Journal of Computational and Applied Mathematics 32 (1990) 387–406
North-Holland

Szegö polynomials associated with Wiener–Levinson filters

William B. JONES *
Department of Mathematics, University of Colorado, Boulder, CO 80309-0426, USA

Olav NJÅSTAD
Department of Mathematics, University of Trondheim (NTH), 7034 Trondheim, Norway

E.B. SAFF *
Department of Mathematics, University of South Florida, Tampa, FL 33620, USA

Received 6 December 1989

Abstract: Szegö polynomials are studied in connection with Wiener–Levinson filters formed from discrete signals $x_N = \{x_N(k)\}_{k=0}^{N-1}$. Our main interest is in the frequency analysis problem of finding the unknown frequencies ω_j, when the signal is a trigonometric polynomial $x_N(k) = \sum_{j=-I}^{I} \alpha_j \, e^{i\omega_j k}$. Associated with this signal is the sequence of monic Szegö polynomials $\{\rho_n(\psi_N;\, z)\}_{n=0}^{\infty}$ orthogonal on the unit circle with respect to a distribution function $\psi_N(\theta)$. Explicit expressions for the weight function $\psi_N'(\theta)$ and associated Szegö function $D_N(z)$ are given in terms of the Z-transform $X_N(z)$ of the signal x_N. Several theorems are given to support the following conjecture which was suggested by numerical experiments: *As N and n increase, the* $2I+1$ *zeros of* $\rho_n(\psi_N;\, z)$ *of largest modulus approach the points* $e^{i\omega_j}$. We conclude by showing that the reciprocal polynomials $\rho_n^*(\psi_N;\, z) := z^n \overline{\rho_n(\psi_N;\, 1/\bar{z})}$ are Padé numerators for Padé approximants (of fixed denominator degree) to a meromorphic function related to $D_N(z)$.

Keywords: Orthogonal polynomials, frequency analysis, digital filter.

1. Introduction

A number of papers have appeared recently on Szegö polynomials $\rho_n(z)$ orthogonal with respect to a distribution function $\psi(\theta)$ [10–12,15,16,21,23]. One reason for interest in Szegö polynomials is their close connection with important problems of digital signal processing [9,14]. The sequences $\{\rho_n\}$ considered in this paper are obtained by forming Wiener–Levinson (linear least squares) filters, starting with a discrete input signal $x = \{x(k)\}_{k=-\infty}^{\infty}$ and N-truncated, causal signals $x_N = \{x_N(k)\}$, where $x_N(k) := 0$ for $k < 0$ and $k \geqslant N$.

Background and preliminary results on Szegö theory and the linear prediction method applied to N-truncated, causal signals $x_N = \{x_N(k)\}$ are described in Section 2. The associated distribution function $\psi_N(\theta)$ and Szegö function $D_N(z)$ are obtained explicitly and the Szegö condition

* Research supported in part by the U.S. National Science Foundation under grants numbered DMS-8700498 and DMS-881-4026.

(cf. (2.12)) is readily seen to hold. Finite impulse response (FIR) filters $g^{(n)} * x_N$ and $h^{(n)} * x_N$ that minimize $\| g^{(n)} * x_N \|_2$ and $\| \delta - h^{(n)} * x_N) \|_2$, respectively, are derived (Theorems 1 and 2) in terms of the Szegö polynomial $\rho_n(\psi_N;\ z)$, where δ denotes the *unit pulse signal.*

Section 3 is used to present results from a numerical-graphical experiment that illustrates typical behavior of the zeros of the $\rho_n(\psi_N;\ z)$, where the input signal $x = \{x(k)\}$ has the special form of a finite Fourier series

$$x(k) = \sum_{j=-I}^{I} \alpha_j \, e^{i\omega_j k}, \quad k = 0, \pm 1, \pm 2, \ldots,$$

where $\alpha_{-j} = \bar{\alpha}_j$, $\omega_{-j} = -\omega_j$ and $0 = \omega_0 < \omega_1 < \cdots < \omega_I < \pi$. These results led us to the conjecture given in Section 4. In support of this conjecture we give: estimates which suggest that $| \rho_n(\psi_N;\ e^{i\omega_j}) |$ is small when n and N are large (Theorem 3), a result (Theorem 6) on the asymptotic behavior of the zeros of $\rho_n(\psi_N;\ z)$, and a proof (Theorem 7) that the weak star limit of $\{(1/N)\ d\psi_N(\theta)\}$ is a point mass distribution with mass $|\alpha_j|^2$ at each of the points $e^{\pm i\omega_j}$, $j = 0, \pm 1, \ldots, \pm I$.

In Section 5 we investigate the approximants $[n/M](z)$ on the Mth row in the Padé table of the rational function $F_N(z) := [D_N(z) D_N^*(z)]^{-1}$, where $D_N^*(z) := z^M \overline{D_N(1/\bar{z})}$ and $M :=$ deg $D_N(z)$. It is shown (Theorem 8) that the polynomial numerator of $[n/M](z)$ is a constant multiple of $\rho_n^*(\psi_N;\ z)$. The convergence of $\{[n/M](z)\}_{n=0}^{\infty}$ is described by Corollary 9, which is a consequence of Theorem 8 and an extension of the Montessus de Ballore theorem. This result provides an independent proof that lim $\rho_n^*(z) = D_N(0)/D_N(z)$, at least for the cases dealt with in Section 4.

2. Applications of Szegö theory and background

Doubly infinite sequences $x = \{x(m)\}_{m=-\infty}^{\infty}$ of real numbers are called (discrete) *signals*, the space of all such signals being denoted by l. A signal x is said to be *causal* if $x(m) = 0$ for $m < 0$. Since our interest is in signals that are observed (measured) we consider *N-truncated, causal signals* $x_N = \{x_N(m)\}$, where

$$x_N(m) := \begin{cases} x(m), & \text{if } 0 \leqslant m \leqslant N-1, \\ 0, & \text{otherwise,} \end{cases} \tag{2.1a}$$

and

$$x_N(0) \neq 0. \tag{2.1b}$$

The *Z-transform* of x_N is given by

$$X_N(z) := \sum_{m=0}^{N-1} x_N(m) z^{-m}.$$

For the absolutely continuous distribution function $\psi_N(\theta)$ defined by

$$\psi_N'(\theta) := \frac{1}{2\pi} \left| X_N(e^{i\theta}) \right|^2 \quad -\pi \leqslant \theta \leqslant \pi,$$

the kth moment μ_k ($k = 0, \pm 1, \pm 2, .\ $) is given by

$$\mu_k := \int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}k\theta}\,\mathrm{d}\psi_N(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}k\theta}\left|\sum_{m=0}^{N-1} x_N(m)\,\mathrm{e}^{-\mathrm{i}m\theta}\right|^2 \mathrm{d}\theta$$

$$= \frac{1}{2\pi}\sum_{m=0}^{N-1}\sum_{j=0}^{N-1} x_N(m)x_N(j)\int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}(j-m-k)\theta}\,\mathrm{d}\theta = \sum_{m=0}^{N-1} x_N(m)x_N(m+k).$$

It follows that the moments μ_k represent the *autocorrelation coefficients of the signal*:

$$\mu_k = \sum_{m=-\infty}^{\infty} x_N(m)x_N(m+k), \quad k = 0, \pm 1, \pm 2, \ldots . \tag{2.4}$$

Moreover, since the μ_k are trigonometric moments, the sequence $\{\mu_k\}$ is *Hermitian positive definite*; that is,

$$\mu_k = \mu_{-k} \quad \text{and} \quad \Delta_k := \det(\mu_{i-j})_{i,j=0}^{k} > 0, \qquad k = 0, 1, 2, \ldots, \tag{2.5}$$

(see, e.g., [1,10,12]). The distribution function $\psi_N(\theta)$ provides an inner product

$$\langle f, g\rangle := \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\overline{g(\mathrm{e}^{\mathrm{i}\theta})}\,\mathrm{d}\psi_N(\theta), \quad f, g \in \Lambda, \tag{2.6}$$

on the space Λ of *Laurent polynomials* $\sum_{k=m}^{n} c_k z^k$, $c_k \in \mathbb{C}$, $-\infty < m \leqslant k \leqslant n < +\infty$. We denote by $\{\rho_n = \rho_n(\psi_N; z)\}_{n=0}^{\infty}$ the sequence of monic Szegö polynomials orthogonal with respect to ψ_N and denote the *reciprocal polynomials* by

$$\rho_n^*(\psi_N; z) := z^n\overline{\rho_n(\psi_N; 1/\bar{z})}, \quad n = 0, 1, 2, \ldots . \tag{2.7}$$

