xrightarrow{w} \hat{X}_1$ (in H_x) if and only if $\{\sum_{n=0}^{k} a_n X_{-n}\}$ is bounded.

Proof. Let $R_k = \hat{X}_1 - \sum_{n=0}^{k} a_n X_{-n}$. Suppose $\{\sum_{n=0}^{k} a_k X_{-k}\}$ is bounded. Then clearly $\{R_k\}$ is bounded; i.e. $\exists\, M > 0$ such that $\sup_k \|R_k\| = M < \infty$. Let $h \in H_x$. For every $k \in \mathbb{N}$, define $h_k = \mathrm{Proj}(h \,|\, H_x(-k-1))$. Since $\bigcap_k H_x(-k-1) = \{0\}$, by regularity of X, we have $h_k \to 0$ (in H_x) as $k \to \infty$. (c.f. Theorem 1.1.7(iv)). On the other hand for each $k \in \mathbb{N}$, $R_k \in H_x(-k-1)$ (using Relations (2.3.2) and (2.2.5)). Therefore, $(h-h_k, R_k) = 0$ for every $k \in \mathbb{N}$. Hence

$$|(h,R_k)| = |(h_k,R_k)| \le \|h_k\| \cdot \|R_k\| \le M\,\|h_k\| \to 0 \text{ as } k \to 0,$$

which means $R_k \xrightarrow{w} 0$ (in H_x) (c.f. Definition 1.1.8). Conversely if $\sum_{n=0}^{k} a_n X_{-n} \xrightarrow{w} \hat{X}_1$, then $\{\sum_{n=0}^{k} a_n X_{-n}\}$ is bounded (c.f. Theorem 1.1.9 (ii)). □

As it was stated earlier, the goal is to find necessary and sufficient condition(s), preferably in term of density functions, to achieve the strong convergence of the series in AR–expansion to $\hat{X}_1$. Theorem 2.4.1 points to the type of conditions to be considered. The following corollary, in part, is the frequency domain analogue of the theorem above.

2.4.2 Corollary. Using the same notation as in Theorem 2.4.1. let $S_n(e^{i\theta}) = \sum_{k=0}^{n} d_k e^{ik\theta}$ for $n \in \mathbb{N}$. Let $\alpha = \sup_n \frac{1}{2\pi} \int_{-\pi}^{\pi} |S_n|^2 f\, d\theta$. Then

(a) $S_n \xrightarrow{w} 1/\varphi$ (in $L^2(f)$) if and only if $\alpha < \infty$.

(b) If $S_n \xrightarrow{\lambda} 1/\varphi$ and $\alpha < \infty$, then $S_n \to 1/\varphi$ (in $L^1(f)$).

Proof. (a) Using the isometry L between H_x and $L^2(f)$ (c.f. Theorem 1.2.20), since $S_n(e^{i\cdot}) = \sum_{k=0}^{n} d_k e^{ik\cdot} = L^{-1}(\sum_{k=0}^{n} d_k X_{-k})$ for $n \in N$, and $L^{-1}(\hat{X}_1) = e^{-i\cdot}[1-c_0/\varphi]$ (Theorem 1.2.19), by Theorem 2.4.1

$$\sum_{k=0}^{n} a_k e^{ik\cdot} \xrightarrow{w} e^{-i\cdot}[1-c_0/\varphi] \text{ if and only if } \Big\|\sum_{k=0}^{n} a_k e^{ik\cdot}\Big\|^2_{L^2(f)} \text{ is bounded.}$$

However, for $n \in \mathbb{N}$ and $\theta \in [-\pi,\pi]$,

$$(2.4.1) \qquad S_{n+1}(e^{i\theta}) = \sum_{k=0}^{n+1} d_k e^{ik\theta} = d_0 - 1/c_0\, e^{i\theta} \sum_{k=0}^{n} a_k e^{ik\theta}$$

$$(d_0 = 1/c_0) \qquad = 1/c_0 \{1 - e^{i\theta} \cdot \sum_{k=0}^{n} a_k e^{ik\theta}\},$$

so that

$$S_{n+1} \xrightarrow{w} 1/\varphi \text{ (in } L^2(f)) \text{ if and only if } \sum_{k=0}^{n} a_k e^{ik\cdot} \xrightarrow{w} e^{-i\cdot} [1 - c_0/\varphi] \text{ in } L^2(f)).$$

