2001

# Some Limit Theorems for Szego Polynomials

Michael Joseph Arciero
*University of Rhode Island*

SOME LIMIT THEOREMS FOR SZEGÖ POLYNOMIALS

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

MATHEMATICS

MICHAEL JOSEPH ARCIERO

UNIVERSITY OF RHODE ISLAND

2001

DOCTOR OF PHILOSOPHY DISSERTATION

OF

MICHAEL JOSEPH ARCIERO

APPROVED:

Dissertation Committee:

Major Professor

DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND

2001

## Abstract

We investigate a variety of convergence phenomena for measures on the unit circle associated with certain discrete time stationary stochastic processes, and for the class of *Szegö* polynomials orthogonal with respect to such measures.

*Szegö* polynomials, which form the basis of *autoregressive* (AR) methods in spectral analysis, are not uniquely defined when the degree is less than the number of points on which the spectral measure is supported; that is, when the spectral measure corresponds to a sum of complex sinusoids, the number of which is less than the degree. We consider the asymptotic behavior of Szegö polynomials of fixed degree for certain sequences of measures which converge weakly to such a sum of point masses.

The sequence of measures can be formed in various ways, one of which is by convolving point mass sums with approximate identities, or *kernels*. In signal processing applications, this corresponds to "windowing" a signal composed of complex sinusoids. The Poisson and Fejër kernels are considered. Another way to form the measures is to add an absolutely continuous measure to a sum of point masses, thus obtaining a spectral measure for sinusoids with additive *noise*, where the noise coloration is described by the density of the absolutely continuous part. We characterize a limit polynomial for several different classes of sequences of measures. Some special cases are used to interpret research done by others in the field.

Situations where the polynomial degree approaches infinity are considered for fixed measures with a rational spectral density. These measures are the spectral measures for autoregressive *moving average* (ARMA) random processes. We study the asymptotic behaviors of the *reflection coefficients*, or constant terms, of the polynomials, and the *zero-distribution measures*, which consist of point masses at each of the polynomial zeros. These analyses help describe the behavior of the "non-signal" zeros observed in some signal processing situations.

## Acknowledgements

The candidate wishes to thank his advisor Lewis Pakula, and his parents, Robert and Barbara Arciero, to whom this thesis is dedicated.

CONTENTS

LIST OF FIGURES

# 1 Introduction

Given a measure on the unit circle there is an associated sequence of polynomials in the complex variable $z$, called *Szegö* polynomials, which are orthogonal on the unit circle with respect to the measure. Szegö polynomials minimize the integral of the squared modulus, over monic polynomials of degree $k$, with respect to the measure. These polynomials have many applications in analysis and applied mathematics, and their properties have been studied extensively. The classic references for much of the early work include Grenander and Szegö ([GS]), and Geronimus ([G1], [G2]).

In recent decades there has been much interest in Szegö polynomials by researchers in signal processing and control theory due to their intimate connection with problems in linear prediction and spectral estimation. The book by Kay ([K]) includes many examples of their use, as well as an abundance of references to the related engineering literature.

Szegö polynomials are not defined when the degree of the polynomial exceeds the number of points on which the measure is supported. Such "point mass" measures can arise, for example, as the *spectral measure* of signals consisting of a finite number of complex sinusoids. However, we can approximate such measures by other absolutely continuous measures, for which the Szegö polynomials *are* defined, and study the polynomials as the approximation improves. Sections 3 and 4 deal with the questions of uniqueness and existence of limits for Szegö polynomials of fixed degree with respect to a sequence, or family, of measures which converges, in some sense, to a sum of point mass measures. A sequence of absolutely continuous measures cannot converge to a finite sum of point masses in the usual sense; that is, in total variation norm on measures. We consider a weaker type of convergence, which can be seen to arise naturally in applications.

Suppose we have a sequence of measures converging in some sense to a sum of point mass measures. Fix $k$, with $k$ greater than the number of point masses, and consider the sequence of Szegö polynomials of degree $k$ corresponding to the converging measures. The convergence of a sequence of measures, in either the usual sense or in the weaker sense we consider, does *not* guarantee convergence of the associated Szegö polynomials if the degree $k$ is less than the number of point masses. We show that any limit point, in polynomial space, of these Szegö polynomials must have zeros at the point mass locations. Any limit point will thus have an " extra" factor with degree equal to the difference of the polynomial degree $k$ and the number of point masses. One such condition under which a unique limit *does* exist is that the Fourier coefficients of the sequence of measures parametrized by $h$, where $h$ approaches zero, depend analytically on $h$.

The two main results of Section 3.2 deal with families of measures formed by convolving the familiar Poisson and Fejér kernels with point mass measures. These kernels are examples of *approximate identities* whose Fourier coefficients have analytic dependence on $h$. In Theorems 3.4 and 3.5 we characterize the limit polynomial for convolution with the Poisson and Fejér kernels, respectively, and show that, in each case, the extra factor is actually a Szegö polynomial with respect to an absolutely continuous measure, which we specify. Furthermore, we obtain the *same* limit polynomial in each case. Comparison of the analytic dependence of the Fourier coefficients of these kernels on $h$, in particular, the agreement of the linear terms in $h$, suggest that this might hold, but the proofs require detailed analysis of exploiting two properties: The orthogonality of the Szegö polynomials and the rate of convergence of zeros of the Szegö polynomials to the point mass locations.

Recent work involving sequences of convergent measures and the associated Szegö polynomials includes that of Pan and Saff ([PS]), Jones, Njåsted and Saff ([JNS]), and Jones, Njåsted and Waadeland ([JNW]). In Section 3.2.2, we use arguments of Pan and Saff to address the convergence rate of the zeros for a general situation where the sequence of measures satisfies two properties, one of which is analytic dependence on $h$. The cases of convolution with the Poisson and Fejér kernels are then considered separately.

We also consider, in Section 4, families of measures formed by adding a multiple, $h$, of a fixed absolutely continuous measure to a sum of point mass measures. We characterize limits of the Szegö polynomials of degree $k$, with $k$ greater than the number of point masses. Again we find that the extra factor is the Szegö polynomial with respect to an absolutely continuous measure. This situation differs from that of convolution of approximate identities with point masses, in that the sequence of measures here *does* converge in total variation norm. We give a method of constructing sequences of measures that converge in total variation norm but whose associated Szegö polynomials do *not* converge.

Another situation which has received attention is where the degree, $k$, of the Szegö polynomials with respect to a *fixed* measure, approaches infinity. The behavior of the constant term, or *reflection coefficient*, $R_k$, of the Szegö polynomial of degree $k$ has been of particular interest. Since $|R_k|^{1/k}$ is the geometric mean of the modulus of the zeros of the Szegö polynomial of degree $k$, information about the reflection coefficients can help describe the asymptotic behavior of the polynomial zeros. The work of Pakula ([P]), Nevai and Totik ([NT]), Saff ([S]), Petersen ([Pe2]), and others addresses this situation. As in the previously described case of polynomials of fixed degree with respect to sequences of measures, the question of existence of limits is a central concern.

In Section 5 we consider absolutely continuous measures whose densities can be expressed as the squared modulus of a rational function on the unit circle. Such measures arise in applications as the spectral measures of certain random processes. Our main result here is Theorem 5.2, which is an extension of a result in [P], where measures whose densities can be expressed as the squared modulus of a *polynomial* are considered. A phenomenon which has been observed in the literature by Kumaresan, in [Ku], and others, is that the zeros of the Szegö polynomials with respect to certain measures appear to accumulate on a circle of a certain radius. This phenomenon is interpreted in [P] as the convergence of a sequence of measures related to the Szegö polynomials, and applies to a general class of measures which includes those with rational densities.

Theorem 5.2 states that, for measures with rational densities, under certain assumptions, $|R_k|^{1/k}$ approaches a limit as $k \to \infty$; this limit being the modulus of the largest zero of the numerator of the density. Results in [P] are then used to draw conclusions about the behavior of the zeros of the Szegö polynomial of degree $k$ as $k \to \infty$. The examples of Section 3.2.5 are special cases of Theorem 5.2. Here, they are also interpreted in the context of Theorems 3.4 and 3.5, and we attempt to make some connections between the situation with Szegö polynomials of fixed degree with respect to sequences of measures, and that of fixed measures with polynomial degree approaching infinity.

A related situation considered by Kumaresan and Tufts in [KT1] and [KT2], that of a signal consisting of damped exponentials, arises in modeling of speech. In Section 5.4, a sequence of measures is formed from a sum of damped exponentials. This sequence converges to a measure with rational density. This is interpreted in light of Theorem 5.2, and observations are made regarding the behavior of the zeros of the associated Szegö polynomials.

# 2 Background

## 2.1 The Moment Problem

The classical trigonometric moment problem (see, eg., [GS], [L], [A]) will serve as a starting point for our discussions. It is stated as follows:

Given a sequence $\{r_\ell\}_{\ell=-\infty}^{\infty}$ of complex numbers, when is a representation of the form

$$r_\ell = \int_{-\pi}^{\pi} \exp(i\ell\theta)\, d\mu(\theta) \tag{1}$$

possible for some positive measure $\mu$?

A necessary condition for (1) to hold is that $\{r_\ell\}$ must be a *positive semi-definite* sequence. That is, given any finite sequence $\{c_\ell\}$ of complex numbers we must have

$$\sum_{j,\ell} c_j \bar{c}_\ell r_{j-\ell} \geq 0. \tag{2}$$

This can be seen upon substitution of (1) into (2). Conversely, if $\{r_\ell\}$ is positive semi-definite there exists a unique positive measure $\mu$ satisfying (1). This result is often referred to as Herglotz's Theorem ([Ka]), but there are also proofs due to Caratheodory, Toeplitz, F. Riesz, and Krein ([A]).

Thus there is a one-to-one correspondence between positive semi-definite sequences and positive measures on the unit circle. The $r_\ell$ are the Fourier coefficients[1], or *trigonometric moments* of $\mu$; $r_\ell = \widehat{\mu}(\ell)$. We will henceforth regard $\mu$ as a measure on the unit circle.

## 2.2 Stationary Processes.

A *stochastic process* is a sequence, $\{X_n\}_{n=-\infty}^{\infty}$, of real or complex random variables on a probability space $(\Omega, M)$. We will assume that for all $n$,

$$\int_\Omega |X_n(\omega)|^2\, dM(\omega) \ < \infty. \tag{3}$$

Let $H_X$ denote the closure in $L^2(\Omega, M)$ of the span of $\{X_n\}$. The process is *wide-sense stationary* if for all integers $\ell$ and $n$,

$$r_\ell := \int_\Omega X_n(\omega) X_{n+\ell}(\omega)\, dM(\omega)$$

[1] The Fourier coefficients are sometimes defined with a normalization factor of $1/2\pi$, or $1/\sqrt{2\pi}$ in front of the integral in (1).

is independent of $n$ and

$$\int_\Omega X_n(\omega)\, dM(\omega) = 0$$

for all $n$. One can show that the sequence $r_\ell$ is positive semi-definite. Thus by Herglotz's theorem there exists a positive measure $\mu$ such that the representation (1) holds. This measure is called the *spectral measure* of the process. A Hilbert space isomorphism between $H_X$ and $L^2(d\mu)$ is established by extending the mapping $X_n(\omega) \to e^{in\theta}$ to all of $L^2(d\mu)$ using linearity and density arguments.

A *time series* , or discrete time *signal*, is a realization of a stochastic process, that is, a sequence $\{x_n\}$ where $x_j = X_j(\omega)$ for some fixed value of $\omega$. The spectral measure describes the spectral, or frequency content of the signal. For example, if $x_n = \sum_{j=1}^m \alpha_j e^{i(n\theta_j + \phi_j)}$ where the $\alpha_j$ are positive and tha $\phi_j$ are independent and unformly distributed on $[-\pi, \pi)$, then the associated spectral measure is $\sum_{j=1}^m \alpha_j \delta_{\theta_j}$, where $\delta_{\theta_j}$ is the point mass measure at $\theta_j$, corresponding to the fact that the signal comprises $m$ complex sampled sinusoids with frequencies $\theta_j$. On the other hand, if $\{X_n\}$ are independent identically distributed Gaussian random variables, the associated spectral measure is a multiple of the uniform Lebesgue measure, and the signal $\{x_n\}$ can be viewed as comprising a *white noise.*

## 2.3 Szegö Polynomials

### 2.3.1 Definition and Properties

Let $\Lambda_k$ denotes the space of monic polynomials of degree $k$, and let $\mu$ be a positive measure on $[-\pi, \pi)$. We will refer to the polynomials $P_k(z, \mu)$, $k = 1, 2, 3, \ldots$ in the complex variable $z$ of degree $k$ satisfying

$$\min_{p \in \Lambda_k} \int_{-\pi}^{\pi} |p(e^{i\theta})|^2 \, d\mu(\theta) = \int_{-\pi}^{\pi} |P_k(e^{i\theta}, \mu)|^2 \, d\mu(\theta) \tag{4}$$

as *Szegö* polynomials. The $P_k(z, \mu)$ are uniquely defined if $\mu$ is supported on more than $k$ points. This will be discussed further in subsequent sections. Szegö polynomials appear widely in the classical literature ([GS],[Ge],[Sz]), and have applications in signal processing, where $\mu$ is the spectral measure of some process ([DG],[JNS],[JN],[K]). The absolutely continuous measures $\mu_k$

defined by

$$d\mu_k(\theta) := \sigma_k^2 |P_k(e^{i\theta}, \mu)|^{-2} d\theta, \tag{5}$$

where $\sigma_k$ are normalization constants, are the basis for autoregressive (AR) spectral estimates. In the AR approach, one uses $\mu_k$ to estimate $\mu$. For $k$ large, $\mu_k$ is "close to" $\mu$ in the sense of the following well-known result, which can be found in [GS]

**Theorem:** *The measures $\mu_n$ in (5) converge to $\mu$ in the weak-star (denoted weak-*) sense [GS], [L]. That is,*

$$\lim_{k\to\infty} \int_{-\pi}^{\pi} f(e^{i\theta})\, d\mu_k(\theta) = \int_{-\pi}^{\pi} f(e^{i\theta})\, d\mu(\theta), \tag{6}$$

*for all continuous $f$ on the unit circle.*

Situations where (6) holds will be studied in Section 3. We remark that weak-* convergence is weaker than convergence in total variation norm on measures (see, for example, Theorem 3.1 and the remarks which follow). The above theorem can be proved using the characterization (23) of Section 3.1 and the Weiner-Kinchin Theorem, which can be found in [K].

Similarly, (4) can be seen to justify the AR approach in frequency estimation, where one uses the arguments of the largest zeros of $P_n(\mu, z)$ as frequency estimates [K]. Let $\mu$ be a measure of *mixed* type. That is

$$\mu = \sum_{j=1}^{M} a_j\, \delta_{\theta_j} + \gamma \quad . \tag{7}$$

where $\theta_j \in [-\pi, \pi)$, $\delta_{\theta_j}$ is the point mass at $\theta_j$ and $\gamma$ is absolutely continuous. The measure $\mu$ is then the spectral measure of a time series comprised of complex sinusoids with additive noise. In order to achieve the minimum in (4) one expects $P_k(\mu, z)$ to have zeros close to $e^{i\theta_j}$ for large $n$ ([S], [PS]). Indeed, if $\gamma = 0$ it is clear that any polynomial with zeros at $z = e^{i\theta_j}$ for $j = 1, 2, 3, ..., m$ will attain the minimum of zero in (4).

The property (4) can also be interpreted from the perspective of linear prediction. If we wish to estimate the random variable $X_n$ in the least squares sense using a linear combination of $X_0, X_1, ..., X_{n-1}$, we can write

$$||X_n - \sum_{\ell=1}^{n-1} a_{n-\ell} X_{n-\ell}||^2 = \int_{-\pi}^{\pi} |e^{in\theta} - \sum_{\ell=1}^{n-1} a_{n-\ell} e^{i(n-\ell)\theta}|^2\, d\mu(\theta).$$

From (4), we see that the prediction coefficients, $a_j$ coincide with the coefficients of $P_k$. In this context $P_k$ represents the *prediction error filter* of order $n$, and the minimum in (4) is called the *prediction error power*[2]

It is easy to see that the $P_k(z,\mu)$ are multiples of the orthonormal polynomials of degree $k$ with respect to $\mu$ obtained by performing the Gram-Schmidt procedure on $1, z, z^2, \ldots$. One simply expands an arbitrary $p \in \mathbf{\Lambda}_k$ in terms of these orthogonal basis polynomials and observes that the minimum in (4) is achieved for a multiple of the $k^{th}$ basis element. Thus $P_k(z,\mu) \perp \mathbf{\Lambda}_{k-1}$, and we have the following orthogonality property, which characterizes $P_k(z,\mu)$ (assuming that it is well-defined).

**Orthogonality Property:** If $p(z)$ is any polynomial of degree less than $k$, then

$$\int_{-\pi}^{\pi} P_k(e^{i\theta},\mu)\,\overline{p}(e^{i\theta})\,d\mu(\theta) = 0. \tag{8}$$

We will make extensive use of this property in the proofs of Theorems 3.4 and 3.5, which are the main results of Section 3.

Given a polynomial $p(z)$, of degree $k$, we define the *reverse* polynomial $p^*(z) := z^k\overline{p}(z^{-1})$, so that

$$P_k^*(z,\mu) := z^k\overline{P}_k(z^{-1},\mu). \tag{9}$$

Thus, for example, if $P_k(z,\mu) = \prod_{j=1}^{k}(z - z_j)$, then $P_k^*(z,\mu) = \prod_{j=1}^{k}(1 - z\overline{z}_j)$ and the zeros of $P_k^*$, are obtained from those of $P_k$ by reflection in the unit circle. It is well-known that if $\mu$ is supported on more than $k$ points, then $P_k(z,\mu)$ is defined and the zeros of $P_k(z,\mu)$ lie in the open unit disk. There are many proofs of this *minimum phase* property (see, for eg, [KP],[S],[L]). That the zeros lie in the closed disk is a consequence of Fejer's Convex Hull Theorem ([Ka]):

*The zeros of the polynomials orthogonal with respect to a measure are contained in the closed convex hull of the support of the measure.*

[2] In the engineering literature, the prediction error power is usually defined as

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

Suppose that $P_k(z,\mu) = \prod(z - w_j)$. Then $P_k^*(z,\mu) = \prod z(1/z - \overline{w}_j) = \prod(1 - z\overline{w}_j)$. Thus the zeros of $P_k^*$ are obtained from those of $P_k$ by reflecting them with respect to the unit circle, and therefore have modulus greater than 1. Note also that

$$|P_k(e^{i\theta},\mu)| = |P_k^*(e^{i\theta},\mu)|. \tag{10}$$

### 2.3.2 Representation of $P_k(z,\mu)$

Let $\mu$ be a finite measure on the unit circle, with moments $\widehat{\mu}(\ell)$. The Toeplitz matrix

$$C_n := \{\widehat{\mu}(i-j,n)\} = \begin{pmatrix} \widehat{\mu}(0) & \widehat{\mu}(1) & \dots & \widehat{\mu}(n) \\ \widehat{\mu}(-1) & \widehat{\mu}(0) & \dots & \widehat{\mu}(n-1) \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{\mu}(-n) & \widehat{\mu}(-n+1) & \dots & \widehat{\mu}(0) \end{pmatrix} \tag{11}$$

is positive semi-definite, and is strictly positive definite for all $n$ if $\log \mu'$ is integrable, where $\mu'$ is the density of the absolutely continuous part of $\mu$ (with respect to Lebesgue measure), and in which case $\mu$ is said to satisfy *Szegö's condition*. See [GS], [H], or [JNS] for further discussion. The following equivalent conditions are well known, and can be found in [JNS].

1. The measure $\mu$ satisfies Szegö's condition; that is,

$$\int \log \mu' \, d\theta > -\infty. \tag{12}$$

2. 

$$\lim_{k\to\infty} \int |P_k(z,\mu)|^2 \, d\mu > 0. \tag{13}$$

3. There exists an analytic function $g(z)$ with no zeros in $|z| < 1$ and $g(0) > 0$ such that

$$\lim_{r\to 1^-} |g(re^{i\theta})|^2 = \mu'(\theta) \quad \text{almost everywhere.} \tag{14}$$

**Spectral Factorization** The function $g(z)$ in (14) is sometimes called the Szegö function for $\mu$. This function is in the Hardy space $H^2$. Furthermore, $|g(re^{i\theta})|^2$ can be factored so that $g(z)$ has no zeros inside the unit circle; that is, $g$ is an *outer* function. This factorization is called the *spectral factorization* of $d\mu/d\theta$. See [GS] or [H] for further discussion. In particular, if $d\mu/d\theta$ is a positive trigonometric polynomial, it can be factored as the square of a polynomial in $z$ of the same degree. This result is due to Riesz and Fejér (see [GS], Sec. 1.12).

The matrix $C_n$ is often referred to as either the *autocorrelation*, or *ACF*, matrix. It is also referred to as the *covariance* matrix. This terminology is due to two methods used to estimate the $\widehat{\mu}(\ell)$ from the data; that is, from a realization of a random process with spectral measure $\mu$ (see, eg. [K] Ch. 7). In this context, the sequence of moments is often referred to as the autocorrelation (ACF) function. This terminology also follows naturally from the relations (1) and (3).

We shall denote, by $D_n(\mu)$, the determinant of $\{\widehat{\mu}(i-j,n)\}$:

$$D_n(\mu) := \det(\{\widehat{\mu}(i-j,n)\}) \tag{15}$$

By the above remarks, $D_n(\mu)$ is non-zero if $\mu$ satisfies Szegö's condition. On the other hand, if $\mu$ is supported on $m$ points, then $D_n(\mu) \neq 0$ for $n < m$, while $D_n(\mu) = 0$ for $n \geq m$.