The polynomials ρ_n and ρ_n^* satisfy the *orthogonality* conditions

$$\langle \rho_n, z^m\rangle = \begin{cases} 0, & 0 \leqslant m \leqslant n-1, \\ \dfrac{\Delta_n}{\Delta_{n-1}} & m = n, \end{cases} \qquad \langle \rho_n^*, z^m\rangle = \begin{cases} \dfrac{\Delta_n}{\Delta_{n-1}} & m = 0, \\ 0, & 1 \leqslant m \leqslant n \end{cases}$$

and the *recurrence relations*

$$\rho_0(\psi_N; z) := \rho_0^*(\psi_N; z) := 1, \tag{2.9a}$$

$$\rho_n(\psi_N; z) = z\rho_{n-1}(\psi_N; z) + \delta_n\rho_{n-1}^*(\psi_N; z), \quad n = 1, 2, 3, \ldots, \tag{2.9b}$$

$$\rho_n^*(\psi_N; z) = \bar{\delta}_n z\rho_{n-1}(\psi_N; z) + \rho_{n-1}^*(\psi_N; z), \quad n = 1, 2, 3, \ldots . \tag{2.9c}$$

The constants δ_n, called *reflection coefficients*, can be computed successively in terms of the moments μ_k by

$$\delta_n = -\frac{\langle z\rho_{n-1}, 1\rangle}{\langle \rho_{n-1}^*, 1\rangle} = -\frac{\sum_{j=0}^{n-1} q_j^{(n-1)}\mu_{-j-1}}{\sum_{j=0}^{n-1} q_j^{(n-1)}\mu_{j+1-n}} \qquad \rho_n(z) =: \sum_{j=0}^{n} q_j^{(n)} z^j \tag{2.10}$$

This procedure for computing the δ_n's is known as *Levinson's algorithm*.

It is well known (e.g., [11]) that all zeros of the $\rho_n(z)$ lie in $|z| < 1$,

$$|\delta_n| < 1 \quad \text{and} \quad \beta_n := \|\rho_n\| = \|\rho_n^*\| = \sqrt{\Delta_n/\Delta_{n-1}}, \qquad n = 1, 2, 3, \tag{2.11a}$$

and

$$\beta_n^2 = \mu_0 \prod_{k=1}^{n} (1 - |\delta_k|^2), \quad n = 1, 2, 3, . \tag{2.11b}$$

It is readily seen that $\psi_N'(\theta)$ satisfies the *Szegö condition*

$$\infty > \int_{-\pi}^{\pi} \psi_N'(\theta)\, \mathrm{d}\theta, \qquad \int_{-\pi}^{\pi} \log \psi_N'(\theta)\, \mathrm{d}\theta > -\infty \tag{2.12}$$

The Szegö function $D_N(z)$ with respect to $\psi_N(\theta)$ is defined by

$$D_N(z) := \sqrt{2\pi}\, \exp\left\{ \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\mathrm{e}^{\mathrm{i}\theta} + z}{\mathrm{e}^{\mathrm{i}\theta} - z} \log \psi_N'(\theta)\, \mathrm{d}\theta \right\} \quad |z| < 1 \tag{2.13}$$

It follows from the well-known theory of Szegö polynomials that

$$\beta := \lim_{n \to \infty} \beta_n > 0, \qquad \sum_{k=1}^{\infty} |\delta_k|^2 < \infty, \tag{2.14}$$

$$D_N(z) = \lim_{n \to \infty} \frac{\beta_n}{\rho_n^*(\psi_N; z)} \in H_2 \text{ (Hardy space)}, \qquad |z| < 1, \tag{2.15a}$$

$$D_N(z) \neq 0 \text{ for } |z| < 1 \quad \text{and} \quad D_N(0) = \beta > 0 \tag{2.15b}$$

and

$$|D_N(\mathrm{e}^{\mathrm{i}\theta})|^2 := \lim_{r \to 1^-} |D_N(r\,\mathrm{e}^{\mathrm{i}\theta})|^2 = 2\pi\psi_N'(\theta), \quad \text{a.e. on } [-\pi, \pi] \tag{2.15c}$$

(see, e.g., [4–7,9,19,24]). Furthermore, one can prove the following assertion:

$$D_N(z) = \pm x_N(0) \prod_{|z_k| \geq 1} (z - z_k) \prod_{|z_k| < 1} (1 - \bar{z}_k z), \tag{2.16}$$

where the z_k's in (2.16) *denote the $N-1$ zeros of the polynomial $z^{N-1}X_N(z)$ and the sign* (± 1) *is chosen so that $D_N(0) > 0$.* This choice of sign is possible since the z_k are either real or else they occur in conjugate pairs. To verify (2.16) we let $\hat{D}_N(z)$ denote the function on the right-hand side of (2.16) and observe that by (2.3)

$$\psi_N'(\theta) = \frac{1}{2\pi} |\hat{D}_N(\mathrm{e}^{\mathrm{i}\theta})|^2, \quad -\pi \leq \theta \leq \pi \tag{2.17}$$

It suffices then to substitute this expression in (2.13) for $\psi_N'(\theta)$ and apply the Poisson integral formula, which holds also for functions of the form $\log |z - \mathrm{e}^{\mathrm{i}\theta_1}|$ (cf. [13, Theorem 8.2]).

Let l_D and l_R denote subsets of the space l of signals. A linear map $T: l_D \to l_R$ is called a *digital filter* if T is *shift invariant*; that is, if $ST = TS$, where S denotes the *shift operator* defined by

$$(Sx)(k) := x(k-1), \quad k = 0, \pm 1, \pm 2, \ldots, \quad x \in l.$$

The *convolution* $h * x$ of sequences h and x, defined by

$$(h * x)(k) := \sum_{m=-\infty}^{\infty} x(m)h(k-m), \quad k = 0, \pm 1, \pm 2,$$

provides a convenient representation of a digital filter. In fact, if $h \in l_1$, then

$$y = Tx := h * x$$

can be shown to be a continuous, linear, shift-invariant mapping of l_∞ into l_∞. Moreover, h is the *unit pulse response*: i.e., $h = T\delta$ where the *unit pulse* $\delta = \{\delta(k)\}$ is defined by $\delta(0) := 1$ and $\delta(k) := 0$ for $k \neq 0$. The *z-transform*

$$H(z) = \sum_{m=-\infty}^{\infty} h(m) z^{-m} \tag{2.21}$$

of h is called the *transfer function* of the filter T since $Y(z) = H(z)X(z)$, where $X(z)$ and $Y(z)$ denote the Z-transforms of the signals x and y, respectively. Since $h \in l_1$, the series (2.21) converges absolutely for $|z| \geqslant 1$. The function $H(e^{i\theta})$ is called the *frequency response function of the filter T*.

Associated with an N-truncated causal signal $x_N = \{x_N(k)\}$ (see (2.1)) is a *linearly predicted signal* $\hat{x}_N = \{\hat{x}_N(k)\}$ given in terms of x_N by

$$\hat{x}_N(k) := \begin{cases} -\sum_{j=1}^{n} g_j^{(n)} x_N(k-j), & g_j^{(n)} \in \mathbb{R}, \ k \geqslant 1, \\ 0, & k \leqslant 0. \end{cases} \tag{2.22}$$

The *residual signal* $\epsilon_N^{(n)} = \{\epsilon_N^{(n)}(k)\} = \{x_N(k) - \hat{x}_N(k)\}_{k=-\infty}^{\infty}$ can be written as a convolution

$$\epsilon_N^{(n)} = g^{(n)} * x_N = \left\{ \sum_{j=0}^{n} g_j^{(n)} x_N(k-j) \right\}_{k=-\infty}^{\infty} \tag{2.23}$$

of $g^{(n)} = \{g_j^{(n)}\}$ and x_N, where $g_0^{(n)} := 1$ and $g_j^{(n)} := 0$ for $j < 0$ and $j > n$. For each $n \geqslant 1$, the polynomial

$$G_n(z) := \sum_{j=0}^{n} g_j^{(n)} z^{-j} \tag{2.24}$$

is the transfer function of the digital filter T_n defined by $\epsilon_N^{(n)} := T_n x_N := g^{(n)} * x_N$. The coefficients $g_j^{(n)}$, $1 \leqslant j \leqslant n$, are determined in such a manner as to minimize the squared l_2-norm

$$\|\epsilon_N^{(n)}\|_2^2 = \sum_{k=-\infty}^{\infty} \left[\epsilon_N^{(n)}(k)\right]^2. \tag{2.25}$$

In fact it follows from (2.23) and (2.4) that

$$
\begin{aligned}
\| \epsilon_N^{(n)} \|_2^2 &= \sum_{k=-\infty}^{\infty} \left[\sum_{j=0}^{n} g_j^{(n)} x_N(k-j) \right]^2 \\
&= \sum_{j=0}^{n} \sum_{m=0}^{n} g_j^{(n)} g_m^{(n)} \sum_{k=-\infty}^{\infty} x_N(k-j) x_N(k-m) \\
&= \sum_{j=0}^{n} \sum_{m=0}^{n} g_j^{(n)} g_m^{(n)} \mu_{j-m} \\
&= \sum_{j=0}^{n} \sum_{m=0}^{n} g_j^{(n)} g_m^{(n)} \int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}(j-m)\theta} \, \mathrm{d}\psi_N(\theta) \\
&= \int_{-\pi}^{\pi} \left| \sum_{j=0}^{n} g_j^{(n)} \mathrm{e}^{-\mathrm{i} j \theta} \right|^2 \mathrm{d}\psi_N(\theta) \\
&= \| G_n \|^2 = \langle G_n, G_n \rangle = \langle \sigma_n, \sigma_n \rangle,
\end{aligned}
$$

where $\sigma_n(z) := z^n G_n(z)$ is a monic polynomial of degree n. From a known property of Szegö polynomials (e.g., [7, Section 2.5]), $\| \sigma_n \|^2 = \langle \sigma_n, \sigma_n \rangle$ attains its minimum value for all monic polynomials $\sigma_n(z)$ of degree n if $\sigma_n(z) = \rho_n(\psi_N; z)$, the nth Szegö polynomial with respect to $\psi_N(\theta)$. This result is summarized by the following theorem.

