Boletín de la Sociedad Matemática Mexicana Vol. 28, No. 2, 1983

# ON AN EXPLICIT REPRESENTATION OF THE LINEAR PREDICTOR OF A WEAKLY STATIONARY STOCHASTIC SEQUENCE*

BY A. G. MIAMEE AND H. SALEHI[1]

## Introduction

Let $Z$ be the set of all integers. $X_k$, $k \in Z$, will denote a univariate weakly stationary stochastic process (WSSP) conidered as a sequence of elements in a Hilbert space $\mathscr{H}$. The main purpose of this note is to establish a spectral necessary and sufficient condition for a series representation of the linear predictor of $X_k$, $k \in Z$, in the form:

$$P_{-1}X_0 = \sum_{k=1}^{\infty} a_k X_{-k} \quad \text{with} \quad \sum_{k=1}^{\infty} |a_k|^2 < \infty,$$

where $P_{-1}$ denotes the orthogonal projection onto the subspace spanned by the past of the process (see Theorem 3.2). Such a series representation was obtained by P. Masani [4, 1960] for a multivariate weakly stationary stochastic process under the additional boundedness assumption on the spectral density of the process. Masani's result for the univariate case is an improvement of that of Akutowicz [1, 1957] who, in addition to the boundedness of spectral density, assumed that the Fourier series of the optimal factor of the spectral density converges absolutely. Although Masani removed this unduly restricted condition on the optimal factor, nevertheless his success in obtaining an autoregressive series for the linear predictor depends heavily on the boundedness of the spectral density. The last assumption is still strong, as was pointed out by Masani himself [4, 1960]. It is in this direction that our note extends the work of Masani, and provides a contribution to the area of predictor theory.

## 1. Preliminaries

Let $X_k$, $k \in Z$, be a univariate WSSP. $F$ will denote the spectral distribution of $X_k$, $k \in Z$. $\gamma$ will stand for the correlation function of $X_k$, $k \in Z$. It is assumed that $F$ is absolutely continuous with density $f$, so that $\gamma$ and $f$ are related by $\gamma(k) = \frac{1}{2\pi} \int_0^{2\pi} e^{-ik\theta} f(\theta)\, d\theta$. $L_{2,f}$ will denote the usual Hilbert space of square integrable functions on $[0, 2\pi]$ with respect to (w.r.t.) the density $f$.

* Presented at the "Workshop on the prediction theory on non-stationary processes and related topics", held at Centro de Investigación en Matemáticas (CIMAT), Guanajuato, México, June 20–26, 1982.

[1] This research was partially supported by the National Science Foundation MCS-7825525. H. Salehi wishes to acknowledge the support of the National Research Council of Canada during his visit at Carleton University during 1978–79.

This work started in 1976–77 while A. G. Miamee was visiting Michigan State University. A version of the work is contained in Research Memorandum 416, November 1980, Department of Statistics and Probability, Michigan State University.

We recall that the space $L_{2,f}$ is isometric to $\sigma(X)$, the subspace generated by $X_k$, $k \in Z$, under the isomorphism map $X_k \leftrightarrow e^{-ik\theta}$. As usual this map between the time and spectral domains will play an important role in this paper. $L_1$, $L_2$ and $L_\infty$ will denote the usual equivalence classes of functions w.r.t. the Lebesgue measure on $[0, 2\pi]$. Let $L_2^{0+} = \{g : g \in L_2 \text{ such that } \int_0^{2\pi} e^{-ik\theta} g(\theta)\, d\theta = 0,\ k \leq -1\}$. For each $n$, $\mathcal{M}_n$ will denote the subspace of $\mathcal{H}$ spanned by the elements $X_k$, $k \leq n$. $P_n$ will denote the orthogonal projection onto the subspace $\mathcal{M}_n$. We assume that $\log f \in L_1$. Let $Y_k$, $k \in Z$, denote the normalized innovation process of $X_k$, $k \in Z$. $Y_k$, $k \in Z$ is a WSSP whose spectral density is the constant 1 on $[0, 2\pi]$. By the Wold decomposition

$$X_0 = \textstyle\sum_{n=0}^{\infty} c_n Y_{-n},\ c_n = (X_0, Y_{-n}),\ n \geq 0 \tag{1.1}$$

and

$$\begin{gathered} \sigma Y_0 = X_0 - P_{-1} X_0 \\ \sigma^2 = \exp\left[\frac{1}{2\pi} \int_0^{2\pi} \log f(\theta)\, d\theta\right] \end{gathered} \tag{1.2}$$

where

$$\begin{gathered} f = \varphi\bar{\varphi} \\ \varphi = \textstyle\sum_{n=0}^{\infty} c_n e^{in\theta},\ \sum_{n=0}^{\infty} |c_n|^2 < \infty,\ c_0 = \sigma > 0. \end{gathered} \tag{1.3}$$

The function $\varphi$ is called the optimal factor of $f$. $\varphi$ has an analytic extension $\varphi_+$ to the unit disc given by

$$\begin{gathered} \varphi_+(z) = \exp\left[\frac{a_0}{2} + \textstyle\sum_{n=1}^{\infty} a_n z^n\right] \\ a_n = \dfrac{1}{2\pi} \int_0^{2\pi} e^{-in\theta} \log f(\theta)\, d\theta. \end{gathered} \tag{1.4}$$

The coefficients $c_n$ in (1.3) may be obtained in terms of the coefficients $a_n$ from

$$\textstyle\sum_{n=0}^{\infty} c_n z^n = \exp\left[\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n z^n\right]. \tag{1.5}$$

The function $\varphi_+(z)$ in (1.4) has no zero inside the unit disc and $\varphi_+(0) = c_0 > 0$. These are equivalent to saying that the factor $\varphi$ occurring in (1.3) satisfies the following conditions

$$\begin{gathered} f = |\varphi|^2,\ \varphi = \textstyle\sum_{n=0}^{\infty} c_n e^{in\theta} \in L_2,\ c_0 > 0, \\ \text{and if } \psi = \textstyle\sum_{n=0}^{\infty} d_n e^{in\theta} \in L_2^{0+} \text{ such that} \\ f = |\psi|^2, \text{ then } c_0 \geq |d_0|. \end{gathered} \tag{1.6}$$

Condition (1.6) expresses a sort of "maximality" for $\varphi$. In some special cases, say when $f$ is a rational function in $e^{i\theta}$ it is easy to compute the optimal factor $\varphi$ by examining the zeros and poles of $f$, and there is no need for determining $\varphi$ through formula (1.5). In this paper we will assume that the optimal factor $\varphi$ is known to us possibly through (1.5). For a function $\psi = \sum_{k=0}^{\infty} b_k e^{ik\theta}$, $\sum_{k=0}^{\infty} |b_k|^2 < \infty$, $\psi^N$ denotes $\sum_{k=0}^{N} b_k e^{ik\theta}$.