The result follows, by the fact that

$$\|S_n\|^2_{L^2(f)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} |S_n|^2 f \, d\theta.$$

(b) Define

$$\mu(A) = (1/2\pi\|f\|_1)) \int_A f \, d\theta; \qquad A \in \mathscr{B}([-\pi,\pi]).$$

Clearly μ is a probability measure on $\mathscr{B}([-\pi,\pi])$, and $L^p(\mu) = L^p(f)$ for $0 < p \le \infty$. Since $\alpha \le M_0 < \infty$, by part (a) $S_n \xrightarrow{w} 1/\varphi$ (in $L^2(\mu)$). This implies that

$$\lim_{n\to\infty} \int_E S_n \, d\mu = \int_E 1/\varphi \, d\mu \qquad E \in \mathscr{B}([-\pi,\pi]).$$

Using this and the fact that

$$\|S_n\|_{L^1(\mu)} \le \|S_n\|_{L^2(\mu)} \le \alpha \le M_0 < \infty,$$

by the moment inequality, we get $S_n \xrightarrow{w} 1/\varphi$ (in $L^1(\mu)$) (c.f. [5], p 291 Theorem 7). On the other hand $S_n \xrightarrow{\mu} 1/\varphi$ because of $S_n \xrightarrow{\lambda} 1/\varphi$ and $\mu << \lambda$. Hence $S_n \to 1/\varphi$ in $L^1(\mu) = L^1(f)$. [c.f. [5], p 295, Theorem 12]. □

§2.5. Necessary and Sufficient Conditions for the Existence and the Uniqueness of the AR–expansion

Theorem 2.5.1 below provides necessary and sufficient conditions for the existence and uniqueness of AR–expansion for $\hat{X}_1$. Of importance in this theorem is condition (c). This condition is slightly stronger than the condition of "*boundedness of partial sums*" of the series in this expansion which guarantees the weak convergence of this series.

2.5.1 Theorem. Let $X = \{X_t;\ t\in\mathbb{Z}\}$ be a regular SSP with spectral density f and optimal factor φ. Let $1/\varphi(z) = \sum_{n=0}^{\infty} d_n z^n$ for $|z| < 1$. Set $a_n = -c_0\, d_{n+1}$, for $n\in\mathbb{N}$, where $c_0 = \sigma$. Let $S_n(z) = \sum_{k=0}^{n} d_k z^k$ for $z\in\mathbb{C}$. Then the following statements are equivalent:

(a) $\sum_{n=0}^{k} a_n X_{-n} \to \hat{X}_1(\text{in } H_x)$ as $k\to\infty$; i.e.

(2.5.1) $$\hat{X}_1 = \sum_{n=0}^{\infty} a_n X_{-n}\ (\text{in } H_x),$$

and $\{a_n\}$ is the unique sequence for which (2.5.1) holds.

(b) The sequence $\{\varphi\cdot S_n\}$ converges in H^2 to 1. (Equivalently $\{\varphi(e^{i\cdot})S_n(e^{i\cdot})\}$ converges in $L^2(\lambda)$ to 1).

(c) $\overline{\lim_{n\to\infty}} \frac{1}{2\pi}\int_{-\pi}^{\pi} |S_n(e^{i\theta})|^2 f(\theta)d\theta \le 1.$

Proof. Define

$$\hat{\psi}(e^{i\theta}) = e^{-i\theta}[1 - c_0/\varphi(e^{i\theta})]; \qquad \theta\in[-\pi,\pi].$$

The following relations will be used in the course of proof:

