The Role of Orthogonal Polynomials in Autoregressive Methods for Noise Cancellation in Acoustic Arrays

25 October, 2004

Michael Arciero
University of New England
Biddeford, Maine, 04002

Army Research Laboratory Advisor:
Thyagaraju Damarla
Signals and Electronic Devices Directorate/Acoustics,
Army Research Laboratory, Adelphi, MD

Abstract

The classical Wiener filter lies at the heart of many engineering problems involving the filtering of noise in digital signals. These problems can be interpreted in the context of linear prediction of sequences of random variables; the problem becomes that of minimizing the noise variance. The least mean square (LMS) algorithm is an approach to noise cancellation which is based on the Wiener filter and is used when the noise statistics are either unknown or changing over time, as is usually the case in applications. Given a sequence of random variables, there is an associated sequence of polynomials which are orthogonal with respect to distributions on the unit circle which are related to the noise process. In the above context, the polynomial coefficients of are equal to the weights of a Wiener filter. We describe this connection and state some results about the asymptotic behavior of the polynomials with respect to sequences of distributions.

KEY WORDS: Orthogonal polynomial, noise cancellation.

CONTACT: Dr. Michael Arciero, University of New England Tel: (207) 283 0171 x2439 email: marciero@une.edu

Thyagaraju Damarla, Ph.D., ARL, Adelphi, MD 20783 Tel: (301) 394-1266 email: rdamarla@arl.army.mil ¡mailto:betsyr@arl.mil¿

1 Introduction

A class of adaptive filtering techniques for noise cancellation, the least-mean-square (LMS) and adaptive prediction algorithms, and their variants, is based on the classical Wiener filter [18], which can be viewed as is a linear operator on stationary time series. Adaptive methods are used when the noise statistics are unknown or non-stationary [3, 16].

Suppose that

$$x(n) = \alpha(n) + s(n); \quad n = 1, 2, 3, ..., \tag{1}$$

is a discrete-time stationary signal, where $s(n)$, is a signal of interest and $\alpha(n)$ is a stationary noise process which is uncorrelated with $s(n)$. As discussed in [10], given a noise *reference*, $y(n)$, which is correlated with $\alpha(n)$ but uncorrelated with $s(n)$, the problem of noise cancellation in the signal $x(n)$

can be viewed in the context of linear prediction. Let $\mathbf{w} := [w_1, w_2, ..., w_k]$ be a vector of filter weights and $\mathbf{y}_n := [y(n), y(n-1), ..., y(n-k+1)]$. The linear combination

$$\sum_{j=0}^{k-1} w_{k+1} y(n-j) = \mathbf{w}\mathbf{y}'_n, \tag{2}$$

which gives the best "predictor" of $x(n)$, in the least-squares sense, minimizes the square error

$$\mathcal{E}(s(n)^2) + \mathcal{E}[(\alpha(n) - (\mathbf{w}\mathbf{y}'))^2], \tag{3}$$

where $\mathcal{E}$ is the expectation operator, with $\alpha(n), s(n)$, and $y(n)$ interpreted as a realization of sequences of random variables. The optimal filter weights are given by the Wiener-Hopf equations

$$\mathbf{w}_{opt} = M_{yy}^{-1} P_{xy} \tag{4}$$

where M_{yy}^{-1} denotes the covariance matrix,

$$M_{yy}(i-j) := \mathcal{E}(y(n+i)y(n+j)),$$

and P_{xy} denotes the vector of cross-correlations,

$$P_{xy} = \mathcal{E}([x(n)y(n), x(n)y(n-1), x(n)y(n-2), ..., x(n)y(n-k+1)]).$$

As discussed in [9], the adaptive prediction algorithm, which is effective with broadband noise, uses a variant of (2), where a linear combination of past samples

$$\sum_{j=m}^{m+k-1} w_{j-m+1} x(n-j) \tag{5}$$

is used to predict $x(n)$ for $m \geq 1$. A noise reference is not used in this case; rather, this technique relies on the de-correlation of broadband components over time. In the case where $\alpha(n)$ is gaussian white noise, for example, it is sufficient to let $m = 1$ in (5) [19]. Equations (3) and (4) become

$$\mathcal{E}[(x(n) - \sum_{j=1}^{k} w_j x(n-j))^2], \tag{6}$$

$$\mathbf{w}_{opt} = M_{xx}^{-1} \mathcal{E}([x(n)x(n-1), x(n)x(n-2), ..., x(n)x(n-k)]). \tag{7}$$