If we replace the last row in the right-hand side of (11) with the vector $(1, z, z^2, ..., z^n)$, we get a matrix whose determinant is a polynomial in $z$. If we define

$$\mathcal{D}_n(z,\mu) := \begin{vmatrix} \widehat{\mu}(0) & \widehat{\mu}(1) & \dots & \widehat{\mu}(n) \\ \widehat{\mu}(-1) & \widehat{\mu}(0) & \dots & \widehat{\mu}(n-1) \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{\mu}(-n+1) & \widehat{\mu}(-n+2) & \dots & \widehat{\mu}(1) \\ 1 & z & \dots & z^n \end{vmatrix}, \tag{16}$$

we obtain the well-known representation ([GS], [JNS])

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

We shall refer to (17) as the *determinant representation* of $P_k(z,\mu)$.

The polynomials $P_k(z,\mu)$ can also be generated by the computationally efficient *Levinson's recursion* ([GS]):

$$P_k(z,\mu) = zP_{k-1}(z,\mu) + R_k(\mu)P^*_{k-1}(z,\mu) \tag{18}$$

where $P_0(z,\mu) = 1$, and the *reflection coefficients* $R_k(\mu)$ are the constant terms defined by

$$R_k(\mu) := P_k(0,\mu). \tag{19}$$

### Conventions

We adopt the conventions throughout this paper that $\zeta = e^{i\theta}$ denotes an arbitrary point on the unit circle, and that all integrals are over $[-\pi, \pi)$ unless indicated otherwise. We will also, as in Section 2.3, define the prediction error power as the minimum in (4), noting that this differs from the usual definition by a factor of $1/2\pi$.

# 3 Sequences of Measures

## 3.1 Introduction

Let $\delta_\theta$ be the point mass measure at $\theta$ and let $\mu$ be a weighted sum of $m$ of these,

$$\mu := \sum_{j=1}^{m} \alpha_j \delta_{\theta_j}, \tag{20}$$

with $\alpha_j > 0$ and the $\theta_j$ distinct. In the next sections we will consider absolutely continuous measures $\mu_h$, on the unit circle such that

$$\lim_{h \to 0} \mu_h = \sum_{j=1}^{m} \alpha_j \delta_{\theta_j}, \tag{21}$$

where the convergence is with respect to the weak-* topology on probability measures characterized in (6). For convenience, we restate this characterization for measures $\mu_h$ parametrized by $h > 0$. The measures $\mu_h$ converge to $\mu$ in the weak-* sense if and only if

$$\lim_{h \to 0} \int f \, d\mu_h = \int f \, d\mu \quad \text{for all continuous } f. \tag{22}$$

It is well-known, and not hard to show, that a necessary and sufficient condition for (22) to hold is that the moments of $\mu_h$ converge to those of $\mu$:

$$\widehat{\mu}_h(\ell) \to \widehat{\mu}(\ell) \ \text{ for all } \ \ell. \tag{23}$$

We will study the family $\{P_k(z, \mu_h)\}_{h>0}$ of Szegö polynomials as $h \to 0$ where $\mu_h$ is an absolutely continuous family which approaches a sum of point masses, as in (21). We will consider cases where $\mu_h$ is obtained by convolution of absolutely continuous measures with point masses of the form (20), and also for measures consisting of a sum of point masses plus an absolutely continuous part.

### The Limit Points of $\{P_k(z, \mu_h)\}$

If $k > m$, then $P_k(z, \mu)$ is not defined, since $D_{k-1}(\mu)$ in the denominator of (17) is zero but note that any polynomial of the form

$$p(z) = Q(z) \prod_{j=1}^{m} (z - e^{i\theta_j}), \tag{24}$$

with $Q \in \mathbf{\Lambda}_{k-m}$ has norm in $\mathcal{L}^2(d\mu)$ equal to zero in (4). We will show that for fixed $k > m$, all limit points, as $h \to 0$, in the space of polynomials of degree $k$, of the family $\{P_k(z, \mu_h)\}$, have the form (24).

**Remark:** If $h$ is a continuous parameter, we will call a decreasing sequence $\{h_\ell\}$ whose limit is zero a *discretization* of $h$. Thus $\mathcal{P}_k(z)$ is a limit point of $\{P_k(z, \mu_h)\}$ if and only if there exists a discretization $h_n$ such that

$$\lim_{n \to \infty} P_k(z, \mu_{h_n}) = \mathcal{P}_k(z). \tag{25}$$

A discretization is simply a subsequence if $h$ is discrete.

Suppose that (21) holds and $k > m$. Equation (17) gives

$$P_k(z, \mu) = \frac{\mathcal{D}_k(z, \mu_h)}{D_{k-1}(\mu_h)}, \tag{26}$$

which is well defined for $h > 0$ if $\mu_h$ is supported on an infinite set. The zeros of the $P_k(z, \mu_h)$ lie inside the open unit disk. Since the $P_k$ are monic, the family $\{P_k(z, \mu_h)\}$ is uniformly bounded in the closed $k-1$ dimensional space $\mathbf{\Lambda}_k$, in any norm we choose. So $\{P_k(z, \mu_h)\}$ has limit points in $\mathbf{\Lambda}_k$.

**Proposition 1** *Suppose (21) holds, for a family of measures $\mu_h$, and suppose that $k > m$. Then all the limit points of $\{P_k(z, \mu_h)\}$ are of the form $Q(z) \prod_{j=1}^{m} (z - e^{i\theta_j})$, where $Q \in \mathbf{\Lambda}_{k-m}$.*

**Proof:** Suppose that $\mathcal{P}_k(z)$ is a limit point of $\{P_k(z, \mu_h)\}$ with (25) holding for some discretization, or subsequence, $\{h_n\}$. For brevity, we define

$$\mathcal{P} := \mathcal{P}_k(z); \quad P_h := P_k(z, \mu_h); \quad \mu := \sum_{j=1}^{m} \alpha_j \delta_{\theta_j}.$$

We need to show that $\mathcal{P}$ has zeros at the $e^{i\theta_j}$. It suffices to show that

$$\int |\mathcal{P}|^2 d\mu = 0.$$

To this end we write

$$\left|\int |\mathcal{P}|^2\,d\mu - \int |P_h|^2\,d\mu_h\right| \leq \left|\int |\mathcal{P}|^2\,d\mu - \int |\mathcal{P}|^2\,d\mu_h\right| + \left|\int |\mathcal{P}|^2\,d\mu_h - \int |P_h|^2\,d\mu_h\right|. \quad (27)$$

The first term on the right-hand side of (27) approaches zero as $h$ approaches zero by weak-* convergence of $\mu_h$. Since $\mu_h \to \mu$ we must have $\mu_h[-\pi, \pi) \leq M$ for some $M > 0$ and all $h$. Thus the second integral on the right-hand side of (27) is less than $M||\,|\mathcal{P}|^2 - |P_h|^2||$, which approaches zero as $h \to 0$ since the uniform convergence, $|P_h|^2 \to |\mathcal{P}|^2$ on $|z| = 1$ follows from uniform convergence of $P_h$ on $|z| = 1$. Thus the Lemma will be proved if we show that $\int |P_h|^2\,d\mu_h \to 0$.

If $Q \in \mathbf{\Lambda}_{k-m}$ is arbitrary, by the minimization property (4) of $P_h$, we have

$$\int |P_h|^2\,d\mu_h \leq \int |Q(z)|^2 \prod_{j=1}^{m} |z - e^{i\theta_j}|^2\,d\mu_h.$$

Again, by weak-star convergence of $\mu_h$, the right-hand side of the above approaches $\int |Q(z)|^2 \prod_{j=1}^{m} |z - e^{i\theta_j}|^2\,d\mu = 0$. □

Thus, under the conditions of Proposition 1, if $\lim_{h\to 0} P_k(z, \mu_h)$ exists, then at least $m$ of the zeros of $P_k(z, \mu_h)$ approach the $e^{i\theta_j}$, and we can write

$$P_k(z, \mu_h) = Q_h(z)\Pi_{j=1}^{m}(z - w_j^{(h)}).$$

We shall suppose, without loss of generality, that $w_j^{(h)} \to e^{i\theta_j}$ for $j = 1, 2, ..., m$. Our aim is to study $Q_h(z)$ as $h \to 0$.

In the context of the frequency estimation problem in signal processing, the $w_j^{(h)}$ are often called signal zeros, while those of $Q_h$ are called non-signal, or extraneous zeros. Information about the zeros of any limit point, $Q$, of $Q_h$ could be useful in discerning which of the zeros of $P_k(z, \mu_h)$ correspond to signal frequencies.

If the moments $\widehat{\mu}_h(\ell)$ are analytic functions of $h$, we can say more.

**Proposition 2** *Suppose (21) holds and that the moments $\widehat{\mu}_h(\ell)$ are functions of $h$ which are analytic on some disk of radius $\varepsilon$ about $h = 0$. Then for $k > m$ there exists $\mathcal{P}_k \in \mathbf{\Lambda}_k$, and $Q \in \mathbf{\Lambda}_{k-m}$ such that*

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

**Proof:** Let $0 < h < \varepsilon$. Since the determinant of a matrix is a polynomial function of its entries, if the $\widehat{\mu}_h(\ell)$ are analytic functions of $h$ then $\mathcal{D}_k(z,\mu_h)$ and $D_k(\mu_h)$ in (15) and (16) will also be an analytic functions of $h$. For $h > 0$, $\mu_h$ satisfies Szegö's condition, so $D_k \neq 0$, and we see from the representation (17) that $P_k(z,\mu_h)$ is an analytic function of $h$ of the form

$$P_k(z,\mu_h) = \frac{T(z)h^p + \beta(h,z)}{Mh^n + \gamma(h)}, \tag{28}$$

where $M$ is a constant, $T(z)$ is a polynomial in $z$, $\beta(h,z)$ is a polynomial in $h$ consisting only of terms with degree larger than $p$, and $\gamma(h)$ is a polynomial in $h$ consisting only of terms with degree larger than $n$. The coefficients of $\beta(h,z)$ are functions of $z$. We must have $n \leq p$, otherwise $P_k(z,\mu_h)$ is unbounded as $h \to 0$. On the other hand, if $n < p$, then $\lim_{h\to 0} P_k(z,\mu_h) = 0$ which cannot happen by Proposition 1. □

## 3.2 Convolution with Approximate Identities

One way to construct measures $\mu_h$ that converge as in (21), is to convolve point mass measures of the form (20) with an *approximate identity*, or *kernel*. If $\nu$ and $\mu$ are two measures, and $\nu$ is absolutely continuous with density $f(\theta)$, the convolution of $\nu$ and $\mu$, written $\nu * \mu$, is the absolutley continuous measure with desity

$$\int f(\theta - t)d\mu(t).$$

In particular, if $\mu = \sum_{j=1}^{m} \alpha_j \delta_{\theta_j}$, then $\nu * \mu$ has density $\sum_{j=1}^{m} \alpha_j f(\theta - \theta_j)$.

We consider the convolution of point masses with the Poisson kernel,

$$\psi_r(\theta) := \frac{1 - r^2}{|\zeta - r|^2} \quad 0 < r < 1, \tag{29}$$

and with the Fejér kernel,

$$\phi_n(\theta) := \begin{cases} \frac{1}{n}\left[\frac{\sin(n\theta/2)}{\sin(\theta/2)}\right]^2 & \theta \neq 0 \\ n & \theta = 0\ . \end{cases} \quad n = 1, 2, 3, \ldots \tag{30}$$

Specifically, we will show that

$$\lim_{r\to 1} P_k(z, \psi_r * \sum_{j=1}^{m} \alpha_j \delta_{\theta_j}) = \lim_{n\to\infty} P_k(z, \phi_n * \sum_{j=1}^{m} \alpha_j \delta_{\theta_j}) \tag{31}$$

Moreover, we will characterize the limit polynomial.

We begin, in Section 3.2.1, by studying some properties of approximate identities, establishing a basic result and giving two simple examples. In Section 3.2.2 we will show that the convergence of signal zeros of Szegö polynomials with respect to convolution of point masses with an approximate identity is of the rate $O(h)$ if the approximate identity satisfies Properties 1 and 2. We will use some arguments from [PS], where a situation in which Property 2 holds is considered. In sections 3.2.3 and 3.2.4, $\psi_r$ and $\phi_n$ are considered separately, Properties 1 and 2 are shown to hold, and, using the $O(h)$ convergence rate of signal zeros, limits are characterized and (31) is shown to hold.

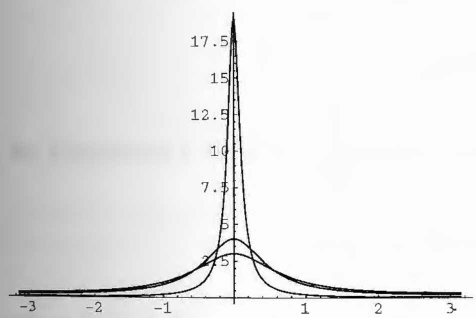


Figure 1: The Poisson kernel, $\psi_r(\theta)$, for $r = 0.5, 0.6, 0.9$.

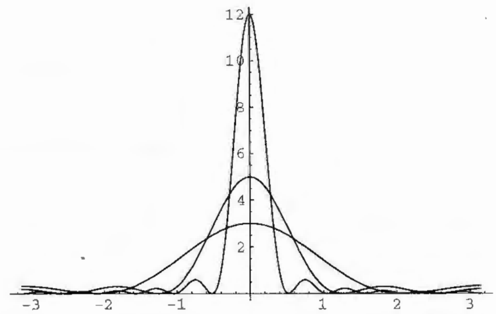


Figure 2: The Fejér kernel, $\phi_n(\theta)$, for $n = 3, 5, 12$.

### 3.2.1 Approximate Identities

The families $\psi_r(\theta)$ and $\phi_n(\theta)$ are examples of *approximate identities* for $\mathcal{L}_1[-\pi, \pi)$. That is, each can be viewed as a family of continuous functions, $\{\mathcal{K}_h(\theta)\}$, where either $h = 1/n$ or $h = 1 - r$, with the following properties:

1. $\mathcal{K}_h(\theta) \geq 0$.

2. $\int \mathcal{K}_h(\theta)\, d\theta = 1$.

3. $\lim_{h \to 0} \mathcal{K}_h(\theta) = 0$ uniformly outside any open interval containing $\theta = 0$.

A well-known property of both $\psi_r$ and $\phi_n$ is reflected in the following, result which can be found in [R].

**Theorem 3.1** *Let $\mathcal{K}_h(\theta)$ be an approximate identity for $\mathcal{L}_1[-\pi, \pi)$, and let $\nu$ be a finite measure on the unit circle. Then the convolution $\nu * \mathcal{K}_h$ converges in the weak-* sense to $\nu$ as $h \to 0$.*

**Remark:** If $\mu$ is a discrete measure, as in (20), and $\nu$ an absolutely continuous measure, then

$$||\mu - \nu|| = ||\mu|| + ||\nu||,$$

where $||\cdot||$ is the total variation norm, which induces the usual topology on measures. Thus, an absolutely continuous family cannot converge in the strong sense (i.e., in total variation norm) to a discrete measure.

Let $\mathcal{K}_h$ be any approximate identity and define

$$\mu_h := \mathcal{K}_h * \sum_{j=1}^{m} \alpha_j \delta_{\theta_j}. \tag{32}$$

By Proposition 1 all limit polynomials of $\{P_k(z, \mu_h)\}_h$ have the form

$$Q(z) \prod_{j=1}^{m} (z - e^{i\theta_j}) \tag{33}$$

where $Q(z) \in \mathbf{\Lambda}_{k-m}$. Since the $P_k(z, \mu_h)$ are polynomials of fixed degree equal $k$ with all zeros inside the unit disk, the convergence of any convergent subsequence is uniform on compact sets in the complex plane, $C$. If $\{h_n\}$ is a discretization such that

$$\lim_{n \to \infty} P_k(z, \mu_{h_n}) = \mathcal{P}_k(z) = Q(z) \prod_{j=1}^{m} (z - e^{i\theta_j}), \tag{34}$$

we can write

$$P_k(z, \mu_{h_n}) = Q_{h_n}(z) \Pi_{j=1}^{m} (z - w_j^{(h_n)}) \tag{35}$$

and suppose, without loss of generality, that $w_j^{(h_n)} \to e^{i\theta_j}$ for $j = 1, 2, ..., m$. Our aim, then, is to study $Q_{h_n}(z)$ as $n \to \infty$.

**Two Examples:**

We compare the Szegö polynomial limits with respect to two approximate identities. Note that since $(f * \delta_0)(\theta) = f(\theta)$ for any $\mathcal{L}^1$ function $f$ this correspond to finding the Szegö polynomial limits with respect to convolution of each identity with the point mass at $\theta = 0$. We find that the limit is the same in each case. Comparing the moments of each approximate identity, we find that they agree to first nonconstant terms when expanded as Taylor series in $h$. (We set $h = 1/n$.)

Let $\mathbf{1}_A$ be the indicator function for the set $A$. Define

$$\begin{aligned} \mathcal{F}_n(\theta) &= n\mathbf{1}_{[-1/2n,1/2n]} \\ \mathcal{G}_n(\theta) &= n(1 - \frac{2\pi}{n} + \frac{1}{n^2})\mathbf{1}_{[-1/2n,1/2n]} + \frac{1}{n}\mathbf{1}_{[-\pi,-1/2n)\cup(1/2n,\pi]}. \end{aligned}$$

The kernels $\mathcal{F}_n$ and $\mathcal{G}_n$ are shown in Figures 3 and 4 for several values of $n$.

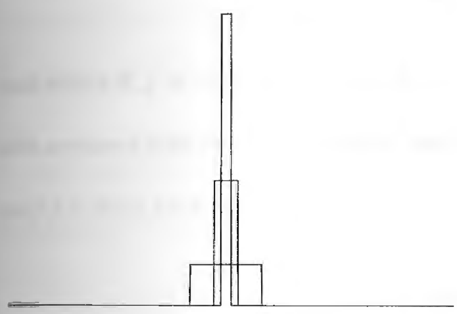

Figure 3: The "stovepipe" kernel $\mathcal{F}_n$ for $n = 1, 3, 7$.

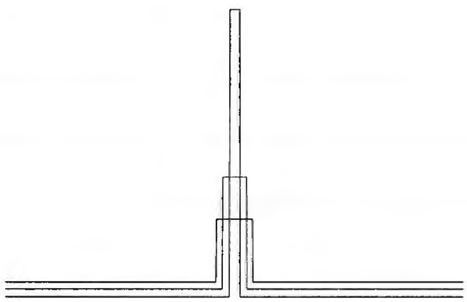

Figure 4: The "tophat" kernel $\mathcal{G}_n$ for $n = 2, 3, 7$

Note that the weak-* limit of both $\mathcal{F}_n$ and $\mathcal{G}_n$ is point mass at $\theta = 0$. By Proposition 1, we know that every limit point of both $\{P_k(z, \mathcal{F}_n)\}$ and $\{P_k(z, \mathcal{G}_n)\}$ have $z - 1$ as a factor. We have the following:

$$\lim_{n\to\infty} P_k(z, \mathcal{F}_n) = (z-1)^k.$$

This follows from Fejér's Convex Hull Theorem (see Section 2.3.1).

Direct computation of $P_k(z, \mathcal{G}_n)$, for example, using *Maple*, for various $n$ and $k$ suggest that $\lim_{n\to\infty} P_k(z, \mathcal{G}_n) = (z-1)^k$ as well. Computing the moments of the two kernels we find, with $h = 1/n$,

$$\widehat{\mathcal{F}}_n(\ell) = 1 - \frac{\ell^2 h^2}{6} + d_1(h, \ell);$$

$$\widehat{\mathcal{G}}_n(\ell) \quad = \quad 1 - \frac{\ell^2 h^2}{6} + d_2(h, \ell)$$

where $d_1(h, \ell)$ and $d_2(h, \ell)$ both contain only terms in $h$ of degree higher that 2. Thus, with suitable parametrization of $h$, the moments of the two kernels agree to first non-constant term in $h$ and $\ell$. We will see that this is the case with the moments of the Poisson and Fejér kernels in 3.2.3 and 3.2.4.

### 3.2.2 Convergence of Signal Zeros

We will show that for a certain class of kernels, the zeros $w_j^{(h_n)}$ converge at the rate (at least) $O(h_n)$. To do this, we will use some of the arguments of Pan and Saff in [PS]. There a discrete time signal $\{x(n)\}_{-\infty}^{\infty}$ is considered, where

$$x(n) = \sum_{j=-I}^{I} \alpha_j e^{in\theta_j},$$

and where $\theta_{-j} = -\theta_j$ and $\alpha_{-j} = \overline{\alpha_j} \neq 0$. Thus the signal consists of $I + 1$ real sinusoids. It is also assumed that the frequencies $\theta_j$ are distinct. Let $\{x_N(n)\}$ be the $N$-*truncated* signal; that is, $x_N(n) = x(n)$ for $n = 0, 1, 2, ..., N-1$, and $x_N(n) = 0$ for $n \geq N$. The Z-transform of the signal is

$$\mathrm{X}_N = \sum_{n=1}^{N-1} x(n) z^{-n}. \tag{36}$$

The measures $\nu_N$, where $d\nu_N = \frac{1}{N}|\mathrm{X}_N|^2 \, d\theta$, are not of the form (32), but they do converge weak-* to a sum of point masses at the signal frequencies ([JNS]):

$$\lim_{N\to\infty} \nu_N := \lim_{N\to\infty} \frac{1}{N}|\mathrm{X}_N|^2 = \sum_{j=-I}^{I} |\alpha_j|^2 \delta_{\theta_j}. \tag{37}$$

If $k > 2I + 1$ the Szegö polynomials $P_k(z, \nu_N)$ do not, in general, approach a limit as $N \to \infty$, but all limit polynomials are of the form (33) for some $Q \in \Lambda_{2I+1-k}$, the zeros of which are necessarily on $|z| \leq 1$. In the proof of Theorem 2.4 in [PS] it is shown that the zeros of any such $Q$ are strictly inside the unit circle. The assertion is that this is sufficient to prove the Theorem 2.4, which states that the $2I + 1$ zeros of largest modulus of $P_k(z, \nu_N)$ approach the $e^{i\theta_j}$. It would thus be possible to discern the $2I + 1$ signal zeros from the extraneous zeros for a large sample (large $N$). But if there are an infinite number of limit points, $P_k(z, \nu_N)$ it can happen that the zeros of limit factors $Q(z)$ get arbitrarily close to the unit circle. What is actually needed to prove Theorem 2.4 is that all limit factors $Q(z)$ have all zeros *uniformly* bounded away from the unit circle.