(2.5.2) $$\sum_{n=0}^{\infty} c_{n+1}e^{in\theta} = e^{-i\theta}\sum_{n=1}^{\infty} c_n e^{in\theta} = e^{-i\theta}\varphi(e^{i\theta})\,[1 - c_0/\varphi(e^{i\theta})]$$
$$= \varphi(e^{i\theta})\,\hat{\psi}(e^{i\theta}), \quad -\pi\le\theta\le\pi$$

and

(2.5.3) $$\sum_{n=0}^{k} a_n e^{in\theta} = -c_0 \sum_{n=0}^{k} d_{n+1}e^{in\theta} = -c_0\, e^{-i\theta}\sum_{n=0}^{k} d_{n+1}\, e^{i(n+1)\theta}$$

$$= -c_0\, e^{-i\theta}\sum_{n=1}^{k+1} d_n\, e^{in\theta} = -c_0\, e^{-i\theta}\,[-1/c_0 + \sum_{n=0}^{k+1} d_n\, e^{in\theta}]$$

$$= e^{-i\theta}\,[1 - c_0 S_{k+1}(e^{i\theta})]; \qquad \theta\in[-\pi,\pi],\ k\in\mathbb{N}.$$

(a $\longleftrightarrow$ b) using the isomorphism L between $L^2(f)$ and H_x (Theorem 1.2.20). We have

$$\sum_{n=0}^{k} a_n e^{in\theta} = L^{-1}\Big(\sum_{n=0}^{k} a_n X_{-n}\Big) \longrightarrow L^{-1}(\hat{X}_1) = \hat{\psi}(e^{i\theta});\ \text{as } k\to\infty,$$

if and only if

$$\int_{-\pi}^{\pi} |\sum_{n=0}^{k} a_n e^{in\theta} - \hat{\psi}(e^{i\theta})|^2 f(\theta)d\theta \longrightarrow 0 \text{ as } k \to \infty$$

if and only if

$$\text{(2.5.4)} \qquad \int_{-\pi}^{\pi} |e^{-i\theta}[\sum_{n=1}^{\infty} c_n e^{in\theta}] - \varphi(e^{i\theta})\sum_{n=0}^{k} a_n e^{in\theta}|^2 d\theta \longrightarrow 0 \text{ as } k \to \infty.$$

Now, since $\varphi \epsilon L^2(\lambda)$ and $\varphi S_n \epsilon L^2(\lambda)$ for $n \epsilon \mathbb{N}$, (2.5.4) holds if and only if

$$\text{(2.5.5)} \qquad \|\sum_{n=0}^{\infty} c_{n+1} e^{in\cdot} - \varphi(e^{i\cdot}) \sum_{n=0}^{k} a_n e^{in\cdot}\|_2 \to 0; \text{ as } k \to \infty.$$

Using relations (2.5.2) and (2.5.3), (2.5.5) is true if and only if

$$\varphi \cdot e^{-i\cdot}[1 - c_0 S_{k+1}] \to e^{-i\cdot} \varphi[1 - c_0/\varphi] \qquad \text{(in } L^2(\lambda))$$

or, equivalently, if and only if

$$\varphi[1 - c_0 S_{k+1}] \to \varphi[1 - c_0/\varphi] \qquad \text{(in } L^2(\lambda))$$

and, since $L^2(\lambda)$ is a Banach space, if and only if

$$\varphi S_{k+1} \to 1 \qquad \text{(in } L^2(\lambda)).$$

The result follow from the fact that for every $n \epsilon \mathbb{N}$, $\varphi S_n \epsilon H^2$ and that the H^2–norm of $\varphi S_n - 1$ equals $\|\varphi S_n - 1\|_2$.

(b $\longleftrightarrow$ c). We first prove that

$$\text{(2.5.6)} \qquad \varphi S_n \to 1 \text{ (in } L^2(\lambda)) \text{ if and only if } \frac{1}{2\pi}\int_{-\pi}^{\pi} |S_n|^2 f d\theta \to 1 \text{ as } n \to \infty.$$

To this end, note that $\varphi S_n \epsilon L^2(\lambda)$ for each $n \epsilon \mathbb{N}$. We have

$$\|1 - \varphi S_n\|_2^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |1 - \varphi(e^{i\theta})S_n(e^{i\theta})|^2 d\theta$$
$$= 1 - 2\text{Re}\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \varphi(e^{i\theta}) S_n(e^{i\theta})d\theta\}$$
$$+ \frac{1}{2\pi}\int_{-\pi}^{\pi} |\varphi(e^{i\theta}) S_n(e^{i\theta})|^2 d\theta; \qquad n \epsilon \mathbb{N},\ \theta \epsilon [-\pi, \pi].$$