## 2. An autoregressive series for the innovation

In this section we obtain a spectral necessary and sufficient condition in order that the innovation process $Y_k$, $k \in Z$, may be expressed as an autoregressive series in terms of the process $X_k$, $k \in Z$. This is an extension of Theorem 1 in [1, 1957]. We start with the following observations: Let $f$ be the spectral density and $\varphi$ its optimal factor ($f$ and $\log f$ are in $L_1$). From $f = \varphi\bar{\varphi}$ it follows that $f^{-1} \in L_1$ whenever $\varphi^{-1}$ is in $L_2$. Conversely suppose $f^{-1} \in L_1$. Then $\varphi^{-1} = \sum_{k=0}^{\infty} d_k e^{ik\theta}$ with $\sum_{k=0}^{\infty} |d_k|^2 < \infty$. This is because $\varphi^{-1} \in L_2$ and $\varphi$ is optimal whose extension $\varphi_+(z)$ has no zeros inside the unit disc. An extension of this result for the matrix-valued case is proved by Masani [4, Lemma 2.7 and Theorem 2.8]. Hence, the two conditions $f^{-1} \in L_1$ and $\varphi^{-1} = \sum_{k=0}^{\infty} d_k e^{ik\theta}$, $\sum_{k=0}^{\infty} |d_k|^2 < \infty$ are equivalent. This observation will be used in the proof of Theorem 2.1 and in Section 3.

THEOREM (2.1.) *Let $X_k$, $k \in Z$, be a* WSSP *with spectral density $f$ such that $\log f \in L_1$. Let $\varphi$ be the optimal factor of $f$. Let $Y_k$, $k \in Z$ be the normalized innovation process of $X_k$, $k \in Z$. Then there exists a sequence $d_k$, $k \geq 0$ such that*

(2.1) $$Y_n = \sum_{k=0}^{\infty} d_k X_{n-k}, \quad \sum_{k=0}^{\infty} |d_k|^2 < \infty$$

*if and only if $f^{-1} \in L_1$ (or equivalently by observation above $\varphi^{-1} \in L_2^{0+}$) and the Fourier series of $\varphi^{-1}$ converges to $\varphi^{-1}$ in $L_{2,f}$. Under these conditions $d_k$ is the kth Fourier coefficient of $\varphi^{-1}$, i.e.,*

$$d_k = \frac{1}{2\pi} \int_0^{2\pi} e^{-k\theta} \varphi^{-1}(\theta)\, d\theta.$$

*Proof.* Let $Y_n = \sum_{k=0}^{\infty} d_k X_{n-k}$, $\sum_{k=0}^{\infty} |d_k|^2 < \infty$. For $n = 0$ we have

$$Y_0 = \sum_{k=0}^{\infty} d_k X_{-k}, \quad \sum_{k=0}^{\infty} |d_k|^2 < \infty.$$

But by (1.1)

$$X_{-k} = \sum_{n=0}^{\infty} c_n Y_{-k-n}, \quad \sum_{n=0}^{\infty} |c_n|^2 < \infty.$$

Hence,

(2.2) $$Y_0 = \sum_{k=0}^{\infty} d_k \sum_{n=0}^{\infty} c_n Y_{-k-n},$$

where the convergence involved is in $\mathscr{H}$. From (2.2) and the isomorphism between $\sigma(Y)$ and $L_{2,f_y} = L_2$ (because the spectral density of $Y_k$, $k \in Z$, is 1) we conclude that

(2.3) $$1 = \lim_{N\to\infty} \sum_{k=0}^{N} [d_k \lim_{M\to\infty} \sum_{n=0}^{M} c_n e^{i(k+n)\theta}],$$

where lim means limit in the space $L_2$, or equivalently

$$(2.4)\qquad \begin{aligned} 1 &= \lim_{N\to\infty} \textstyle\sum_{k=0}^{N} [d_k e^{ik\theta} \lim_{M\to\infty} \sum_{n=0}^{M} c_n e^{in\theta}] \\ &= \lim_{N\to\infty} \textstyle\sum_{k=0}^{N} d_k e^{ik\theta}\varphi(\theta). \end{aligned}$$

But $\sum_{k=0}^{\infty} | d_k |^2 < \infty$ implies that

$$(2.5)\qquad \lim_{N\to\infty} \textstyle\sum_{k=0}^{N} d_k e^{ik\theta} \quad \text{exists} \quad = \chi(\theta) \in \mathrm{L}_2^{0+}.$$

We note that $\varphi = \sum_{k=0}^{\infty} c_k e^{ik\theta} \in L_2^{0+}$. Therefore the $n$th Fourier coefficient of $X^N\varphi = \sum_{k=0}^{N} d_k e^{ik\theta}\varphi$ is $\sum_{k=0}^{n} d_k c_{n-k}$ if $0 \le n \le N$ and $\sum_{k=0}^{N} d_k c_{n-k}$ if $n \ge N + 1$. Hence by (2.4) and Parseval's identity we have $| c_0 d_0 - 1 |^2 + \sum_{n=0}^{N} | \sum_{k=0}^{n} d_k c_{n-k} |^2 \to 0$, as $N \to \infty$. This implies that

$$(2.6)\qquad \begin{aligned} c_0\, d_0 &= 1 \\ \textstyle\sum_{k=0}^{n} d_k c_{n-k} &= 0,\ n \ge 1. \end{aligned}$$