When $k = 2I + 1$,

$$\lim_{N\to\infty} P_k(z, \nu_N) = P_k(z, \sum_{j=-I}^{I} |\alpha_j|^2 \delta_{\theta_j}) = \prod_{j=-I}^{I} (z - e^{i\theta_j})$$

and it is shown in [PS] that the rate of this convergence, as well as that of the moments of $\nu_N$, its prediction error power, and the zeros of $P_k(z, \nu_N)$, is $O(1/N)$.

For the situation addressed here, note that the measure in (20) corresponds to $m$ *complex sinusoids*. Also, by Proposition 2, if the moments $\widehat{\mu}_h(\ell)$ are analytic functions of $h$, as is the case for the Poisson and Fejér kernels, $\lim_{h\to 0} P_k(z, \mu_h)$ exists, and the problem of (asymptotically) discerning signal zeros from extraneous zeros does not arise. We will, however, use the proof of Pan and Saff to show that the zeros of any limit factor $Q(z)$ lie strictly inside the unit circle.

Let $\rho_{h,k}$ denote the prediction error power for $P_k(z, \mu_h)$ defined as the minimum in (4) of Section 2.3.1.[3] With (32) we have

$$\rho_{h,k} = \int |P_k(\zeta, \mu_h)|^2 \, d\mu_h \tag{38}$$

$$= \sum_{j=1}^{m} \alpha_j \int |P_k(\zeta, \mu_h)|^2 \mathcal{K}_h(\theta - \theta_j) \, d\theta. \tag{39}$$

As we will later show, the following properties hold for the measures when $\mathcal{K}_h = \psi_r$ and $\mathcal{K}_h = \phi_n$. They will be assumed here for otherwise arbitrary $\mathcal{K}_h$, with $\mu_h$ defined in (32).

**Property 1** *The approximate identity $\{\mathcal{K}_h(\theta)\}$ (and therefore $\mu_h$) has moments, $\widehat{\mathcal{K}}_h(\ell)$, which are analytic functions of $h$ at $h = 0$.*

**Property 2** *Let $\rho_{h,k}$ be as defined in (38). There exist constants $c$ and $C$, with $0 < c < C$, such that for all $h > 0$ and all $k > m$,*

$$ch \le \rho_{h,k} \le Ch. \tag{40}$$

[3] Recall our convention of omitting the factor of $1/2\pi$ in the definition of the prediction error power.

Assuming that Property 1 holds, by Proposition 2 $\lim_{h\to 0} P_k(z,\mu_h)$ exists. Using the notation of (34) and (35) we can write

$$\mathcal{P}_k(z) := \lim_{h\to 0} P_k(z,\mu_h) = Q(z)\prod_{j=1}^{m}(z - e^{i\theta_j}), \tag{41}$$

and

$$P_k(z,\mu_h) = Q_h(z)\Pi_{j=1}^{m}(z - w_j^{(h)}) \tag{42}$$

where $w_j^{(h_n)} \to e^{i\theta_j}$ for $j = 1, 2, ..., m$. We will use some of the arguments of Pan and Saff to show that the rate of convergence of the signal zeros $w_j^{(h)}$ in (42) is at least $O(h)$.

We can use the relation

$$|P_{l+1}(0,\mu_h)|^2 = 1 - \frac{\rho_{h,l+1}}{\rho_{h,l}}, \tag{43}$$

found in [GS] and elsewhere, to bound the reflection coefficients, $P_{l+1}(0,\mu_h)$, away from the unit circle uniformly in both $k$ and $h$. The following is an immediate consequence of (40) and (43).

**Lemma 3.1** *Suppose that $\mu_h$ has Property 2. Then for all $k > m$ and $h > 0$*

$$|P_k(0,\mu_h)|^2 \le 1 - \frac{c}{C} < 1. \tag{44}$$

We now study the convergence of $P_k(z,\mu_h)$ in (41). If $\mathcal{K}_h$ has moments which are analytic functions of $h$, by Proposition 2 and its proof (see eq. (28)), we can write

$$P_k(z,\mu_h) = \frac{T(z)h^p + \beta(h,z)}{Mh^p + \gamma(h)}, \tag{45}$$

where $M$ is a constant, $T(z)$ is a polynomial in $z$, and $\beta(h,z)$ and $\gamma(h)$ are polynomials in $h$ consisting only of terms with degree larger than $p$. Evidently, $\mathcal{P}_k(z) = \frac{T(z)}{M}$. This yields

$$\begin{aligned} |P_k(z,\mu_h) - \mathcal{P}_k(z)| &= \left|\frac{M\beta(h,z) - T(z)\gamma(h)}{M^2h^p + M\gamma(h)}\right| \\ &= \frac{|\tau(h,z)|}{|M^2 + M\kappa(h)|} \end{aligned}$$

where $\tau(h,z)$ and $\kappa(h)$ are polynomials in $h$ consisting only of terms of degree at least 1.

Thus we have

**Theorem 3.2** *Suppose that Property 1 holds. Then there exists a constant $N > 0$, such that, for $h > 0$*

$$|P_k(z, \mu_h) - \mathcal{P}_k(z)| \leq Nh \tag{46}$$

*for all $|z| \leq 1$.*

**Remark:**

Recall that the denominator in the determinant representation for $P_k(z, \mu_h)$, (17), is the $k \times k$ matrix $D_{k-1}(\mu_h)$. Equation (40) with the representation

$$\rho_{h,k-1} = \frac{D_{k-1}(\mu_h)}{D_{k-2}(\mu_h)}, \tag{47}$$

found in [G+S], p. 71, give

$$chD_{k-2}(\mu_h) \leq D_{k-1}(\mu_h) \leq ChD_{k-2}(\mu_h)$$

for some constants $c$ and $C$, and for all $h > 0$. It follows immediately that

$$c_1 h^{k-m} D_{m-1}(\mu_h) \leq D_{k-1}(\mu_h) \leq C_1 h^{k-m} D_{m-1}(\mu_h)$$

for constants $c_1$ and $C_1$. Since $\lim_{h \to 0} D_{m-1}(\mu_h) = D_{m-1}(\sum_{j=1}^{m} \alpha_j \delta_{\theta_j}) > 0$ (see the remarks accompanying (15)), we have

$$c_2 h^{k-m} \leq D_{k-1}(\mu_h) \leq C_2 h^{k-m} \tag{48}$$

for constants $c_2$ and $C_2$. From the proof of Proposition 2, we must also have

$$c_3 h^{k-m} \leq \mathcal{D}_k(z, \mu_h) \leq C_3 h^{k-m}. \tag{49}$$

We conclude from the above and (17) that $p = k - m$ is the smallest integer for which (45) holds.

Before we address the convergence of the signal zeros, $w_j^{(h)}$, we will need to show that all the zeros of $Q(z)$ in (41) lie in the *open* unit disk. To do this we use the proof of Theorem 2.4 in [PS].

**Theorem 3.3** *Assume that Property 2 holds. Then all the zeros of the limit factor $Q(z)$ in (41) lie strictly inside the open unit disk.*

**Proof: (Pan and Saff)** We prove this by induction on $k \geq m$. The theorem is vacuously true for $k = m$. Assume that it is true for $k > m$. Using the Levinson recursion we can write

$$P_{k+1}^*(z, \mu_h) = P_k^*(z, \mu_r) + z\overline{P_{k+1}^*(0, \mu_h)} P_k(z, \mu_h). \tag{50}$$

By (41) we have

$$\lim_{h \to 0} P_{k+1}(z, \mu_h) = \mathcal{P}_{k+1}(z) = R(z)\Pi_{j=1}^m (z - e^{i\theta_j}) \tag{51}$$

for some $R(z)$ in $\Lambda_{k+1-m}$. Define $\Pi(z) := \Pi_{j=1}^m (z - e^{i\theta_j})$. Then (41), (50), and (51) yield

$$\begin{aligned} \mathcal{P}_{k+1}^*(z) &= \Pi^*(z) Q^*(z) + z\overline{\mathcal{P}_{k+1}(0)} \Pi(z) Q(z) \\ &= \Pi^*(z) \left( \frac{\Pi(z)}{\Pi^*(z)} Q^*(z) + z\overline{\mathcal{P}_{k+1}(0)} \Pi(z) Q(z) \right) \end{aligned} \tag{52}$$

$$\tag{53}$$

By (51) and (53),

$$R^*(z) = \frac{\Pi(z)}{\Pi^*(z)} Q^*(z) + z\overline{\mathcal{P}_{k+1}(0)} \Pi(z) Q(z).$$

Define $\tau(z) := \frac{\Pi(z)}{\Pi^*(z)}$. Writing

$$\tau(z) = \prod_{j=1}^m \frac{z - e^{i\theta_j}}{1 - ze^{-i\theta_j}} = \prod \frac{1}{e^{-i\theta_j}} \frac{z - e^{i\theta_j}}{e^{i\theta_j} - z},$$

we see that $\tau(z)$ is a constant of modulus 1 on $C - \cup_{j=1}^m e^{i\theta_j}$ with removable singularities at the $e^{i\theta_j}$. We can then write

$$R^*(z) = \tau Q^*(z) + z\overline{\mathcal{P}_{k+1}(0)} Q(z), \tag{54}$$

where $\tau = \Pi \frac{1}{e^{-i\theta_j}}$. The theorem will be proved if we show that all the zeros of $R^*(z)$ lie in $|z| > 1$. We do this by contradiction. Suppose that $R(z_0) = 0$ with $|z_0| \leq 1$. We have, from (54),

$$|Q^*(z_0)| = |z_0 \overline{\mathcal{P}_{k+1}(0)} Q(z_0)|.$$

By the induction hypothesis, $Q^*$ has all its zeros outside the circle, so $z_0$ and $Q^*$ are both different from zero. Thus, by Lemma 3.1

$$1 = \left| z_0 \overline{\mathcal{P}_{k+1}(0)} \frac{Q(z_0)}{Q^*(z_0)} \right| \leq \left| \mathcal{P}_{k+1}(0) \frac{Q(z_0)}{Q^*(z_0)} \right| < \left| \frac{Q(z_0)}{Q^*(z_0)} \right|.$$

Since $\frac{Q(z_0)}{Q^*(z_0)}$ is a Blaschke product, $|\frac{Q(z_0)}{Q^*(z_0)}| \leq 1$. This contradiction completes the proof. □

Finally, we have the following

**Lemma 3.2** *Suppose that Properties 1 and 2 hold. Then the zeros of $P_k(z, \mu_h)$ approach the $e^{i\theta_j}$ the the rate $O(h)$. That is, there exists a constant, $N_j$, such that*

$$|w_j^{(h)} - e^{i\theta_j}| \leq N_j h. \tag{55}$$

**Proof:** It follows from (41) and Theorem 3.2 that

$$|P_k(e^{i\theta_j}, \mu_h)| \leq Nh, \text{ for } j = 1, 2, ..., m. \tag{56}$$

Write

$$|w_j^{(h)} - e^{i\theta_j}| = \left| \frac{P_k(e^{i\theta_j}, \mu_h)}{Q_h(e^{i\theta_j})\Pi_{p\neq j}^m (w_j^{(h)} - e^{i\theta_p})} \right|. \tag{57}$$

By Theorem 3.3, $\lim_{h\to 0} Q_h(e^{i\theta_j}) \neq 0$. Since $w_j^{(h)} \to e^{i\theta_j}$ for $j = 1, 2, ..., m$, the denominator of (57) converges to a non-zero constant, so (56) and (57) give (55). □

### 3.2.3 The Poisson Kernel

We now let $h = 1 - r$ and consider the convolution of the Poisson kernel with the sum of point masses in (20). In the present context (32) becomes

$$\mu_r := \psi_r * \sum_{j=1}^m \alpha_j \delta_{\theta_j}. \tag{58}$$

We will show that Property 1 holds for $\psi_r$. The main result of this section is Theorem 3.4, which characterizes the limit $\mathcal{P}_k$ in (41) for the measure $\mu_r$. First, we consider the convergence of the moments, $\widehat{\psi}_r$ and $\widehat{\mu}_r$.

Recall that the moments of $\psi_r$ are

$$\hat{\psi}_r(\ell) = r^{|\ell|}. \tag{59}$$

Note $\psi_r$ thus has Property 1. By Proposition 2, $\lim_{r\to 1} P_k(z, \mu_r)$ exists, and (42) and (41) become, respectively

$$P_k(z, \mu_r) = Q_r(z)\Pi_{j=1}^m (z - w_j^{(r)}), \tag{60}$$

and

$$\mathcal{P}_k(z) = \lim_{r \to 1} P_k(z, \mu_r) = Q(z) \prod_{j=1}^{m} (z - e^{i\theta_j}). \tag{61}$$

**Lemma 3.3** *For all $\ell$ and $r < 1$, we have*

$$|\hat{\psi}_r(\ell) - 1| \leq |\ell|(1 - r) \tag{62}$$

**Proof:** Write

$$|\hat{\psi}_r(\ell) - 1| = |r^{|\ell|} - 1| = (1 + r + r^2 + \ldots + r^{|\ell| - 1})(1 - r) \leq |\ell|(1 - r).$$

□

Define

$$\mu := \sum_{j=1}^{m} \alpha_j \delta_{\theta_j}. \tag{63}$$

Since $\widehat{\mu}(\ell) = \sum_{j=1}^{m} \alpha_j e^{i\ell\theta_j}$, the moments of $\mu * \psi_r(\theta)$ are

$$\widehat{(\mu * \psi_r)}(\ell) = \sum_{j=1}^{m} \alpha_j r^{|\ell|} e^{i\ell\theta_j}, \tag{64}$$

and we have, for all $\ell$ and $r < 1$,

$$|\widehat{(\mu * \psi_r)}(\ell) - \widehat{\mu}(\ell)| = (1 - r^{|\ell|})|\sum_{j=1}^{m} \alpha_j e^{i\ell\theta_j}|.$$

So by Lemma 3.3 we have

**Corollary 3.1**

$$|\widehat{(\mu * \psi_r)}(\ell) - \widehat{\mu}(\ell)| \leq (1 - r)|\ell| \sum_{j=1}^{m} \alpha_j. \tag{65}$$

We can express the convolution of $\psi_r$ and $\mu$ explicitly:

$$\begin{aligned} d\mu_r(\theta) := (\mu * \psi_r)(\theta) &= (1 - r^2) \sum_{j=1}^{m} \frac{\alpha_j}{|\zeta - re^{i\theta_j}|^2} \, d\theta & (66) \\ &= (1 - r^2) \frac{\sum_{j=1}^{m} \alpha_j \Pi_{p \neq j}^{m} |\zeta - re^{i\theta_p}|^2}{\Pi_{j=1}^{m} |\zeta - re^{i\theta_j}|^2} \, d\theta. & (67) \end{aligned}$$

Define, for $j = 1, 2, ..., m$,

$$X_j(z) := \frac{1}{(z - re^{i\theta_j})}\ ;\ \ X_j^*(z) := \frac{1}{(1 - zre^{-i\theta_j})}.$$

Then $|X_j(\zeta)| = |X_j^*(\zeta)|$. (Compare with $P_k^*$ in (9) and accompanying remarks.) Using (66), the prediction error power in (38) can then be written

$$\rho_{k,r} = (1 - r^2) \sum_{j=1}^{m} \alpha_j \int |P_k^*(\zeta, \mu_r)|^2 |X_j^*(\zeta)|^2 \, d\theta. \tag{68}$$

Recall that $\rho_{k,r}$ is the minimum in (4) for the measure $\mu_r$. We now prove that $\psi_r$ has Property 2.

**Lemma 3.4** *For all $r < 1$ and all $k > m$,*

$$(1 - r^2) \sum_{j=1}^{m} \alpha_j \le \rho_{r,k} \le 4^{m-1} (1 - r^2) \sum_{j=1}^{m} \alpha_j. \tag{69}$$

**Proof:** Since $P_k^*(z, \mu_r)$ and $X_j^*(z)$ are analytic in $|z| < 1 + \varepsilon$ for some $\varepsilon > 0$, each integrand in the right-hand side of (68) is subharmonic in that region. Thus

$$1 = |P_k^*(0, \mu_r)|^2 |X_j^*(0)|^2 \le \int_{-\pi}^{\pi} |P_k^*(\zeta, \mu_r)|^2 |X_j^*(\zeta)|^2 \, d\theta. \tag{70}$$

This, with (68) proves the left-hand side inequality in (69).

To finish the proof, since $\rho_{k,r}$ is the minimum in (4) and $z^{k-m} \prod_{j=1}^{m} (z - re^{i\theta_j}) \in \mathbf{\Lambda}_k$, the representation (67), gives

$$\begin{aligned}
\rho_{k,r} &\le \int |\zeta^{k-m}|^2 \Pi_{j=1}^{m} |\zeta - re^{i\theta_j}|^2 \, d\mu_r \\
&= (1 - r^2) \int |\zeta^{k-m}|^2 \Pi_{j=1}^{m} |\zeta - re^{i\theta_j}|^2 \frac{\sum_{j=1}^{m} \alpha_j \Pi_{p \ne j}^{m} |\zeta - re^{i\theta_p}|^2}{\Pi_{j=1}^{m} |\zeta - re^{i\theta_j}|^2} \, d\theta \\
&= (1 - r^2) \int \sum_{j=1}^{m} \alpha_j \Pi_{p \ne j}^{m} |\zeta - re^{i\theta_p}|^2 \, d\theta \\
&\le (1 - r^2)(2^2)^{m-1} \sum_{j=1}^{m} \alpha_j \\
&= 4^{m-1} (1 - r^2) \sum_{j=1}^{m} \alpha_j.
\end{aligned}$$

This completes the proof. □

Since $\psi_r$ satisfies Properties 1 and 2, Lemma 3.2 and (55) gives

$$|w_j^{(r)} - e^{i\theta_j}| \le N_j(1-r). \tag{71}$$

Using (67) and the orthogonality property (8), we have, for any $T \in \mathbf{\Lambda}_{k-1}$,

$$\begin{aligned} 0 &= (1-r^2)\int P_k(\zeta,\mu_r)\overline{T(\zeta)}\ \ d(\mu * \psi_r)(\theta) && (72) \\ &= \int \Pi_{j=1}^m(\zeta - w_j^{(r)})Q_r(\zeta)\overline{T(\zeta)}\frac{\sum_{j=1}^m \alpha_j \Pi_{p\neq j}^m|\zeta - re^{i\theta_p}|^2}{\Pi_{j=1}^m|\zeta - re^{i\theta_j}|^2}\ d\theta. && (73)\end{aligned}$$

In particular, if $T(z) = q(z)\Pi_{j=1}^m(z - re^{i\theta_j})$, with $q(z) \in \mathbf{\Lambda}_{k-m-1}$ arbitrary, (73) gives

$$\int \prod_{j=1}^m \frac{\zeta - w_j^{(r)}}{(\zeta - re^{i\theta_j})}Q_r(\zeta)\overline{q(\zeta)}\sum_{j=1}^m \alpha_j\Pi_{p\neq j}^m|\zeta - re^{i\theta_p}|^2\, d\theta = 0\,. \tag{74}$$

In studying the limit factor $Q(z)$ of $P_k(z,\mu_r)$ in (41) we would like to take limits under the integral sign in (74). The next result concerns the convergence of the rational function in the integrand of (74), and is the key idea used in the proof of Theorem 3.4, our first main result.

**Lemma 3.5** *Let $w_j^{(r)}$ satisfy (55). The function $\Pi_{j=1}^m\frac{\zeta - w_j^{(r)}}{(\zeta - re^{i\theta_j})}$ converges to* 1 *in $\mathcal{L}_1[-\pi,\pi]$ as $r \to 1$.*

**Proof:** Clearly, $\Pi_{j=1}^m\frac{\zeta - w_j^{(r)}}{(\zeta - re^{i\theta_j})}$ converges pointwise to 1 for $\zeta \neq e^{i\theta_j}$, $j = 1, 2, ..., m$. We show that the function is uniformly bounded on $|z| = 1$ for $r < 1$. The lemma will then follow from the Lebesgue Bounded Convergence Theorem.

Using Lemma 3.2 and the fact that $|\zeta - re^{i\omega}| \ge |e^{i\omega} - re^{i\omega}| = 1 - r$, we have

$$\left|\frac{\zeta - w_j^{(r)}}{\zeta - re^{i\theta}} - 1\right| = \left|\frac{w_j^{(r)} - re^{i\theta_j}}{\zeta - re^{i\theta}}\right| \le \frac{N_j(1-r)}{1-r} = N_j.$$

So $\left|\frac{\zeta - w_j^{(r)}}{\zeta - re^{i\theta}}\right| \le 1 + N_j$. Thus $|\Pi_{j=1}^m\frac{\zeta - w_j^{(r)}}{(\zeta - re^{i\theta_j})}| \le \Pi_{j=1}^m(1+N_j)$. This proves the lemma. □

We come now, to the first main result of this work. Lemma 3.5 will allow us to let $r$ approach 1 under the integral sign in (74), and we can now characterize the limit polynomial, $\mathcal{P}_k$, in (41). We will see that the "extra" factor $Q(z)$, in (41), is actually a Szegö polynomial of degree $k - m$ with respect to an absolutely continuous measure, which we specify.