If the process $x(n)$ is not completely deterministic, the matrix M_{xx} will be *positive definite*, that is, $\mathbf{a}' M_{xx} \mathbf{a} > 0$ for any vector $\mathbf{a}$ of length k. It follows that, for any k, M_{xx} defines an inner product on the space of vectors of length k:

$$< \mathbf{a}, \mathbf{a} >:= \mathbf{a} M_{xx} \mathbf{a}'.$$

It follows that with the above inner product, a Hilbert space isomorphism exists between the vector space of dimension $k+1$ and the square-integrable functions on the unit circle in the complex plane with respect to a certain distribution $\mu(\theta)$ on $[-\pi, \pi)$. (See, for example, [1] for details.) With $f(z) = a_0 z^k + a_1 z^{k-1} + ... + a_k z + a_k$ and $g(z) = b_0 z^k + b_1 z^{k-1} + ... + b_k z + b_k$ we have

$$< \mathbf{a}, \mathbf{b} >= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\theta}) \overline{g(e^{i\theta})} \, d\mu(\theta). \tag{8}$$

The distribution μ is called the *spectral density* of the process $x(n)$, and contains information about the spectral content. The Fourier coefficients of μ are the elements of the matrix M_{xx}. We have

$$\widehat{\mu}(\ell) := \frac{1}{2\pi} \int e^{-i\ell\theta} d\mu(\theta) = \mathcal{E}(x(n+i)x(n+j)). \tag{9}$$

More generally, we will refer to μ as a *measure*, which will allow for singular components where a density does not exist. For example, if $x(n)$ consists solely of a sinusoidal signal with normalized frequency ω cycles/sample, μ is the point mass, or delta-function at the frequency ω. On the other hand, if $x(n)$ is a white noise, μ is the uniform density on $[-\pi, \pi)$.

In this context, (6) becomes

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|e^{ik\theta}-\sum_{j=1}^{k}w_j e^{ij\theta}\right|^2 d\mu(\theta). \tag{10}$$

It follows that the polynomial which minimizes (10) is $z^k + \sum_{j=1}^{k} w_j z^j$, where the coefficients, w_j are given in (7). This polynomial is called the *Szegö polynomial* of degree k with respect to the measure μ, which we will denote $P_k(z, \mu)$. We define the *prediction error power* as the minimum in(10),

$$\rho_k(\mu) := \frac{1}{2\pi}\int_{-\pi}^{\pi}|P_k(e^{i\theta}, \mu)|^2\, d\mu. \tag{11}$$

Szegö polynomials form the basis for *autoregressive* methods in digital filtering and spectral analysis [3, 6]. For example, the measure μ can be estimated by the density $\frac{1}{2\pi}|P_k(e^{i\theta}, \mu)|^{-2}$. Szegö polynomials are also used as frequency estimates. If $x(n)$ consists of $m < k$ sinusoids with frequencies ω_j, with additive noise, one would expect m of the zeros of $P_k(z, \mu)$ to be close to the $e^{i\theta_j}$, in order that the minimum in (10) is attained.

In [10] an approach is described presented where the frequencies of sinusoidal components of $x(n)$ which result from vehicle self-noise are estimated, then cancelled from the signal. In this paper we present some results about the behavior of the polynomial zeros which are used as estimates, with respect to measures μ_h, which converge in the *weak-star* topology on measures to a sum of point masses,

$$\lim_{h\to 0}\mu_h = \mu_\delta := \sum_{j=1}^{m}\alpha_j\delta_{\theta_j} \quad \text{weak} - \text{star} \tag{12}$$

where $h > 0$ is a discrete or continuous parameter, the α_j are arbitrary complex numbers, and δ_{θ_j} is the point mass at θ_j. Weak-star convergence can be characterized as follows. We have

$$\lim_{h\to 0}\mu_h = \mu \quad \text{weak} - \text{star}$$

if and only if

$$\lim_{h\to 0}\mu_h\int_{-\pi}^{\pi} f(\theta)d\mu_h(\theta) = \int_{-\pi}^{\pi} f(\theta)d\mu(\theta)$$

for all continuous f on $[-\pi, \pi)$. As we shall see, such measures arise naturally when forming estimates from the samples $x(n)$.