Theorem 1. *Let $x_N = \{x_N(m)\}$ be a given N-truncated causal signal* (2.1) *and let $\psi_N(\theta)$ be defined by* (2.3). *For each $n \geqslant 1$, let $\epsilon_N^{(n)} := g^{(n)} * x_N$, where $g^{(n)} = \{ g_j^{(n)} \} \in l$, $g_j^{(n)} = 0$ for $j < 0$ and $j > n$, and*

$$
G_n(z) := \sum_{j=0}^{n} g_j^{(n)} z^{-j}, \quad g_0^{(n)} := 1, \quad g_j^{(n)} \in \mathbb{R} \tag{2.26}
$$

Then

$$
\min_{g^{(n)}} \| \epsilon_N^{(n)} \|_2 = \min_{g^{(n)}} \| g^{(n)} * x_N \|_2 = \beta_n = \sqrt{\Delta_n / \Delta_{n-1}} \tag{2.27}
$$

is attained by $G_n(z) = z^{-n} \rho_n(\psi_N; z)$, where $\{\rho_n\}$ is the sequence of monic Szegö polynomials with respect to the distribution function $\psi_N(\theta)$.

Moreover, as $n \to \infty$,

$$
\min_{g^{(n)}} \| \epsilon_N^{(n)} \|_2 = \beta_n \searrow \beta = D_N(0) = x_N(0) \cdot \prod_{|z_k| \geqslant 1} |z_k|, \tag{2.28}
$$

where the z_k denote the zeros of $X_N(z)$ such that $|z_k| \geqslant 1$ and where the product is replaced by one if there are no such zeros.

Theorem 2. *Let $x_N = \{x_N(m)\}$ be a given N-truncated, causal signal and let $\psi_N(\theta)$ be defined by* (2.3). *For each $n \geqslant 1$, let $h^{(n)} = \{ h_j^{(n)} \} \in l$ be such that $h_0^{(n)} \neq 0$ and $h_j^{(n)} = 0$ for $j < 0$ and $j > n$ and let*

$$
H_n(z) := \sum_{j=0}^{n} h_j^{(n)} z^{-j}, \quad h_j^{(n)} \in \mathbb{R}
$$

Then

$$\min_{h^{(n)}} \left\| \delta - \left(h^{(n)} * x_N\right) \right\|_2 = \sqrt{1 - \left(\frac{x_N(0)}{\beta_n}\right)^2} \tag{2.30}$$

and this minimum is attained by choosing

$$H_n(z) = \frac{x_N(0)}{\beta_n^2} z^{-n} \rho_n(\psi_N;\ z), \tag{2.31}$$

where $\{\rho_n\}$ *is the sequence of monic Szegö polynomials with respect to the distribution function* $\psi_N(\theta)$. *Here* δ *denotes the unit pulse sequence (see definition following* (2.20)).

Moreover, as $n \to \infty$,

$$\min_{h^{(n)}} \left\| \delta - \left(h^{(n)} * x_N\right) \right\|_2 \searrow \sqrt{1 - \left(\frac{x_N(0)}{\beta}\right)^2} = \sqrt{1 - \prod_{|z_k| \geqslant 1} |z_k|^{-2}}, \tag{2.32}$$

where the z_k *denote the zeros of* $X_N(x)$ *such that* $|z_k| \geqslant 1$ *and where the product is replaced by* 1 *if there are no such zeros.*

Proof. Let c_n and $g^{(n)} = \{g_j^{(n)}\} \in l$ be chosen so that

$$h^{(n)} = c_n g^{(n)}, \quad c_n \neq 0, \quad g_0^{(n)} := 1, \quad g_j^{(n)} \in \mathbb{R}, \tag{2.33}$$

and let $G_n(z)$ be defined by (2.24). We then obtain

$$\begin{aligned}
\left\| \delta - \left(h^{(n)} * x_N\right) \right\|_2^2 &:= \sum_{k=-\infty}^{\infty} \left[\delta(k) - \left(h^{(n)} * x_N\right)(k)\right]^2 \\
&= \sum_{k=-\infty}^{\infty} [\delta(k)]^2 - 2c_n \sum_{k=-\infty}^{\infty} \delta(k)\left(g^{(n)} * x_N\right)(k) + c_n^2 \left\| g^{(n)} * x_N \right\|_2^2 \\
&= 1 - 2c_n \sum_{j=0}^{n} g_j^{(n)} x_N(-j) + c_n^2 \left\| g^{(n)} * x_N \right\|_2^2 \\
&= 1 - 2c_n x_N(0) + c_n^2 \left\| g^{(n)} * x_N \right\|_2^2.
\end{aligned}$$

We minimize this in two steps: first with respect to $g^{(n)}$ and then with respect to c_n. An application of Theorem 1 gives

$$\begin{aligned}
\min_{g^{(n)}} \left\| \delta - \left(h^{(n)} * x_N\right) \right\|_2^2 &= 1 - 2c_n x_N(0) + c_n^2 \beta_n^2 \\
&= 1 - \left(\frac{x_N(0)}{\beta_n}\right)^2 + \beta_n^2 \left(c_n - \frac{x_N(0)}{\beta_n^2}\right)^2
\end{aligned}$$

It is then easily seen that the minimum of this with respect to c_n is

$$\min_{h^{(n)}} \left\| \delta - \left(h^{(n)} * x_N\right) \right\|_2^2 = 1 - \left(\frac{x_N(0)}{\beta_n}\right)^2$$

and this is attained by choosing $c_n = x_N(0)/\beta_n^2$, $G_n(z) = z^{-n} \rho_n(\psi_N;\ z)$; hence (2.31) holds.

Equation (2.32) follows immediately from (2.30) and (2.16). $\square$

3. Frequency analysis for trigonometric polynomial signals

In this section we consider a numerical-graphical example obtained by applying the well-known linear prediction method and Levinson algorithm described in Section 2 to a signal $x = \{x(k)\}$ of the form

$$\sum_{j=-I}^{I} \alpha_j e^{i\omega_j k} \quad k = 0, \pm 1, \pm 2, \tag{3.1a}$$

where

$$0 = \omega_0 < \omega_1 < \omega_2 < \quad < \omega_I < \pi, \qquad \omega_{-j} \quad -\omega_j, \tag{3.1b}$$

$$\alpha_0 \in \mathbb{R}, \qquad 0 \neq \alpha_j = \bar{\alpha}_{-j} \in \mathbb{C}, \quad j = 1, 2, \ldots, I. \tag{3.1c}$$

Our goal is to approximate the (unknown) ω_j's from the data $\{x(k)\}_{k=0}^{N-1}$. Actually, for this illustration we add to the signal a small noise component $\eta = \{\eta(k)\}$ so that

$$x(k) = u(k) + \eta(k), \quad k = 0, \pm 1, \pm 2, \ldots, \tag{3.2a}$$

where

$$u(k) = \sum_{j=-4}^{4} \alpha_j e^{i\omega_j k} = \sum_{j=1}^{4} 2\lambda_j \sin \omega_j k, \quad k = 0, \pm 1, \pm 2, \ldots, \tag{3.2b}$$

and we take $I = 4$, $\alpha_j = -i\lambda_j$, where the λ_j and ω_j are given in Table 1.

The sequence $\eta = \{\eta(k)\}$ with $\eta(0) \neq 0$ is formed from white noise, the $\eta(k)$ being random numbers normalized to have a mean of zero and variance 0.02. The noise component was added in order to simulate an actual observed signal. The level of noise is small enough so that the "observed" signal x behaves essentially the same as the "true" signal u. The signal we actually work with is the 400-truncated, causal signal x_{400}. The choice of $N = 400$ was made so that we can consider Szegö polynomials $\rho_n(\psi_N;\, z)$ with n in the range $1 \leqslant n \leqslant N - 2I = 392$. We note that although $u_{400}(0) = 0$, we have $x_{400}(0) = \eta(0) \neq 0$; therefore (2.1b) holds.