**Theorem 3.4** *Let $\mu_r$ be given in (58)*

$$\mathcal{P}_k(z) := \lim_{r\to 1} P_k(z,\mu_r) = P_{k-m}(z,\nu)\prod_{j=1}^{m}(z-e^{i\theta_j}), \tag{75}$$

*where $\nu$ is the absolutely continuous measure with*

$$\frac{d\nu}{d\theta} = \sum_{j=1}^{m} \Pi_{p\neq j}^{m}\alpha_j|\zeta - e^{i\theta_p}|^2. \tag{76}$$

**Proof:** We need to show that the factor $Q(z)$, in (41), is the Szegö polynomial $P_{k-m}(z,\nu)$. We will show that $Q(z)$ has the orthogonality property (8), which characterizes $P_{k-m}(z,\nu)$.

Consider the factors in the integrand of (74). With the exception of $\Pi_{j=1}^{m}\frac{\zeta - w_j^{(r)}}{(\zeta - re^{i\theta_j})}$, which converges in $\mathcal{L}_1$, all converge uniformly on $|z|=1$. Thus the entire integrand in (74) converges in $\mathcal{L}_1$ to

$$Q(z)\overline{q(z)}\sum_{j=1}^{m}\Pi_{p\neq j}^{m}\alpha_j|\zeta - e^{i\theta_p}|^2 .$$

Now replace the integrand in (74) with its $\mathcal{L}_1$ limit to obtain

$$\int Q(z)\overline{q(z)}\sum_{j=1}^{m}\Pi_{p\neq j}^{m}\alpha_j|\zeta - e^{i\theta_p}|^2\, d\theta = 0. \tag{77}$$

and obtain Since $q(z) \in \mathbf{\Lambda}_{k-m-1}$ is arbitrary, $Q$ has the desired orthogonality property and thus

$$Q(z) = P_{k-m}(z, \sum_{j=1}^{m}\Pi_{p\neq j}^{m}\alpha_j|\zeta - e^{i\theta_p}|^2).$$

This, along with the definition of $Q(z)$ in (41) proves the theorem. □

### 3.2.4 The Fejér Kernel

We now let $h = 1/n$ and consider the convolution of the Fejér kernel with the sum of point masses $\mu$. Here, (32) becomes

$$\mu_n := \phi_n * \sum_{j=1}^{m}\alpha_j\delta_{\theta_j}. \tag{78}$$

Recall that the Fourier coefficients of $\phi_n$ are given by

$$\widehat{\phi}_n(\ell) = \begin{cases} 1-\frac{|\ell|}{n} & |\ell|\le n \\ 0 & |\ell| > n\ . \end{cases} \tag{79}$$

Note that $\phi_n$ has Property 1. Thus, as in the last section, Proposition 2 applies. Equations (42) and (41) become, respectively

$$P_k(z,\mu_n) = Q_n(z)\Pi_{j=1}^m(z-w_j^{(n)}) \tag{80}$$

$$\mathcal{P}_k(z) = \lim_{n\to\infty} P_k(z,\mu_n) = Q(z)\prod_{j=1}^m(z-e^{i\theta_j}). \tag{81}$$

**The Fourier Coefficients of $\phi_n$**

We will see that the Poisson and Fejér kernels have a similar character. In fact we will show that the Szegö polynomials with respect to either of these kernels have the same limit; that is, the limit in (41) is the same limit found in Theorem 3.4 with a change of parameter from the continuous $r\to 1$ in the case of the Poisson kernel, to the discrete $n\to\infty$ in the present case.

One starting point for comparison is the respective Fourier coefficients. To compare those of $\psi_r$ with (79) we substitute $r=1-1/n$ in (59) to obtain

$$\begin{aligned}\widehat{\psi}_r(\ell) &= (1-\frac{1}{n})^{|\ell|}\\ &= 1-h|\ell|+\gamma(n)\end{aligned} \tag{82}$$

where $\gamma$ contains only terms in $1/n$ of order larger than 1. Upon comparison with (79), we see that the moments of the two kernels agree up to linear terms in $h=1/n=1-r$.

Again, define $\mu:=\sum_{j=1}^m \alpha_j\delta_{\theta_j}$. Since $\hat{\mu}(\ell)=\sum_{j=1}^m \alpha_j e^{i\ell\theta_j}$, the moments of $\mu*\phi_n$ are

$$\widehat{(\mu*\phi_n)}(\ell)=\begin{cases}(1-\frac{|\ell|}{n})\sum_{j=1}^m\alpha_je^{i\ell\theta_j} & |\ell|\le n\\ 0 & |\ell|>n\ .\end{cases} \tag{83}$$

Thus we have, with $h=1/n$

$$|\widehat{(\mu*\phi_n)}(\ell)-\sum_{j=1}^m\alpha_je^{i\ell\theta_j}|=\begin{cases}h|\ell| & |\ell|\le n\\ 0 & |\ell|>n\end{cases} \tag{84}$$

as the analog of (65). We see that $\widehat{(\mu*\psi_r)}(\ell)$ of (64) also agrees up to linear term in $h$ with $\widehat{(\mu*\phi_n)}(\ell)$. We can thus consider $\widehat{(\mu*\psi_r)}(\ell)$ and $\widehat{(\mu*\phi_n)}(\ell)$ as polynomials in $h=1/n$, with the latter as linear approximations of the former. In either case (that is, for $h=1-r$ with $\mu_r$ defined in (58) or for $h=1/n$ with $\mu_n$ defined in (78)), we see from (15) and (16) that both $\mathcal{D}_k(z,\mu_h)$

and $D_{k-1}(\mu_h)$ are polynomial functions of $h$. With the determinant representation (17), equations (31) and (28), in light of these remarks, suggest that only the constant and linear terms, $1 - h|\ell|$, of (82), contribute in the limit $\mathcal{P}_k(z)$ found in Theorem 3.4.

We will prove directly that $\lim_{n\to\infty} P_k(z,\mu_n) = \mathcal{P}_k(z)$ by exploiting the analytical properties of the Fejér kernel and the orthogonality of the $P_k(z,\phi_n)$. A key point will be the convergence of the signal zeros, $w_j^{(n)} \to e^{i\theta_j}$, which has as counterpart the rate of convergence in (71).

**The Measures $\phi_n * \sum \alpha_j \delta_{\theta_j}$**

From (30) it is easy to see that

$$\phi_n(\theta) = \frac{2}{n}\frac{(1-\cos n\theta)}{|\zeta - 1|^2}, \tag{85}$$

and hence that

$$d\mu_n = \frac{2}{n}\frac{\sum_{j=1}^m \alpha_j(1-\cos n(\theta-\theta_j))\prod_{p\neq j}^m |\zeta - e^{i\theta_p}|^2}{\prod_{j=1}^m |\zeta - e^{i\theta_j}|^2}\, d\theta. \tag{86}$$

From (30) we also have the representation

$$\begin{aligned}
\phi_n &= \frac{1}{n}\left[\frac{e^{in\theta/2} - e^{-in\theta/2}}{e^{i\theta/2} - e^{-i\theta/2}}\right]^2 \\
&= \frac{1}{n}\left|\frac{e^{in\theta} - 1}{e^{i\theta} - 1}\right|^2 \\
&= \frac{1}{n}\left|\frac{\zeta^n - 1}{\zeta - 1}\right|^2 && (87)\\
&= \frac{1}{n}|\zeta^{n-1} + \zeta^{n-2} + \cdots + 1|^2. && (88)
\end{aligned}$$

The above yield two more representations for $\mu_n$. From (87) and (20) follows

$$\begin{aligned}
d\mu_n &= \frac{1}{n}\sum_{j=1}^m \alpha_j \left|\frac{e^{in(\theta-\theta_j)} - 1}{e^{i(\theta-\theta_j)} - 1}\right|^2 d\theta \\
&= \frac{1}{n}\sum_{j=1}^m \alpha_j \left|\frac{\zeta^n - e^{i\theta_j}}{\zeta - e^{i\theta_j}}\right|^2 d\theta, && (89)
\end{aligned}$$

and from (88) and (20) follows

$$\begin{aligned}
d\mu_n &= \frac{1}{n}\sum_{j=1}^m \alpha_j |e^{i(n-1)(\theta-\theta_j)} + e^{i(n-2)(\theta-\theta_j)} + \cdots + 1|^2\, d\theta \\
&= \frac{1}{n}\sum_{j=1}^m \alpha_j |e^{i(n-1)\theta_j}\zeta^{(n-1)} + e^{i(n-2)\theta_j}\zeta^{(n-2)} + \cdots + 1|^2\, d\theta \\
&= \frac{1}{n}\sum_{j=1}^m \alpha_j |\zeta^{(n-1)} + e^{i\theta_j}\zeta^{(n-2)} + \cdots + e^{i(n-1)\theta_j}|^2\, d\theta\,. && (90)
\end{aligned}$$

Now, the density in (90) is a non-negative trigonometric polynomial, and has spectral factorization (see Sec. 2.3.2)

$$|g_n(z)|^2 = \frac{1}{n}\sum_{j=1}^{m} \alpha_j |z^{(n-1)} + e^{i\theta_j} z^{(n-2)} + \cdots + e^{i(n-1)\theta_j}|^2 \tag{91}$$

where $g_n(z)$ is a polynomial of degree $n-1$ with all its zeros outside the unit circle, which is uniquely defined with the requirement that $g_n(0) > 0$.

We can now show that Property 2 holds for the Fejér kernel. In the present context, (38) becomes

$$\rho_{n,k} = \int |P_k(\zeta, \mu_n)|^2 \, d\mu_n \tag{92}$$

$$= \int_{-\pi}^{\pi} |P_k^*(\zeta, \mu_n)|^2 |g_n(\zeta)|^2 \, d\theta, \tag{93}$$

where we have used (91) and the properties of $P_k^*(\zeta, \mu_n)$. Since the integrand in (93) is subharmonic, (90), (91), and (93) give

$$\rho_{n,k} \geq |P_k^*(0, \mu_n)|^2 |g_n(0)|^2 = \frac{1}{n}\sum_{j=1}^{m} \alpha_j.$$

This proves one half of the following analog of Lemma 3.4. The remainder of the proof is similar to that of Lemma 3.4, and is omitted. Note, however, the additional factor of 4.

**Lemma 3.6** *For all $n = 1, 2, 3, \ldots$ and all $k > m$,*

$$\frac{1}{n}\sum_{j=1}^{m} \alpha_j \leq \rho_{n,k} \leq \frac{4^m}{n}\sum_{j=1}^{m} \alpha_j. \tag{94}$$

By Lemma 3.2 and (55) we now have

$$|w_j^{(n)} - e^{i\theta_j}| \leq \frac{N_j}{n} \tag{95}$$

for constants $N_j > 0$.

### A Limit for $P_k(z, \mu_n)$

By the orthogonality condition (8) characterizing $P_k(z, \mu_n)$, we have, using (80) and (86),

$$\int Q_n(\zeta)\overline{T(\zeta)} \prod_{j=1}^{m} (\zeta - w_j^{(n)}) \frac{\sum_{j=1}^{m} \alpha_j (1 - \cos n(\theta - \theta_j)) \prod_{p \neq j}^{m} |\zeta - e^{i\theta_p}|^2}{\prod_{j=1}^{m} |\zeta - e^{i\theta_j}|^2} \, d\theta = 0$$

for $T(z) \in \Lambda_{k-1}$ arbitrary. In particular, if $T(z) = t(z)\prod_{j=1}^{m}(z - e^{i\theta_j})$ with $t(z) \in \mathbf{\Lambda}_{k-m-1}$ arbitrary we have

$$\int \gamma(n, \zeta)\, d\theta = 0, \tag{96}$$

where we define

$$\gamma(n, \zeta) := Q_n(\zeta)\overline{t(\zeta)}\frac{\overline{\prod_{j=1}^{m}(\zeta - w_j^{(n)})}}{\prod_{j=1}^{m}(\zeta - e^{i\theta_j})}\sum_{j=1}^{m}\alpha_j(1 - \cos n(\theta - \theta_j))\prod_{p \neq j}|\zeta - e^{i\theta_p}|^2. \tag{97}$$

We write down some simple bounds for future reference which hold for all $n = 1, 2, ...$; and $j = 1, 2, ..., m$. As a result of the relationship between chord length and arc length between points on the unit circle we have

$$\begin{aligned} \frac{|\theta - \theta_j|}{2} &< |\zeta - e^{i\theta_j}| \\ &= |e^{i(\theta - \theta_j)} - 1| \leq |\theta - \theta_j| \quad \text{for} \quad |\theta - \theta_j| < \pi\,. \end{aligned} \tag{98}$$

As a result of (95) and the fact that $|w_j^{(n)}| \leq 1$ we have

$$|\zeta - w_j^{(n)}| \leq |\zeta - e^{i\theta_j}| + |w_j^{(n)} - e^{i\theta_j}| \leq |\theta - \theta_j| + \frac{N_j}{n}$$

and

$$|\zeta - w_j^{(n)}| < 2.$$

Thus

$$|\zeta - w_j^{(n)}| \leq \min\{2, |\theta - \theta_j| + \frac{N_j}{n}\}. \tag{99}$$

To prove Theorem 3.4 of the last section, we used the $\mathcal{L}_1$ convergence of the integrand to exchange integration and limit operations in (74). In particular, the factors $\frac{\zeta - w_j^{(r)}}{\zeta - re^{i\theta_j}}$ converge pointwise except (possibly) at the $\theta_j$, where they are bounded uniformly in $r$. The other factors of the integrand converges uniformly.

The present case, with (96) as the analog of (74), will require a different approach. Factors which appear in $\gamma(n, \zeta)$ of the form

$$Z_j(n, \theta) := \frac{(\zeta - w_j^{(n)})(1 - \cos n(\theta - \theta_j))}{\zeta - e^{i\theta_j}}, \tag{100}$$

for $j = 1, 2, ..., m$, do not converge in $\mathcal{L}_1$, due to the $\cos n(\theta - \theta_j)$ term. We must also carefully consider the behavior of $Z_j(n, \theta)$ near $\theta_j$. Since

$$\lim_{\theta \to \theta_j} \frac{1 - \cos n(\theta - \theta_j)}{\zeta - e^{i\theta_j}} = 0,$$

we have

$$\lim_{\theta \to \theta_j} Z_j(n, \theta) = 0.$$

So for each $n = 1, 2, ...,$ and $j = 1, 2, ..., m$, $Z_j(n, \theta)$ can be made continuous by defining $Z_j(n, \theta_j) = 0$. We now show that the $Z_j(n, \theta)$ are also bounded uniformly in $n$.

**Lemma 3.7** *Let the convergence $w_j^{(n)} \to e^{i\theta_j}$ be as in (95) for $j = 1, 2, ..., m$. There exist constants, $M_1, M_2, ...,$ such that for all $\theta$, and for all $n = 1, 2, ...,$ such that*

$$\left| (\zeta - w_j^{(n)}) \frac{1 - \cos n(\theta - \theta_j)}{\zeta - e^{i\theta_j}} \right| \leq M_j \tag{101}$$

**Proof:** Assume without loss of generality that $\theta_j = 0$. The function $Z_j(n, \theta)$ defined above takes the form $Z_j(n, \theta) = (\zeta - w_j^{(n)}) \frac{1 - \cos n\theta}{\zeta - 1}$. Suppose first that $0 < |\theta| \leq \frac{1}{n}$. Using (98), and (99) we obtain

$$\left| (\zeta - w_j^{(n)}) \frac{1 - \cos n\theta}{\zeta - 1} \right| \leq \frac{2}{n}(N_j + 1) \left| \frac{1 - \cos n\theta}{\theta} \right|. \tag{102}$$

Now observe that

$$0 < \max_{|\theta| < 1/n} \left\{ \left| \frac{1 - \cos n\theta}{\theta} \right| \right\} = \max_{|\theta| < 1} \left\{ \frac{n(1 - \cos \theta)}{\theta} \right\} < n, \tag{103}$$

where the inequality follows since $\frac{1 - \cos \theta}{\theta} < 1$ for all $\theta$. Equations (102) and (103) now give

$$\left| (\zeta - w_j^{(n)}) \frac{1 - \cos n\theta}{\zeta - 1} \right| < 2(N_j + 1) \quad \text{for} \quad |\theta| \leq 1/n. \tag{104}$$

Now suppose that $|\theta| > 1/n$. With (95) and (98) we obtain

$$\begin{aligned} \left| \frac{\zeta - w_j^{(n)}}{\zeta - 1} \right| \leq 1 + \frac{|w_j^{(n)} - 1|}{|\zeta - 1|} &\leq 1 + \frac{2N_j(\frac{1}{n})}{\theta} \\ &= 1 + 2N_j, \quad \text{for} \quad 1/n \leq |\theta|. \end{aligned}$$

Thus

$$\left|(\zeta - w^{(n)})\frac{1-\cos n\theta}{\zeta - 1}\right| \leq 2(1+2N_j). \tag{105}$$

Equations (104) and (105) now give (101) with $M_j = 2 + 4N_j$. □

Lemma 3.5 of the last section addressed the $\mathcal{L}_1$ convergence of the factor $\frac{\prod_{j=1}^{m}(\zeta - w_j^{(r)})}{\prod_{j=1}^{m}|\zeta - re^{i\theta_j}|^2}$ of (74). The factor $\frac{\prod_{j=1}^{m}(\zeta - w_j^{(n)})}{\prod_{j=1}^{m}(\zeta - e^{i\theta_j})}$ of the present section is neither bounded, nor does it converge in $\mathcal{L}_1$. However, we have the following result, which is one of the key ideas used in the proof Theorem 3.5, the second main result of this work.

**Lemma 3.8** *Let $I_j(\delta) = (\theta_j - \delta, \theta_j + \delta)$ and let $\mathcal{X}_\delta$ be the indicator function for the set $[\pi, \pi) \setminus \cup_{j=1}^{m} I_j(\delta)$. Let the convergence $w_j^{(n)} \to e^{i\theta_j}$ be as in (95) for $j = 1, 2, ..., m$. Then*

$$\lim_{n\to\infty} \frac{\prod_{j=1}^{m}(\zeta - w_j^{(n)})}{\prod_{j=1}^{m}(\zeta - e^{i\theta_j})} \mathcal{X}_{n^p} = 1 \quad \text{in} \quad \mathcal{L}_1(d\theta) \quad \text{for} \quad -1 \leq p < 0 \ .$$

**Proof:** Since $\frac{\prod_{j=1}^{m}(\zeta - w_j^{(n)})}{\prod_{j=1}^{m}(\zeta - e^{i\theta_j})} \mathcal{X}_{n^p} \to 1$ pointwise for $p < 0$, except at the $\theta_j$, we need only show that the function is bounded for $-1 \leq p < 1$. The lemma will then follow from Lebesgue's Bounded Convergence Theorem.

Suppose that $-1 \leq p < 0$. Clearly, it suffices to show that $\frac{\zeta - w_j^{(n)}}{\zeta - e^{i\theta_j}} \mathcal{X}_{n^p}$ is bounded for each $j = 1, 2, ..., m$. With (98), the convergence in (95) gives

$$\begin{aligned}
\left|\frac{\zeta - w_j^{(n)}}{\zeta - e^{i\theta_j}}\right| \mathcal{X}_{n^p} &\leq \left\{1 + \left|\frac{w_j^{(n)} - e^{i\theta_j}}{\zeta - e^{i\theta_j}}\right|\right\} \mathcal{X}_{n^p} \\
&\leq 1 + \frac{N_j/n}{n^p/2} \\
&= 1 + \frac{2N_j}{n^{p+1}} \\
&\leq 1 + 2N_j.
\end{aligned}$$

□

We will need one more simple lemma before proceeding.

**Lemma 3.9** *Let $f_n \to f$ in $\mathcal{L}_1(d\theta)$, and let $g_n$ be a bounded sequence with $|g_n(\theta)| \leq M$ for all $n$ and $\theta$. Then*

$$\lim_{n\to\infty} \int f_n g_n \, d\theta = \lim_{n\to\infty} \int f g_n \, d\theta.$$

*provided the limits exist.*

**Proof:** We have

$$\begin{aligned} \left| \int f_n g_n \, d\theta - \int f g_n \, d\theta \right| &\leq \int |f_n g_n - f g_n| \, d\theta \\ &\leq M \int |f_n - f| \, d\theta. \end{aligned}$$

Letting $n \to \infty$ gives the lemma. □

We can now state and prove our second main result; that the limit $\mathcal{P}_k(z)$ defined in (81) is the same as that found in the case of the Poisson kernel of the last section.

**Theorem 3.5** *Let $\mu_n$ be given in (78). Then*

$$\mathcal{P}_k(z) := \lim_{n\to\infty} P_k(z, \mu_n) = P_{k-m}(z,\nu) \prod_{j=1}^{m} (z - e^{i\theta_j}), \tag{106}$$

*where*

$$\frac{d\nu}{d\theta} = \sum_{j=1}^{m} \alpha_j \prod_{p\neq j}^{m} |\zeta - e^{i\theta_p}|^2. \tag{107}$$

**Proof:** By the orthogonality characterization, (8), of the $P_k(z,\nu)$, the result will be proved if we can show that

$$\int Q(\zeta)\overline{t(\zeta)} \sum_{j=1}^{m} \alpha_j \prod_{p\neq j}^{m} |\zeta - e^{i\theta_p}|^2 \, d\theta = 0, \tag{108}$$

where $Q(\zeta)$ is the uniform limit of $Q_n(\zeta)$ defined in (81) and (80), and $t(z) \in \mathbf{\Lambda}_{k-m-1}$ is arbitrary.