However,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \varphi(e^{i\theta}) S_n(e^{i\theta})\, d\theta = c_0 d_0 = 1, \qquad n \epsilon \mathbb{N}.$$

Since $\varphi S_n \epsilon L^2(\lambda)$, and

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |\varphi(e^{i\theta})S_n(e^{i\theta})|^2\, d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} |S_n(e^{i\theta})|^2\, f(\theta)d\theta, \qquad n\epsilon\mathbb{N},$$

it follows that

$$\|1 - \varphi\, S_n\|_2^2 = 1 - 2 + \frac{1}{2\pi}\int_{-\pi}^{\pi} |S_n(e^{i\theta})|^2\, f(\theta)d\theta, \qquad n\epsilon\mathbb{N},$$

which proves (2.5.6). As a result of this

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |S_n|^2\, f\, d\theta \geq 1$$

which implies

$$\underline{\lim}_{n\to\infty}\ \frac{1}{2\pi}\int_{-\pi}^{\pi} |S_n|^2\, f\, d\theta \geq 1.$$

Therefore, by (2.5.6) and the fact the $\underline{\lim}\ a_n \leq \overline{\lim}\ a_n$ for any sequence $\{a_n\}$, $\varphi\, S_n \to 1$ (in $L^2(\lambda)$) if and only if $\overline{\lim}_{n\to\infty}\ \frac{1}{2\pi}\int_{-\pi}^{\pi} |S_n(e^{i\theta})|^2 f(\theta)\, d\theta \leq 1.$

The ***uniqueness*** of AR–expansion follows from Theorem 2.3.1. □

The following example is due to Topsoe [see [12], p 92], to demonstrate that a process with spectral density $f(\theta) = |1+e^{i\theta}|^2$, $-\pi \leq \theta \leq \pi$, does not admit an AR–expansion for $\hat{X}_1$. Our Theorem 2.5.1 confirms his conclusion.

Example. Let $X = \{X_t;\ t\epsilon\mathbb{Z}\}$ be any SSP with spectral density $f(\theta) = |1 + e^{i\theta}|^2$, $-\pi \leq \theta \leq \pi$. Note that f is continuous, $\log f \in L^1(\lambda)$, its optimal factor is $\varphi(z) = 1 + z$ for $|z| \leq 1$, and

$$1/\varphi(z) = \sum_{k=0}^{\infty} (-1)^k z^k; \qquad |z| < 1.$$

Therefore

$d_n = (-1)^n$ and $S_n(z)\varphi(z) = 1 + (-1)^n z^{n+1}$, for every $n\epsilon\mathbb{N}$. Hence

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |S_n\varphi|^2\, d\theta = 2,$$

which using 2.5.1(c), shows AR–expansion does not exist for $\hat{X}_1$.

Theorem 2.5.1 subsumes the results of Wiener–Masani, and that of Miamee–Salehi. Note that the latter authors assume the conditions $(1/f)\epsilon L^1(\lambda)$ and $S_n \to 1/\varphi$ (in $L^2(f)$) to achieve $\hat{X}_1 = \sum_{n=0}^{\infty} a_n X_{-n}$ for some $\{a_n\}$ with $\sum_{n=0}^{\infty} |a_n|^2 < \infty$ and conversely. As we'll see later (see example following Corollary 2.6) the condition $(1/f)\epsilon L^1(\lambda)$ by no means is necessary to have the AR–expansion of $\hat{X}_1$.

