PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 91, Number 2, June 1984

THE HELSON-SARASON-SZEGO THEOREM AND THE ABEL SUMMABILITY OF THE SERIES FOR THE PREDICTOR

MOHSEN POURAHMADI[1]

ABSTRACT. It is shown that the best linear least squares predictor of a stationary stochastic process has a mean Abel summable series representation in the time domain if its density satisfies the condition of the Helson-Sarason-Szego theorem. This provides an answer to an open question of Wiener and Masani (1958) in prediction theory.

1. Introduction. Let H denote the Hilbert space of complex-valued random variables with zero expectations and finite variances. The inner product in H is defined by $(X, Y) = EX\overline{Y}, X, Y \in H$.

An important problem in the theory of linear least squares prediction of a purely nondeterministic weakly stationary stochastic process (WSSP) $\{X_n\}_{n=-\infty}^{+\infty}$ in H is to find conditions on its spectral density w such that the best linear least squares predictor of X_ν ($\nu \geq 1$), denoted by $\hat{X}_\nu$, based on the past and present observations, i.e. X_k, $k \leq 0$, can be written as $\hat{X}_\nu = \sum_{k=0}^{\infty} a_k X_{-k} = \lim_{N\to\infty} \sum_{k=0}^{N} a_k X_{-k}$, where $a_k = \sum_{j=0}^{k} c_{\nu+j} d_{k-j}$ with c_k and d_k being the kth Fourier (Taylor) coefficients of ϕ and ϕ^{-1}, respectively, with ϕ being the optimal factor of w, cf. [**1**, **8**]. If such a representation for $\hat{X}_\nu$ exists, then we say that the best linear least squares predictor of the process has a mean-convergent autoregressive series representation in the time domain [**8**].

It was first shown by Wiener and Masani [**8**] in 1958 that if

$$0 < c \leq w \leq d < \infty, \quad \text{a.e. (Leb.)}, \tag{1}$$

then $\hat{X}_\nu$ has a mean-convergent autoregressive series representation. Later, Masani [**5**] weakened the severe boundedness condition in (1) and replaced it by

$$w \in L^\infty \quad \text{and} \quad w^{-1} \in L^1. \tag{2}$$

For $1 \leq p \leq \infty$, L^p (H^p) denotes the usual Lebesgue (Hardy) space of functions on the unit circle in the complex plane and for a density function w, $L^2(w)$ stands for the Hilbert space of measurable functions f on the unit circle with $\int_{-\pi}^{\pi} |f|^2 w \, d\theta < \infty$.

Masani's condition $w \in L^\infty$ is also restrictive. By using a deep result of Helson and Szego [**2**] the author has shown [**6**] that if

$$w = e^{u+\tilde{v}}, \tag{3}$$

Received by the editors December 10, 1982 and, in revised form, September 27, 1983.

1980 *Mathematics Subject Classification*. Primary 60G25; Secondary 42A24.

[1]This research was supported in part by the NSF Grant MCS-8301240

where u and v are bounded real-valued functions with $\|v\|_\infty < \pi/2$, then $\hat{X}_\nu$ has a mean-convergent autoregressive series representation in the time domain. It is easy to check that densities w exist which satisfy (3) but for which $w \notin L^\infty$, cf. [**6**]. This shows that the boundedness condition in (2) can be relaxed.

Both conditions (2) and (3) require that $w^{-1} \in L^1$. Motivated by a yet open question of Wiener and Masani [**8**, p. 123], concerning the summability of the best linear least squares predictor of a purely nondeterministic WSSP, in the following theorem we show that the condition $w^{-1} \in L^1$ can be relaxed and replaced by a weaker condition, if one replaces the requirement of mean-convergence of the linear predictor by the mean Abel summability of the series involved.