Let $I_j(\delta) = (\theta_j - \delta, \theta_j + \delta)$, where $\delta > 0$ is small so that the $I_j$ are disjoint. Let $\mathcal{X}_\delta(\theta)$ be the indicator function for $[-\pi,\pi) \setminus \cup_{j=1}^{m} I_j(\delta)$. Then (96) gives

$$0 = \int \gamma(n,\zeta) \, d\theta = \int \gamma(n,\zeta)\mathcal{X}_\delta(\theta) \, d\theta + \int_{\cup_{j=1}^{m} I_j(\delta)} \gamma(n,\zeta) \, d\theta \, . \tag{109}$$

We will use Lemma 3.7 to show that $\gamma(n,\zeta)$ is bounded on $\cup_{j=1}^{m} I_j(\delta)$ uniformly in $n$, and therefore that the first integral on the right hand side of (109) is small if $\delta$ is small.

First, we rewrite $\gamma(n,\zeta)$. For $s \in \{1,2,...,m\}$, we can write

$$\frac{\prod_{j=1}^{m}(\zeta - w_j^{(n)})}{\prod_{j=1}^{m}(\zeta - e^{i\theta_j})} = \left(\frac{\zeta - w_s^{(n)}}{\zeta - e^{i\theta_s}}\right) \frac{\prod_{j\neq s}^{m}(\zeta - w_j^{(n)})}{\prod_{j\neq s}^{m}(\zeta - e^{i\theta_j})}$$

and

$$\sum_{j=1}^{m} \alpha_j(1-\cos n(\theta-\theta_j)) \prod_{p\neq j} |\zeta - e^{i\theta_p}|^2 =$$
$$\alpha_s(1-\cos n(\theta-\theta_s)) \prod_{p\neq s} |\zeta - e^{i\theta_p}|^2 \quad + \quad \sum_{j\neq s} \alpha_j(1-\cos n(\theta-\theta_j)) \prod_{p\neq j} |\zeta - e^{i\theta_p}|^2.$$

In the sum appearing on the right hand side of the above, $j \neq s$. Thus each of the products which appear in the terms of the sum contain the factor $|\zeta - e^{i\theta_s}|^2$. Therefore this sum can be written

$$|\zeta - e^{i\theta_s}|^2 \sum_{j\neq s} \alpha_j(1-\cos n(\theta-\theta_j)) \prod_{p\neq j,s} |\zeta - e^{i\theta_p}|^2.$$

Hence

$$\gamma(n,\zeta) =$$
$$Q_n(\zeta)\overline{t(\zeta)} \frac{\prod_{j\neq s}^{m}(\zeta - w_j^{(n)})}{\prod_{j\neq s}^{m}(\zeta - e^{i\theta_j})} \left\{ \left(\frac{\zeta - w_s^{(n)}}{\zeta - e^{i\theta_s}}\right) \alpha_s(1-\cos n(\theta-\theta_s)) \prod_{p\neq s} |\zeta - e^{i\theta_p}|^2 \right.$$
$$\left. + \quad (\zeta - w_s^{(n)})(\overline{\zeta - e^{i\theta_s}}) \sum_{j\neq s} \alpha_j(1-\cos n(\theta-\theta_j)) \prod_{p\neq j,s} |\zeta - e^{i\theta_p}|^2 \right\}. \tag{110}$$

Now suppose that $\theta \in I_s(\delta)$. Lemma 3.7 with (98) and (99) now give

$$|\gamma(n,\zeta)| \leq |Q_n(\zeta)\overline{t(\zeta)}| \frac{(2^{m-1})}{(\frac{\delta}{2})^{m-1}} \left\{ (M_s)2^{2(m-1)} + (2)(2) \sum_{j\neq s} \alpha_j (2)(2^{2(m-2)}) \right\}. \tag{111}$$

Now $Q_n(\zeta)\overline{t(\zeta)}$ is bounded uniformly in $n$ and $\theta$ by uniform convergence, so we see from (111) that $\gamma$ is bounded uniformly in $n$ for $\theta \in I_s(\delta)$ for $s = 1,2,...,m$. Since $\gamma$ is clearly bounded uniformly on $[-\pi,\pi) \setminus \cup I_j(\delta)$, there exists $M$ such that

$$|\gamma(n,\zeta)| \leq M \quad \text{for all} \quad n \quad \text{and} \quad \theta. \tag{112}$$

It follows from (112) and (109) that

$$\left| \int \gamma(n,\theta)\mathcal{X}_\delta \, d\theta \right| = \left| \int_{\cup I_j(\delta)} \gamma(n,\theta)\, d\theta \right| \leq Mm\delta.$$

With $\delta = \frac{1}{n}$ we have

$$\lim_{n\to\infty} \int \gamma(n,\theta)\mathcal{X}_{1/n}\, d\theta = 0. \tag{113}$$

Note that the integrand $\gamma(n,\theta)\mathcal{X}_{1/n}$ does not converge in $\mathcal{L}_1$ since the factors $Z(n,\theta)$ do not converge, as previously observed. We can, however, express $\gamma(n,\theta)\mathcal{X}_{1/n}$ as the sum of two terms, one of which does converge in $\mathcal{L}_1$.

To this end we write

$$\begin{aligned}\gamma(n,\theta) &= Q_n(\zeta)\overline{t(\zeta)}\prod_{j=1}^{m}\frac{(\zeta - w_j^{(n)})}{(\zeta - e^{i\theta_j})}\sum_{j=1}^{m}\alpha_j\prod_{p\neq j}^{m}|\zeta - e^{i\theta_p}|^2 \\ &- Q_n(\zeta)\overline{t(\zeta)}\prod_{j=1}^{m}\frac{(\zeta - w_j^{(n)})}{(\zeta - e^{i\theta_j})}\sum_{j=1}^{m}\alpha_j\cos n(\theta-\theta_j)\prod_{p\neq j}^{m}|\zeta - e^{i\theta_p}|^2.\end{aligned} \tag{114}$$

By Lemma 3.8 and the uniform convergence of $Q_n$ we see that

$$Q_n(\zeta)\overline{t(\zeta)}\prod_{j=1}^{m}\frac{(\zeta - w_j^{(n)})}{(\zeta - e^{i\theta_j})}\mathcal{X}_{1/n} \longrightarrow Q(\zeta)\overline{t(\zeta)} \quad \text{in} \quad \mathcal{L}_1. \tag{115}$$

Thus, regarding the first term on the right hand side of (114), we have

$$\begin{aligned}\lim_{n\to\infty}\int Q_n(\zeta)\overline{t(\zeta)}\prod_{j=1}^{m}\frac{(\zeta - w_j^{(n)})}{(\zeta - e^{i\theta_j})}\sum_{j=1}^{m}\alpha_j\prod_{p\neq j}^{m}|\zeta - e^{i\theta_p}|^2\,\mathcal{X}_{1/n}\,d\theta \\ = \int Q(\zeta)\overline{t(\zeta)}\sum_{j=1}^{m}\alpha_j\prod_{p\neq j}^{m}|\zeta - e^{i\theta_p}|^2\,d\theta\,.\end{aligned} \tag{116}$$

Regarding the second term on the right hand side of (114), we use Lemma 3.9 with (115) and the fact that $\sum_{j=1}^{m}\alpha_j\cos n(\theta-\theta_j)\prod_{p\neq j}^{m}|\zeta - e^{i\theta_p}|^2$ is uniformly bounded in $n$ and $\theta$ to write

$$\begin{aligned}\lim_{n\to\infty}\int Q_n(\zeta)\overline{t(\zeta)}\prod_{j=1}^{m}\frac{(\zeta - w_j^{(n)})}{(\zeta - e^{i\theta_j})}\sum_{j=1}^{m}\alpha_j\cos n(\theta-\theta_j)\prod_{p\neq j}^{m}|\zeta - e^{i\theta_p}|^2\mathcal{X}_{1/n}\,d\theta \\ = \lim_{n\to\infty}\int Q(\zeta)\overline{t(\zeta)}\sum_{j=1}^{m}\alpha_j\cos n(\theta-\theta_j)\prod_{p\neq j}^{m}|\zeta - e^{i\theta_p}|^2\,d\theta\,.\end{aligned} \tag{117}$$

**Claim:** *The limit, (117), is zero.*

To prove the claim, let $f(\theta) = Q(\zeta)\overline{t(\zeta)}\prod_{p\neq j}^{m}|\zeta - e^{i\theta_p}|^2$, and let $\mathcal{R}e\{f\}$ and $\mathcal{I}m\{f\}$ be the real and imaginary parts of $f$, respectively. Then

$$\sum_{j=1}^{m}\alpha_j\int \mathcal{R}e\{f(\theta)\}e^{in(\theta-\theta_j)}\,d\theta = \sum_{j=1}^{m}\alpha_j e^{in\theta_j}\widehat{\mathcal{R}e\{f\}}(n). \tag{118}$$

Since $\mathcal{R}e\{f\}$ is an $\mathcal{L}^2$ function, by the Riemann-Lebesgue Lemma,

$$\lim_{n\to\infty}\widehat{\mathcal{R}e\{f\}}(n) = 0. \tag{119}$$

From (118) and (119) we have

$$\lim_{n\to\infty}\sum_{j=1}^{m}\alpha_j\int \mathcal{R}e\{f(\theta)\}(\cos n(\theta-\theta_j)+i\sin n(\theta-\theta_j))\,d\theta=0. \tag{120}$$

Setting real and imaginary parts of (120) equal to zero we obtain

$$\lim_{n\to\infty}\sum_{j=1}^{m}\alpha_j\int \mathcal{R}e\{f(\theta)\}\cos n(\theta-\theta_j)\,d\theta=0. \tag{121}$$

Similarly, one can show that

$$\lim_{n\to\infty}\sum_{j=1}^{m}\alpha_j\int \mathcal{I}m\{f(\theta)\}\cos n(\theta-\theta_j)\,d\theta=0. \tag{122}$$

The claim follows from adding (121) and (122).

With the claim and equations (113), (114), and (116), we see that (116) is zero. Thus (108) holds and the theorem is proved. □

The following is immediate.

**Corollary 3.2**

$$\lim_{r\to 1}P_k(z,\psi_r*\sum_{j=1}^{m}\alpha_j\delta_{\theta_j})=\lim_{n\to\infty}P_k(z,\phi_n*\sum_{j=1}^{m}\alpha_j\delta_{\theta_j}).$$

### 3.2.5 Special Cases and Related Results

In this section we will consider the measure $\nu(\theta)$ of Theorems 3.4 and 3.5 and relate these results to a result in [P] concerning the limit of the reflection coefficients of Szegö polynomials, with respect to a measure whose density is the squared modulus of a polynomial, as $k\to\infty$. We then consider convolution of $m$ point masses with either the Poisson or Fejér kernel, for $m=1$ and $m=2$. For $m=1$ we exhibit the limit polynomial $\mathcal{P}_k$ characterized in Theorems 3.4 and 3.5. For the case $m=2$, we factor the density $\frac{d\nu}{d\theta}$ of Theorems 3.4 and 3.5 as the squared modulus of a linear function. Using results in [P] we relate the modulus of the zero of this function to the distribution of the zeros of $P_{k-2}(z,\nu)$ as $k\to\infty$. The explicit form of the limit $\mathcal{P}_k$ is given for a "degenerate" case. The $m=2$ situation, in light of Theorems 3.4 and 3.5 and the results in [P], is used to interpret an example of Petersen in [Pe2].

The density $\frac{d\nu}{d\theta}$ of Theorems 3.4 and 3.5 is a non-negative trigonometric polynomial of degree $m-1$. Recall, from Section 2.3.2, that this density can be factored as the square of a polynomial $g(z)$ of degree $m-1$ on the unit circle, with no zeros inside the unit circle. Thus, if

$$d\nu/d\theta = \sum_{j=1}^{m} \alpha_j \prod_{p\neq j}^{m} |\zeta - e^{i\theta_p}|^2 , \tag{123}$$

then there exists $g(z)$ of the form

$$|g(z)|^2 = c \prod_{j=1}^{m-1} |(z - v_j)|^2$$

such that

$$|g(\zeta)|^2 = d\nu/d\theta = \sum_{j=1}^{m} \alpha_j \prod_{p\neq j}^{m} |\zeta - e^{i\theta_p}|^2 . \tag{124}$$

It is easy to see that for arbitrary $z_0$, $|\zeta - z_0|^2 = |(\zeta - z_0)(\zeta\overline{z}_0 - 1)|$. Thus, we have the spectral factorization of $\nu$ on the unit circle:

$$\begin{aligned} d\nu/d\theta = |g(\zeta)|^2 &= c \prod_{j=1}^{m-1} |(\zeta - v_j)|^2 \\ &= c \prod_{j=1}^{m-1} |(\zeta - v_j)(\zeta\overline{v}_j - 1)|, \end{aligned} \tag{125}$$

where $c$ is a positive constant. Now the zeros of $c\prod_{j=1}^{m-1}(z - v_j)(z\overline{v}_j - 1)$ are symmetric with respect to the unit circle, and those of modulus equal 1 are of even multiplicity. We therefore can specify that

$$g(z) = \sqrt{c} \prod_{j=1}^{m-1} (z - v_j), \tag{126}$$

so that $|v_j| \leq 1$ for $j = 1, 2, ... m-1$.

Densities of the form (125), and the associated Szegö polynomials are considered by in [P]. There, the asymptotic behavior as $k \to \infty$ of the zeros of $P_k(z, \nu)$ is studied. Suppose that (123) and (126) hold, and assume the following:

1. The $v_j$ are distinct.

2. There is a unique $v_j$ of maximum modulus; without loss of generality

$$1 \geq r = |v_1| > |v_j|, \quad j = 2, 3, ..., m-1.$$

Lemma 5 of [P] states that if $\nu$ has a density that can be factored in the form (126) on the unit circle where Assumptions 1 and 2 hold, then

$$\lim_{n\to\infty} |P_k(0,\nu)|^{1/k} = r. \tag{127}$$

Assumptions 1 and 2 are key considerations here. As discussed in [P], if they do not hold it may only be the case that

$$\overline{\lim}_{k\to\infty}|P_k(0,\nu)|^{1/k} = r.$$

The reflection coefficient $|P_k(0,\nu)|^{1/k}$ is the geometric mean of the zeros of $P_k(z,\nu)$, and (127) gives information about the modulus of the zeros of $P_k$ for large $k$.

It has been observed ([Ku], [S]) that for several different processes, including damped sinusoids (which we consider in Section 5.4) the zeros of polynomials used in AR estimation tend to become uniformly distributed on circles of various radii when the polynomial degree is large. In an attempt to interpret this observed phenomenon, Pakula, in [P], defines the *zero-distribution* measure, $\frac{1}{k}\sum_{j=1}^{k}\delta_{w_j}$, consisting of point masses of weight $1/k$ at each of the zeros, $w_1, w_2, ..., w_k$, of $P_k(z,\nu)$. This is a measure on the unit disk. We have the following (Theorem 4 of [P])

**Theorem 3.6 (Pakula)** *Suppose (127) holds. Then the zero distribution measures of $P_k(z,\nu)$ converge in the weak-* sense to the uniform measure on the circle of radius $r$.*

Recall that the main results of Sections 3.2.3 and 3.2.4 depend on an assumption analogous to Assumption 1 above; the $\theta_j$ of the point mass measure defined in (20) of Section 3.1 were assumed distinct. This was used in the proof of Lemma 3.2.

We now consider the zeros of $g$ and the the limit factor $P_{k-m}(z,\nu)$ in (75) and (106) more closely.

Denote, by $R_k$ the $k^{th}$ reflection coefficient for the limit polynomial $\mathcal{P}_k$ in (75) and (106). Since $\mathcal{P}_k$ has $m$ zeros of modulus 1, we have $|R_k|^{1/k} = |P_{k-m}(0,\nu)|^{1/k}$, so that

$$\lim_{k\to\infty} |R_k|^{1/k} = \lim_{k\to\infty} |P_{k-m}(0,\nu)|^{1/k} = \lim_{k\to\infty} |P_{k-m}(0,\nu)|^{1/k-m}. \tag{128}$$

So, in studying the behavior of the $R_k$, we need only consider the reflection coefficients of $P_n(z,\nu)$.

A density of the form (125) is called a *moving average* spectral density. Suppose that $g(z) = \sum_{j=0}^{m-1} b_j z^j$. The following nonlinear relationship between the coefficients, $b_j$ and the moments of $\nu$ can be found in [G+S], sec.1.12, p.21, and in [K], sec. 5.4, p.116).

$$\widehat{\nu}(\ell) = \begin{cases} c\sum_{j=0}^{m-1-\ell} \overline{b}_\ell b_{\ell+j} & 0 \le \ell \le m-1 \\ 0 & \ell \ge m \end{cases} \tag{129}$$

We will use (129) to perform the spectral factorization for measure $\nu$ of Theorems 3.4 and 3.5 for the case $m = 2$. First, we consider the simplest case.

**The Case** $m = 1$. We consider the case $m = 1$, corresponding to one point mass convolved with the approximate identity $\mathcal{K}_h$; either the Fejér or the Poisson kernel. With $\alpha := \alpha_1$, $\omega := \theta_1$ (32) becomes

$$\mu_h = \alpha\delta_\omega * \mathcal{K}_h,$$

and $\nu\, d\theta = \alpha\, d\theta$, that is, a constant multiple of Lebesgue measure. Since $1, z, z^2, \ldots$ are monic and orthogonal on the unit circle with respect to Lebesgue measure, Theorems 3.4 and 3.5 give

$$\lim_{h\to 0} P_k(z,\mu_h) = (z - e^{i\omega})P_{k-1}(z, \alpha\, d\theta) = z^{k-1}(z - e^{i\omega}).$$

Thus, all the non-signal zeros approach the origin.

**The Case** $m = 2$. With $m = 2$, the measure $\mu_h$ in (32) becomes

$$\mu_h = (\alpha_1\delta_{\theta_1} + \alpha_2\delta_{\theta_2}) * \mathcal{K}_h. \tag{130}$$

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

$$\mu_h = (\alpha\delta_0 + (1-\alpha)\delta_\omega) * \mathcal{K}_h. \tag{131}$$

In this case, the measure $\nu$ defined in (75) and (106) becomes

$$\nu\, d\theta = \alpha|\zeta - e^{i\omega}|^2 + (1-\alpha)|\zeta - 1|^2\, d\theta.$$

From (124) and (125) we have

$$\begin{aligned} |g(\zeta)|^2 &= \alpha|\zeta - e^{i\omega}|^2 + (1-\alpha)|\zeta - 1|^2 \qquad (132)\\ &= c|\zeta - v_0|^2. \qquad (133) \end{aligned}$$

If we write $g(z) = b_0 + b_1 z$ and obtain the Fourier coefficients $\widehat{\nu}(\ell)$, we can solve the system (129) for $b_0$ and $b_1$. In this simple case, we solve a quadratic equation and obtain

$$v_0 = \frac{1 \pm \sqrt{2\alpha(1-\cos\omega)(1-\alpha)}}{1-\alpha+\alpha e^{-i\omega}}. \tag{134}$$

Specifying that $|v_0| \leq 1$ leads to $g(z) = \sqrt{c}(z - v_0)$ where

$$v_0 = \frac{1 - \sqrt{2\alpha(1-\cos\omega)(1-\alpha)}}{1-\alpha+\alpha e^{-i\omega}}. \tag{135}$$

Figures 5 and 6 show how $v_0$ and varies with $\alpha$ for various values of $\omega$.

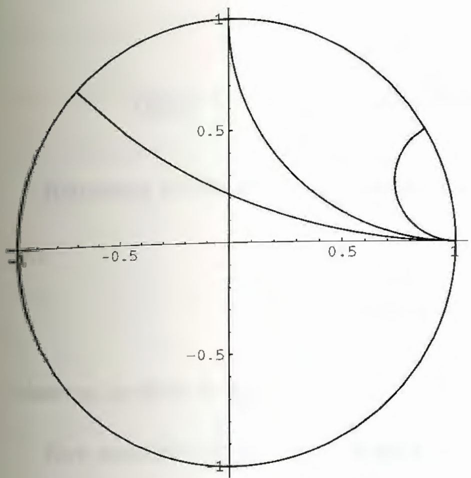


Figure 5: The zero $v_o$ as a function of $\alpha \in [0,1]$ for $\omega = \pi/6, \pi/2$, and $3\pi/4$.

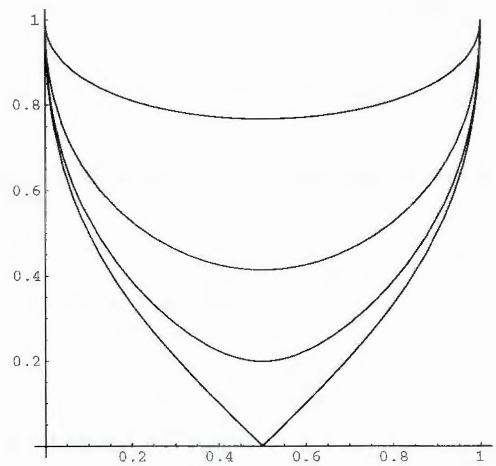


Figure 6: The modulus of $v_o$ as a function of $\alpha \in [0,1]$ for $\omega = \pi/6, \pi/2, 3\pi/4$, and $\pi$.

The case $\alpha = 1/2$, $\omega = \pi$ may be considered "degenerate" in the sense that, from (132), we see that $|g(\zeta)|^2 = 2$ is constant. That is, $g(z)$ is constant. So this situation is like the case $m = 1$, where we saw that all the the extraneous zeros of $P_k(z, \mu_h)$ approach the origin, and Theorems 3.4 and 3.5 give

$$\lim_{h \to 0} P_k(z, \mu_h) = z^{k-2}(z+1)(z-1).$$

Since $g$ is constant, it has no zeros. This is reflected in the fact that $v_0$ is undefined in (135). However, we have

$$\lim_{\alpha \to 1/2} v_0 = 0.$$

Figure 7 shows $v_0$, which is real, as a function of $\alpha$ for $\alpha \in [0,1]$. This includes the "degenerate" case $\alpha = 1/2$.