Using (3.2) as starting data, we have computed autocorrelation coefficients (moments μ_k) by (2.4) and then reflection coefficients δ_k and norms $\beta_k := \| \rho_k \|$ by the Levinson algorithm based on (2.9) and (2.10). Some of the results are given in Table 2. When computing the δ_k we also constructed the Szegö polynomials $\rho_k(\psi_N;\, z)$ and obtained their zeros $z_j^{(k)}$, $j = 1, 2, \ldots, k$, $k = 1, 2, \ldots, 50$. In Table 3 we give $|z_j^{(k)}|$ and arg $z_j^{(k)}$, $k = 10, 20, 30, 40, 50$, for each of the four zeros lying nearest to the unit circle $|z| = 1$. From these results it is clear that the 4 angles ω_j used to form the input x_{400} in (3.2) can be approximated with 3 or 4 significant digits. In Figs. 1 and 2 are shown plots of all of the zeros of $\rho_k(\psi_N;\, z)$ for $k = 4, 6, 8, 10, 20, 30, 40, 50$. The actual zeros occur at the endpoints of lines radiating from the origin. It is remarkable that even

Table 1
Parameters λ_j and ω_j

j	0	1	2	3	4

Table 2
Reflection coefficients δ_k and norms $\beta_k = \| \rho_k \|$ computed from the Levinson algorithm

k	δ_k	β_k
	1.00000	143.7
	0.67266	106.3
	0.83772	58.1
	−0.59796	46.5
	0.63989	35.8
	−0.26756	34.5
	0.27821	33.1
	0.31942	31.4
	0.00186	31.4
	0.33703	29.5
	0.60999	23.4
20	−0.05775	17.0
30	0.07575	16.5
40	−0.01212	16.4
50	−0.00874	16.3

Table 3
Zeros $z_j^{(k)}$ of $\rho_k(z)$ near $|z|=1$

	k	$\|z_j^{(k)}\|$	$\arg z_j^{(k)}$	ω_j
$j=1$				
$j=2$	10	0.955..		$\frac{1}{3}\pi \doteq 1.047198$
	20	0.995..		
	30	0.997..		
	40	0.997..		
	50	0.9971		
$j=3$	10	0.981...		$\frac{1}{6}\pi \doteq 0.523599$
	20	0.996...		
	30	0.997...		
	40	0.997...		
	50	0.997...		
$j=4$	10	0.999...		$\frac{3}{4}\pi \doteq 2.356194$
	20	0.997...		
	30	0.997...		
	40	0.997...		
	50	0.997...		

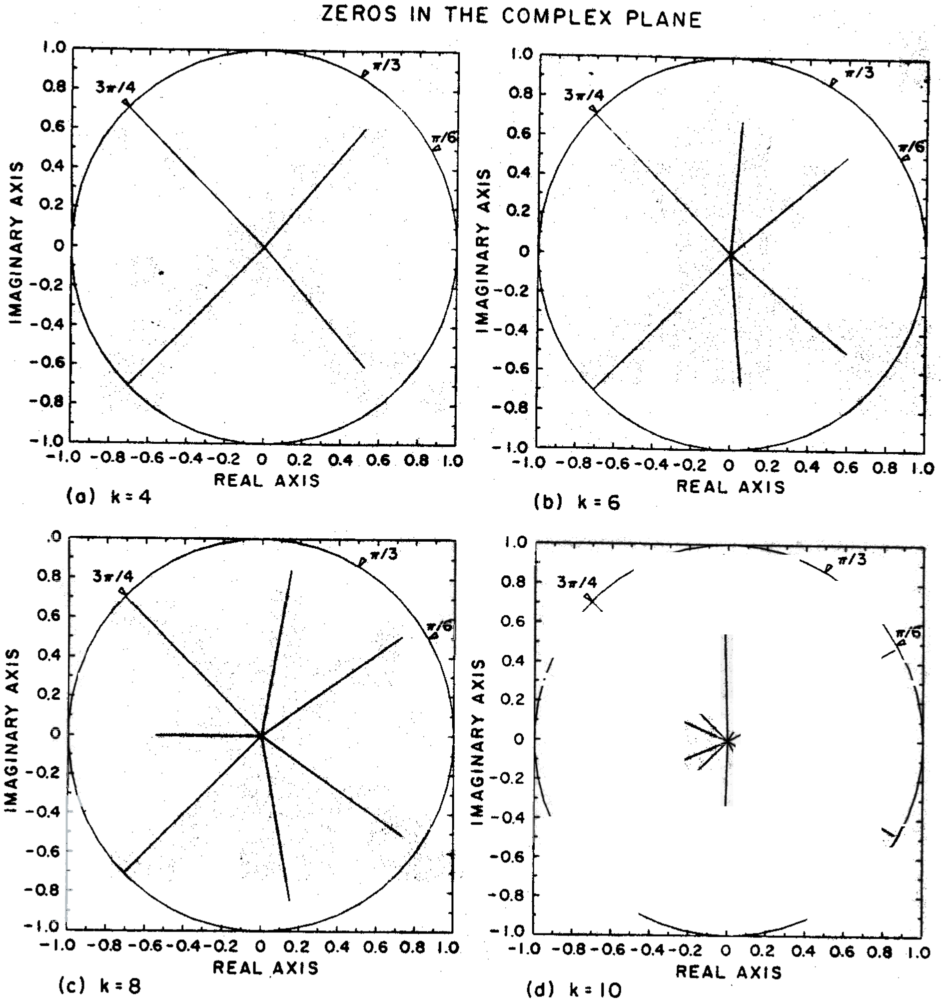
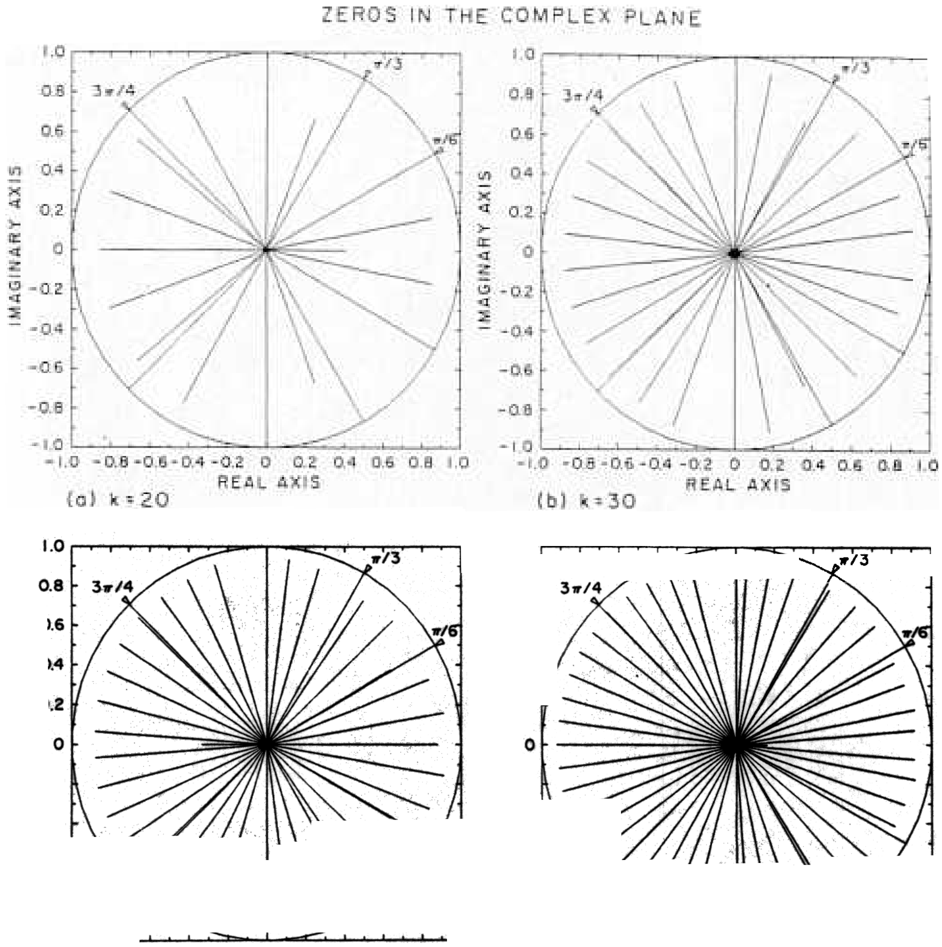


Fig. 1. The zeros $z_j^{(k)}$ of the Szegö polynomials $\rho_k(z)$ are shown as the end points of lines radiating from the origin. The $\rho_k(z)$ were constructed from input data (3.2) and Table 1.

at $k = 4$, two of the zeros of $\rho_4(z)$ are very close to the points $e^{i\omega_j}$ with $\omega_j = \pm \frac{3}{4}\pi$. With $k = 10$, all of the points $e^{i\omega_j}$ are being approximated by zeros of $\rho_{10}(z)$. It is also notable that for each k, the zeros of $\rho_k(\psi_N; z)$ are fairly uniformly distributed by argument around the circle, except for some gaps that appear near the critical points $e^{i\omega_j}$. As k increases the amplitudes $|z_j^{(k)}|$ appear to increase, with certain zeros (apparently) converging quickly to the $e^{i\omega_j}$.