In general, the corresponding limit

$$\lim_{h\to 0} P_k(z, \mu_h) \tag{13}$$

need not exist for fixed $k > m$. Indeed, it is shown by example in [8] that even strong convergence of μ_h does not guarantee convergence of $P_k(z, \mu_h)$. On the other hand, any limit point of$\{P_k(z, \mu_h)\}_h$ has m zeros at the $e^{i\theta_j}$. We prove existence of the limit (13) for a general class of measures for which (12) holds, and characterize this limit explicitly for two cases: when μ_h is formed by convolving μ_δ with the Poisson kernel, and when μ_h is formed by adding to μ_δ an absolutely continuous measure.

In the last decade, several papers have appeared which address the asymptotic behavior of Szegö polynomials, with respect to measures for which a special case of (12) holds, in the context of frequency estimation [4, 5, 12, 13, 14, 15]. It is shown in [4] that (12) holds for the periodogram, ψ_N, of N samples

of a real trigonometric signal. Specifically, if $x(n) = \sum_{j=-M}^{M} \beta_j e^{in\theta_j}$, where $\beta_{-j} = \overline{\beta}_j$ and $\theta_j = \theta_{-j}$, and $\psi_N(\theta) = \frac{1}{2\pi N} |\sum_{k=0}^{N-1} x(k) e^{-ik\theta}|^2$, then we have the weak-star limit

$$\lim_{N\to\infty} \psi_N = \sum_{j=-M}^{M} |\beta_j|^2 \delta_{\theta_j}. \tag{14}$$

It is shown by example in [15] that $P_k(z, N\psi_N)$ do not, in general, converge as $N \to \infty$ for $k > 2M+1$. The behavior of the $k-2M-1$ "extraneous zeros", those that do not approach signal locations, has been one area of interest. In [15] it is shown that the extraneous zeros of any limit point of $\{P_k(z, \psi_N)\}_N$ (note that $P_k(z, N\psi_N) = P_k(z, \psi_N)$) lie strictly inside the open unit disk. It would thus seem feasible to distinguish signal zeros from extraneous zeros, for large N, as those of largest modulus[1].

The problem of non-uniqueness of polynomial limits is addressed in [5, 12, 14] by modifying the measures ψ_N of [15], in such a way as to ensure convergence of the $P_k(z, \psi_N)$. In the *R-process* [5] and the *V-process* [14], the ℓth moment, $\frac{1}{N}\widehat{\psi}_N(\ell)$ is multiplied, respectively, by $r^{|\ell|}$ and r^{ℓ^2}, which are recognized as the moments of the Poisson kernel and the *wrapped* Gaussian. Limits are taken, first as $N \to \infty$, then as $r \uparrow 1$. By weak-star convergence (14), we see that $N = \infty, r < 1$ corresponds to convolution of point masses with these two approximate identities. Thus, the characterization of the limit polynomial of the *R-process* will follow as a special case of our results here.

We also briefly consider the behavior as $k \to \infty$ of the constant terms, or *reflection coefficients*

$$R_k(\mu) := P_k(0, \mu), \tag{15}$$

which give information about the region of accumulation of the extraneous zeros. We apply our results and a result in [11] to find $\lim_{k\to\infty} \lim_{h\to 0} P_k(0, \mu_h)$ for the case where $m = 2$ and μ_h is formed by convolution with the Poisson kernel. A result in [13] is seen as a special case.

Szegö polynomials can be characterized by the following orthogonality property.

$$\int_{-\pi}^{\pi} P_k(e^{i\theta}, \mu)\, \overline{p}(e^{i\theta})\, d\mu(\theta) = 0 \tag{16}$$

for any polynomial p of degree less than k. If μ is supported on strictly greater than k points, the zeros of $P_k(z, \mu)$ lie in the open unit disk; this is the *minimum phase* property [6], and is a special case of Fejér's Convex Hull Theorem.