Since both $\varphi$ and $\chi$ are in $L_2^{0+}$, relation (2.6) implies that $\varphi\chi$ and the constant function 1 have the same Fourier coefficients. Hence

$$(2.7)\qquad \varphi\chi = 1 \text{ a.e.}$$

By (2.7) $\chi = \varphi^{-1}$. This and (2.5) imply that $\varphi^{-1} \in L_2^{0+}$,

$$| \varphi^{-1} |^2 = f^{-1} \in L_1 \quad \text{and} \quad d_k = \frac{1}{2\pi} \textstyle\int_0^{2\pi} \varphi^{-1}(\theta) e^{ik\theta}\, d\theta.$$

(2.4) means

$$\begin{aligned} &\lim_{N\to\infty} \textstyle\int_0^{2\pi} |[\varphi^{-1}]^N(\theta)\varphi(\theta) - 1 |^2\, d\theta = 0, \text{ or equivalently} \\ &\lim_{N\to\infty} \textstyle\int_0^{2\pi} |[\varphi^{-1}]^N(\theta) - \varphi^{-1}(\theta) |^2 | \varphi(\theta) |^2\, d\theta = 0, \text{ that is} \\ &\lim_{N\to\infty} \textstyle\int_0^{2\pi} |[\varphi^{-1}]^N(\theta) - \varphi^{-1}(\theta) |^2 f(\theta)\, d\theta = 0, \end{aligned}$$

so that the Fourier series of $\varphi^{-1}$ converges to $\varphi^{-1}$ in $L_{2,f}$.

Conversely let us assume that $f^{-1} \in L_1$ or equivalently $\varphi^{-1} = \sum_{k=0}^{\infty} d_k e^{ik\theta} \in L_2^{0+}$ by the observation in the opening paragraph of Section 2, and

$$\textstyle\int_0^{2\pi} | [\varphi^{-1}]^N(\theta) - \varphi^{-1}(\theta) |^2 f(\theta)\, d\theta \to 0, \quad \text{as} \quad N \to \infty.$$

Then by reversing the order of the chain of equalities just proved we have

$$\lim_{N\to\infty} \textstyle\sum_{k=0}^{N} [d_k e^{ik\theta} \lim_{M\to\infty} \sum_{n=0}^{N} c_n e^{in\theta}] = 1.$$

From isomorphism between $\sigma(Y)$ and $L_2$ we conclude that

$$\textstyle\sum_{k=0}^{\infty} [d_k \sum_{n=0}^{\infty} c_n Y_{-k-n}] = Y_0.$$

But $\sum_{n=0}^{\infty} c_n Y_{-k-n} = X_{-k}$. Hence $\sum_{k=0}^{\infty} d_k X_{-k} = Y_0$, and therefore $Y_n = \sum_{k=0}^{\infty} d_k X_{n-k}$. Obviously $\sum_{k=0}^{\infty} | d_k |^2 < \infty$, because $| \varphi^{-1} |^2 = f^{-1} \in L_1$.

## 3. An autoregressive series for the linear predictor

Using the result of last section we will give a spectral necessary and sufficient condition for autoregressive series representation for the linear predictor. Our result extends the work of Akutowicz [1, 1957] and those of Masani in the univariate case [4, 1960]. Akutowicz assumes that the Fourier series of $\varphi$ converges absolutely and Masani assumes that the spectral density $f$ is bounded. Both Akutowicz and Masani give only sufficient conditions under which they obtain an explicit formula for the linear predictor. Our result will improve their set of sufficient conditions as well as will provide a necessary set of conditions for the representation of the linear predictor as an autoregressive series.

In view of Theorem 2.1 we explicitly write down the following Assumption.

ASSUMPTION (3.1).

(i) $f^{-1} \in L_1$ *or equivalently* $\varphi^{-1} = \sum_{k=0}^{\infty} d_k e^{ik\theta}$ *with* $\sum_{k=0} |d_k|^2 < \infty$, *and*

(ii) $\lim_{N\to\infty} \int_0^{2\pi} |[\varphi^{-1}]^N(\theta) - \varphi^{-1}(\theta)|^2 f(\theta)\, d\theta = 0.$

We note that when $f$ is bounded if $f^{-1} \in L_1$, then $\lim_{N\to\infty} \int_0^{2\pi} |[\varphi^{-1}]^N - \varphi^{-1}|^2 f(\theta)\, d\theta = 0$. Also when $\sum_{k=0}^{\infty} |d_k| < \infty$ the conditions, 3.1(i)–(ii) follow. We state this result as a proposition below. However we mention that under $f^{-1} \in L_1$ or even $\sum_{k=0}^{\infty} |d_k| < \infty$ the condition $f$ in $L_\infty$ is not necessary for the validity of 3.1(ii) as seen from example 5.1.

PROPOSITION (3.2). *Let* $\sum_{k=0}^{\infty} |d_k| < \infty$. *Then* $f^{-1} \in L_1$ *and* $\lim_{N\to\infty} \int_0^{2\pi} |[\varphi^{-1}]^N - \varphi^{-1}|^2 f(\theta)\, d\theta = 0$.

*Proof.*

$$\sum_{k=0}^{\infty} |d_k| < \infty \Rightarrow \sum_{k=0}^{\infty} |d_k|^2 < \infty \Leftrightarrow \varphi^{-1} \in L_2 \Leftrightarrow f^{-1} \in L_1.$$

Next

$$|[\varphi^{-1}]^N(\theta) - \varphi^{-1}(\theta)|^2 f(\theta)$$
$$= |[\varphi^{-1}]^N(\theta)|^2 - \varphi(\theta)[\varphi^{-1}]^N(\theta) - \bar{\varphi}(\theta)[\bar{\varphi}^{-1}]^N(\theta) + 1.$$

Because $\varphi$ and $\varphi^{-1}$ are in $L_2$, both the integrals $\int_0^{2\pi} \varphi(\theta)[\varphi^{-1}]^N(\theta)\, d\theta$ and $\int_0^{2\pi} \bar{\varphi}(\theta)[\varphi^{-1}]^N(\theta)\, d\theta$ converge to $2\pi$, as $N \to \infty$.