4. Conjecture and supportive theory

The following conjecture deals with the asymptotic behavior of zeros of Szegö polynomials $\rho_n(\psi_N, z)$ formed from trigonometric polynomial signals of the form (3.1). The conjecture is suggested by the example described in Section 3. Several theorems and other arguments that support the conjecture are given in this section.



Conjecture (C). Let $\{\rho_n(\psi_N, z)\}$ denote the sequence of monic Szegö polynomials with respect to an absolutely continuous distribution function $\psi_N(\theta)$ defined by (2.3), where $X_N(z)$ is the Z-transform of a trigonometric polynomial signal of the form (3.1). Then it is conjectured that, as n and N tend to infinity (in a manner to be determined), with $1 \leqslant n \leqslant N-2I$, the $2I+1$ zeros of $\rho_n(\psi_N, z)$ of largest modulus will approach the critical points $e^{i\omega_j}$, $j = 0, \pm 1, \ldots, \pm I$. (We have assumed above that $\alpha_0 \neq 0$ so that $e^{i\omega_0} = e^0 = 1$ is a critical point. If $\alpha_0 = 0$, then there are only $2I$ critical points, since $e^{i0} = 1$ should not be counted; hence the statement of the conjecture should be adjusted accordingly.)

Theorem 3. *Let $x = \{x(k)\}$ be a given signal of the form* (3.1) *and let $x_N = \{x_N(k)\}$ denote the corresponding N-truncated causal signal. There exists a constant A, dependent upon N and the α_j*

and ω_j *for* $-I \leqslant j \leqslant I$, *but independent of* n, *such that*

$$|\rho_n(\psi_N, e^{i\omega_j})| = |G_n(e^{i\omega_j})| \leqslant A \sum_{k=n}^{n+2I} |\epsilon_N^{(n)}(k)|, \quad j = -I, -I+1, \ldots, I, \tag{4.1}$$

provided $1 \leqslant n \leqslant N - 2I$, *where* $G_n(z) = \sum_{j=0}^{n} g_j^{(n)} z^{-j} = z^{-n}\rho_n(\psi_N;\ z)$ *and* $\epsilon_N^{(n)} = g^{(n)} * x_N$.

Our proof of Theorem 3 makes use of the following lemma.

Lemma 4. *Let* $g \in l_1$ *be given and let* T *denote the digital filter defined by*

$$\epsilon := Tx = g * x, \quad \text{with } x \text{ of the form } (3.1). \tag{4.2}$$

Let $G(z) = \sum_{m=-\infty}^{\infty} g(m) z^{-m}$ *denote the Z-transform of* $g = \{g(m)\}$ *(also the transfer function of* T*). Then:*

(A) *For* $k = 0, \pm 1, \pm 2, \ldots,$

$$\epsilon(k) := (Tx)(k) = \sum_{j=-I}^{I} \alpha_j G(e^{i\omega_j})\, e^{i\omega_j k}. \tag{4.3}$$

(B) *Let* p *be a given integer. Then there exist complex constants* A_{jk} *(depending only upon* p *and the* α_j *and* ω_j, $-I \leqslant j \leqslant I$*) such that*

$$G(e^{i\omega_j}) = \sum_{k=p}^{p+2I} A_{j,k}\epsilon(k), \quad -I \leqslant j \leqslant I. \tag{4.4}$$

Proof. (A) By (2.18), (2.20), (3.1a) and (4.3),

$$\begin{aligned}
\epsilon(k) &= \sum_{m=-\infty}^{\infty} x(m) g(k-m) = \sum_{m=-\infty}^{\infty} g(m) x(k-m) \\
&= \sum_{m=-\infty}^{\infty} g(m) \sum_{j=-I}^{I} \alpha_j\, e^{i\omega_j(k-m)} \\
&= \sum_{j=-I}^{I} \alpha_j \left[\sum_{m=-\infty}^{\infty} g(m) (e^{i\omega_j})^{-m} \right] e^{i\omega_j k} \\
&= \sum_{j=-I}^{I} \alpha_j G(e^{i\omega_j})\, e^{i\omega_j k}.
\end{aligned}$$

(B) This follows from an application of Cramer's rule to the system of $2I+1$ linear equations (4.3), with $p \leqslant k \leqslant p + 2I$, in the $2I+1$ unknowns $a_j := \alpha_j G(e^{i\omega_j})$, $-I \leqslant j \leqslant I$. The determinant of the system is the Vandermonde determinant

$$\begin{vmatrix}
1 & 1 & & 1 \\
Z_{-I} & Z_{-I+1} & & Z_I \\
\vdots & & & \\
(Z_{-I})^{2I} & (Z_{-I+1})^{2I} & & (Z_I)^{2I}
\end{vmatrix}$$

divided by $\exp[\mathrm{i}p\Sigma_{j=-I}^{I}\omega_j]$, where $Z_j := \mathrm{e}^{\mathrm{i}\omega_j}$. This determinant does not vanish since, by (3.1b), the Z_j are all distinct. □

Proof of Theorem 3. For each $n \geqslant 1$ let T_n denote the digital filter

$$T_n x := g^{(n)} * x, \quad x \in l_\infty.$$

Since x, $x_N \in l_\infty$ it follows that, if

$$\epsilon^{(n)} := T_n x := g^{(n)} * x \quad \text{and} \quad \epsilon_N^{(n)} := T_n x_N := g^{(n)} * x_N,$$

then

$$\epsilon^{(n)}(k) = \sum_{m=0}^{n} g_m^{(n)} x(k-m), \quad k = 0, \pm 1, \pm 2, \ldots, \tag{4.6b}$$

and

$$\epsilon_N^{(n)}(k) = \sum_{m=0}^{n} g_m^{(n)} x_N(k-m), \quad k = 0, \pm 1, \pm 2, \ldots \tag{4.6c}$$

By the definition of x_N and (4.6b), (4.6c) we obtain

$$\epsilon^{(n)}(k) = \epsilon_N^{(n)}(k) \quad \text{for } n \leqslant k \leqslant N-1.$$

Therefore by Lemma 4(B) it follows that

$$G_n(\mathrm{e}^{\mathrm{i}\omega_j}) = \sum_{k=n}^{n+2I} A_{j,k}^{(n)} \epsilon^{(n)}(k) = \sum_{k=n}^{n+2I} A_{j,k}^{(n)} \epsilon_N^{(n)}(k), \quad -I \leqslant k \leqslant I,$$

provided $n + 2I \leqslant N$. Thus

$$\left| G_n(\mathrm{e}^{\mathrm{i}\omega_j}) \right| \leqslant \max \left| A_{j,k}^{(n)} \right| \sum_{k=n}^{n+2I} \left| \epsilon_N^{(n)}(k) \right|, \quad -I \leqslant k \leqslant I,$$

provided $1 \leqslant n \leqslant N - 2I$. Here we hold N fixed and $A := \max \ A_{j,k}^{(n)} \,|$, where j, k and n take on the allowed values. □

Remarks. The relations (2.27) and (4.1) (in Theorem 3) suggest that the numbers $|G_n(\mathrm{e}^{\mathrm{i}\omega_j})| = |\rho_n(\mathrm{e}^{\mathrm{i}\omega_j})|$ get smaller as n increases. This suggests a close proximity to the $\mathrm{e}^{\mathrm{i}\omega_j}$ of zeros of $\rho_n(z)$. Additionally, in Theorem 2 if we let x_N be a signal of the form (3.1), then $H_n(z)$ minimizes $\| \delta - (h^{(n)} * x_N) \|_2$. Therefore, since $1 - H_n(z) X_N(z)$ is the Z-transform of $\delta - (h^{(n)} * x_N)$ where

$$X_N(z) := \sum_{k=0}^{N-1} x_N(k) z^{-k} = \sum_{k=0}^{N-1} x(k) z^{-k}, \tag{4.9}$$

one can expect $1/H_n(z)$ to be a good approximation of $X_N(z)$ for large n; that is,

$$X_N(z) \approx \frac{1}{H_n(z)} = \frac{\beta_n^2}{x_N(0)} \frac{z^n}{\rho_n(\psi_N; z)}. \tag{4.10}$$

Setting $X_\infty(z) := \sum_{k=0}^{\infty} x(k) z^{-k}$ and using (3.1), we obtain, for $z| > 1$,

$$X_\infty(z) = \sum_{k=0}^{\infty}\left[\sum_{j=-I}^{I} \alpha_j \, \mathrm{e}^{\mathrm{i}\omega_j k}\right] z^{-k}$$

$$\sum_{j=-I}^{I} \alpha_j \sum_{k=0}^{\infty}\left(\frac{\mathrm{e}^{\mathrm{i}\omega_j}}{z}\right)^k = \sum_{j=-I}^{I} \frac{\alpha_j z}{z-\mathrm{e}^{\mathrm{i}\omega_j}}. \tag{4.11}$$

Since $X_\infty(z)$ is a rational function of z with simple poles at the points $\mathrm{e}^{\mathrm{i}\omega_j}$, $-I \leqslant j \leqslant I$, and since $X_N(z) \to X_\infty(z)$ as $N \to \infty$, (4.10) suggests that there will be zeros of $\rho_n(z)$ near the $\mathrm{e}^{\mathrm{i}\omega_j}$ when n is large.