2. THEOREM. *Let $\{X_n\}_{n=-\infty}^{+\infty}$ be a purely nondeterministic WSSP with the spectral density function w and $\hat{X}_\nu$ the best linear least squares predictor of X_ν $(\nu \geq 1)$ based on X_k, $k \leq 0$. Then, $\hat{X}_\nu = \lim_{r\to 1} \sum_{k=0}^{\infty} r^k a_k X_{-k}$, if*

$$w = |P|^2 e^{u+\tilde{v}}, \tag{4}$$

where u and v are as in (3) and P is an analytic polynomial of some degree n such that all of its roots lie on the unit circle in the complex plane.

3. Remarks. (a) Densities satisfying (3) and (4) were first studied by Helson and Szego [**2**] and Helson and Sarason [**3**] in connection with the following important problem. Let $H(X,0) = \overline{\text{sp}}\{X_k; k \leq 0\}$ and $H^n(X) = \overline{\text{sp}}\{X_k; k \geq n\}$, $n = 1, 2, \ldots$. As a measure of the cosine of the angle between the two subspaces $H(X,0)$ and $H^n(x)$ in H define $\rho_n = \sup|(X,Y)|$, where the supremum is taken over all X, Y, as elements of the unit balls of $H(X,0)$ and $H^n(X)$, respectively. It is easy to see that $0 \leq \rho_n \leq 1$. For fixed $n \geq 1$, it is said that $H(X,0)$ and $H^n(X)$ are at *positive angle* if $p_n < 1$. The problem of characterizing WSSP's or densities w such that $\rho_n < 1$ is solved in [**2**] for $n = 1$ and in [**3**] for $n > 1$. The results of these authors, known as the Helson-Sarason-Szego theorem, state that $\rho_n < 1$, if and only if $w = |P|^2 e^{u+\tilde{v}}$, where P is an analytic polynomial of degree $n-1$ and u, v are as before. It is easy to check that a w satisfying (4) does not necessarily have the property $w^{-1} \in L^1$.

(b) Densities w satisfying (4) have the property that $w \in L^1$ and $|P|^2/w \in L^1$. It is of interest to know whether the converse is true. A positive answer will provide a condition weaker than (4) for the Abel summability of the linear predictor. Also it is of interest to know whether the mere assumption of pure nondeterminism, that is $w \in L^1$ and $\log w \in L^1$, will imply the Abel summability of the linear predictor.

Proof of Theorem 2 follows from the next three lemmas.

4. LEMMA. *Let $\{X_n\}, w$ and $\hat{X}_\nu$ be as in Theorem 2, ϕ the optimal factor of w, ϕ_ν the spectral isomorph of $\hat{X}_\nu$, $H^2(w) = \overline{\text{sp}}\{e^{in\theta}; n \geq 0\}$ in $L^2(w)$, $P_r(\theta) = (1-r^2)/(1 - 2r\cos\theta + r^2)$, and $\phi_r(\theta) = (P_r * \phi)(\theta)$, $0 \leq r < 1$. Assume that*

$$\int_{-\pi}^{\pi} P_r(\theta - x) \left| \frac{\phi(x)}{\phi_r(x)} \right|^2 dx \leq C \tag{5}$$

for all $0 \leq r < 1$ and $-\pi \leq \theta \leq \pi$, where C is constant. Then:

- (a) *ϕ_ν is of Nevanlina class and so has Fourier (Taylor) coefficients a_k, $k \geq 0$ (for the definition of a_k's see §1).*
- (b) *The series for the linear predictor of X_ν is mean Abel summable, if and only if $\lim_{r\to 1} \|\sum_{k=0}^{\infty} r^k a_k e^{ik\theta} - \phi_\nu\|_{L^2(w)} = 0$.*

PROOF. Since ϕ_ν is the spectral isomorph of $\hat{X}_\nu$, it belongs to $H^2(w)$ and $\phi_\nu \phi \in H^2$. Hence ϕ_ν is of Nevanlina class [**9**, p. 271] and (a) follows. A similar argument shows that all functions in $H^2(w)$ are of Nevanlina class, this fact is needed in Lemma 5. (b) is an immediate consequence of the isomorphism between the time and spectral domains of WSSP; cf. [**1**, p. 563].