$$\frac{1}{2\pi} \int_0^{2\pi} \sum_{k,n=0}^{N} d_k \bar{d}_n e^{-i(n-k)\theta} f(\theta)\, d\theta = \sum_{k=0}^{N} \sum_{n=0}^{N} d_k \bar{d}_n a_{k,n},$$

where $a_{n,k} = \dfrac{1}{2\pi} \int_0^{2\pi} e^{-in\theta} e^{ik\theta} f(\theta)\, d\theta.$

Let $S_{M,N} = \sum_{k=0}^{M} \sum_{n=0}^{N} d_k \bar{d}_n a_{k,n}$. $S_{M,N}$ converges as $M$ and $N \to \infty$, because

$$\sum_{k=0}^{\infty} \sum_{n=0}^{\infty} | d_k \bar{d}_n a_{n,k} | \leq \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} | d_k | \, | d_n | \frac{1}{2\pi} \int_0^{2\pi} | e^{ik\theta} f(\theta) | \, d\theta$$

$$\leq \frac{1}{2\pi} \int_0^{2\pi} | f(\theta) | \, d\theta \, (\sum_{k=0}^{\infty} | d_k |)(\sum_{n=0}^{\infty} | d_n |) < \infty.$$

Hence the sum $\sum_{k,n=0}^{N} d_k \bar{d}_n a_{n,k}$ has a limit, as $N \to \infty$, and its limit is $\sum_{k=0}^{\infty} d_k (\sum_{n=0}^{\infty} \bar{d}_n a_{n,k})$. The inside sum

$$\sum_{n=0}^{\infty} \bar{d}_n a_{n,k} = \frac{1}{2\pi} \int_0^{2\pi} [\bar{\varphi}]^{-1}(\theta) e^{ik\theta} f(\theta) \, d\theta = \frac{1}{2\pi} \int_0^{2\pi} e^{ik\theta} \bar{\varphi}(\theta) \, d\theta = \bar{c}_k.$$

Therefore,

$$\sum_{k,n=0}^{\infty} d_k \bar{d}_n a_{n,k} = \sum_{n=0}^{\infty} d_k \bar{c}_k = \frac{1}{2\pi} \int_0^{2\pi} \varphi^{-1}(\theta) \varphi(\theta) \, d\theta = 2\pi.$$

Hence

$$\lim_{N \to \infty} \int_0^{2\pi} | [\varphi^{-1}]^N(\theta) - \varphi^{-1}(\theta) |^2 f(\theta) \, d\theta = 2\pi - 2\pi - 2\pi + 2\pi = 0,$$

which completes the proof.

THEOREM (3.3). *With the notation of Theorem* 2.1 *we have that* $P_{-1}X_0 = \sum_{k=1}^{\infty} a_k X_{-k}$ *with* $\sum_{k=1}^{\infty} | a_k |^2 < \infty$ *if and only if Assumption* 3.1 *is satisfied. When the above is the case,* $a_k = -c_0 d_k$.

*Proof.* Since $\sigma Y_0 = X_0 - P_{-1}X_0$, we have

$$\text{(A)} \qquad \sigma^{-1}(X_0 - P_{-1}X_0) = \sum_{k=0}^{\infty} a_k X_{-k}$$

with $\sum_{k=0}^{\infty} | a_k |^2 < \infty$, if and only if Assumption 3.1 is satisfied; and then $a_k = d_k$ (cf. Theorem 2.1). But $d_0 = 1/c_0 = 1/\sigma$, so in this case (A) is equivalent to $d_0(X_0 - P_{-1}X_0) = \sum_{k=0}^{\infty} d_k X_{-k}$. Thus

$$P_{-1}X_0 = -c_0 \sum_{k=1}^{\infty} d_k X_{-k}$$

if and only if Assumption 3.1 is satisfied; proving the theorem.

We are ready to show that the algorithm given by Masani for $\hat{X}_n = P_0 X_n$, $n \geq 1$, is valid under our Assumption 3.1. Masani was able to establish his result under the boundedness of the spectral density $f$.

THEOREM (3.4). *With the notation of Theorem* 2.1 *under the Assumption* 3.1 *with* $\hat{X}_n = P_0 X_n$, $n \geq 1$, *we have*

$$\hat{X}_n = \sum_{k=0}^{\infty} d_{n,k} X_{-k}, \; d_{n,k} = -(c_0 d_{n+k} + \cdots + c_{n-1} d_{k+1})$$

$$= \sum_{k=0}^{\infty} e_{n,k} X_{-k}, \; e_{n,k} = (c_n d_k + \cdots + c_{n+k} d_0).$$

*Proof.* For the sake of the proof we will put $d_k = 0$ if $k \leq -1$. For $N \geq n - 1$

we note that

$$[(\sum_{k=0}^{n-1} c_k e^{ik\cdot})\varphi^{-1}]^N = \sum_{k=0}^{n-1} c_k e^{ik\cdot}(\varphi^{-1})^{N-k}.$$

Therefore

$$\begin{aligned} &|[(\sum_{k=0}^{n-1} c_k e^{ik\cdot})\varphi^{-1}]^N - (\sum_{k=0}^{n-1} c_k e^{ik\cdot})\varphi^{-1}|^2 \\ &= |\sum_{k=0}^{n-1} c_k e^{ik\cdot}(\varphi^{-1})^{N-k} - (\sum_{k=0}^{n-1} c_k e^{ik\cdot})\varphi^{-1}|^2 \\ &= |\sum_{k=0}^{n-1} c_k e^{ik\cdot}[(\varphi^{-1})^{N-k} - \varphi^{-1}]|^2 \\ &\le \sum_{k=0}^{n-1} |c_k|^2 \sum_{k=0}^{n-1} |(\varphi^{-1})^{N-k}|^2. \end{aligned}$$