By a theorem of Jentzsch [25, Section 7.8, p.238] we can easily deduce the fact that, for every $\epsilon > 0$, there exists an N such that $X_N(z)$ has a zero in the annulus $1-\epsilon < |z| < 1+\epsilon$. The following theorem is an extension of this result.

Theorem 5. *Let $x = \{x(k)\}$ be a signal of the form* (3.1) *and let*

$$X_N(z) = \sum_{k=0}^{N-1} x(k) z^{-k}, \quad x(0) \neq 0. \tag{4.12}$$

Then for each $\epsilon > 0$, there exists an integer $N(\epsilon)$ such that, for every $N \geqslant N(\epsilon)$, all but at most $4I$ zeros of $X_N(z)$ lie in the annulus $1-\epsilon < |z| < 1+\epsilon$.

Proof. Observe that

$$s_N(z) := X_N\left(\frac{1}{z}\right) = \sum_{j=-I}^{I} \alpha_j \frac{1-(z\,\mathrm{e}^{\mathrm{i}\omega_j})^N}{1-z\,\mathrm{e}^{\mathrm{i}\omega_j}} \tag{4.13}$$

is the $(N-1)$th Taylor section of

$$F(z) := X_\infty\left(\frac{1}{z}\right) = \sum_{j=-I}^{I} \frac{\alpha_j}{1-z\,\mathrm{e}^{\mathrm{i}\omega_j}}, \tag{4.14}$$

so that $\{s_n(z)\}$ converges locally uniformly in $|z| < 1$. Let ρ be chosen such that $1/(1+\epsilon) < \rho < 1$ and $F(z) \neq 0$ for z on the circle C defined by $|z| = \rho$. Let $M := \min[\,|F(z)| : |z| = \rho]$. Then $M > 0$ and there exists an integer $N_0(\epsilon)$ such that $|s_N(z)| \geqslant |F(z)| - |F(z) - s_N(z)| \geqslant M - \frac{1}{2}M = \frac{1}{2}M$ for $|z| = \rho$ and $N \geqslant N_0(\epsilon)$. The rational functions $F(z)$ and $s_N(z)$ are all analytic in $|z| < 1$ (the poles of F occurring on the unit circle $|z| = 1$). It follows that $\{s_N'(z)/s_N(z)\}$ converges to $F'(z)/F(z)$ uniformly on C and hence

$$\lim_{N\to\infty} \frac{1}{2\pi\mathrm{i}} \int_C \frac{s_N'(z)}{s_N(z)}\,\mathrm{d}z = \frac{1}{2\pi\mathrm{i}} \int_C \frac{F'(z)}{F(z)}\,\mathrm{d}z.$$

Thus for every $N \geqslant N_0(\epsilon)$, $s_N(z)$ has the same number of zeros in $|z| < \rho$ as $F(z)$. But the rational function $F(z)$ has at most $2I$ zeros in $\mathbb{C}$, and hence in $|z| \leqslant 1/(1+\epsilon)$. Since

$\rho > 1/(1+\epsilon)$, we obtain *Assertion* (a): *For every* $N \geqslant N_0(\epsilon)$, $s_N(z)$ *has at most* $2I$ *zeros in* $|z| \leqslant 1/(1+\epsilon)$, *if the zeros are counted according to their multiplicities.*

Now we consider the functions

$$t_N(z) := \frac{s_N(z)}{z^N}, \quad N = 1, 2, 3, \ldots, \tag{4.15}$$

which are analytic and uniformly bounded on every closed subset of $E := \bar{\mathbb{C}} \setminus [z : |z| \leqslant 1]$, where $\bar{\mathbb{C}} := \mathbb{C} \cup [\infty]$. Furthermore, every limit function of the normal family $\{t_N(z)\}_1^\infty$ in E is of the form

$$\sum_{j=-I}^{I} \frac{\alpha_j \lambda_j}{1 - z\,\mathrm{e}^{\mathrm{i}\omega_j}}, \tag{4.16}$$

where the λ_j are constants of unit modulus. In fact, if K is a closed subset of E and $\{t_{N_k}(z)\}_1^\infty$ is a subsequence converging for $z \in K$, then by (4.13) and (4.15),

$$\lim_{N_k \to \infty} t_{N_k}(z) = -\lim_{N_k \to \infty} \sum_{j=-I}^{I} \frac{\alpha_j (\mathrm{e}^{\mathrm{i}\omega_j})^{N_k}}{1 - z\,\mathrm{e}^{\mathrm{i}\omega_j}}, \quad z \in K.$$

Since $\{(\mathrm{e}^{\mathrm{i}\omega_1})^{N_k}\}$ is bounded, there exists a subsequence $\{(\mathrm{e}^{\mathrm{i}\omega_1})^{N_{k,1}}\}$ converging to a limit λ_1, with $|\lambda_1| = 1$. Continuing in this manner to extract converging subsequences for each ω_j from the preceding subsequences, we obtain the λ_j in (4.16). Clearly every function of the form (4.16) has at most $2I+1$ zeros in $\bar{\mathbb{C}}$ (including the zero at $z = \infty$), and hence in $|z| \geqslant 1/(1-\epsilon)$. We now prove *Assertion* (b): *There exists an* $N_1(\epsilon)$ *such that, for every* $N \geqslant N_1(\epsilon)$, $t_N(z)$ *has at most* $2I+1$ *zeros* (*and hence* $s_N(z)$ *has at most* $2I$ *zeros*) *in* $|z| \geqslant 1/(1-\epsilon)$. For, assume that there exists a subsequence $\{t_{N_i}(z)\}$ such that, for every N_i, $t_{N_i}(z)$ has at least $2I+2$ zeros in $|z| \geqslant 1/(1-\epsilon)$. Then there exists a subsequence $\{t_{N_{i_j}}(z)\}$ of the normal family $\{t_{N_i}(z)\}$ converging locally uniformly on every closed subset K of E to a rational function of the form (4.16). By the same type of argument used above for $\{s_N(z)\}$, we can conclude that there exists an $N_2(\epsilon)$ such that, for every $N_{i_j} \geqslant N_2(\epsilon)$, $t_{N_{i_j}}(z)$ has at most $2I+1$ zeros in $|z| \geqslant 1/(1-\epsilon)$. This contradicts the above assumption and therefore proves Assertion (b).

We have shown that for $N \geqslant N(\epsilon) := \max(N_0(\epsilon), N_1(\epsilon))$, there are at most $4I$ zeros of $s_N(z)$ that are not in the annulus $1/(1+\epsilon) \leqslant |z| \leqslant 1/(1-\epsilon)$, from which the theorem follows. □

The following two theorems give information concerning the asymptotic behavior of the zeros of $\rho_n(\psi_N, z)$ which has been observed in Figs. 1 and 2.

Theorem 6. *Let* $x = \{x(k)\}$ *be a given signal of the form* (3.1), *and let* $N \geqslant 1$ *be fixed. Let* r_N *denote the distance from the unit circle to the nearest zero of* $X_N(z)$ *defined by* (4.12). *Then the zeros of the Szegö polynomials* $\rho_n(\psi_N; z)$ *will be asymptotically* (*as* $n \to \infty$) *equally spaced on the circle*

$$z| = \frac{1}{1 + r_N}$$

for a subsequence of integers n.

Remarks. (1) A proof of Theorem 6 follows from a recent result of Mhaskar and Saff [17,21] and the observation that $|z| = 1 + r_N$ is the largest disk centered at the origin throughout which the function $[D_N(z)]^{-1}$ is analytic.

(2) By the term, "asymptotically (as $n \to \infty$) equally spaced" used in the theorem we mean that the normalized distributions associated with the zeros of $\rho_n(\psi_N; z)$ converge in the weak-star topology to the uniform distribution on the circle $|z| = 1/(1 + r_N)$ (cf. [15]).