2.5.2. Corollary (*Masani*). If $(1/f)\epsilon L^1(\lambda)$ and $f\epsilon L^{\infty}(\lambda)$, then for some sequence $\{a_n\}$

$$\hat{X}_1 = \sum_{n=0}^{\infty} a_n X_{-n} \text{ (in } H_x).$$

Proof. Let Ess. sup $f = \alpha$. By assumption $\alpha < \infty$. Since $(1/f)\epsilon L^1(\lambda)$, we have $(1/\varphi)\epsilon L^2(\lambda)$ and $\|S_n - 1/\varphi\|_2 \to 0$. Therefore, as $n \to \infty$

$$\int_{-\pi}^{\pi} |\varphi S_n - 1|^2 d\lambda = \int_{-\pi}^{\pi} |S_n - 1/\varphi|^2 \cdot |\varphi|^2 d\lambda \leq \alpha \int_{-\pi}^{\pi} |S_n - 1/\varphi|^2 d\lambda \to 0$$

which implies condition (b) in Theorem 2.5.1. Hence (a) in this theorem follows. □

2.5.3. Corollary (*Wiener–Masani*). If $0 < c \leq f \leq d < \infty$ for some c and d, then there exists a sequence $\{a_n\}$ such that

$$\hat{X}_1 = \sum_{n=0}^{\infty} a_n X_{-n} \text{ (in } H_x).$$

Proof. An immediate consequence of Corollary 2.5.2.

2.5.4. Corollary. (*Miamee–Salehi*). $\hat{X}_1 = \sum_{n=0}^{\infty} a_n X_{-n}$ (in H_x) for some sequence $\{a_n\}$ with $\sum_{n=0}^{\infty} |a_n|^2 < \infty$ if and only if

(i) $(1/f)\ \epsilon L^1(\lambda)$

and

(ii) The Fourier series of $1/\varphi$ converges to $1/\varphi$ in $L^2(f)$.

When (i) and (ii) are true, $a_n = -c_0 d_{n+1}$, where $c_0 = \sigma$ and $\{d_n\}$ is the sequence of coefficients in $1/\varphi(\theta) = \sum_{k=0}^{\infty} d_k e^{ik\theta}$.

Proof. Suppose $\hat{X}_1 = \sum_{n=0}^{\infty} a_n X_{-n}$ (in H_x). Then by Theorem 2.3.1, $a_n = -c_0 d_{n+1}$ for each $n \in \mathbb{N}$, and by part (b) of Theorem 2.5.1 condition (ii) holds. Assumption $\sum_{n=0}^{\infty} |a_n|^2 < \infty$ easily implies that $1/\varphi \in L^2(\lambda)$ and therefore $1/f \in L^1(\lambda)$.

On the other hand if $1/f \in L^1$ then $\sum_{n=0}^{\infty} |a_n|^2 < \infty$ for $a_n = -c_0 d_{n+1}$, $n \in \mathbb{N}$, and assuming condition (ii) we get (c) in Theorem (2.5.1). Thus $\hat{X}_1 = \sum_{n=0}^{\infty} a_n X_{-n}$ (in H_x). □

2.5.5. Corollary (*Pourahmadi*). The positivity of the angle between the past – present and the future of the SSP X is sufficient to guarantee the existence of an AR–expansion of $\hat{X}_1$.

Proof. Since by assumption $\rho(f) < 1$ (c.f. Definition 1.2.22), using Theorem 1.2.23 we have

(a) $(1/f) \in L^1(\lambda)$, and

(b) the Fourier series of $1/\varphi$, the reciprocal of the optimal factor, converges to $1/\varphi$ in $L^2(f)$. The result follows from Corollary 2.5.4. □

Theorem 2.5.1 has further consequences. Two of them are stated below.

2.5.6. Corollary. Let $X = \{X_t; t \in Z\}$ be a regular SSP with optimal factor φ. Let $1/\varphi(z) = \sum_{n=0}^{\infty} d_n z^n$ for $|z| < 1$. If $\hat{X}_1 = \sum_{n=0}^{\infty} a_n X_{-n}$ converges for some sequence $\{a_n\}$, then

(2.5.7) $$S_n \chi_{[f>\alpha]} \to 1/\varphi \, \chi_{[f>\alpha]} \quad (\text{in } L^2(\lambda))$$

for every $\alpha > 0$.