We also have the well-known representation [1]

$$P_k(z, \mu) = \frac{\mathcal{D}_k(z, \mu)}{D_{k-1}(\mu)}. \tag{17}$$

where

$$\mathcal{D}_n(z, \mu) := \begin{vmatrix} \widehat{\mu}(0) & \widehat{\mu}(1) & \dots & \widehat{\mu}(n) \\ \widehat{\mu}(-1) & \widehat{\mu}(0) & \dots & \widehat{\mu}(n-1) \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{\mu}(-n+1) & \widehat{\mu}(-n+2) & \dots & \widehat{\mu}(1) \\ 1 & z & \dots & z^n \end{vmatrix}, \tag{18}$$

and

$$D_n(\mu) := \begin{vmatrix} \widehat{\mu}(0) & \widehat{\mu}(1) & \dots & \widehat{\mu}(n-1) \\ \widehat{\mu}(-1) & \widehat{\mu}(0) & \dots & \widehat{\mu}(n-1) \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{\mu}(-n+1) & \widehat{\mu}(-n+2) & \dots & \widehat{\mu}(0) \end{vmatrix}. \tag{19}$$

[1] If the polynomial sequence has an infinite number of limit points, the zeros of which accumulate near the unit circle, it may not be possible to make this distinction. This possibility has not been excluded in the literature.

The $P_k(z,\mu)$ can also be computed using *Levinson's recursion* [1]:

$$P_{k+1}(z,\mu) = zP_k(z,\mu) + R_k(\mu)P_k^*(z,\mu), \tag{20}$$

where, for a polynomial $p(z)$, of degree k, $p^*(z)$ denotes the *reverse polynomial*: $p^*(z) := z^k\overline{p}(1/z)$. The zeros of p^* are obtained from those of p by reflection in the unit circle.

The proofs of the results in this paper can be found in [8].

2 Preliminary Results

In this section we give results concerning the existence of limits and rates of convergence of Szegö polynomials for a general class of weakly convergent measures. An important similarity which exists between the measures we will consider and the measures ψ_N of (14) is that both give rise to reflection coefficients which are uniformly bounded away from (inside) the unit circle, and subsequently, that extraneous zeros of limit polynomials lie strictly inside the unit circle. In establishing these results here, we adapt some of the arguments of [15].

Suppose that a $h > 0$ is either a continuous or discrete parameter and $\{\mu_h\}$ is a family of absolutely continuous measures on $[-\pi,\pi)$ for which the weak-star convergence (12) holds. Note that an absolutely continuous family cannot converge strongly (i.e., in total variation norm) to a sum of point masses. Both numerator and denominator in the expression for $P_k(z,\mu_\delta)$, as defined by (17) are zero for $k > m$, thus $P_k(z,\mu_\delta)$ is not well defined. On the other hand, it follows from the minimum phase property that the monic family $\{P_k(z,\mu_h)\}$ is uniformly bounded, and thus contains limit points.

Proposition 2.1 *Suppose (12) holds for a family of absolutely continuous measures μ_h. Then for $k \geq m$ all the limit points of $\{P_k(z,\mu_h)\}$, as $h \to 0$, are of the form*

$$Q(z)\prod_{j=1}^{m}(z - e^{i\theta_j}), \tag{21}$$

where Q is a monic polynomial of degree $k-m$ with all zeros in $|z| \leq 1$.

If the reflection coefficients, $R_k(\mu_h)$, are bounded in modulus uniformly away from (inside) the unit circle, it can be shown that the zeros of any limit factor, Q, of Proposition 2.1, will lie strictly inside the open unit disk. This is shown in the proof of Theorem 2.4 of [15] for point masses associated with real signals. The proof is readily adapted for the complex case. See [8] for details. We have the following.

Proposition 2.2 *Suppose that for $k > m$ and $h > 0$ the reflection coefficients for a family of absolutely continuous measures μ_h satisfy*

$$|R_k(\mu_h)| \leq c < 1. \tag{22}$$

for some constant c. Then the zeros of any limit factor Q, in (21), are strictly less than one in modulus.