(3) Since from Theorem 5 we have $r_N \to 0$ as $N \to \infty$, we see from Theorem 6 that, for certain n and N suitably large, $\rho_n(\psi_N; z)$ will have zeros arbitrarily close to the unit circle.

Theorem 7. *Let $x = \{x(k)\}$ be a given signal of the form* (3.1), *let $x_N = \{x_N(k)\}$ denote its N-truncated, causal signal and let $X_N(z)$ and $\psi_N(\theta)$ be defined by* (2.2) *and* (2.3). *Then, as $N \to \infty$,*

$$\frac{1}{N}\,\mathrm{d}\psi_N(\theta) = \frac{1}{2\pi N}\left|X_N(\mathrm{e}^{\mathrm{i}\theta})\right|^2 \mathrm{d}\theta \xrightarrow{*} \sum_{j=-I}^{I} |\alpha_j|^2 \delta_{\mathrm{e}^{\mathrm{i}\omega_j}}, \tag{4.17}$$

where δ_w denotes the delta distribution with mass 1 *at $z = w$ and the convergence in* (4.17) *is in the sense of the weak-star topology.*

Proof. It suffices to show that for every function $f(z)$ continuous on $|z| = 1$ we have

$$\lim_{N\to\infty} \frac{1}{N}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\psi_N(\theta) = \sum_{j=-I}^{I} |\alpha_j|^2 f(\mathrm{e}^{\mathrm{i}\omega_j}).$$

We let $\epsilon > 0$ be chosen so that the interval $[\omega_s - \epsilon, \omega_s + \epsilon]$ does not contain ω_j for $j \neq s$.

$$\begin{aligned}
&\frac{1}{N}\int_{\omega_s-\epsilon}^{\omega_s+\epsilon} f(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\psi_N(\theta)\\
&= \frac{1}{2\pi N}\int_{\omega_s-\epsilon}^{\omega_s+\epsilon} f(\mathrm{e}^{\mathrm{i}\theta})\left|X_N(\mathrm{e}^{\mathrm{i}\theta})\right|^2 \mathrm{d}\theta\\
&= \frac{1}{2\pi N}\int_{\omega_s-\epsilon}^{\omega_s+\epsilon} f(\mathrm{e}^{\mathrm{i}\theta})\left|\sum_{j=-I}^{I}\alpha_j \frac{1-\mathrm{e}^{\mathrm{i}N\omega_j}z^{-N}}{1-\mathrm{e}^{\mathrm{i}\omega_j}z^{-1}}\right|^2 \mathrm{d}\theta \quad (z = \mathrm{e}^{\mathrm{i}\theta})\\
&= \frac{1}{2\pi N}\int_{\omega_s-\epsilon}^{\omega_s+\epsilon} f(\mathrm{e}^{\mathrm{i}\theta})\left(\sum_{j=-I}^{I}\alpha_j \frac{1-\mathrm{e}^{\mathrm{i}N(\omega_j-\theta)}}{1-\mathrm{e}^{\mathrm{i}(\omega_j-\theta)}}\right)\left(\sum_{m=-I}^{I}\alpha_m \frac{1-\mathrm{e}^{-\mathrm{i}N(\omega_m-\theta)}}{1-\mathrm{e}^{-\mathrm{i}(\omega_m-\theta)}}\right)\mathrm{d}\theta
\end{aligned} \tag{4.19}$$

For the second equality in (4.19) we have used the summation formula for terminating geometric series. Consider the term in (4.19) when $j = m = s$: that is,

$$\begin{aligned}
&\frac{1}{2\pi N}\int_{\omega_s-\epsilon}^{\omega_s+\epsilon} f(\mathrm{e}^{\mathrm{i}\theta})\,|\alpha_s|^2\left|\frac{1-\mathrm{e}^{\mathrm{i}N(\omega_s-\theta)}}{1-\mathrm{e}^{\mathrm{i}(\omega_s-\theta)}}\right|^2 \mathrm{d}\theta\\
&= \frac{|\alpha_s|^2}{2\pi N}\int_{-\epsilon}^{\epsilon} f(\mathrm{e}^{\mathrm{i}(\omega_s-\theta)})\left|\frac{1-\mathrm{e}^{\mathrm{i}N\theta}}{1-\mathrm{e}^{\mathrm{i}\theta}}\right|^2 \mathrm{d}\theta = \frac{|\alpha_s|^2}{2\pi N}\int_{-\epsilon}^{\epsilon} f(\mathrm{e}^{\mathrm{i}(\omega_s-\theta)})\left(\frac{\sin\frac{1}{2}N\theta}{\sin\frac{1}{2}\theta}\right)^2 \mathrm{d}\theta.
\end{aligned}$$

From well-known properties of the Fejér kernel [8] we have

$$\lim_{N\to\infty}\frac{|\alpha_s|^2}{2\pi N}\int_{-\epsilon}^{\epsilon} f(\mathrm{e}^{\mathrm{i}(\omega_s-\theta)})\left(\frac{\sin\frac{1}{2}N\theta}{\sin\frac{1}{2}\theta}\right)^2 \mathrm{d}\theta = |\alpha_s|^2 f(\mathrm{e}^{\mathrm{i}\omega_s})$$

Furthermore, the remaining terms on the right-hand side of (4.19) are of order $O(1/N)$ if neither j nor m equals s and of order $O(1/\sqrt{N})$ if j or m (but not both) equals s. For the latter estimate we apply the Schwarz inequality. Hence

$$\lim_{N\to\infty} \frac{1}{N}\int_{\omega_s-\epsilon}^{\omega_s+\epsilon} f(e^{i\theta})\, d\psi_N(\theta) = |\alpha_s|^2 f(e^{i\omega_s})$$

and so (4.18) follows by suitably partitioning the interval $[-\pi, \pi]$. □

Remark. It follows from (4.18) that, for N large,

$$\int_{-\pi}^{\pi} \left|\rho_n(\psi_N, e^{i\theta})\right|^2 d\psi_N(\theta) \approx \sum_{j=-I}^{I} N|\alpha_j|^2 \left|\rho_n(\psi_N, e^{i\omega_j})\right|^2. \tag{4.20}$$

Since the Szegö polynomial $\rho_n(\psi_N, z)$ has the property that it minimizes the integral on the left-hand side of (4.20) with respect to all monic polynomials of degree n, the right-hand side of (4.20) suggests that the critical points $e^{\pm i\omega_j}$ that are approximated best by a zero of ρ_n will be those associated with the largest $|\alpha_j|^2$. This is indeed the case in the example given in Section 3 (see Table 3 and Figs. 1 and 2). For that example, $|\alpha_1| = |\alpha_2| = |\alpha_3| = 1$, $|\alpha_4| = 10$, and $\omega_4 = \frac{3}{4}\pi$. It is clear that the critical points $e^{\pm 3\pi i/4}$ are approximated best by zeros of the ρ_n.

5. Related Padé approximants

In the final section we investigate properties of the Padé approximants for the function

$$F_N(z) := \frac{1}{D_N(z)D_N^*(z)}, \tag{5.1}$$

where $D_N(z)$ denotes the Szegö function (see (2.16)) and $D_N^*(z) := z^M\overline{D_N(1/\bar{z})}$, $M :=$ degree of $D_N(z)$.

Theorem 8. *Let $x_N = \{x_N(m)\}$ be a given N-truncated, causal signal* (2.1), *and let $\psi_N(\theta)$ and $D_N(z)$ denote the associated distribution function* (2.3) *and Szegö function* (2.16), *respectively. Let M denote the degree of $D_N(z)$, $0 \leqslant M \leqslant N-1$, and let $[n/M]$ be the (classical) Padé approximant of type (n, M) to the rational function $F_N(z)$ (with $2M$ poles) defined by* (5.1). *Then there exists a polynomial $Q_{n,M}(z) \in \mathscr{P}_M$ (of degree $\leqslant M$) such that*

$$[n/M](z) = \frac{\rho_n^*(\psi_N, z)}{Q_{n,M}(z)}, \tag{5.2}$$

where $\rho_n^* = \rho_n^*(\psi_N; z)$ denotes the reciprocal Szegö polynomial (2.7) *with respect to $\psi_N(\theta)$.*

Proof. Let $P_n \in \mathscr{P}_n$ and $Q_M\ (\not\equiv 0) \in \mathscr{P}_M$ satisfy

$$Q_M F_N - P_n = O(z^{M+n+1}), \quad \text{as } z \to 0.$$

Then $Q_M - P_n D_N D_N^* = O(z^{M+n+1})$ and hence

$$\int_{|z|=1} \frac{Q_M(z)}{z^{M+k}}\, dz - \int_{|z|=1} \frac{P_n(z)D_N(z)D_N^*(z)}{z^{M+k}}\, dz = 0 \quad \text{for } k \leqslant n+1$$