The following lemma is a slightly different version of Theorem 1 of Rosenblum [**7**].

5. LEMMA. *The Fourier series of all functions in $H^2(w)$ are Abel summable in the norm of $L^2(w)$, if and only if ϕ (ϕ is the optimal factor of w) satisfies* (5).

6. LEMMA. (a) *If $w_0 = e^{u+\tilde{v}}$, with u and v as in* (3), *then w_0 satisfies* (5).

(b) *If w_0 is as in* (a) *and $w = |Q|^2 w_0$, where Q is a polynomial of some degree n, then w satisfies* (5).

PROOF. (a) It follows from [**2**, p. 131] that if $w_0 = e^{u+\tilde{v}}$, then the Fourier series of all functions in $H^2(w_0)$ are convergent in the norm of $L^2(w_0)$. Thus, from the regularity of the Abel summability method, it follows that the Fourier series of all functions in $H^2(w_0)$ are Abel summable in the norm of $L^2(w_0)$. That $w_0 = |\phi|^2$ (ϕ being the optimal factor of w_0) satisfies (5) is a consequence of Lemma 5. Therefore, there exists a constant C_1 such that $\int_{-\pi}^{\pi} P_r(\theta - r)|\phi(x)/\phi_r(x)|^2\,dx \le C_1$, for $0 \le r < 1$ and $-\pi \le \theta \le \pi$.

(b) By an application of the Fejér-Riesz theorem, it can be assumed that the polynomial Q is an analytic polynomial of degree n with no zero inside the open unit disc. Then it is easy to show that with $Q_r = P_r * Q$, $|Q(\theta)/Q_r(\theta)| \le C_2$, for some constant C_2 and all $0 \le r < 1$ and $-\pi \le \theta \le \pi$. Since $w = |Q|^2 w_0 = |Q\phi|^2$ and both Q and ϕ are analytic functions, it follows that $P_r * (Q\phi) = Q_r \phi_r$ and $\int_{-\pi}^{\pi} P_r(\theta - x)|Q\phi(x)/Q_r\phi_r|^2\,dx \le C_1 C_2^2$, for all $0 \le r < 1$ and $-\pi \le \theta \le \pi$.

REFERENCES

1. J. L. Doob, *Stochastic processes*, Wiley, New York, 1953.
2. H. Helson and G. Szego, *A problem in prediction theory,* Ann. Mat. Pura Appl. (4) **51** (1960), 107–138.
3. H. Helson and D. Sarason, *Past and future,* Math. Scand. **21** (1967), 5–16.
4. A. N. Kolmogorov, *Stationary sequences in a Hilbert space*, Bull. Math. Univ. Moscow **2** (1941), 1–40.
5. P. R. Masani, *The prediction theory of multivariate stochastic processes.* III, Acta Math. **104** (1960), 141–162.
6. M. Pourahmadi, *A matricial extension of the Helson-Szego Theorem and its application in multivariate prediction,* J. Multivariate Anal. (to appear).
7. M. Rosenblum, *Summability of Fourier Series in $L^p(d\mu)$,* Trans. Amer. Math. Soc. **105** (1962), 32–42.
8. N. Wiener and P. Masani, *The prediction theory of multivariate stochastic processes*, Acta Math. **99** (1958), 93–137.
9. A. Zygmund, *Trigonometric series.* I, Cambridge Univ. Press, London and New York, 1959.

DEPARTMENT OF MATHEMATICAL SCIENCES, NORTHERN ILLINOIS UNIVERSITY, DEKALB, ILLINOIS 60115

Current address: Center for Stochastic Processes, Department of Statistics, University of North Carolina, Chapel Hill, North Carolina 27514