We wish to consider measures whose moments can be expanded as a power series about $h = 0$. This property, which ensures convergence of $P_k(z,\mu_h)$, holds for the Poisson kernel and the wrapped Gaussian, and thus for the underlying measures of the R- and V- processes of [5, 14].

Proposition 2.3 *Suppose (12) holds for a family of absolutely continuous measures, μ_h and suppose that for each ℓ, $\widehat{\mu}_h(\ell)$ has a power series representation about $h = 0$. Then for $k > m$ there exists a monic polynomial, Q, of degree $k-m$ such that*

$$\lim_{h\to 0} P_k(z,\mu_h) =: \mathcal{P}_k(z) = Q(z)\prod_{j=1}^{m}(z - e^{i\theta_j}). \tag{23}$$

Assuming (23) holds we can write

$$P_k(z,\mu_h) = Q_h(z)\prod_{j=1}^{m}(z - w_j^{(h)}), \tag{24}$$

where, without loss of generality, $w_j^{(h)} \to e^{i\theta_j}$ for $j = 1, 2, ..., m$, and $Q_h(z)$ converges to a monic polynomial $Q(z)$ of degree $k - m$ as $h \to 0$.

The following two results are analogs of equation (2.4) and Corollary 2.3 [15], respectively.

Corollary 2.1 *Suppose that the moments, $\widehat{\mu}_h(\ell)$, of a family of absolutely continuous measures satisfying (12), have power series representations about $h = 0$. Then there exists a constant $c > 0$, such that, for $k > m$*

$$|P_k(z,\mu_h) - \mathcal{P}_k(z)| \le ch, \tag{25}$$

where $\mathcal{P}_k$ is defined in (23).

Corollary 2.2 *Suppose that for a family of absolutely continuous measures satisfying (12) with the θ_j distinct, the $\widehat{\mu}_h(\ell)$ have power series representations about $h = 0$ and the reflection coefficients satisfy (22). Then there exist constants, K_j, for $j = 1, 2, ..., m$, such that*

$$|w_j^{(h)} - e^{i\theta_j}| \le K_j h, \tag{26}$$

where $w_j^{(h)}$ are defined in (24).

3 Convolution with Poisson Kernel

One way to construct measures μ_h for which (12) holds is to convolve point mass measures with an *approximate identity*. For a discussion of approximate identities and their properties see [2]. In particular, denote by $\psi_r(\theta)$ the Poisson kernel:

$$\psi_r(\theta) := \frac{1 - r^2}{|e^{i\theta} - r|^2}, \quad \theta \in [-\pi, \pi), \quad 0 < r < 1, \tag{27}$$

and define

$$\mu_r := \psi_r * \sum_{j=1}^{m} \alpha_j \delta_{\theta_j}\,, \tag{28}$$

where $\alpha_j > 0$ and $\theta_j \in [-\pi, \pi)$ distinct for $j = 1, 2, ..., m$. We can express the convolution (28) as

$$d\mu_r(\theta) = (1 - r^2)\sum_{j=1}^{m} \frac{\alpha_j}{|\zeta - re^{i\theta_j}|^2}\, d\theta \tag{29}$$

$$= (1 - r^2)\frac{\sum_{j=1}^{m} \alpha_j \Pi_{p\neq j}^{m} |\zeta - re^{i\theta_p}|^2}{\Pi_{j=1}^{m} |\zeta - re^{i\theta_j}|^2}\, d\theta, \tag{30}$$

where, for notational brevity, we henceforth define

$$\zeta := e^{i\theta}.$$

Since $\widehat{\psi}_r(\ell) = r^{|\ell|}$, the moments $\widehat{\mu}_r$ are polynomial functions of $h = 1 - r$, by Proposition 2.3, $\lim_{r\to 1} P_k(z,\mu_r)$ exists and (24) becomes

$$P_k(z,\mu_r) = Q_r(z)\prod_{j=1}^{m}(z - w_j^{(r)}), \tag{31}$$

where $w_j^{(r)} \to e^{i\theta_j}$ for $j = 1, 2, ...m$ and $\lim_{r\to 1} Q_r(z)$ is a monic polynomial of degree $k - m$.