We also have

$$\int_{|z|=1} \frac{Q_M(z)}{z^{M+k}}\, dz = 0 \quad \text{for } k \geqslant 2$$

Therefore

$$\int_{|z|=1} \frac{P_n(z) D_N(z) D_N^*(z)}{z^{M+k}}\, dz = 0 \quad \text{for } 2 \leqslant k \leqslant n+1.$$

For $|z| = 1$, we have $z = e^{i\theta}$, $dz = iz\, d\theta$ and $D_N^*(z) = z^M \overline{D_N(z)}$ and hence

$$\int_{-\pi}^{\pi} \frac{P_n(z) D_N(z) \overline{D_N(z)}}{z^{k-1}}\, d\theta = 0 \quad \text{for } 2 \leqslant k \leqslant n+1 \quad (z = e^{i\theta}).$$

Thus by (2.15c)

$$\int_{-\pi}^{\pi} \frac{P_n(z)}{z^{k-1}} |D_N(z)|^2\, d\theta = 2\pi \int_{-\pi}^{\pi} \frac{P_n(z)}{z^{k-1}}\, d\psi_N(\theta) = 0 \quad \text{for } 2 \leqslant k \leqslant n+1 \quad (z = e^{i\theta}).$$

It follows that $\hat{P}_n(z) := z^n \overline{P_n(1/\bar{z})} \in \mathscr{P}_n$ satisfies

$$0 = \int_{-\pi}^{\pi} \overline{P_n(z)} z^{k-1}\, d\psi_N(\theta) = \int_{-\pi}^{\pi} \hat{P}_n(z) z^{k-1-n}\, d\psi_N(\theta),$$

$$\text{for} \quad 2 \leqslant k \leqslant n+1 \quad (z = e^{i\theta}).$$

Therefore,

$$\langle \hat{P}_n, z^m \rangle = \int_{-\pi}^{\pi} \hat{P}_n(z) \overline{z^m}\, d\psi_N(\theta) = 0 \quad \text{for } 0 \leqslant m \leqslant n-1 \quad (z = e^{i\theta}),$$

and hence $\hat{P}_n(z) = c_n \rho_n(z)$ for some constant c_n. Thus

$$\rho_n^*(z) := z^n \overline{\rho_n(1/\bar{z})} = \bar{c}_n^{-1} z^n \overline{\hat{P}_n(1/\bar{z})} = \bar{c}_n^{-1} P_n(z)$$

provided $c_n \neq 0$; that is, provided $P_n(z) \not\equiv 0$. If $P_n(z) \equiv 0$, then by (5.3), $Q_M = O(z^{M+n+1})$ which contradicts the fact that $0 \not\equiv Q_M \in \mathscr{P}_M$. Thus $Q_{n,M} = \bar{c}_n^{-1} Q_M$. □

Corollary 9. *With the hypotheses of Theorem* 8,

$$\lim_{n\to\infty} [n/M](z) = \lim_{n\to\infty} \frac{\rho_n^*(z)}{Q_{n,M}(z)} = \frac{1}{D_N(z) D_N^*(z)}$$

locally uniformly on the set $|z| < R_N \setminus \{$*zeros of* $D_N^*(z)\}$, *where*

$$R_N := \min_{k=1}^{M} [\, |\zeta_k| \colon D_N(\zeta_k) = 0] \geqslant 1.$$

Moreover, the M zeros of $Q_{n,M}(z)$ *approach the M zeros of* $D_N^*(z)$.

Proof. We apply an extended version of the Montessus de Ballore theorem. The extension appears in an implicit form in [18,20] (see also [2]) and explicitly in [22]. □

We conclude with the following remarks. Since $\rho_n^*(0) = 1$, it follows from Corollary 9 that

$$\lim_{n\to\infty} \frac{1}{Q_{n,M}(0)} = \frac{1}{D_N(0) D_N^*(0)} \quad \text{and} \quad \lim_{n\to\infty} Q_{n,M}(z) = c D_N^*(z),$$

for some constant c. Hence $D_N(0)D_N^*(0) = cD_N^*(0)$ and so $c = D_N(0)$. Therefore by this and Corollary 9

$$\lim_{n\to\infty} \rho_n^*(z) = \lim_{n\to\infty} Q_{n,M}(z \frac{\rho_n^*(\psi_N;\, z)}{Q_{n,M}(z)} = \frac{D_N(0)D_N^*(z)}{D_N(z)D_N^*(z)} = \frac{D_N(0)}{D_N(z)};$$

that is,

$$\lim_{n\to\infty} \rho_n^*(\psi_N;\, z) = \frac{D_N(0)}{D_N(z)} \quad \text{for } |z| < R_N. \tag{5.6}$$

This result is in agreement with (2.15a) and demonstrates the connection between the Montessus de Ballore theorem for Padé approximants and the Szegö theory of asymptotics for orthogonal polynomials with respect to nonnegative trigonometric weights.

Acknowledgement

The authors wish to thank Kevin D. Jones for able computing assistance related to the tables and graphs.

References

[1] N.E. Akhiezer, *The Classical Moment Problem and Some Related Questions in Analysis* (translated by N. Kemmer) (Hafner, New York, 1965).
[2] G.A. Baker Jr, *Essentials of Padé Approximants* (Academic Press, New York, 1975).
[3] C.K. Chui, Approximation by double least square inverses, *J. Math. Anal. Appl.* **75** (1980) 149–163.
[4] P.L. Duren, *Theory of H^p Spaces* (Academic Press, New York, 1970).
[5] Ya.L. Geronimus, Polynomials orthogonal on a circle and their applications, *Amer. Math. Soc. Transl.* **104** (1954).
[6] Ya.L. Geronimus, *Orthogonal Polynomials* (Consultants Bureau, New York, 1961).
[7] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications* (Univ. of California Press, Berkeley, CA, 1958).
[8] D. Jackson, *Fourier Series and Orthogonal Polynomials*, Carus Math. Monographs **6** (Math. Assoc. America, Washington, DC, 1941).
[9] W.B. Jones and O. Njåstad, Applications of Szegö polynomials to digital signal processing, *Rocky Mountain J. Math.*, to appear.
[10] W.B. Jones, O. Njåstad and W.J. Thron, Continued fractions associated with the trigonometric and other strong moment problems, *Constr. Approx.* **2** (1986) 197–211.
[11] W.B. Jones, O. Njåstad and W.J. Thron, Schur fractions, Perron–Carathéodory fractions and Szegö polynomials, a survey, in: W.J. Thron, Ed., *Analytic Theory of Continued Fractions II*, Lecture Notes in Mathematics **1199** (Springer, New York, 1986) 127–158.
[12] W.B. Jones, O. Njåstad and W.J. Thron, Moment theory, orthogonal polynomials, quadrature and continued fractions associated with the unit circle, *Bull. London Math. Soc.* **21** (1989) 113–152.
[13] J.D. Markel and A.N. Gray Jr, *Linear Prediction of Speech* (Springer, New York, 1976).
[14] A.I. Markushevich, *Theory of Functions of a Complex Variable, Vol. II* (translated from Russian by R.A. Silverman) (Prentice-Hall, Englewood Cliffs, NJ, 1965).
[15] A. Máté, P. Nevai and V. Totik, Asymptotics for the ratio of leading coefficients of orthogonal polynomials on the unit circle, *Constr. Approx.* **1** (1985) 63–69.
[16] A. Máté, P. Nevai and V. Totik, Extensions of Szegö's theory of orthogonal polynomials, II, III, *Constr. Approx.* **3** (1987) 51–72, 73–96.

[17] H.N. Mhaskar and E.B. Saff, On the distribution of zeros of polynomials orthogonal on the unit circle, *J. Approx. Theory*, to appear.

[18] B. Parlett, Global convergence of the basic QR algorithm on Hessenberg matrices, *Math. Comp.* **22** (1968) 803–817.

[19] W. Rudin, *Real and Complex Analysis* (McGraw-Hill, New York, 2nd ed., 1974).

[20] E.B. Saff, On the row convergence of the Walsh array for meromorphic functions, *Trans. Amer. Math. Soc.* **146** (1969) 241–257.

[21] E.B. Saff, Orthogonal polynomials from a complex perspective, I.C.M. #89-022, Dept. Math., Univ. South Florida, Tampa, FL.

[22] A. Sidi, Quantitative and constructive aspects of the generalized Koenig's and de Montessus's theorems for Padé approximants, *J. Comput. Appl. Math.* **29** (3) (1990) 257–291.

[23] J. Szabados, On some problems connected with orthogonal polynomials on the complex unit circle, *Acta Math. Acad. Sci. Hungar.* **33** (1979) 197–210.

[24] G. Szegö, *Orthogonal Polynomials*, Colloquium Publications **23** (Amer. Math. Soc., New York, 1959).

[25] E.C. Titchmarsh, *The Theory of Functions* (Oxford Univ. Press, Oxford, 2nd ed., 1939).