EXTENSIONS OF ROSENBLATT'S RESULTS ON THE ASYMPTOTIC BEHAVIOR OF THE PREDICTION ERROR FOR DETERMINISTIC STATIONARY SEQUENCES

NIKOLAY M. BABAYAN\textsuperscript{a} MAMIKON S. GINOVYAN\textsuperscript{b} AND MURAD S. TAQQU\textsuperscript{b}\textsuperscript{*}

\textsuperscript{a}Department of Mathematical Cybernetics, Russian-Armenian University, Yerevan, Armenia

\textsuperscript{b}Department of Mathematics and Statistics, Boston University, Boston, MA, USA

One of the main problems in prediction theory of discrete-time second-order stationary processes \(X(t)\) is to describe the asymptotic behavior of the best linear mean squared prediction error in predicting \(X(0)\) given \(X(t), -n \leq t \leq -1\), as \(n\) goes to infinity. This behavior depends on the regularity (deterministic or non-deterministic) of the process \(X(t)\). In his seminal article ‘Some purely deterministic processes’ (J. of Math. and Mech., 6(6), 801–10, 1957), Rosenblatt has described the asymptotic behavior of the prediction error for deterministic processes in the following two cases: (i) the spectral density \(f\) of \(X(t)\) is continuous and vanishes on an interval, (ii) the spectral density \(f\) has a very high order contact with zero. He showed that in the case (i) the prediction error behaves exponentially, while in the case (ii), it behaves like a power as \(n \to \infty\). In this article, using an approach different from the one applied in Rosenblatt’s article, we describe extensions of Rosenblatt’s results to broader classes of spectral densities. Examples illustrate the obtained results.

Received 14 April 2020; Revised 29 October 2020; Accepted 10 November 2020

Keywords: Asymptotic behavior of the prediction error; deterministic stationary process; Rakhmanov’s theorem; Rosenblatt’s theorems; transfinite diameter; Verblunsky’s coefficients.

MOS subject classification: 60G10, 60G25, 62M15, 62M20.

1. INTRODUCTION

1.1. The prediction problem

Let \(X(t), t \in \mathbb{Z} := \{0, \pm 1, \ldots\}\), be a centered discrete-time second-order stationary process. The process is assumed to have an absolutely continuous spectrum with spectral density function \(f(\lambda), \lambda \in [-\pi, \pi]\). The ‘finite’ linear prediction problem is as follows.

Suppose we observe a finite realization of the process \(X(t): \{X(t), -n \leq t \leq -1\}, n \in \mathbb{N} := \{1, 2, \ldots\}\). We want to make an one-step ahead prediction, that is, to predict the unobserved random variable \(X(0)\), using the linear predictor \(Y = \sum_{k=1}^{n} c_k X(-k)\).

The coefficients \(c_k, k = 1, 2, \ldots, n\), are chosen so as to minimize the mean-squared error: \(\mathbb{E}[X(0) - Y]^2\), where \(\mathbb{E}[\cdot]\) stands for the expectation operator. If such minimizing constants \(\hat{c}_k := \hat{c}_{k,n}\) can be found, then the random variable \(\hat{X}_n(0) := \sum_{k=1}^{n} \hat{c}_k X(-k)\) is called the best linear one-step ahead predictor of \(X(0)\) based on the observed finite past: \(X(-n), \ldots, X(-1)\). The minimum mean-squared error:

\[\sigma^2_n(f) := \mathbb{E}[X(0) - \hat{X}_n(0)]^2 \geq 0\]

is called the best linear one-step ahead prediction error of \(X(t)\) based on the past of length \(n\).

\textsuperscript{*}Corresponding to: Murad S. Taqqu

This article is dedicated to the memory of Professor Murray Rosenblatt.

© 2020 John Wiley & Sons Ltd
One of the main problems in prediction theory of second-order stationary processes, called the ‘direct’ prediction problem is to describe the asymptotic behavior of the prediction error $\sigma^2_n(f)$ as $n \to \infty$. This behavior depends on the regularity nature (deterministic or non-deterministic) of the observed process $X(t)$.

Observe that $\sigma^2_{n+1}(f) \leq \sigma^2_n(f)$, $n \in \mathbb{N}$, and hence the limit of $\sigma^2_n(f)$ as $n \to \infty$ exists. Denote by $\sigma^2(f) := \sigma^2_\infty(f)$ the prediction error of $X(0)$ by the entire infinite past: $X(t)$, $t \leq -1$.

From the prediction point of view it is natural to distinguish the class of processes for which we have error-free prediction by the entire infinite past, that is, $\sigma^2(f) = 0$. Such processes are called deterministic or singular. Processes for which $\sigma^2(f) > 0$ are called non-deterministic.