It can be shown, using the lines of argument of Lemmas 3.2 and 3.3 of [15], that (22) holds for the measures μ_r. The details, which are similar, are omitted.

Lemma 3.1 *Let μ_r be given in (28). Then the reflection coefficients satisfy*

$$|R_k(\mu_r)| \leq \left(1 - \frac{1}{4^{m-1}}\right)^{1/2}. \tag{32}$$

Outline of Proof: Define $X_j(z) := \frac{1}{(z - re^{i\theta_j})}$, for $j = 1, 2, ..., m$. Using (29), the prediction error power (11) can then be written

$$\rho_k(\mu_r) = \frac{(1-r^2)}{2\pi} \sum_{j=1}^{m} \alpha_j \int_{-\pi}^{\pi} |P_k(\zeta, \mu_r)|^2 |X_j(\zeta)|^2 \, d\theta \tag{33}$$

$$= \frac{(1-r^2)}{2\pi} \sum_{j=1}^{m} \alpha_j \int_{-\pi}^{\pi} |P_k^*(\zeta, \mu_r)|^2 |X_j^*(\zeta)|^2 \, d\theta, \tag{34}$$

where P_k^* and X_j^* are formed by replacing a factor $z - a$ with $1 - z\overline{a}$. (P_k^* is the *reverse polynomial* for P_k.) From subharmonicity of P_k^* and X_j^* we have $\int |P_k^*(\zeta, \mu_r)|^2 |X_j^*(\zeta)|^2 \, d\theta \geq 1$. From (11) and (30) it follows that

$$\rho_{k,r} \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |\zeta^{k-m}|^2 \Pi_{j=1}^{m} |\zeta - re^{i\theta_j}|^2 \, d\mu_r \leq \frac{4^{m-1}(1-r^2)}{2\pi} \sum_{j=1}^{m} \alpha_j.$$

These estimates with (33) and (34) now give

$$\frac{1-r^2}{2\pi} \sum_{j=1}^{m} \alpha_j \leq \rho_k(\mu_r) \leq \frac{4^{m-1}(1-r^2)}{2\pi} \sum_{j=1}^{m} \alpha_j. \tag{35}$$

This, and the relation [1] $|R_{k+1}(\mu_r)|^2 = 1 - \rho_k(\mu_r)/\rho_{k+1}(\mu_r)$, now give (32).

The rate of convergence of *signal zeros* now follows from Lemma 3.1 and Corollary 2.2.

Corollary 3.1 *Let $w_j^{(r)}$ be given in (31) and μ_r defined in (28) where the θ_j are distinct. Then there exist constants K_j, for $j = 1, 2, ..., m$, such that*

$$|w_j^{(r)} - e^{i\theta_j}| \leq K_j(1 - r). \tag{36}$$

We will need one more result before characterizing the limit in (23) for the measures μ_r.

Lemma 3.2 *Let $w_j^{(r)}$ satisfy (36) for constants K_j, $j = 1, 2, ..., m$. The function $\prod_{j=1}^{m} \frac{\zeta - w_j^{(r)}}{\zeta - re^{i\theta_j}}$ converges to 1 in $\mathcal{L}_1[-\pi, \pi]$ as $r \to 1$.*

The following is the main result of this paper.

Theorem 3.1 *Let μ_r be given in (28). Then*

$$\lim_{r\to 1} P_k(z, \mu_r) = P_{k-m}(z, \nu) \prod_{j=1}^{m} (z - e^{i\theta_j}), \tag{37}$$

where ν is the absolutely continuous measure with

$$\frac{d\nu}{d\theta} = \sum_{j=1}^{m} \Pi_{p\neq j}^{m} \alpha_j |\zeta - e^{i\theta_p}|^2. \tag{38}$$

We remark that the assumption that the θ_j are distinct is crucial; it is used in the proof of Corollary 2.2, upon which Lemma 3.1, the key idea in the proof of Theorem 3.1, depends.