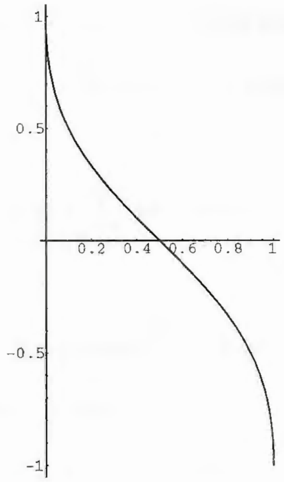


Figure 7: The zero of Fig 5, which is real for $\omega = \pi$, plotted for $\alpha \in (0, 1)$.

Returning to the general situation for $m = 2$, with $\mu_h$ defined in (131), Theorems 3.4 and 3.5 give

$$\lim_{h \to 0} P_k(z, \mu_h) = (z - e^{i\omega})(z - 1)P_{k-2}(z, |\zeta - v_0|^2 \, d\theta).$$

where $v_0$ is given in (135).

Now consider the situation with $k \to \infty$. With (128), Lemma 5 [P] gives

$$\lim_{k \to \infty} |R_k|^{1/k} = |v_0| = \frac{1 - \sqrt{2\alpha(1 - \cos\omega)(1 - \alpha)}}{\sqrt{1 + 2(1 - \alpha)^2(1 - \cos\omega)}}. \tag{136}$$

Furthermore, by Theorem 4 of [P], the zero distribution measures for $P_{k-2}(z, \nu)$ converge weak-* to the uniform measure on $|z| = |v_0|$. Now the two zeros of the limit polynomial $\mathcal{P}_k$ (from Theorems 3.4 and 3.5) on the unit circle will not contribute asymptotically to the zero-distribution measure for $\mathcal{P}_k$. Thus it follows from Theorem 4 of [P] that

> The zero-distribution measures for $\mathcal{P}_k$ converge in the weak-* sense to the uniform measure on the circle $|z| = |v_0|$, where $v_0$ is given by (135).

An example considered by Petersen [Pe2] can be seen as a special case of the above situation. Recall the real signal

$$x(n) = \sum_{j=-I}^{I} \alpha_j e^{in\theta_j}, \tag{137}$$

and the measures $\nu_N = \frac{1}{N}|X_N|^2$ formed from the Z-transform of the signal, defined in (36), (37), and the remarks preceding. We have $\alpha_0 > 0$, $\alpha_{-j} = \overline{\alpha}_j$, and $\theta_{-j} = -\theta_j$. Thus the signal can be represented as

$$x(n) = \alpha_0 + \sum_{j=1}^{I} 2|\alpha_j| \cos(n\theta_j + \gamma_j)$$

where the $\gamma_j = arg(\alpha_j)$.

Let $\widehat{\nu}_N(\ell)$ be the $\ell^{th}$ moment for the measure $\nu_N$. Jones, Njästed, and Waadeland [JNW], in an effort to address the non-uniqueness of limit, as $N \to \infty$ of the $P_k(z, \nu_N)$, define the *R-process* by replacing $\widehat{\nu}_N(\ell)$ with $r^{|\ell|}\widehat{\nu}_N(\ell)$ for $0 < r < 1$. The measure whose moments are thus defined is the convolution of $\nu_N$ with the density whose moments are $r^{|\ell|}$: the Poisson kernel $\psi_r$ defined in (29). That is, if

$$\mathcal{X}_{r,N} := \psi_r * \nu_N, \tag{138}$$

then

$$\widehat{\mathcal{X}}_{r,N}(\ell) = r^{|\ell|}\widehat{\nu}_N(\ell). \tag{139}$$

By (37) and the convergence of moments of a weak-* convergent sequence, (139) gives

$$\mathcal{X}_{r,\infty} := \lim_{N\to\infty} \mathcal{X}_{r,N} = \psi_r * \sum_{j=-I}^{I} |\alpha_j|^2 \delta_{\theta_j}, \tag{140}$$

which is of the form (58). Thus the R-process, in the limiting case $N = \infty$, can be seen as a specialization to real signals, of (58), and therefore, a special case of (32). We remark that in [JNS], a limit for the polynomials $P_k(z, \mathcal{X}_{r,N})$, for fixed $k$, as $r \to 1$ and $N \to \infty$ is neither exhibited nor characterized, as we do here in Theorem 3.4 for the case $N = \infty$. Furthermore, that a limit, merely exists as $r \to 1$ for the case $N = \infty$ is easily established by Proposition 1, which is quite simple and general. On the other hand, Petersen, in [Pe1], does prove the existence of limit as *both* $N \to \infty$ and $r \to 1$ in a prescribed manner. The fact that the signal is real is exploited in the proof, and the limit is not characterized.

Now consider the reflection coefficients $R_k$, of $P_k(z, \mathcal{X}_{r,\infty})$ approaching infinity. Shortly, we collect two immediate consequences of Theorem 3.4 and Lemma 5, [P] for the case where the limits are taken in the order $N \to \infty$, $r \to 1$, $k \to \infty$. In [Pe2], Petersen finds an explicit form for the reflection coefficients, where the limits are taken in this order, for what is seen as a special instance of the $m = 2$ case just considered in (130). Petersen considers a signal of the form

$$x(n) = \frac{1}{\sqrt{2}}e^{in\omega} + \frac{1}{\sqrt{2}}e^{-in\omega} = \sqrt{2}\cos n\omega, \tag{141}$$

and uses the R-process with $N = \infty$. The associated measure is, from (140)

$$\mathcal{X}_{r,\infty} := \psi_r * (\frac{1}{2}\delta_\omega + \frac{1}{2}\delta_{-\omega}). \tag{142}$$

Thus, $\mathcal{X}_{r,\infty} = \mu_h$ of (130), where $\alpha_1 = \alpha_2 = 1/2, \theta_1 = -\theta_2$, and $\mathcal{K}_h = \psi_r$. The moments of (141) are

$$\widehat{\mathcal{X}}_{r,\infty}(\ell) = r^\ell \cos \ell\omega\,.$$

It is shown (Proposition 2, [Pe]) that

$$\lim_{k\to\infty}\lim_{r\to 1} |P_k(0, \mathcal{X}_{r,\infty})|^{1/(k-2)} = \frac{|\cos\omega|}{1+\sin\omega}.$$

Using (128) and the definition of $R_k$ this can be written as

$$\lim_{k\to\infty} |R_k|^{1/k} = \frac{|\cos\omega|}{1+\sin\omega}. \tag{143}$$

This is readily seen to agree with (136). Letting $\theta_1 = \omega$, and $\theta_2 = -\omega$ in (130), equation (135) 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. (136) becomes

$$\begin{aligned}\lim_{k\to\infty} |R_k|^{1/k} &= \frac{1-\sin\omega}{\sqrt{1-\sin^2\omega}} \\ &= \frac{1-\sin\omega}{|\cos\omega|},\end{aligned}$$

which, if we define $v_0|_{\omega=\pi/2} = 0$, is equivalent to (143).

To summarize and place the example [Pe2] in the present context, observe that, with $\mathcal{X}_{r,\infty}$ defined in (142) and $\mu_r = \mathcal{X}_{r,\infty}$, the measure $\nu$, of Theorem 3.4, is simply a rotation of the measure (134) with $\alpha = 1/2$. (Recall that we assumed, without loss of generality, that $\theta_1 = 0$ and defined $\omega := \theta_2$ in (130).) The measure $\nu$ has density $|g(\zeta)|^2$ which is a rotation of (132), which can be factored as $\sqrt{c}|\zeta - v_0|^2$.

### The General Case; a Conjecture

Suppose that $m$ is arbitrary and let $\nu$ of Theorems 3.4 and 3.5 have density $d\nu = |g(\zeta)|^2$ with $g(z)$ given by (126). Let $R_k(h)$ denote the $k^{th}$ reflection coefficient of $P_k(z, \mu_h)$, and denote, by $\mu(h,k)$, the zero-distribution measure for $P_k(z, \mu_h)$. Equations (75) and (106) become

$$\lim_{h\to 0} P_k(z,\mu_h) = \prod_{j=1}^{m}(z - e^{i\theta_j})P_{k-m}(z, c\prod_{j=1}^{m-1}|\zeta - v_j|^2),$$

where we can assume that all the $v_j$ are contained in the closed unit disk. If, additionally, the $v_j$ are distinct and $|v_1| > |v_i|$ for $i = 2, 3, ..., m-1$, Lemma 5 [P], with (128) and the remarks that follow now give

**Corollary 3.3**

$$\lim_{k\to\infty}\lim_{h\to 0}|R_k(h)|^{1/k} = |v_1|.$$

By Theorem 3.6, we can thus conclude that

**Corollary 3.4**

$$\lim_{k\to\infty}\lim_{h\to 0}\mu(h,k) = \gamma$$

*where $\gamma$ is the uniform measure on $z = |v_1|$.*

The restriction on the $v_j$ can be stated a different way. The $v_j$ are a continuous function of the $\alpha_j$ and $\theta_j$. Vectors of the form $(\alpha_1, \alpha_2, ..., \alpha_m, \theta_1, \theta_2, ..., \theta_m)$ which result in non-distinct $v_j$ or multiple $v_j$ of maximum modulus form a set of measure zero in $\Re^{2m}$. Thus the above statements hold for almost every choice of the signal zeros $\theta_j$, and masses $\alpha_j$. We note that the above results apply to the generalization of the example in [Pe2], considered previously, to $m$ *complex* sinusoids.

We wish to consider the situation where $h \to 0$ and $k \to \infty$ simultaneously. The above remarks suggest that if $k \to \infty$ slowly enough, the measures $\mu(h,k)$ may converge. We make the following

**Conjecture:** *Let $\gamma$ be the uniform measure on $z = |v_1|$. For almost every choice of the signal zeros $\theta_j$, and masses $\alpha_j$,*

$$\lim \mu(h,k) = \gamma,$$

*where the limit is taken as $h \to 0$ and $k \to \infty$ in a manner to be determined.*

The asymptotic behavior of reflection coefficients for measures with rational densities will be studied in Section 5. We will also consider the reflection coefficients for a related measure associated with damped sinusoidal signals in Section 5.4.

# 4 A Mixed Spectral Measure

Another way to construct measures $\mu_h$ where (21) holds is to take a discrete measure of the form (20) and add an absolutely continuous measure which converges to zero as $h \to 0$. Let

$$\mu_h = \sum_{j=1}^{m} \alpha_j \delta_{\theta_j} + h\gamma \tag{144}$$

where $\gamma$ is an absolutely continuous measure. Here, the "mixed" measure $\mu_h$ is the spectral measure of a sum of complex sinusoids with additive noise, where $\gamma$ is the density of the noise process. In the case of white noise, for example, $\gamma$ is the uniform measure on the circle. We remark that, in contrast to measures obtained by convolution of point masses with approximate identities discussed in Section 3.2, measures of the form (144) actually converge in the strong sense; that is, in total variation norm, thus $\mu_h$ clearly satisfies (21). By Proposition 1, for $k > m$ all limit points of $P_k(z, \mu_h)$ are of the form (24). We remark that the fact that (21) holds for a measure of mixed type says little about the existence of a limit of the associated Szegö polynomials. We will comment further on this shortly. However in the present case we have

**Theorem 4.1** *Let $\mu_h$ be given in (144). As $h \to 0$, $P_k(z, \mu_h)$ approaches a unique limit, $\mathcal{P}_k$;*

$$\mathcal{P}_k(z) := \lim_{h \to 0} P_k(z, \mu_n) = P_{k-m}(z, \nu) \prod_{j=1}^{m} (z - e^{i\theta_j}), \tag{145}$$

*where*

$$d\nu(\theta) = \prod_{j=1}^{m} |\zeta - e^{i\theta_j}|^2 \, d\gamma(\theta) \,. \tag{146}$$

**Proof:** By the orthogonality property, (8), which characterizes the $P_k(z, \mu_h)$ we have, for any $q \in \Lambda_k$

$$\int P_k(\zeta, \mu_h) \overline{q(\zeta)} \, d\mu_h = 0.$$

With (144), we have

$$\sum_{j=1}^{m} \alpha_j P_k(e^{i\theta_j}, \mu_h) \overline{q(e^{i\theta_j})} + h \int P_k(\zeta, \mu_h) \overline{q(\zeta)} \, d\gamma = 0. \tag{147}$$

In particular, (147) holds for $q(z) = \Pi_{j=1}^{m}(z - e^{i\theta_j})r(z)$ where $r \in \mathbf{\Lambda}_{k-m-1}$. Thus, for all $h > 0$ we have

$$\int P_k(\zeta, \mu_h)\overline{r(\zeta)} \prod_{k=1}^{m} \overline{(\zeta - e^{i\theta_j})}\, d\gamma = 0. \tag{148}$$

Suppose that $h_n$ is a discretization such that

$$\lim_{n\to\infty} P_k(z, \mu_h) = Q(z) \prod_{j=1}^{m}(z - e^{i\theta_j}).$$

By simple compactness arguments, there are such convergent sequences (see Section 3.1). The convergence in the above limit is uniform on $|z| = 1$. So, since (148) holds for all $h$, we have

$$0 = \lim_{h\to 0} \int P_k(\zeta, \mu_h)\overline{r(\zeta)} \prod_{j=1}^{m} \overline{(\zeta - e^{i\theta_j})}\, d\gamma \tag{149}$$

$$= \int \prod_{j=1}^{m}(\zeta - e^{i\theta_j})Q(\zeta)\overline{r(\zeta)} \prod_{j=1}^{m} \overline{(\zeta - e^{i\theta_j})}\, d\gamma \tag{150}$$

$$= \int Q(\zeta)\overline{r(\zeta)} \prod_{j=1}^{m} |\zeta - e^{i\theta_j}|^2\, d\gamma. \tag{151}$$

Since $r \in \mathbf{\Lambda}_{k-m-1}$ is arbitrary, $Q \perp \mathbf{\Lambda}_{k-m-1}$ with respect to the measure $\Pi_{j=1}^{m}|\zeta - e^{i\theta_j}|^2\, d\gamma$. By the orthogonality property (8), $Q$ is the unique Szegö polynomial; $Q(z) = P_{k-m}(z, \Pi_{j=1}^{m}|z - e^{i\theta_j}|^2 d\gamma)$.

□

Theorem 4.1 suggests a way to construct $\mu_h$, where $\mu_h \to \sum \alpha_j \delta_{\theta_j}$ strongly, but such that $P_k(z, \mu_h)$ does not converge. Let

$$\mu_h = \sum_{j=1}^{m} \alpha_j \delta_{\theta_j} + h\gamma_h$$

where $\gamma_h$, is a non-convergent family of probability measures, and suppose that $\{h_{\ell_i}\}_{i=0}^{\infty}$ and $\{h_{n_i}\}_{i=0}^{\infty}$ are two discretizations of h such that

$$\lim_{i\to\infty} \gamma_{\ell_i} = \Gamma_1 \quad \text{weak} - *$$

and

$$\lim_{i\to\infty} \gamma_{n_i} = \Gamma_2 \quad \text{weak} - *.$$

Then by Theorem 4.1,

$$\lim_{i\to\infty} P(z, \mu_{h_{\ell_i}}) = P_{k-m}(z, \prod_{j=1}^{m} |\zeta - e^{i\theta_j}|^2\, d\Gamma_1) \prod_{j=1}^{m}(z - e^{i\theta_j}),$$

while

$$\lim_{n\in\Omega_2} P(z,\mu_{h_{n_i}}) = P_{k-m}(z, \prod_{j=1}^{m} |\zeta - e^{i\theta_j}|^2 \, d\Gamma_2) \prod_{j=1}^{m} (z - e^{i\theta_j}).$$

Clearly, using this construction, examples where an infinite number of limit points of $P_k(z, \mu_h)$ exist can be constructed.

# 5 Reflection Coefficients for Rational Densities

## 5.1 Introduction and Statement of Main Result

In this section we consider a fixed spectral measure with rational spectral density, and study the asymptotic behavior of the reflection coefficients, $R_k = P_k(0, \mu)$ as $k \to \infty$. This was considered in Section 3.2.5 for measures with densities of the form $\prod_{j=1}^{m} |\zeta - w_j|^2$. Lemma 5, [P], and Theorem 3.6 (Theorem 4 in [P]) describe asymptotic behavior of the reflection coefficients and zero distribution measures, respectively, of $P_k(z, \mu)$ in this case. We seek to generalize these results to measures with rational densities. The main result of this section is an extension of Lemma 5 in [P] regarding $\lim |R_n|^{1/n}$. The idea of the proof is sketched in [P]. The significance of our result is that the $\lim |R_n|^{1/n}$ exists, and not merely the $\limsup |R_n|^{1/n}$. As a corollary we will have an extension of Theorem 3.6 to rational spectral densities.

The behavior of reflection coefficients is considered in [S] and [NT], where we have the following

**Theorem 5.1 (Nevai and Totik):** *Let $\mu$ satisfy Szegö's condition and suppose that $g(z)$ is the function defined in (14). Then $\limsup |R_n|^{1/n} = r$ where $r$ is the smallest number such that $1/g(z)$ has an analytic continuation to the disk $|z| < 1/r$.*

For example, if

$$\begin{aligned} d\mu &= \left|\frac{p(\zeta)}{q(\zeta)}\right|^2 d\theta \\ &= \frac{\prod_{j=1}^{\ell} |\zeta - w_j|^2}{\prod_{j=1}^{m} |\zeta - v_j|^2} d\theta, \end{aligned} \tag{152}$$

then

$$g(z) = C \frac{\prod (z - \overline{w}_k^{-1})}{\prod (z - \overline{v}_k^{-1})}, \tag{153}$$

and $1/g(z)$ is analytic on $|z| < 1/r$ where $r = \max |w_j|$. Thus $\limsup |R_n|^{1/n} = r$.

We now state the main result of this section.

**Theorem 5.2** *Suppose that* $\mu$ *has a rational density of the form (152). Let* $R_k(\mu)$ *be the* $k^{th}$ *reflection for the measure* $\mu$*, as defined in (19). We assume, without loss of generality, that all the zeros of* $p(z)$ *and* $q(z)$ *lie within the closed unit disk, and that* $w_j \neq 0$ *for* $j = 1, 2, \ldots, \ell$*. We further assume the following:*

1. *The zeros* $w_j$ *are distinct.*

2. *There is a unique* $w_j$ *of maximum modulus, this being strictly less than 1:*

$$1 > |w_1| > |w_j|, j = 2, 3, ..., \ell.$$

*Then*

$$\lim_{k\to\infty} |R_k(\mu)|^{1/k} = |w_1|. \tag{154}$$

The proof of the theorem, though elementary, is quite technical, and we will need some results about the form of the inverse of the autocorelation matrix defined in Section 2.3.2 for the measure $\mu$ defined in (152). Therefore for clarity, before proving the theorem we will sketch the main ideas.

**Sketch of Proof:** Let $n = k + \ell$ and define

$$\begin{aligned} B(z) : &= P_k(z,\mu)p(z) \\ &= z^n + b_1 z^{n-1} + \ldots + b_n. \end{aligned} \tag{155}$$

By the minimization characterization, (4), of $P_k(z,\mu)$, $B(z)$ satisfies

$$\int |B(\zeta)|^2 \, |q(\zeta)|^{-2} d\theta = \min_{A\in\lambda_{k+\ell}} \int |A(\zeta)|^2 \, |q(\zeta)|^{-2} d\theta \tag{156}$$

where $\lambda_{k+\ell} \subseteq \mathbf{\Lambda}_{k+\ell}$ consists of monic polynomials having $\ell$ zeros at the $w_j$. Since $|b_n| = |R_k(\mu) \prod_{j=1}^{\ell} w_j|$, it is easy to see that

$$\begin{aligned} \lim_{n\to\infty} |b_n|^{1/n} &= \lim_{n\to\infty} |R_k(\mu)|^{1/n} \\ &= \lim_{k\to\infty} |R_k(\mu)|^{1/k}. \end{aligned} \tag{157}$$

So it suffices to show that

$$\lim_{n\to\infty} |b_n|^{1/n} = |w_1|. \tag{158}$$

The goal of the proof is to obtain a suitable expression for $b_n$. In fact, we will see (eq. (198)) that we can write

$$b_n = \sum_{i=1}^{\ell} a_i w_i^n \tag{159}$$

where the $a_i$ are bounded. The $a_i$ depend on the coefficients of the polynomials $p$ and $q$ appearing in the density $d\mu$, as well as th $w_j$ and $n$. With the aid of Theorem 5.1 we then show that $a_1 \neq 0$. To show that $\lim_{n\to\infty} |b_n|^{1/n} = |w_1|$ is then straightforward, and the theorem follows from (157).

To get an expression of the form (159), we recast the minimum property, (156), for $B$, in matrix form. Define

$$W := \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & w_1^{n-1} & w_1^{n-2} & \dots & 1 \\ 0 & w_2^{n-1} & w_2^{n-2} & \dots & 1 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & w_\ell^{n-1} & w_\ell^{n-2} & \dots & 1 \end{pmatrix}, \tag{160}$$

$$\mathbf{d} := [1 \ -w_1^n \ -w_2^n \ \dots \ -w_\ell^n]^T, \tag{161}$$

and

$$\mathbf{b} := [1 \ b_1 \ b_2 \ \dots \ b_n]^T \tag{162}$$

where we suppress the $n$-dependence for the vectors $\mathbf{d}$ and $\mathbf{b}$, and for the matrix $W$. In light of (165) and the remarks of sec. 5.2.1, the condition (156), is satisfied if and only if $\mathbf{b}C_n\mathbf{b}^*$ is a minimum, subject to $W\mathbf{b} = \mathbf{d}$, over all vectors $[1 \ a_1 \ a_2 \ \dots \ a_n]$, where $C_n$ is the covariance matrix for the measure $d\mu = |q(\zeta)|^{-2}d\theta$, as defined in (11). It follows that

$$\mathbf{b} = C_n^{-1}W^*(WC_n^{-1}W^*)^{-1}\mathbf{d}. \tag{163}$$

We now study the form of the matrix $C_n^{-1}$.