Hence

$$\frac{1}{2\pi}\int_0^{2\pi} |[(\sum_{k=0}^{n-1} c_k e^{ik\cdot})\varphi^{-1}]^N - (\sum_{k=0}^{n-1} c_k e^{ik\cdot})\varphi^{-1}|^2 f\, d\theta$$

$$\le \sum_{k=0}^{n-1} |c_k|^2 \sum_{k=0}^{n-1} \frac{1}{2\pi}\int_0^{2\pi} |(\varphi^{-1})^{N-k} - \varphi^{-1}|^2 f\, d\theta.$$

But by Assumption 3.1, for each $k$, $0 \le k \le n-1$, the term

$$\frac{1}{2\pi}\int_0^{2\pi} |(\varphi^{-1})^{N-k} - \varphi^{-1}|^2 f\, d\theta \to 0; \qquad \text{as } N \to \infty.$$

It follows that $[(\sum_{k=0}^{n-1} c_k e^{ik\cdot})\varphi^{-1}]^N$ converges to $(\sum_{k=0}^{n-1} c_k e^{ik\cdot})\varphi^{-1}$ in $L_{2,f}$. But the $k$th Fourier coefficient of $(\sum_{k=0}^{n-1} c_k e^{ik\cdot})\varphi^{-1}$ is $\sum_{\ell=0}^{n-1} c_\ell d_{k-\ell}$. Therefore by isomorphism map between $\sigma(X)$ and $L_{2,f}$ we conclude that $\sum_{k=0}^{\infty} (\sum_{\ell=0}^{n-1} c_\ell d_{k-\ell})X_{-k}$ converges in $\mathscr{H}$ and corresponds to $(\sum_{k=0}^{n-1} c_k e^{ik\cdot})\varphi^{-1}$. Now since $0 = \int_0^{2\pi}(\sum_{k=0}^{n-1} c_k e^{ik\cdot})\bar{\varphi}e^{-i\ell\cdot}\, d\theta$ for $\ell \ge$ n, we have

$$0 = \int_0^{2\pi} \left[\frac{\sum_{k=0}^{n-1} c_k e^{ik\cdot}}{\varphi}\right]\varphi\bar{\varphi}e^{-i\ell\cdot}\, d\theta = \int_0^{2\pi} \left[\frac{\sum_{k=0}^{n-1} c_k e^{ik\cdot}}{\varphi}\right] f e^{-i\ell\cdot}\, d\theta.$$

Hence $(\sum_{k=0}^{\infty} \sum_{\ell=0}^{n-1}) c_\ell d_{k-\ell} X_{-k} \perp X_{-i}$ in $\mathscr{H}$ for $i \ge n$. ($\perp$ stands for orthogonality). That is

$$X_0 + \sum_{k=1}^{\infty} (\sum_{\ell=0}^{n-1} c_\ell d_{k-\ell})X_{-k} \perp X_{-i}, \qquad i \ge n.$$

This implies

$$X_0 + \sum_{k=n}^{\infty} (\sum_{\ell=0}^{n-1} c_\ell d_{k-\ell})X_{-k} \perp X_{-i}, \qquad i \ge n,$$

because if $1 \le k \le n-1$ then in

$$\sum_{\ell=0}^{n-1} c_\ell d_{k-\ell} = [c_0 d_k + c_1 d_{k-1} + \cdots + c_k d_0] + [c_{k+1} d_{-1} + \cdots + c_{n-1} d_{k-(n-1)}]$$

the second bracket is zero since $0 = d_{-1} = \cdots = d_{k-(n-1)}$ and the first bracket is zero since it is the $k$th Fourier coefficient of $\varphi \cdot 1/\varphi = 1$ which is zero. Hence

$$X_0 - [-\sum_{k=n}^{\infty} (\sum_{\ell=0}^{n-1} c_\ell d_{k-\ell})X_{-k}] \perp X_{-i} i \ge n.$$

That is to say $\hat{X}_n = -\sum_{k=n} (\sum_{\ell=0} c_\ell d_{k-\ell})X_{n-k}$ which after a change of variables becomes $\hat{X}_n = -\sum_{k=0}^{\infty} (\sum_{\ell=0}^{n-1} c_\ell d_{n+k-\ell})X_{-k}$. Since $\varphi \cdot 1/\varphi = 1$ we have $0 = \sum_{\ell=0}^{k} c_\ell d_{k-\ell}$ (except for $k = 0$, $c_0 d_0 = 1$). Hence if $n + k \neq 0$ then $\sum_{\ell=0}^{n+k} c_\ell d_{n+k-\ell} = 0$ implying that $-\sum_{\ell=0}^{n-1} c_\ell d_{n+k-\ell} = \sum_{\ell=n}^{n+k} c_\ell d_{n+k-\ell}$.

This completes the proof.

## 4. Recursive expression for the linear predictor of lag *n*.

In this section we will obtain two additional expressions for computing the linear predictor $\hat{X}_n$. These expressions are partly recursive and may shed some additional light on the problem. For any pair of integers $m$ and $n$, let us write $\hat{X}(n/m)$ for the best linear predictor of $X_n$ based on $X_k$, $k \leq m$, and let $\check{X}(n/m) = X_n - \hat{X}(n/m)$. We will assume that Assumption 3.1 is satisfied, and we will write $a_{k+1}$ for $e_{1,k} = d_{1,k}$ occurring in Theorem 3.4.

THEOREM (4.1)
(a) $\hat{X}(n + m + 1/n) = \sum_{k=0}^{m-1} a_{k+1}\hat{X}(n + m - k/n) + \sum_{k=m}^{\infty} a_{k+1}X_{n-k+m}$ *for* $m \geq 1$.

(b) *Let* $Q_0 = 0$, $Q_1 = 1, \ldots, Q_n = a_1Q_{n-1} + a_2Q_{n-2} + \cdots + a_{n-1}Q_1$. *Then for* $m \geq 1$, $\hat{X}(n + m/n) = \sum_{k=0}^{\infty} (Q_m a_{k+1} + \cdots + Q_1 a_{k+m})X_{n-k}$.

(c) *With* $Q_0, Q_1, \ldots, Q_n, \ldots$ *as above we have* $\hat{X}(n/m) = \hat{X}(n/0) + \sum_{k=0}^{m-1} Q_{m-k}\check{X}(k + 1/k)$, $0 \leq m \leq n$.