Proof. Let $\hat{X}_1 = \sum_{n=0}^{\infty} b_n X_{-n}$ for some sequence $\{b_n\}$. By Theorem 2.3.1, $b_n = -c_0 d_{n+1}$ for $n \in \mathbb{N}$, and by part (b) of Theorem 2.5.1, $S_n \to 1/\varphi$ in $L^2(f)$ which implies $|S_n|^2 \to 1/f$ in $L^1(f)$. As a result $\{|S_n|^2\}$ is uniformly integrable w.r.t. the probability measure defined by

$$\mu(A) = (1/2\pi)\|f\|_1 \int_A f\,d\theta; \qquad A \in \mathscr{B}([-\pi,\pi]).$$

So, for every $\alpha > 0$, as $c \to \infty$ we have

$$0 \longleftarrow \sup_n \int_{[|S_n|^2>c]} |S_n|^2 f\,d\theta \geq \sup_n \int |S_n|^2 f\, \chi_{[|S_n|^2>c]\cap[f>\alpha]}\, d\theta$$

$$\geq \alpha\cdot\sup_n \int |S_n|^2 \chi_{[|S_n|^2>c]\cap[f>\alpha]}\, d\theta$$

$$= \alpha\cdot \sup_n \int |S_n\chi_{[f>\alpha]}|^2 \chi_{[|S_n|^2>c]}\, d\theta$$

$$\geq \alpha\cdot \sup_n \int |S_n\chi_{[f>\alpha]}|^2 \chi_{[|S_n\chi_{[f>\alpha]}|^2>c]}\, d\theta.$$

As a result, for each $\alpha \in (0,+\infty)$, $\{|S_n\chi_{[f>\alpha]}|^2\}$ is uniformly integrable with respect to λ. Since $S_n \xrightarrow{\lambda} 1/\varphi$ (because $S_n\varphi \xrightarrow{\lambda} 1$ by (b) in Theorem 2.5.1), we get

$$S_n\chi_{[f>\alpha]} \longrightarrow 1/\varphi\, \chi_{[f>\alpha]} \quad (\text{in } L_2(\lambda)). \qquad \square$$

In view of condition (2.5.7) of this corollary, it is tempting to conjecture that the condition "$S_n \to 1/\varphi$ (in $L^2(\lambda)$)" would hold under the assumption of existence of the AR–expansion of $\hat{X}_1$. Had this been true, one would get $(1/f) \in L^1(\lambda)$. However, as it is shown (see also [14] p 317), the example $f(\theta) = |1+e^{i\theta}|^{2\lambda}$ for any $1/2 \leq \lambda < 1$ demonstrates that the condition "$(1/f)L^1\epsilon(\lambda)$" is not necessary for the existence of AR–expansion of $\hat{X}_1$.

Example. Consider SSP $X = \{X_t;\ t \in \mathbb{Z}\}$ with spectral density $f(\theta) = |1+e^{i\theta}|^{2\lambda}$, $-\pi \leq \theta \leq \pi$, and $1/2 \leq \lambda < 1$. $\hat{X}_1$ has an AR–expansion (c.f.

[14], p 317). However $(1/f) \notin L^1(\lambda)$. Clearly $\varphi(z) = (1+z)^{\lambda}$ for $|z| \leq 1$, and using Taylor expansion, we have

$$1/\varphi(z) = \sum_{k=0}^{\infty} \frac{(\lambda)_k}{k!} (-1)^k z^k, \text{ where } (\lambda)_k = \lambda(\lambda+1)\ldots(\lambda+k-1).$$

Therefore by Theorem 2.5.1 we have

$$\hat{X}_1 = \sum_{k=0}^{\infty} [(c_0(-1)^{k+1} (\lambda)_k)/(k!)] X_{-k} \qquad \text{(in } H_x\text{)}. \quad \square$$

The corollary below, whose proof is deduced from Theorem 2.5.1, puts the conditions on the density function f.

2.5.7 Corollary. With the same notation as in Theorem 2.5.1, let $1/f \in L^p(\lambda)$ for some p, $1 \leq p < \infty$ and $f \in L^q$ for $1 < q \leq \infty$ where $1/p + 1/q = 1$ then $\hat{X}_1 = \sum_{n=0}^{\infty} a_n X_{-n}$, and $\{a_n\}$ is the unique sequence with this property.