## 5.2 The Autocorrelation Matrix

### 5.2.1 Eigenvalues

Let $\mu$ be an absolutely continuous measure on the unit circle with $d\mu = f(\theta)d\theta$ and suppose that

$$0 < m \leq f(\theta) \leq M < \infty. \tag{164}$$

Let $C_n$ be the ACF matrix defined in (11) of section 2.3.2. We briefly consider an isometry between the space of polynomials of degree $n$ in $\mathcal{L}^2(d\mu)$ and $\mathcal{R}^{n+1}$.

To each polynomial $A(z)$, of degree $n$, associate the vector of coefficients, $\mathbf{a} = [a_0\ a_1\ \ldots\ a_n]^T$ so that $A(z) = a_0 z^n + a_1 z^{n-1} + \ldots + a_n$. If $A(z)$ and $B(z)$ are two polynomials of degree $n$ we can express the $\mathcal{L}^2(d\mu)$ inner product (see, for example, [L]) as

$$\begin{aligned} < A(z), B(z) >_{\mathcal{L}^2(d\mu)} &= \int A(\zeta)\overline{B(\zeta)}d\mu \\ &= \int \sum_{m=0}^{n}\sum_{\ell=0}^{n} a_m b_\ell^* e^{i(\ell-m)\theta} d\mu(\theta) \\ &= \sum_m \sum_\ell a_m b_\ell^* \int e^{i(\ell-m)\theta} d\mu(\theta) \\ &= \sum_m \sum_\ell a_m b_\ell^* \widehat{\mu}(\ell - m) \\ &= \mathbf{b}^* C_n \mathbf{a}. \end{aligned} \tag{165}$$

Thus, the space of polynomials of degree $n$ in $\mathcal{L}^2(d\mu)$ is isometrically isomorphic to $\mathcal{R}^{n+1}$ with the inner product defined in (165), and we have

$$||A(z)||^2 = \mathbf{a}^* C_n \mathbf{a}.$$

Using results in [GS, sec. 5.2] we find bounds for the eigenvalues of $C_n$ with $\mu$ bounded as above. Suppose that $\int |A(\zeta)|^2\, d\theta = 1$. Interpreting this in light of (165), with $d\mu = d\theta$ and $C_n$ as the identity matrix (or simply by direct computation), we have $\mathbf{a}^*\mathbf{a} = 1$. Now (164) and (165) give

$$m < \mathbf{a}^* C_n \mathbf{a} < M. \tag{166}$$

If $\mathbf{e}$ is an eigenvector of $C_n$ normalized so that $\mathbf{e}^*\mathbf{e} = 1$ with corresponding eigenvalue $\lambda$. Then

$$\mathbf{e}^* C_n \mathbf{e} = \mathbf{e}^* \lambda \mathbf{e} = \lambda. \tag{167}$$

So that

$$m \leq \lambda \leq M. \tag{168}$$

Note that these bounds are independent of $n$.

### 5.2.2 Inverses for a Class of ACF Matrices

Let $q(z) = \prod_{j=1}^{m} |z - v_j|^2$ and let $C_n$ be the ACF matrix defined in (11) of Section 2.3.2 for the measure $|q(\zeta)|^{-2}d\theta$. We will need some results for the form of $C_n^{-1}$. Spectral measures of the form $|q(\zeta)|^{-2}d\theta$ correspond to autoregressive processes of order $m$. The properties of $C_n^{-1}$ have been studied extensively ([K], [KVM], [Si]). Write $q(z) = z^m + q_1 z^{m-1} + \ldots + q_m$, and denote, by $\rho_n$, the prediction error power defined in (4):

$$\rho_n := \int |P_n(z, |q(\zeta)|^{-2})|^2 |q(\zeta)|^{-2} d\theta.$$

For $n > m$, we have $\rho_n = \rho_m$ (see, eg, [K], p.176). The following representation is due to Burg ([B]).

$$C_n^{-1} = Q_n \Delta_n^{-1} Q_n^* \tag{169}$$

$$Q_n = \begin{pmatrix} 1 & q_1^* & q_2^* & \ldots & q_m^* & 0 \ldots & & & 0 \\ 0 & 1 & q_1^* & q_2^* & \ldots & q_m^* & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & & & & & \ddots & \vdots \\ 0 & \ldots & 0 & 1 & q_1^* & \ldots & & q_m^* & 0 \\ 0 & \ldots & & 0 & 1 & q_1^* & & \ldots & q_m^* \\ \vdots & & & & \ddots & \ddots & & & \vdots \\ 0 & \ldots & & & & & 0 & 1 & q_1^* \\ 0 & \ldots & & & & & & 0 & 1 \end{pmatrix}, \tag{170}$$

and

$$\Delta_n = \begin{pmatrix} 1 & 0 & \ldots & & 0 \\ 0 & \rho_1 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ & & 0 & \rho_m & 0 \\ 0 & \ldots & & 0 & \rho_m \end{pmatrix}. \tag{171}$$

Note that $C_n, Q_n$, and $\Delta_n$ are $(n+1) \times (n+1)$ matrices whose elements depend on $n$.

Define $q_\xi := 0$ for $\xi > n$ and $\xi < 0$. We will denote the $\{i, j\}$ entry of a matrix $A$ by $\{A\}_{i,j}$. The Toeplitz-like character of $C_n^{-1}$ is reflected in the following, due to Trench ([T]):

$$\{C_n^{-1}\}_{i+1,j+1} = \{C_n^{-1}\}_{i,j} + \frac{1}{\rho_n}(q_i q_j^* - q_{n-i}^* q_{n-j}). \tag{172}$$

It follows from (172) that

$$\{C_n^{-1}\}_{i+1,j+1} = \{C_n^{-1}\}_{i,j} \quad \text{for} \quad m < i < n-m,\ m < j < n-m. \tag{173}$$

Thus $C_n^{-1}$ has a Toeplitz structure in its central portion. From the representation (169) we see that in general, the elements of $C_n^{-1}$ are products of vectors of the form $[1\ q_1\ \dots\ q_m]$ and its conjugate, possibly truncated, prepended and/or appended with zeros, and divided by the $\rho_j$, and have the form

$$\{C_n^{-1}\}_{i,j} = [0, \dots 0,\ \rho_{i-1}^{-1},\ q_1^*\rho_i^{-1}, \dots q_m^*\rho_{m+i-1}^{-1},\ 0 \dots 0][0 \dots 0\ 1,\ q_1, \dots q_m,\ 0 \dots 0]^T. \tag{174}$$

where $\rho_0 := 1$. Here, the first $i-1$ entries of the first vector on the right-hand side are zero, as are the last $n+1-i-m$ entries. The first $j-1$ entries of the second vector are zero, as are the last $n+1-j-m$ entries. More precisely, if we define $q_\xi = 0$ for $\xi < 0$ and $\xi > m$, we have the following, due to Siddiqui [Sd], which can also be found in [K], p. 176:

$$\{C_n^{-1}\}_{i,j} = \frac{1}{\rho_m}\sum_{\xi=1}^{n}(q_{i-\xi}q_{j-\xi}^* - q_{n-i+\xi}q_{n-j+\xi}^*) \quad \text{for} \quad i \geq j. \tag{175}$$

From the form (174) we see that each $\{C_n^{-1}\}_{i,j}$ has at most $m+1$ (possibly nondistinct) nonzero terms of the form $q_\xi q_\zeta^*$. From (175) we have

$$\{C_n^{-1}\}_{i,j} = 0 \quad \text{for} \quad |i-j| \geq m+1. \tag{176}$$

Since $C_n^{-1}$ is Hermitian, it follows from (176) that $C_n^{-1}$ is a band matrix with bandwidth $2m+1$. From (175) one can also show that

$$\{C_{n+1}^{-1}\}_{i,j} = \{C_n^{-1}\}_{i,j} \quad \text{for } i \text{ or } j < n-(m+1) \tag{177}$$

and

$$\{C_{n+1}^{-1}\}_{i+1,j+1} = \{C_n^{-1}\}_{i,j} \quad \text{for } i \text{ or } j \geq m+1. \tag{178}$$

Additionally, the matrix $C_n^{-1}$ is *persymmetric*, that is, it is symmetric with respect to the principal cross-diagonal (see [K]). This can be also be proved using (175), and is reflected in the relation

$$\{C_n^{-1}\}_{i,j} = \{C_n^{-1}\}_{n+1-i,n+1-j}. \tag{179}$$

Now (178) and (179) give

$$\{C_{n+1}^{-1}\}_{n+2,n+2-j} = \{C_n^{-1}\}_{n+2,n+1-j}. \tag{180}$$

Thus, $\{C_n^{-1}\}_{n+1,n+1-j}$ is constant in $n$. The salient points, for our purposes, of preceding discussion can be summarized as follows: $C_n^{-1}$ is a persymmetric band matrix of bandwidth $2m+1$. Each element, $\{C_n^{-1}\}_{i,j}$, is a sum of at most $m+1$ not necessarily distinct terms of the form $q_\xi q_\zeta^*$. Also, for $n > m+1$, the vectors $\tau_j$ do not change with $n$. This is evident from the form (174) and from (177) and (178).

## 5.3 Reflection Coefficients Revisited

We now proceed with the proof of the main result of the section.

**Proof of Theorem 5.2:**

We first consider the form of $C_n^{-1}W^*$. In light of (160), for $i,j>1$ we can represent the $i,j$ entry of $W^*$ as $\{W^*\}_{i,j} = (w_{j-1}^*)^{n-(i-1)}$, so that the elements of $C_n^{-1}W^*$ are of the form

$$\sum_\xi \{C_n^{-1}\}_{i,\xi}(w_{j-1}^*)^{n-(\xi-1)} \quad \text{for} \quad i,j>1. \tag{181}$$

We have

$$C_n^{-1}W^* =$$

$$\begin{pmatrix} \{C_n^{-1}\}_{1,1} & \sum_{j=1}^n (w_1^*)^{n-j}\{C_n^{-1}\}_{1,j+1} & \cdots & \sum_{j=1}^n (w_\ell^*)^{n-j}\{C_n^{-1}\}_{1,j+1} \\ \{C_n^{-1}\}_{2,1} & \sum_{j=1}^n (w_1^*)^{n-j}\{C_n^{-1}\}_{2,j+1} & \cdots & \sum_{j=1}^n (w_\ell^*)^{n-j}\{C_n^{-1}\}_{2,j+1} \\ \vdots & \vdots & & \vdots \\ \{C_n^{-1}\}_{n+1,1} & \sum_{j=1}^n (w_1^*)^{n-j}\{C_n^{-1}\}_{n+1,j+1} & \cdots & \sum_{j=1}^n (w_\ell^*)^{n-j}\{C_n^{-1}\}_{n+1,j+1} \end{pmatrix}. \tag{182}$$

From the observations of Section 5.2.2, the summation in any element of (182) is over at most $2m+1$ terms, each in turn consisting of at most $m+1$ terms of the form $\rho_i q_j q_k^*$. Moreover, the elements of $C_n^{-1}W^*$, considered as polynomials in the $q_j$ and $q_j^*$, have products of the $\rho_j$ and powers of the $w_j^*$ as coefficients, and as $n$ increases it is only the powers of the $w_j^*$ that may change.

Indeed, from the form of $C_n^{-1}$ and inspection of the the matrix $W$, for $i$ and $j$ fixed, the powers of $w_\xi^*$ appearing as coefficients increase with $n$. Since each of the $w_\xi$ have modulus less than 1, each $\{C_n^{-1}W^*\}_{i,j}$ is the partial sum of a convergent geometric sequence.

On the other hand, if $i$ is fixed, then for $n > i+m$ the last $i$ rows of $C_n^{-1}W^*$ are constant in $n$. In particular, by (176), $\{C_n^{-1}\}_{n+1,j} = 0$ for $j \le n+1-m$, so that the terms in each of the sums in the last row of (182) are non-zero only for $j > n+1-m$. We can thus express the last row of

(182) as

$$[0, \quad \sum_{j=0}^{m}(w_1^*)^j\{C_n^{-1}\}_{n+1,n+1-j}, \quad \ldots, \quad \sum_{j=0}^{m}(w_\ell^*)^j\{C_n^{-1}\}_{n+1,n+1-j}].$$

By (180), $\{C_n^{-1}\}_{n+1,n+1-j}$ is constant in $n$. It follows that each element of the last row of $C_n^{-1}W^*$ is constant in $n$ and we can now express the last row of $C_n^{-1}W^*$ as

$$[0 \quad \alpha_1 \quad \alpha_2 \quad \ldots \quad \alpha_\ell] \tag{183}$$

where $\alpha_1, \alpha_2, \ldots, \alpha_\ell$ are constants.

Define $\mathbf{q} = (q_1, q_2, \ldots, q_m, q_1^*, q_2^*, \ldots, q_m^*)$. Using the above arguments we write, suppressing the dependence upon the $\rho_j$,

$$C_n^{-1}W^* = \begin{pmatrix} \{C_n^{-1}\}_{1,1} & T_1(w_1,\mathbf{q}) & \ldots & T_1(w_\ell,\mathbf{q}) \\ \{C_n^{-1}\}_{2,1} & T_2(w_1,\mathbf{q}) & \ldots & T_2(w_\ell,\mathbf{q}) \\ \vdots & \vdots & & \vdots \\ \{C_n^{-1}\}_{m+1,1} & T_{m+1}(w_1,\mathbf{q}) & \ldots & T_{m+1}(w_\ell,\mathbf{q}) \\ 0 & T_{m+2}(w_1,\mathbf{q}) & \ldots & T_{m+2}(w_\ell,\mathbf{q}) \\ \vdots & \vdots & & \vdots \\ 0 & T_{n+1}(w_1,q) & \ldots & T_{n+1}(w_\ell,q) \end{pmatrix}, \tag{184}$$

where $T_i(w_j,\mathbf{q})$ is a polynomial in the elements of $\mathbf{q}$ whose coefficients are powers of $w_j^*$. Keep in mind that $T_{n+1}(w_j,q) = \alpha_j$.

Multiplying $C_n^{-1}W^*$ on the left by $W$ we obtain

$$WC_n^{-1}W^* =$$

$$\begin{pmatrix} \{C_n^{-1}\}_{1,1} & T_1(w_1,\mathbf{q}) & \ldots & T_1(w_\ell,\mathbf{q}) \\ \sum_{i=2}^{m+1} w_1^{n+1-i}\{C_n^{-1}\}_{i,1} & \sum_{i=2}^{n+1} w_1^{n+1-i}T_i(w_1,\mathbf{q}) & \ldots & \sum_{i=2}^{n+1} w_1^{n+1-i}T_i(w_\ell,\mathbf{q}) \\ \vdots & \vdots & & \vdots \\ \sum_{i=2}^{m+1} w_\ell^{n+1-i}\{C_n^{-1}\}_{i,1} & \sum_{i=2}^{n+1} w_\ell^{n+1-i}T_i(w_1,\mathbf{q}) & \ldots & \sum_{i=2}^{n+1} w_\ell^{n+1-i}T_i(w_\ell,\mathbf{q}) \end{pmatrix}. \tag{185}$$

Note that the sums in the first column are from $i = 2$ to $m+1$ since, by (176), $\{C_n^{-1}\}_{i,1} = 0$ for $i > m+1$. Thus, for $j = 1, 2, \ldots, \ell$ the smallest power of $w_j$ appearing in $\{WC_n^{-1}W^*\}_{j+1,1}$ is $n-m$. Denoting $U := WC_n^{-1}W^*$we can write

$$\begin{aligned} U_{j+1,1} &= \sum_{i=2}^{m+1} w_j^{n+1-i}\{C_n^{-1}\}_{i,1} \\ &= w_j^{n-m}\sum_{i=2}^{m+1} w_j^{m+1-i}\{C_n^{-1}\}_{i,1} \\ &= w_j^n\beta_j \end{aligned} \tag{186}$$

where $\beta_j$ is a polynomial in the elements of $\mathbf{q}$. Since, by (177), $\{C_n^{-1}\}_{i,1}$ is constant in $n$, $\beta_j$ is also constant in $n$. The remaining elements of $U$ either do not change with $n$, or are partial sums of convergent geometric series which converge to polynomials in the elements of $\mathbf{q}$. Thus $U_\infty := \lim_{n\to\infty} U$ exists.

**Claim 1:** *The eigenvalues of $WC_n^{-1}W^*$ are bounded away from zero as $n \to \infty$.*

**Proof of Claim 1:** The zeros of $q(z)$ lie strictly inside the unit circle, so there exist positive constants $M_1$ and $M_2$ such that $M_1 < |q(\zeta)|^{-2} < M_2$. By (168), if $\lambda$ is an eigenvalue of $C_n^{-1}$, then

$$\frac{1}{M_2} < \lambda < \frac{1}{M_1}. \tag{187}$$

Since $C_n^{-1}$ is also Hermitian and the zeros, $w_j$, of $p(z)$ are distinct, we see that $WC_n^{-1}W^*$ is a positive definite Hermitian matrix for all $n$. By the Rayleigh-Ritz theorem, to prove Claim 1 it suffices to show that there exists a positive constant $\varepsilon$ such that

$$\frac{\mathbf{t}^*WC_n^{-1}W^*\mathbf{t}}{\mathbf{t}^*\mathbf{t}} > \varepsilon \quad \text{for all } \mathbf{t} \text{ and all } n. \tag{188}$$

Of course, it would also follow that the limit matrix is invertible. We write

$$\frac{\mathbf{t}^*WC_n^{-1}W^*\mathbf{t}}{\mathbf{t}^*\mathbf{t}} = \frac{(W^*\mathbf{t})^*C_n^{-1}(W^*\mathbf{t})}{(W^*\mathbf{t})^*(W^*\mathbf{t})}\,\frac{\mathbf{t}^*WW^*\mathbf{t}}{\mathbf{t}^*\mathbf{t}}. \tag{189}$$

Rayleigh-Ritz and (188) applied to the first factor on the right-hand side of (189) yield

$$\frac{(W^*\mathbf{t})^*C_n^{-1}(W^*\mathbf{t})}{(W^*\mathbf{t})^*(W^*\mathbf{t})} \geq \frac{1}{M_2}.$$

Thus, (188) will be proved if we show that there exists a positive constant $\varepsilon$ such that

$$\frac{\mathbf{t}^*WW^*\mathbf{t}}{\mathbf{t}^*\mathbf{t}} > \varepsilon \quad \text{for all } \mathbf{t} \text{ and all } n. \tag{190}$$

Again, using Rayleigh-Ritz, (190) will hold if all eigen values of $WW^*$ are bounded away from zero, or, equivalently, if $\lim_{n\to\infty} WW^*$ is invertible.

This is similar to the situation addressed in [P]. Note that $WW^*$ is positive definite and Hermitian. We see that

$$WW^* = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & J & \\ 0 & & & \end{pmatrix}, \tag{191}$$

where $J$ is an $\ell \times \ell$ matrix with $J_{i,j} = \sum_{\xi=0}^{n-1} w_i^{n-\xi}(w_\xi^*)^{n-\xi}$. Thus $\lim_{n\to\infty} J_{i,j} = \frac{1}{1-w_i w_j^*}$. That $J_\infty := \lim_{n\to\infty} J$ is invertible is proved in the proof of Lemma 5, [P]. For completeness, we sketch the argument here. Define $f_j(z) = \frac{1}{1-zw_j^*}$. If $J_\infty$ were not invertible, there would exist constants $c_j$ such that function

$$f(z) = \sum_{j=1}^{m} c_j f_j(z)$$

has $\ell$ distinct zeros at the $w_j$. This is impossible since $f(z)$ can be expressed as a rational function with numerator of degree stricly less than $\ell$. Thus $J_\infty$ is invertible. It follows that $\lim_{n\to\infty} WW^*$ is also invertible, thus (190) holds and the claim is proved.

Now consider the elements of the first column of $U^{-1}$. Let $U_i$ be the matrix $U$ with the $i$th column replaced with the vector $[1\ 0\ 0\ \ldots\ 0]^T$. By Cramer's rule, we have

$$U_{i,1}^{-1} = \frac{\det U_i}{\det U}. \tag{192}$$

Let $H_i$ be the matrix obtained by striking the first row and $i$th column from $U_i$, and let $V(i,j)$ be the matrix obtained by striking the first column and $j$th row from $H_i$. For $i \geq 2$ we may write

$$\begin{aligned} \det U_i &= (-1)^{i+1} \det H_i \\ &= (-1)^{i+1} \sum_{j=1}^{\ell} U_{j+1,1} \det V(i,j) \\ &= (-1)^{i+1} \sum_{j=1}^{\ell} w_j^n \beta_j \det V(i,j) \end{aligned} \tag{193}$$

where we have used (186).

Now using (183), the first element of the last row of $C_n^{-1}W^*$, which we denote by $\Theta_1$, is

$$\begin{aligned} \Theta_1 &:= (C_n^{-1}W^*U^{-1})_{\ell+1,1} && (194) \\ &= \frac{1}{\det U} \sum_{i=1}^{\ell} \alpha_i (-1)^{i+1} \sum_{j=1}^{\ell} w_j^n \beta_j \det V(i,j) \\ &= \frac{1}{\det U} \sum_{j=1}^{\ell} w_j^n \beta_j \sum_{i=1}^{\ell} \alpha_i (-1)^{i+1} \det V(i,j) \\ &= \frac{1}{\det U} \sum_{j=1}^{\ell} w_j^n \gamma_j && (195) \end{aligned}$$

where the $\gamma_j$ converge to polynomials in the elements of $\mathbf{q}$.

In general, the elements $U_{i,j}^{-1}$ can be expressed as

$$U_{i,j}^{-1} = \frac{\tau(i,j)}{\det U} \tag{196}$$

where $\tau(i,j)$ is a sum of determinants of principal minors of $U$, so the $\tau(i,j)$ likewise converge to polynomials in the elements of $\mathbf{q}$.