Note. The term ‘deterministic’ here is not used in the usual sense of absence of randomness. Instead determinism of a process means that there is an extremely strong dependence between the successive random variables forming the process, yielding error-free prediction when using the entire infinite past (for more about this term see Bingham, 2012, and Grenander and Szegő, 1958, p. 176.)

Define the relative prediction error $\delta_n(f) := \sigma^2_n(f) - \sigma^2(f)$, and observe that $\delta_n(f)$ is non-negative and tends to zero as $n \to \infty$. But what about the speed of convergence of $\delta_n(f)$ to zero as $n \to \infty$? The article deals with this question. Specifically, the prediction problem we are interested in is to describe the rate of decrease of $\delta_n(f)$ to zero as $n \to \infty$, depending on the regularity nature (deterministic or non-deterministic) of the observed process $X(t)$.

The prediction problem stated above goes back to classical works of Kolmogorov (1941a, 1941b), Szegő (1915, 1967) and Wiener (1949). It was then considered by many authors for different classes of non-deterministic processes (see, e.g., Baxter, 1962, Devinatz, 1964, Golinskii, 1974, Golinskii and Ibragimov, 1971, Grenander and Rosenblatt, 1954, 1957, Grenander and Szegő, 1958, Nelson and Szegő, 1960, Ibragimov, 1964, Ibragimov and Solec, 1968, Inoue, 2002, Pourahmadi, 2001, Rozanov, 1967, and reference therein). More references can be found in the survey articles Bingham (2012) and Ginovian (1999).

We focus in this article on deterministic processes, that is, when $\sigma^2(f) = 0$. This case is not only of theoretical interest, but is also important from the point of view of applications. For example, as pointed out by Rosenblatt (1957), situations of this type arise in Neumann’s theoretical model of storm-generated ocean waves. Such models are also of interest in meteorology (see Fortus, 1990).

Only few works are devoted to the study of the speed of convergence of $\delta_n(f) = \sigma^2_n(f)$ to zero as $n \to \infty$, that is, the asymptotic behavior of the prediction error for deterministic processes. One needs to go back to the classical work of Rosenblatt (1957). Using the technique of orthogonal polynomials on the unit circle (OPUC), Rosenblatt investigated the asymptotic behavior of the prediction error $\sigma^2_n(f)$ for deterministic processes in the following two cases:

(a) the spectral density $f(\lambda)$ is continuous and vanishes on an entire arc of the unit circle,
(b) the spectral density $f(\lambda)$ is positive away from zero and has a very high order of contact with zero at $\lambda = 0$.

Later the problem (a) was studied by Babayan (1984, 1985), (see also Davison, 1965 and Fortus, 1990), where some generalizations and extensions of Rosenblatt’s results have been obtained.

Some notation. Throughout the article we will use the following notation and conventions.

The standard symbols $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, and $\mathbb{C}$ denote the sets of natural, integer, real, and complex numbers, respectively. Also, we denote $\mathbb{Z}_+ := \{0, 1, 2, \ldots\}$, $\Lambda := [-\pi, \pi]$, $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$, $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$. The letters $C$, $c$, $M$, and $m$ with or without indices are used to denote positive constants, the values of which can vary from line to line. For a set $E$ we denote the closure of $E$. By $L^p(\mu) := L^p(\mathbb{T}, \mu)$ $(p \geq 1)$ we denote the weighted Lebesgue space with respect to the measure $\mu$, and by $(\cdot, \cdot)_{p, \mu}$ and $\| \cdot \|_{p, \mu}$ we denote the inner product and the norm in $L^p(\mu)$, respectively. In the special case where $\mu$ is the Lebesgue measure ($\mu_1$), we will use the notation $L^p$, $(\cdot, \cdot)_{p}$, and $\| \cdot \|_{p}$, respectively. By $G(h)$ we denote the geometric mean of a function $h$. For two functions $f(\lambda) \geq 0$ and $g(\lambda) \geq 0$ we will write $f(\lambda) \sim g(\lambda)$ as $\lambda \to \lambda_0$ if $\lim_{\lambda \to \lambda_0} f(\lambda)/g(\lambda) = 1$, and $f(\lambda) \approx g(\lambda)$ as $\lambda \to \lambda_0$ if $\lim_{\lambda \to \lambda_0} f(\lambda)/g(\lambda) = c > 0$. We will use similar notation for sequences: for two sequences $\{a_n \geq 0, n \in \mathbb{N}\}$ and $\{b_n > 0, n \in \mathbb{N}\}$, we will write $a_n \sim b_n$ if $\lim_{n \to \infty} a_n/b_n = 1$, $a_n \approx b_n$ if $\lim_{n \to \infty} a_n/b_n = c > 0$, and $a_n = O(b_n)$ if $a_n/b_n$ is bounded, and $a_n = o(b_n)$ if $a_n/b_n \to 0$ as $n \to \infty$. We will use the abbreviations: OPUC for ‘orthogonal polynomials on the unit circle’, PACF for ‘partial autocorrelation function’, and ‘a.e.’ for ‘almost everywhere’ (with respect to
the Lebesgue measure). We will assume that all the relevant objects are defined in terms of Lebesgue integrals, and so are invariant under change of the integrand on a null set.