Proof. Let $1 \leq p < \infty$. Since $1/f \in L^p(\lambda)$, we have $1/\varphi \in L^{2p}(\lambda)$ which implies $S_n \to 1/\varphi$ in $L^{2p}(\lambda)$ (c.f. [8]; p 50, Theorem 1); i.e.

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |S_n - 1/\varphi|^{2p} d\theta \to 0 \qquad \text{as } n \to \infty.$$

This, in turn, using Holder and Minkowski's inequalities implies

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} ||S_n|^2 - 1/f|^p d\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} ||S_n|^2 - |1/\varphi|^2|^p d\theta$$

$$\leq \{\frac{1}{2\pi} \int_{-\pi}^{\pi} |S_n + 1/\varphi|^{2p} d\theta\}^{\frac{1}{2}} \{\frac{1}{2\pi} \int_{-\pi}^{\pi} |S_n - 1/\varphi|^{2p} d\theta\}^{\frac{1}{2}}$$

$$\leq \{[\frac{1}{2\pi} \int_{-\pi}^{\pi} |S_n|^{2p} d\theta]^{\frac{1}{2p}} + [\frac{1}{2\pi} \int_{-\pi}^{\pi} |1/\varphi|^{2p} d\theta]^{\frac{1}{2p}}\}^p \times$$

$$\{\frac{1}{2\pi} \int_{-\pi}^{\pi} |S_n - 1/\varphi|^{2p} d\theta\}^{\frac{1}{2}}$$

$$\leq \{2 \|1/f\|_p\}^{\frac{p}{2}} \|S_n - 1/\varphi\|_{2p}^{\frac{1}{4p}} \to 0.$$

Now, since $1/p + 1/q = 1$ and $f \in L^q$, using Holder's inequality once again, we get

$$|S_n|^2 f \to 1 \ \ (\text{in } L^1(\lambda)).$$

Hence condition (c) in Theorem 2.5.1. Consequently (a) in this theorem follows.

After the completion of this thesis it was brought to our attention that a result based on Hunt, Muckenhoupt and Wheeden Theorem [1973], more general than Corollary 2.5.7 without the uniqueness of coefficients in the resulting AR–expansion, had also been obtained by P. Bloomfield in Annals of Prob.; 1985, Vol. 13, No. 1, 226–233.

§2.6. Invertible Processes and Connection with AR–expansion

2.6.1 Definition. A regular process $X = \{X_t;\ t\in\mathbb{Z}\}$ with innovation process $\zeta = \{\zeta_t;\ t\in\mathbb{Z}\}$ (c.f. Theorem 1.2.14) is said to be *invertible* if there exists a sequence $\{e_j\}$ such that

$$\zeta_t = \sum_{j=0}^{\infty} e_j X_{t-j} \ (\text{in } H_x) \qquad t\in\mathbb{Z}.$$

There is a strong relationship between the innovation process ζ of the process X and $\{\hat{X}_t;\ t\in\mathbb{Z}\}$ via Relation 2.5.1. This relationship is explored in the following theorem.

2.6.2 Theorem (*Invertibility*). Let $X = \{X_t;\ t\in\mathbb{Z}\}$ be a regular SSP with the optimal factor φ. Let $1/\varphi(z) = \sum_{n=0}^{\infty} d_n z^n$ for $|z| < 1$. Set $a_n = -c_0 d_{n+1}$ for $n\in\mathbb{N}$, where $c_0 = \sigma$. The following statements are equivalent

(a) $\sum_{n=0}^{k} a_n X_{-n} \to \hat{X}_1 \ (\text{in } H_x)$ as $k \to \infty$; i.e.

(2.6.1) $$\hat{X}_1 = \sum_{n=0}^{\infty} a_n X_{-n}$$

and $\{a_n\}$ is the unique sequence for which (2.6.1) holds

(b)

(2.6.2) $$\zeta_t = \sum_{k=0}^{\infty} d_k X_{t-k} \text{ (in } H_x); \qquad t \in \mathbb{Z}$$

and $\{d_n\}$ is the unique sequence with this property.