In light of the remarks in Section 1, we see that the *R-process* of [5] is equivalent to taking the limit on the left-hand side of (37), where μ_δ is the measure on the right-hand side of (14), with $\alpha_j = \overline{\alpha}_{-j}$ and $\theta_j = -\theta_{-j}$; that is, for point masses associated with real sinusoidal signals. The strength of Theorem 3.1 is that it characterizes this limit for a general class of measures μ_δ.

Reflection coefficients; $m = 2$, $k \to \infty$.

We consider the density ν of Theorem 3.1 in light of results in [11] concerning the limit of the reflection coefficients and *zero-distribution* measures of Szegö polynomials with degree $k \to \infty$. For $m = 2$, the density is easily factored as the squared modulus of a linear function. Results in [11] are used to interpret in [13].

With $m = 2$, the measure μ_r in (28) becomes

$$\mu_r = (\alpha_1 \delta_{\theta_1} + \alpha_2 \delta_{\theta_2}) * \psi_r. \tag{39}$$

We will assume without loss of generality that $\theta_1 = 0$ and define $\omega := \theta_2$, $\alpha_1 := \alpha \in (0, 1)$, and $\alpha_2 = 1 - \alpha$. We then consider

$$\mu_r = (\alpha\delta_0 + (1-\alpha)\delta_\omega) * \psi_r.$$

The density ν defined in (38) becomes $\nu\, d\theta = \alpha|\zeta - e^{i\omega}|^2 + (1-\alpha)|\zeta - 1|^2\, d\theta$. The *spectral factorization* [1, 6] is easily found as

$$d\nu/d\theta = c|\zeta - v_0|^2. \tag{40}$$

where c is a positive constant and

$$v_0 = \frac{1 \pm \sqrt{2\alpha(1-\cos\omega)(1-\alpha)}}{1 - \alpha + \alpha e^{-i\omega}}. \tag{41}$$

Specifying that $|v_0| \leq 1$ gives

$$v_0 = \frac{1 - \sqrt{2\alpha(1-\cos\omega)(1-\alpha)}}{1 - \alpha + \alpha e^{-i\omega}}. \tag{42}$$

In [11] it is shown that for densities of the form $d\nu = c\prod_{j=1}^{m-1} |(\zeta - v_j)|^2$ (where without loss of generality, [2] $|v_j| \leq 1$), that if the v_j are distinct and there is a unique v_j of maximum modulus, then

$$s := \lim_{k\to\infty} |R_k(\nu)|^{1/k} = \max_j\{|v_j|\}. \tag{43}$$

It is further shown in [11] that the *zero-distribution measure* measures, $\frac{1}{k}\sum_{j=1}^{k} \delta_{w_j}$, consisting of point masses of weight $1/k$ at each of the zeros, $w_1, w_2, ..., w_k$, of $P_k(z, \nu)$, converge in the weak-star sense to the uniform measure on the circle of radius s.

Note that Theorem 3.1 gives the convergence

$$\lim_{r\to 1} |R_k(\mu_r)| = |R_{k-m}(\nu)|. \tag{44}$$

In the present context, (43) becomes

$$\lim_{k\to\infty} |R_k(\nu)|^{1/k} = \frac{1 - \sqrt{2\alpha(1-\cos\omega)(1-\alpha)}}{\sqrt{1 + 2(1-\alpha)^2(1-\cos\omega)}}. \tag{45}$$

A result of Petersen is seen as a special case of (45). In light of the remarks following Theorem 3.1, equation (2.16) [13] can be written

$$\lim_{k\to\infty}\lim_{r\to 1}|R_k(\mu_r)|^{1/(k-2)} = \frac{|\cos\omega|}{1+\sin\omega}, \tag{46}$$

where

$$\mu_r = \psi_r * (\frac{1}{2}\delta_\omega + \frac{1}{2}\delta_{-\omega}).$$

With (44), (46) is equivalent to

$$\lim_{k\to\infty}|R_k(\nu)|^{1/k} = \frac{|\cos\omega|}{1+\sin\omega}. \tag{47}$$

This is readily seen as a rotation of (45) for the case $\alpha = 1/2$. Letting $\theta_1 = \omega$, and $\theta_2 = -\omega$ in (39), (42) becomes $v_0 = \frac{1-2(\sin\omega)\sqrt{\alpha-\alpha^2}}{e^{i\omega}-2i\alpha\sin\omega}$. So with $\alpha = 1/2$, eq. (45) becomes $\lim_{k\to\infty}|R_k|^{1/k} = \frac{1-\sin\omega}{\sqrt{1-\sin^2\omega}} = \frac{1-\sin\omega}{|\cos\omega|}$, which, if we define $v_0|_{\omega=\pi/2} = 0$, is equivalent to (47).