The last relation in (c) tells us how to update our linear predictor $\hat{X}_n$ in terms of the new innovations as more observations become available.

*Proof.*

(a) Since by Theorem 3.3, $\hat{X}(n + m + 1/n + m) = \sum_{k=0}^{\infty} a_{k+1}X_{n+m-k}$ we therefore have

$$X_{n+m+1} - \sum_{k=0}^{m-1} a_{k+1}\hat{X}(n + m - k/n) - \sum_{k=m}^{\infty} a_{k+1}X_{n+m-k}$$
$$= [X_{n+m+1} - \hat{X}(n + m + 1/n + m)] + \sum_{k=0}^{m-1} a_{k+1}[X_{n+m-k} - \hat{X}(n + m - k/n)].$$

Since for each $-1 \leq k \leq m - 1$, $X_{n+m-k} - \hat{X}(n + m - k/n)$ is orthogonal to $X_\ell$, $\ell \leq n$ and $\sum_{k=0}^{m-1} a_{k+1}\hat{X}(n + m - k/n) - \sum_{k=m}^{\infty} a_{k+1}X_{n+m-k}$ is in $\sigma(X_\ell, \ell \leq n)$ we conclude that

$$\hat{X}(n + m + 1/n) = \sum_{k=0}^{m-1} a_{k+1}\hat{X}(n + m - k/n) + \sum_{k=m}^{\infty} a_{k+1}X_{n+m-k}$$

proving (a).

(b) The proof is given by induction. For $m = 1$, (b) reduces to Theorem 3.4. Assume that for each $j$, $1 \leq j \leq m$ we have

$$\hat{X}(n + j/n) = \sum_{\ell=0}^{\infty} (Q_j a_{\ell+1} + \cdots + Q_1 a_{\ell+j})X_{n-\ell}.$$

Then for each $k$, $0 \leq k \leq m - 1$,

$$\hat{X}(n + m - k/n) = \sum_{\ell=0}^{\infty} (Q_{m-k}a_{\ell+1} + \cdots + Q_1 a_{m-k+\ell})X_{n-\ell}$$

Since by (a)

$$\hat{X}(n + m + 1/n) = \textstyle\sum_{k=0}^{m-1} a_{k+1}\hat{X}(n + m - k/n) + \sum_{k=m}^{\infty} a_{k+1}X_{n+m-k}$$

it follows that

$$\begin{aligned}\hat{X}(n + m + 1/n) &= \textstyle\sum_{k=0}^{m-1} a_{k+1} \sum_{\ell=0}^{\infty} (Q_{m-k}a_{\ell+1} + \cdots + Q_1 a_{m-k+\ell})X_{n-\ell} \\ &\quad + \textstyle\sum_{\ell=m}^{\infty} a_{\ell+1}X_{n+m-\ell} \\ &= a_1[\textstyle\sum_{\ell=0}^{\infty} (Q_m a_{\ell+1} + \cdots + Q_1 a_{m+\ell})X_{n-\ell}] \\ &\quad + a_2[\textstyle\sum_{\ell=0}^{\infty} (Q_{m-1}a_{\ell+1} + \cdots + Q_1 a_{m+\ell-1})X_{n-\ell}] \\ &\quad \vdots \\ &\quad + a_{m-1}[\textstyle\sum_{\ell=0}^{\infty} (Q_2 a_{\ell+1} + Q_1 a_{\ell+2})X_{n-\ell}] \\ &\quad + a_m[\textstyle\sum_{\ell=0}^{\infty} (Q_1 a_{\ell+1})X_{n-\ell}] + \sum_{\ell=m}^{\infty} a_{\ell+1}X_{n+m-\ell} \\ &= \textstyle\sum_{\ell=0}^{\infty} (Q_{m+1}a_{\ell+1} + \cdots + Q_2 a_{m+\ell})X_{n-\ell} + \sum_{\ell=m}^{\infty} a_{\ell+1}X_{n+m-\ell}\end{aligned}$$

But

$$\textstyle\sum_{\ell=0}^{\infty} Q_1 a_{m+\ell+1}X_{n-\ell} = \sum_{\ell=m}^{\infty} a_{\ell+1}X_{n+m-\ell}$$

so that

$$\hat{X}(n + m + 1/n) = \textstyle\sum_{\ell=0}^{\infty} (Q_{m+1}a_{\ell+1} + \cdots + Q_1 a_{\ell+m+1})X_{n-\ell}.$$

Therefore by induction we can complete the proof.

(c) We have seen that

$$\begin{aligned}\hat{X}(n + m/n) &= \textstyle\sum_{k=0}^{\infty} (Q_m a_{k+1} + \cdots + Q_1 a_{k+m})X_{n-k} \\ &= \textstyle\sum_{k=0}^{\infty} Q_m a_{k+1}X_{n-k} + \sum_{k=0}^{\infty} (Q_{m-1}a_{k+2} + \cdots + Q_1 a_{k+m})X_{n-k}.\end{aligned}$$

Therefore

$$\begin{aligned}\hat{X}(n + m/n + 1) &= \hat{X}((n + 1) + (m - 1)/n + 1) \\ &= \textstyle\sum_{k=0}^{\infty} (Q_{m-1}a_{k+1} + \cdots + Q_1 a_{k+m-1})X_{n+1-k} \\ &= (Q_{m-1}a_1 + \cdots + Q_1 a_{m-1})X_{n+1} + \textstyle\sum_{k=1}^{\infty} (Q_{m-1}a_{k+1} + \cdots + Q_1 a_{k+m-1})X_{n+1-k}.\end{aligned}$$

It follows that

$$\hat{X}(n + m/n + 1) - (Q_{m-1}a_1 + \cdots + Q_1 a_{m-1})X_{n+1} = \hat{X}(n + m/n) - \textstyle\sum_{k=0}^{\infty} Q_m a_{k+1}X_{n-k}.$$

This means that

$$\begin{aligned}\hat{X}(n + m/n + 1) &= \hat{X}(n + m/n) - Q_m \textstyle\sum_{k=0}^{\infty} a_{k+1}X_{n-k} + Q_m X_{n+1} \\ &= \hat{X}(n + m/n) + Q_m(X_{n+1} - \hat{X}(n + 1/n)) \\ &= \hat{X}(n + m/n) + Q_m\check{X}(n + 1/n).\end{aligned}$$