Let $\Theta_j$ denote the $j$th element of the last row of $C_n^{-1}W^*$, with $\Theta_1$ defined in (195). Using (183) and (196), we can write

$$\begin{aligned} \Theta_j &= \frac{1}{\det U}\sum_{i=1}^{\ell}\alpha_i\tau(i+1,j) \\ &= \frac{\kappa_{j-1}}{\det U} \quad \text{for } j \geq 2 \end{aligned} \tag{197}$$

where $\kappa_{j-1}$ is again a polynomial in the elements of $\mathbf{q}$. Note that with this notation, the indices on the $\kappa_\xi$ run from 1 to $\ell$. With $\mathbf{b}$ and $\mathbf{d}$ defined in (162) and (161), (163) now gives $b_n = [\Theta_1\ \Theta_2\ \ldots\ \Theta_{\ell+1}][1\ -w_1^n\ w_2^n\ \ldots\ -w_\ell^n]^T$ so that

$$b_n = \frac{1}{\det U}\sum_{i=1}^{\ell}\Gamma_i w_i^n \tag{198}$$

where $\Gamma_i = \gamma_i - \kappa_i$. By the Claim 1, $\det U$ is bounded away from zero, thus it converges to a positive constant. The $\Gamma_i$ converge to polynomials in the elements of $\mathbf{q}$.

**Claim 2:** $\Gamma_1 \neq 0$.

**Proof of Claim 2:** Suppose to the contrary that $\Gamma_d \neq 0$, for some $d > 1$, but $\Gamma_i = 0$, for $i = 1, 2, .., d-1$. Also suppose, without loss of generality, that $|w_d| \geq |w_i|$ for $i = d+1, \ldots, \ell$. Then it is easy to see from (198) that $|b_n| \leq K|w_d|^n$ for some $K > 0$. But then $\limsup |b_n|^{1/n} \leq |w_d| < |w_1|$. This, with (157) contradicts Theorem 5.1 and the claim is proved.

Now write

$$|b_n| = \frac{|w_1|^n}{\det U}\left|\Gamma_1 + \sum_{j=2}^{\ell}\Gamma_j(w_j/w_1^n)\right|. \tag{199}$$

As argued in the proof of Lemma 4 [P], since $|w_j| < |w_1|$ for $j = 2, 3, \ldots, \ell$,

$$\lim_{n\to\infty}\left(\frac{\Gamma_1 + \sum_{j=2}^{\ell}\Gamma_j(w_j/w_1)^n}{\Gamma_1}\right) = 1,$$

and it follows, taking $n$th roots, that

$$\lim_{n\to\infty}(\Gamma_1 + \sum_{j=2}^{\ell}\Gamma_j(w_j/w_1)^n)^{1/n} = 1.$$

This, with (199) gives of **q**,

$$\lim_{n\to\infty}|b_n|^{1/n} = |w_1|,$$

which, along with (158), proves the theorem.

Recall the *zero-distribution* measure of $P_k(z,\mu)$, defined in the remarks preceding Theorem 3.6, which assigns mass $1/k$ at each of the zeros of $P_k(z,\mu)$. From Theorems 5.2 and 3.6 follows

**Corollary 5.1** *Let $\mu$ satisfy the hypotheses of Theorem 5.2. Then the zero distribution measures of $P_k(z,\mu)$ converge in the weak-* sense to the uniform measure on the circle of radius $r$.*

**Example:** The following example shows that Theorem 5.2 is false if we relax the requirement that there is a unique $w_j$ of maximum modulus. Let

$$d\mu(\theta) = |(\zeta - re^{i\theta_1})(\zeta + re^{i\theta_2})|^2\delta\theta$$

where $|\theta_1 - \theta_2| = \pi$. Then by direct computation using *Mathematica*, we find

$$|R_n| = \begin{cases} r^n/\sum_{j=0}^{n} r^{2n} & n \text{ even} \\ 0 \quad n \text{ odd} \end{cases}.$$

Thus $\overline{\lim}|R_n|^{1/n} = r$ while $\underline{\lim}|R_n|^{1/n} = 0$.

## 5.4 A Signal Consisting of Damped Sinusoids

We now consider a signal consisting of damped sinusoids. Let

$$x_n = \sum_{j=1}^{K}\alpha_j v_j^n \quad \text{for } n = 1, 2 \tag{200}$$

where $|v_j| < 1$ and $\alpha_j$ are complex. We can write $\alpha_j = a_j e^{i\omega_j}$ and $v_j = \rho_j e^{i\theta_j}$, for $j = 1, 2, \ldots, K$, where the $a_j$ are real amplitudes, $\rho_j < 1$ are damping factors, and the $\theta_j$ and $\omega_j$ are frequencies and phases, respectively. Several methods exist for exactly determining amplitudes, damping factors, frequencies, and phases for the signal (200) given at least $K$ of the $x_n$ (see [K], p. 224). Modern techniques such as the covariance method are variations of Prony's method, which dates to 1795. Several methods have been investigated for determining signal zeros for damped sinusoids when noise is present ([KT1, KT2]). Our interest here is in the mathematical properties of the Szegö polynomials, and to illuminate the behavior of AR techniques, rather than to propose alternate means of determining signal zeros in applications.

We consider a related sequence of measures derived from the signal (200) and illustrate some connections with the results of Section 5. Recall that $\zeta = e^{i\theta}$ represents an arbitrary point on the unit circle. We form an analog of the *periodogram* (see, eg. [K], Sec 4.3) by defining

$$d\mu_N(\theta) := |X_N(\zeta)|^2 \tag{201}$$

where $X_N(z)$ is the $z$-transform of the signal (200)

$$X_N(z) = \sum_{n=0}^{N} x_n z^{-n}. \tag{202}$$

The moments of $\mu_N$ are then

$$\begin{aligned}\widehat{\mu}_N(k) &= \int \zeta^k \sum_{m=0}^{N} x_m \zeta^{-m} \sum_{n=0}^{N} \overline{x}_n \zeta^n d\theta \\ &= \sum_{m=0}^{N} \sum_{n=0}^{N} \left( \int \zeta^{k+n-m} d\theta \right) x_m \overline{x}_n \\ &= \sum_{m=0}^{N-k} x_m \overline{x}_{m+k}. \end{aligned} \tag{203}$$

The $\widehat{\mu}_N(k)$ are seen to be multiples of the ACF function estimates used in the autocorrelation method (see [K], Sec 7.3). Given a signal $x_0, x_1, \ldots,$ of the form (200), this method is equivalent to using the arguments of the zeros of $P_k(z, \mu_N)$ as estimates of the frequencies, $\theta_j$. Multiplying (201) by $1/N$ yields the standard form for the periodogram, which is the basis for classical spectral estimation techniques ([K], Ch. 4). A factor of $1/N$ in (201) would in turn result in a factor of $1/N$ in (203).

The reason for not normalizing is that we wish to consider the behavior of $\mu_N$ as $N \to \infty$. Since $|v_j| < 1$, $X_N$ is bounded in $N$. On the other hand, $(1/N)\mu_N \to 0$. More precisely, from

(202) we have

$$
\begin{aligned}
X_N(z) &= \sum_{m=0}^{N} z^{-m} \sum_{j=1}^{K} \alpha_j v_j^m \\
&= \sum_{j=1}^{K} \sum_{m=0}^{N} z^{-m} \alpha_j v_j^m \\
&= \sum_{j=1}^{K} \sum_{m=0}^{N} \alpha_j (v_j z^{-1})^m,
\end{aligned}
$$

so that

$$
\lim_{N\to\infty} X_N(z) = \sum_{j=1}^{K} \frac{\alpha_j}{1 - v_j z^{-1}}. \tag{204}
$$

Thus

$$
\begin{aligned}
\lim_{N\to\infty} |X_N(\zeta)|^2 &= \left| \frac{\sum_{j=1}^{K} \alpha_j \prod_{m\neq j}(1 - v_m \zeta^{-1})}{\prod_{j=1}^{K}(1 - v_j \zeta^{-1})} \right|^2 \\
&= \left| \frac{\sum_{j=1}^{K} \alpha_j \prod_{m\neq j}(\zeta - v_m)}{\prod_{j=1}^{K}(\zeta - v_j)} \right|^2 .
\end{aligned} \tag{205}
$$

The numerator in (205) is a polynomial of degree $K-1$, and can be factored as

$$
d\mu := \lim_{N\to\infty} d\mu_N = C \left| \frac{\prod_{j=1}^{K-1}(\zeta - w_j)}{\prod_{j=1}^{K}(\zeta - v_j)} \right|^2 d\theta, \tag{206}
$$

where $C > 0$ and where, without loss of generality, the $|w_j| \leq 1$ (see Sec. 2.3.2). The $w_j$ depend on the both the signal zeros $v_j$ and the $\alpha_j$. Since $\alpha_j = a_j e^{i\omega_j}$, the $w_j$ depend on the phases of the frequency components of the signal. Write

$$
d\mu = C \left| \frac{p(\zeta)}{q(\zeta)} \right|^2 d\theta \tag{207}
$$

where $p(z) = \prod_{j=1}^{K-1}(\zeta - w_j)$ and $q(z) = \prod_{j=1}^{K}(\zeta - v_j)$. Thus, the asymptotic situation is the same as that considered in Sec. 5. In particular, if there is a unique $w_j$ of maximum modulus, which is less than 1, Theorem 5.2 holds, and by Corollary 5.1, the zero distribution measures of $P_k(z, \mu_N)$ converge in the weak-* sense to the uniform measure on $|z| = r$. We make the following conjecture:

**Conjecture:** *Let $x_n$ and $\mu_N$ be defined in (200), (201), and (202), and assume that the $v_j$ and $w_j$ in (206) are distinct and less than one in modulus, with $1 > r = |w_1| > |w_j|$ for $j = 2, 3, \ldots, K-1$. Then as $N$ and $k$ approach infinity in a manner to be determined, the zero distribution measures of $P_k(z, \mu_N)$ converge in the weak-* sense to the uniform measure on $|z| = r$.*

Note that the conjecture is true if we first let $N \to \infty$ while holding $n$ constant.

**Convergence of zeros of $P_n(z, \mu)$**

We now consider convergence of the zeros of $P_n(z, \mu)$ to signal zeros $v_j$. Suppose that the hypotheses of the the above conjecture hold, and that $|v_j| > r$ for $j = 1, 2, ..., d$, and $|v_j| \leq r$ for $j = d + 1, \ldots, K$. We will expand upon the remarks on p. 575, [P], and show that that as $n \to \infty$ the $d$ largest zeros of $P_n(z, \mu)$ converge to the $v_j$ for $j = 1, 2, ..., d$.

Let $k_n$ be defined so that $\phi_n(z) := k_n P_n(z, \mu)$ has $\mathcal{L}^2(d\mu)$ norm equal 1, so that the $\phi_n(\zeta)$ are orthonormal with respect to $\mu$, and let $\ell_n := \phi_n(0)$ be the constant. Then the reflection coefficients are $R_n = l_n/k_n$ and $k_n^{-2} = ||P_n(\zeta, \mu)||^2$ is the prediction error power defined in Section 2.3. Since $\mu$ satisfies Szegö's condition (see Sec. 2.3.2),

$$\lim_{n \to \infty} k_n > 0, \tag{208}$$

so that

$$\lim_{n \to \infty} |R_n|^{1/n} = \lim_{n \to \infty} |l_n|^1/n \tag{209}$$

if the limit exists. Recall, also from Section 2.3, the reverse polynomial, $\phi_n^*(z) := z^n \overline{\phi}_n(z^{-1})$. The following hold ([GS], Secs. 2.2, 3.4):

$$\sum_{j=0}^{n} \overline{l}_j \phi_j(z) = k_n \phi_n^*(z), \tag{210}$$

$$\lim_{n \to \infty} k_n \phi_n^*(z) = \frac{1}{g(0)g(z)} \quad \text{for } |z| < 1, \tag{211}$$

$$\lim_{n \to \infty} z^{-n} \phi_n(z) = \frac{1}{\overline{g}(z^{-1})} \quad \text{for } |z| > 1 \tag{212}$$

where $g(z)$ is defined in Section 2.3.2. Here,

$$g(z) = C \frac{\prod(z - \overline{w}_k^{-1})}{\prod(z - \overline{v}_k^{-1})},$$

where $C$ is defined so that $g(0) > 0$.

Let $1/r > \rho > 1$, and let $L = \max |1/\overline{g}(z^{-1})|$ on $|z| = |\rho\zeta| = \rho > 1$. Then, by (212), given $\varepsilon > 0$ there exists $M$ such that,

$$|\phi_n(\rho\zeta)| < (L + \varepsilon)\rho^n \quad \text{for } n > M.$$

By the maximum modulus principle the above bound holds for all $|z| \le \rho$; that is

$$|\phi_n(z)| < (L+\varepsilon)\rho^n \quad \text{for} |z| \le \rho \text{ and } n > M. \tag{213}$$

By Theorem 5.2 and (208), $\lim_{n\to\infty} |R_N|^{1/n} = r$. So (209) and (213) give

$$\lim |l_n \phi_n(z)|^{1/n} < \lim r|\rho^n(L+\varepsilon)|^{1/n} < 1 \tag{214}$$

for $|z| \le \rho < 1/r$. Thus, by (210), (214), and the root test $k_n\phi_n^*(z)$ converges uniformly on $|z| \le \rho$. The limit function is analytic, and, by (211), is the analytic continuation of $1/g(0)g(z)$.

We further restrict $\rho$ so that only the $d$ smallest zeros of $1/g(0)g(z)$ fall inside $|z| < \rho$. Since $k_n\phi_n^*(z) \to 1/g(0)g(z)$ uniformly on $|z| \le \rho$, for large $n$ $k_n\phi_n^*(z)$ will have no zeros on $|z| = \rho$. By Hurwitz' Theorem, $k_n\phi_n^*(z)$ has exactly $d$ zeros inside $|z| < \rho$ for large $n$. Since the zeros of $\phi^*(z)$ are those of $\phi(z)$ reflected in the unit circle, letting $\rho \to 1/r$ yields the following: $\phi_n(z)$ has exactly $d$ zeros in $|z| > r$; these approach $v_1, ..., v_d$.

In the context of estimating the signal zeros, $v_j$, via autocorrelation method, the preceding remarks suggest that for large $N$, the largest zeros of $P_n(z, \mu_N)$ will be close to those $v_j$ which lie outside $|z| = r = \max w_j$ in (206). Since the $w_j$ depend on the $\alpha_j$, and thus upon the relative phases $\omega_j$, in (200), it would be of interest to determine the exact nature of the dependence of the $w_j$ upon the phases, $\omega_j$. For example, what conditions on the phases guarantee that one or more of the $v_j$ lie outside $r$? There do not seem to be general results in this area, but the following corollary of Lucas' Theorem can be found in [M].

**Theorem:** (Lucas) *Suppose $\alpha_j > 0$ for $j = 1, 2, \ldots, K$. Then the zeros of (204) lie inside the closed convex hull of the $v_j$.*

## LIST OF CITATIONS


[A] Akheizer, N. *The Classical Moment Problem*, London: Oliver and Boyd, 1965.

[B] Burg, J. "Maximum Entropy Spectral Analysis", Ph.D. dissertation, Stanford University, 1975.

[D] Delsarte, P. and Genin, Y. "On the Role of Orthogonal Polynomials on the Unit Circle in Digital Signal Processing Applications", in *Orthogonal Polynomials*, (P. Nevai, ed.), Kluwer Academic, pp. 115-133, 1990.

[G1] Geronimus, I.A. *Orthogonal Polynomials: Estimates, Asymptotic Formulas, and Series of Polynomials Orthogonal on the Unit Circle and Interval*, New York: Consultants Bureau, 1961.

[G2] Geronimus, I.A. *Polynomials Orthogonal on a Circle and Interval*, New York: Pergamon Press, 1960.

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

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

[JN] Jones, W.B., Njåsted, O. "Applications of Szegö Polynomials to Digital Signal Processing", *Rocky Mountain Journal of Mathematics* Vol. 21, No.1, pp. 387-486, Winter, 1991.

[JNS] 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.

[JNW] 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.

[KVM] Kailath, T., Vieira, A., and Morf, M. " Inverses of Toeplitz Operators, Innovations, and Orthogonal Polynomials", *SIAM Review.*, Vol. 20, pp. 106-119, 1978.

[Ka] Katznelson, Y. *An Introduction to Harmonic Analysis*, New York: Dover, 1968.

[K] Kay, S. *Modern Spectral Estimation*, New Jersey: Prentice-Hall 1988.

[KP] Kay, S. and Pakula, L. " Simple Proofs of the Minimum Phase Property of the Prediction Error Filter", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-31, No. 2, p.501, April 1983.

[Ku] Kumaresan, R. "On the Zeros of the Linear Prediction Error Filter for Deterministic Signals", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-31, No. 1, pp. 217-220, Feb 1983.

[KT1] Kumaresan, R. and Tufts, D. "Estimating Parameters of Exponentially Damped Sinusoids and Pole-Zero Modeling in Noise", *IEEE Transactions on Speech, Acoustics, and Signal Processing*, Vol. ASSP-30, pp. 833-840, Dec. 1982.

[KT2] Kumaresan, R. and Tufts, D. "Accurate Parameter Estimation of Noisy Speech-Like Signals", *Conference Proceedings ICASSP*, Paris, 1982.

[L] Landau, H. J. "Maximum Enropy and the Moment Problem", *Bulletin of the American Mathematical Society*, Vol. 16, No. 1, pp. 47-77, January, 1987.

[M] Marden, M. *Geometry of Polynomials*, Providence, R.I.: American Mathematical Society, 1966.

[NT] Nevai, P. and Totik, V. "Orthogonal Polynomials and Their Zeros", *Acta Scientiarum Mathematicarum*, 53 pp. 99-104, 1989.

[P] 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.

[Pe1] 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.

[Pe2] 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.

[PS] 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.

[R] Rudin, W. *Real and Complex Analysis*, McGraw-Hill, 1966.

[S] Saff, E.B. "Orthogonal Polynomials From a Complex Perspective" in *Orthogonal Polynomials* (P.Nevai, ed.), Kluwer, pp. 363-393, 1990.

[Si] Siddiqui, M.M. "On the Inversion of the Sample Covariance Matrix in a Stationary Autoregressive Process", *Annals of Mathematical Statistics*, Vol. 58, pp.585-588, 1958.

[T] Trench, W. "An Algorithm for the Inversion of Finite Toeplitz Matrices", *SIAM Journal*, 12, pp. 515-522, 1964.

BIBLIOGRAPHY

Akheizer, N. *The Classical Moment Problem*, London: Oliver and Boyd, 1965.

Burg, J. "Maximum Entropy Spectral Analysis", Ph.D. dissertation, Stanford University, 1975.

Delsarte, P. and Genin, Y. "On the Role of Orthogonal Polynomials on the Unit Circle in Digital Signal Processing Applications", in *Orthogonal Polynomials*, (P. Nevai, ed.), Kluwer Academic, pp. 115-133, 1990.

Geronimus, I.A. *Orthogonal Polynomials: Estimates, Asymptotic Formulas, and Series of Polynomials Orthogonal on the Unit Circle and Interval*, New York: Consultants Bureau, 1961.

Geronimus, I.A. *Polynomials Orthogonal on a Circle and Interval*, New York: Pergamon Press, 1960.

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

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

Jones, W.B., and Njåsted, O. "Applications of Szegö Polynomials to Digital Signal Processing", *Rocky Mountain Journal of Mathematics* Vol. 21, No.1, Winter, pp.387-436, 1991.

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.

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., pp. 141-151, 1994.

Kailath, T., Vieira, A., and Morf, M. " Inverses of Toeplitz Operators, Innovations, and Orthogonal Polynomials", *SIAM Review.*, Vol. 20, pp. 106-119, 1978.

Katznelson, Y. *An Introduction to Harmonic Analysis*, New York: Dover, 1968.

Kay, S. *Modern Spectral Estimation*, New Jersey: Prentice Hall 1988.

Kay, S. and Pakula, L. " Simple Proofs of the Minimum Phase Property of the Prediction Error Filter", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-31, No. 2, p. 501, April 1983.

Kumaresan, R. "On the Zeros of the Linear Prediction Error Filter for Deterministic Signals", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-31, No. 1, pp.217-220, Feb 1983.

Kumaresan, R. and Tufts, D. "Estimating Parameters of Exponentially Damped Sinusoids and Pole-Zero Modeling in Noise", *IEEE Transactions on Speech, Acoustics, and Signal Processing*, Vol. ASSP-30, pp. 833-840, Dec. 1982.

Kumaresan, R. and Tufts, D. "Accurate Parameter Estimation of Noisy Speech-Like Signals", *Conference Proceedings ICASSP*, Paris, 1982.

Landau, H. J. "Maximum Enropy and the Moment Problem", *Bulletin of the American Mathematical Society,* Vol. 16, No. 1, pp. 47-77, January, 1987.

Marden, M. *Geometry of Polynomials,* Providence, R.I.: American Mathematical Society, 1966.

Nevai, P. and Totik, V. "Orthogonal Polynomials and Their Zeros", *Acta Scientiarum Mathematicarum,* 53 pp. 99-104, 1989.

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.

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.

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.

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.

Rudin, W. *Real and Complex Analysis,* McGraw-Hill, 1966.

Saff, E.B. "Orthogonal Polynomials From a Complex Perspective" in *Orthogonal Polynomials* (P.Nevai, ed.), Kluwer, pp. 363-393, 1990.

Siddiqui, M.M. "On the Inversion of the Sample Covariance Matrix in a Stationary Autoregressive Process", *Annals of Mathematical Statistics,* Vol. 58, pp.585-588, 1958.

Trench, W. "An Algorithm for the Inversion of Finite Toeplitz Matrices", *SIAM Journal,* 12, pp. 515-522, 1964.