References

[1] Grenander, U. and Szegö, S. *Toeplitz Forms and their Applications*, California: University of California Press, 1958.

[2] Hoffman, K. *Banach Spaces of Analytic Functions*, New Jersey: Prentice Hall, 1962.

[3] Haykin, S. *Adaptive Filter Theory*, 3rd ed. Prentice Hall, NJ, 1996

[4] Jones, W.B., Njåsted, O., and Saff, E.B. "Szegö Polynomials associated with Weiner-Levinson filters", *Journal of Computational and Applied Mathematics*, 32, pp. 387-406, 1990.

[5] Jones, W.B., Njåsted, Waadeland, H. "An alternative way of using Szegö polynomials in frequency analysis", in *Continued Fractions and Orthogonal Functions: Theory and Applications*, Marcel Dekker, Inc. 141-151, 1994.

[6] Kay, S. *Modern Spectral Estimation*, Prentice Hall, NJ. 1988.

[7] Honig, M and Messerschmidt, D. *Adaptive Filters, Structures, Algorithms, and Applications*, Kluwer, MA. 1986.

[8] Arciero, M. "Limits for Szegö Polynomials in Frequency Analysis" *Journal of Mathematical Analysis and Applications* accepted for publication, Sept. 2004.

[9] Arciero, M. "Some Algorithms for the Attenuation of Broadband Noise" Proceedings, Eleventh Annual ARL and USMA Technical Symposium, Aberdeen Proving Ground, Maryland, November 2003.

[10] Arciero, M. "An Approach to Attenuation of Broadband Platform Noise" Proceedings, Tenth Annual ARL and USMA Technical Symposium, Aberdeen Proving Ground, Maryland, November 2002.

[11] Pakula, L. "Asymptotic zero distribution of orthogonal polynomials in sinusoidal frequency estimation", *IEEE Transactions on Information Theory*, Vol. IT 33, No.4, pp. 569-576, July 1987.

[12] Petersen, V. "A combination of two methods in frequency analysis: The *R(N)*-Process", *Orthogonal Functions, Moment Theory, and Continued Fractions*, (Campinas 1996), *Lecture Notes in Pure and Applied Mathematics, 199*, New York: Dekker, pp. 387-398, 1998.

[13] Petersen, V. "Zeros of Szegö polynomials used in frequency analysis", *Orthogonal Functions, Moment Theory, and Continued Fractions*, (Campinas 1996), *Lecture Notes in Pure and Applied Mathematics, 199*, New York: Dekker, pp. 399-408, 1998.

[14] Petersen, V. "Modification of a method using Szegö polynomials in frequency analysis: the V-process", *Journal of Computational and Applied Mathematics*, Volume 133, Issues 1-2 August 2001, pp. 535-544.

[15] Pan, K. and Saff, E.B. "Asymptotics for zeros of Szego polynomials", *Journal of Approximation Theory*, Vol. 71, No. 3, pp. 239-251, Dec. 1992.

[16] Widrow, B., McCool, J.M, Larimore, M.G,Johnson, C.R, " Adaptive noise cancelling: principles and applications", *Proc. IEEE*, vol. 63, no. 12, pp.1692-1716, Dec, 1975.

[17] Widrow, B., Stearns, S. *Adaptive Signal Processing*, Prentice Hall, NJ., 1985.

[18] Wiener, N. *Extrapolation, Interpoation, and Smoothing of Stationary Time Series*, MIT Press, Cambridge, MA, 1949.

[19] Zeidler, J. "Performance analysis of LMS adaptive prediction filters", *Proc. IEEE*, vol. 78 no. 12, pp. 1781-1806, Dec. 1990