Continuing in this manner we get

$$\hat{X}(n+m/n+2) = \hat{X}(n+m/n+1) + Q_{m-1}\check{X}(n+2/n+1)$$
$$= \hat{X}(n+m/n) + Q_m\check{X}(n+1/n) + Q_{m-1}\check{X}(n+2/n+1).$$

Hence for $0 \le m \le n$ we get

$$\hat{X}(n+m/n+\ell) = \hat{X}(n+m/n) + \textstyle\sum_{k=0}^{\ell-1} Q_{m-k}\check{X}(n+k+1/n+k).$$

That is

$$\hat{X}(n/m) = \hat{X}(n/0) + \textstyle\sum_{k=0}^{m-1} Q_{m-k}\check{X}(k+1/k),\ 0 \le m \le n.$$

This completes the proof.

## 5. An example and further discussion

*Example* (5.1). In the following we give a class of unbounded spectral densities where our Assumption 3.1 is satisfied. Because of unboundedness of these densities Masani's result [4, 1960] cannot be invoked to obtain an autoregressive series representation for the linear predictor. For $\frac{1}{4} \le \lambda < \frac{1}{2}$ these densities are not even in $L_2$. However by alluding to our Theorem 3.4 we can give an explicit representation for $\hat{X}_n$. The class of spectral densities is as follows: Let $0 < \lambda < \frac{1}{2}$, and $f(\theta) = |1 - e^{i\theta}|^{-2\lambda}$. Obviously $f$ is not bounded. The optimal factor of $f$ is $\varphi(\theta) = (1 - e^{-i\theta})^{-\lambda}$. Let $(a)_k$ stand for $a(a+1)\cdots(a+k-1)$ and $(a)_0 = 1$. Then

$$\varphi(\theta) = \textstyle\sum_{k=0}^{\infty} \dfrac{(\lambda)_k}{k!} e^{ik\theta} \quad \text{in} \quad L_2$$

$$\varphi^{-1}(\theta) = \textstyle\sum_{k=0}^{\infty} \dfrac{(-\lambda)_k}{k!} e^{ik\theta} \quad \text{in} \quad L_2.$$

It is known that $\sum_{k=0}^{\infty} \left|\dfrac{(-\lambda)_k}{k!}\right| < \infty$ (c.f. [2] p. 1), hence by Proposition 3.2 our Assumption 3.1 is satisfied. However to obtain a rate at which $(\varphi^{-1})^N - \varphi^{-1}$ converges to zero in $L_{2,f}$, as $N \to \infty$, we proceed as follows. Let $N$ be any nonnegative integer. Simple calculations show that

$$\frac{1}{2\pi} \int_0^{2\pi} [(\varphi^{-1})^N\varphi - 1](\theta)e^{i\ell\theta}\, d\theta$$

$$= \begin{cases} 0, & \ell \le n \\ \sum_{k=0}^{N} \dfrac{(-\lambda)_k}{k!} \dfrac{(\lambda)_{N+j-k}}{(N+j-k)!}, & \ell = j + N,\ j \ge 1. \end{cases}$$

Note that for $j \ge 1$,

$$(\lambda)_{N+j-k} = \frac{(-1)^k(\lambda)_{N+j}}{(1-\lambda-N-j)_k} \quad \text{and} \quad (-N-j)_k = (-1)^k \frac{(N+j)!}{(N+j-k)!}.$$

Therefore for $j \geq 1$ we have

$$\sum_{k=0}^{N} \frac{(-\lambda)_k}{k!} \frac{(\lambda)_{N+j-k}}{(N+j-k)!} = \sum_{k=0}^{N} \frac{(-\lambda)_k(\lambda)_{N+j}(-N-j)_k(-N)_k}{k!(N+j)!(1-\lambda-N-j)_k(-N)_k}.$$

By applying Saalschütz's theorem [2, p. 9] to

$$\sum_{k=0}^{N} \frac{(\lambda)_{N+j}(-\lambda)_k(-N-j)_k(-N)_k}{(N+j)!k!(1-\lambda-N-j)_k(-N)_k}$$

we have that for

$$j \geq 1, \sum_{k=0}^{N} \frac{(-\lambda)_k}{k!} \frac{(\lambda)_{N+j-K}}{(N+j-k)!} = \frac{(\lambda)_{N+j}}{(N+j)!} \frac{(\lambda)_N(j)_N}{(-N)_N(j+\lambda)_N}$$
$$= \frac{(\lambda-N)_N}{(-N)_N} \frac{(\lambda)_j}{\Gamma(j)(N+j)}.$$

Therefore

$$\frac{1}{2\pi} \int_0^{2\pi} |(\varphi^{-1})^N\varphi - 1|^2 \, d\theta = \sum_{j=0}^{\infty} \left| \sum_{k=1}^{N} \frac{(-\lambda)_k}{k!} \frac{(\lambda)_{N+j-k}}{(N+j-k)!} \right|^2 =$$
$$\left| \frac{(\lambda-N)_N}{(-N)_N} \right|^2 \sum_{j=1}^{\infty} \left| \frac{(\lambda)_j}{\Gamma(j)(N+j)} \right|^2$$
$$= \left| \frac{\lambda(1-\lambda)_N}{N!} \right|^2 \sum_{j=0}^{\infty} \left| \frac{(\lambda+1)_j}{j!} \right|^2 \frac{1}{(N+j+1)^2}.$$

Using Stirling's formula we can show that

$$\sum_{j=0}^{\infty} \left| \frac{(\lambda+1)_j}{j!} \right|^2 \frac{1}{(N+j+1)^2} = o\left(\frac{1}{N^{1-2\lambda}}\right) \quad \text{and} \quad \frac{(1-\lambda)_N}{N} = o\left(\frac{1}{N^{\lambda}}\right).$$

Hence we conclude that $\frac{1}{2\pi} \int_0^{2\pi} |(\varphi^{-1})^N\varphi^{-1}|^2 \, d\theta = o(N^{-1})$. This shows that our Assumption 3.1 is satisfied, and in addition it provides us with a rate of convergence.