Proof. Let (a) be true. Then, using Theorem 1.2.14, we have

$$\begin{aligned}\zeta_1 &= (1/c_0)\,[X_1 - \hat{X}_1] \\ &= (1/c_0)\,[X_1 + c_0 \sum_{k=0}^{\infty} d_{k+1} X_{-k}] \\ &= d_0 X_1 + \sum_{k=1}^{\infty} d_k X_{1-k} \\ &= \sum_{k=0}^{\infty} d_k X_{1-k},\end{aligned}$$

and, using shift operators of X,

$$\zeta_t = \sum_{k=0}^{\infty} d_k X_{t-k}; \qquad t \in \mathbb{Z}.$$

The uniqueness of coefficients $\{d_k\}$ in (2.6.2) follows from the uniqueness of coefficients $\{a_n\}$ in (2.6.1). (c.f. Theorem 2.3.1). Hence (b).

The implication (b) $\rightarrow$ (a) follows from above relations and the moving average theorem (c.f. Theorem 1.2.14). □

§2.7 Sufficiency of the Condition $(1/f) \in L^1(\lambda)$

As it was mentioned earlier, among the Masani's set of conditions "$(1/f) \in L^1(\lambda)$ and $f \in L^{\infty}(\lambda)$" which guarantee *an* AR–expansion of $\hat{X}_1$, can be reduced (c.f. Theorem 2.5.1). Still the question whether $(1/f) \in L^1(\lambda)$ implies the existence of the AR–espansion, raised by Miamee–Salehi remains open. This problem is under consideration by the author.

REFERENCES

1. Bauer, *H.*: *Probability Theory and Elements of Measure Theory*; Second edition, Academic Press, 1981.

2. Billingsley, P.: *Probability and Measure.* second edition, John Wiley and Sons (1986).

2'. Billingsley, P.: *Convergence of Probability Measures*, John Wiley and Sons (1968).

3. Brockwell, P.J. and Davis, R.A.: *Time Series Theory and Methods*; first edition; Springer–Verlag (1987).

4. Carleson, L.: *On the Convergence and Growth of Partial Sums of Fourier Series*, Acta Math. 116 (1966), 135–157.

5. Dunford, N. and Schwartz, J.T.: *Linear Operators*; Interscience Publisher, Inc. (1967).

6. Fabian, V. and Hannan, J.: *Intro. to Probability and Math. Statistics*; first edition, John Wiley & Sons (1985).

7. Helson, H. and Szegö, G.: *A Problem in Prediction Theory*; Ann. Mat. Pura Appl. 51, 107–138 (1960).

8. Katznelson, Y.: *An Introduction to Harmonic Analysis*; Dover (1976).

9. Makagon, A. and Salehi, H.: *Stationary Fields with Positive Angle*, Journal of Multivariate Analysis, 106–125 (1986).

10. Masani, P.: *The Prediction Theory of Multivariate Stochastic Processes*; III. Acta Math. 104, 141–162 (1960).

11. Masani, P.: *Recent Trends in Multivariate Prediction Theory*; In: P.R. Krishnaiah (ed.) Multivariate analysis, 351–382. New York: Academic Press (1966).

12. Miamee, A.G. and Salehi, H.: *On an Explicit Representation of the Linear Predictor of a Weakly Stationary Stochastic Sequence*; Bol. Soc. Math. Mexicana 28, 81–93 (1983).

13. Pourahmadi, M.: *The Helson–Szegö Theorem and the Abel Summability of the Series for the Predictor*; Proc. Am. Math. Soc. 91, 306–308 (1984).

14. Pourahmadi, M.: *Autoregressive Representation of Multivariate Stationary Stochastic Processes*; Probab. Th. Rel. Fields 80, 31–322. Springer–Verlag (1988).

15. Rozanov, Yu. A.: *Stationary Random Processes* (translated from Russian); Holden Day (1967).

16. Rudin, W.: *Real and Complex Analysis*; Third edition, McGraw–Hill (1987).

17. Wiener, N. and Masani, P.: The Prediction Theory of Multivariate Stochastic Processes. Part I. Acta. Math. 98, 111–150, 1957); Part II. Acta Math 99, 93–137 (1958).