In this case by Theorem 3.4 we have

$$e_{n,k} = \sum_{j=0}^{k} \frac{(-\lambda)_j}{j!} \frac{(\lambda)_{n+k-j}}{(n+k-j)!}.$$

Again by applying Saalschütz's theorem as above we have

$$\sum_{j=0}^{k} \frac{(-\lambda)_j}{j!} \frac{(\lambda)_{n+k-j}}{(n+k-j)!} = \frac{(\lambda-k)_k}{(-k)} \frac{(\lambda)_n}{\Gamma(n)(k+n)}$$

which by a simple calculation is equal to

$$\frac{(1-\lambda)_k}{k!} \frac{(\lambda)_n}{(k+n)(n-1)!} = \frac{\Gamma(1-\lambda+k)}{\Gamma(\lambda)\Gamma(1-\lambda)\Gamma(1+k)} \frac{\Gamma(\lambda+n)}{(k+n)\Gamma(n)}.$$

$\left(\text{For } n = 1 \text{ this is just } \frac{\lambda(1-\lambda)_k}{(k+1)!} = \frac{\lambda}{(k+1)!}\frac{\Gamma(1-\lambda+k)}{\Gamma(\lambda)\Gamma(1-\lambda)}\right)$. Hence

$$\hat{X}_n = P_0X_n = \frac{(\lambda)_n}{(n-1)!}\sum_{k=0}^{\infty}\frac{(1-\lambda)_k}{(k+n)k!}X_{-k}.$$

In this case

$$\| X_n - \hat{X}_n \|^2 = \sum_{k=0}^{n-1}\left(\frac{(\lambda)_k}{k!}\right)^2 \longrightarrow \frac{\Gamma(1-2\lambda)}{[\Gamma(1-\lambda)]^2},$$

as $n \longrightarrow \infty$, as one expects.

*Remark* (5.2). In the discussion of linear predictor we have restricted ourselves to the case where $\hat{X}_1 = P_0X_1$ may have a representation in the form $P_0X_1 = \sum_{k=0}^{\infty} a_kX_{-k}$ with $\sum_{k=0}^{\infty} | a_k |^2 < \infty$. The general question when $P_0X_1$ has such a representation without the condition $\sum_{k=0}^{\infty} | a_k |^2 < \infty$ remains open. In case $f^{-1}$ is not in $L_1$ there may not exist any sequence $a_k$, $k \geq 0$, such that $\sum_{k=0}^{\infty} a_kX_{-k} = P_0X_1$. This, for example, is the case for a process with spectral density $f(\theta) = | 1 - e^{i\theta} |^2$, as shown by Topsøe in [7, 1977]. Other conditions such as $f^{-1/2}$ being in $L_1$ may be relevant here. As a specific example let $f(\theta) = | 1 - e^{i\theta} |^{1+\lambda}$, $0 \leq \lambda < 1$. Could the linear predictor for a process with this spectral density have the representation $P_0X_1 = \sum_{k=0}^{\infty} a_kX_{-k}$ for a sequence $a_k$, $k \geq 0$? Another question worthy of study is the following: Let $\varphi$ the optimal factor of $f$ and $\varphi^{-1}$ be in $L_2$. Then, does $(\varphi^{-1})^N$ converge to $\varphi^{-1}$ in $L_{2,f}$? A positive answer to this question will reduce our Assumption 3.1 to merely $f^{-1} \in L_1$. In case the answer is negative, we may ask whether conditions such as $f^{-1} \in L^1$ and $f \in L^2$ would imply that $(\varphi^{-1})^N$ converges to $\varphi^{-1}$ in $L_{2,f}$. Recently M. Pourahmadi [6] has studied the representation $P_0X_1 = \sum_{k=0}^{\infty} a_kX_{-k}$ in connection with the angle between the past and future of the process $X_k$, $K \in Z$. His main contribution seems to be that if the past and future subspaces of $X_k$, $K \in Z$, are at positive angle, then Theorem 3.4 holds. Analytic condition on $f$ under which the angle between the past and future is positive is characterized in [3]. The relationship between the analytic condition for the positivity of the angle and conditions $f^{-1} \in L$ & $f \in L_\infty$, etc. deserves further investigation. Recursive algorithm for the determination of $c_k$, $k \geq 0$ and $d_k$, $k \geq 0$, the Taylor coefficients of $\varphi$ and $\varphi^{-1}$ respectively are derived in terms of the Fourier coefficients of $\log f$ in [5]. However the question of representing $P_0X_1$ as a convergent series $\sum_{k=0}^{\infty} a_kX_{-k}$ is not exploited in [5].

ISFAHAN UNIVERSITY OF TECHNOLOGY
ISFAHAN, IRAN

MICHIGAN STATE UNIVERSITY
EAST LANSING, MICHIGAN, U.S.A.

REFERENCES

[1] E. J. AKUTOWICZ, *On an explicit formula in linear least squares prediction.* Math. Scand. **5** (1957), 261–266.

[2] W. N. BAILEY, Generalized hypergeometric series. Stechert-Hafner Service Agency Inc., New York, 1964.

[3] H. HELSON AND D. SARASON, *Past and future*, Math. Scand. **21**(1967), 5–16.

[4] P. MASANI, *The prediction theory of multivariate stochastic processes*, III. Acta Math. **104**(1960), 141–162.

[5] Z. A. PIRANASHVILI, N. P. ASLANIDI AND T. G. SULABRIDZE, *Explicit formulas of optimal linear prognosis and its mean square error for a stationary sequence* (Russian). Bull. Acad. Sci. Georgian SSR, **93**(1979), 37–40.

[6] M. POURAHMADI, *Multivariate prediction theory and matricial Helson-Szegö Theorem.* To appear.

[7] E. TOPSØE, *Banach algebra methods in prediction theory*, Manuscripta Math., **23**(1977), 19–25.

[8] N. WIENER AND P. MASANI. *The prediction theory of multivariate stochastic processes*, II. Acta Math. **99**(1958), 93–137.