We start by describing Rosenblatt’s results concerning the asymptotic behavior of the prediction error $\sigma^2_n(f)$, obtained in Rosenblatt (1957) for the above stated cases (a) and (b).

1.2. Rosenblatt’s results about speed of convergence

For the case (a) above, Rosenblatt proved in Rosenblatt (1957) that the prediction error $\sigma^2_n(f)$ decreases to zero exponentially as $n \to \infty$. More precisely, Rosenblatt proved the following theorem.

**Theorem A.** (Rosenblatt’s first theorem). Let the spectral density $f$ of a discrete-time stationary process $X(t)$ be positive and continuous on the arc $(\pi/2 - \alpha, \pi/2 + \alpha)$, $0 < \alpha < \pi$, and zero elsewhere. Then the prediction error $\sigma^2_n(f)$ approaches zero exponentially as $n \to \infty$. More precisely, the following asymptotic relation holds:

$$\sigma^2_n(f) \approx (\sin(\alpha/2))^{2n+1} \quad \text{as} \quad n \to \infty.$$  

(1.1)

Thus, when the spectral density $f$ is continuous and vanishes on an entire interval, then the prediction error $\sigma^2_n(f)$ approaches zero with a sufficiently high speed, namely as a geometric progression with common ratio $\sin^2(\alpha/2) < 1$. Notice that (1.1) implies that

$$\lim_{n \to \infty} \sqrt[n]{\sigma^2_n(f)} = \sin(\alpha/2).$$  

(1.2)

Concerning the case (b) above, Rosenblatt proved in Rosenblatt (1957) that the prediction error $\sigma^2_n(f)$ decreases to zero like a power, that is, $\sigma^2_n(f) \approx n^{-\alpha}$ ($a > 0$) as $n \to \infty$. More precisely, the deterministic process $X(t)$ considered in Rosenblatt (1957) has the spectral density

$$f_a(\lambda) := \frac{\alpha^{(2\lambda-\pi)\phi(\lambda)}}{\cos \lambda (\pi \phi(\lambda))}, \quad f_a(-\lambda) = f_a(\lambda), \quad 0 \leq \lambda \leq \pi,$$  

(1.3)

where $\phi(\lambda) = (a/2) \cot \lambda$ and $a$ is a positive parameter. We have

$$f_a(\lambda) \sim 2 \exp[-a\pi |\lambda|] \sin(\lambda) \quad \text{as} \quad \lambda \to 0.$$  

(1.4)

Using the technique of OPUC and Szegő’s results, Rosenblatt (1957) proved the following theorem.

**Theorem B.** (Rosenblatt’s second theorem). Suppose that the process $X(t)$ has spectral density $f_a$ given by (1.3).

Then the following asymptotic relation for the prediction error $\sigma^2_n(f_a)$ holds:

$$\sigma^2_n(f_a) \sim \frac{\Gamma^2((a + 1)/2)}{\pi 2^{2-a}} n^{-a} \quad \text{as} \quad n \to \infty.$$  

(1.5)

In this article, using an approach different from the one applied in Rosenblatt (1957), we extend Theorems A and B to broader classes of spectral densities.

Concerning Theorem A, we describe an extension of the asymptotic relation (1.2) to the case of several arcs, without having to stipulate continuity of the spectral density $f$ (Theorem 3.1).

As for the extension of Theorem B, we first prove that if the spectral density $f$ is such that the sequence $\sigma_n(f)$ is weakly varying (a term defined in Section 4.1) and if, in addition, $g$ is the spectral density of a non-deterministic process satisfying some conditions, then the sequences $\sigma_n(fg)$ and $\sigma_n(f)$ have the same asymptotic behavior, up to some positive multiplicative factor (Theorem 4.1). This allows us to derive the asymptotic behavior of $\sigma_n(fg)$ from that of $\sigma_n(f)$.
Using this result, we obtain the following extension of Theorem B: if the spectral density $f$ has the form $f = f_0 g$, where $f_0$ is as in (1.3), then $\sigma'_n(f) \approx n^{-\alpha}$ as $n \to \infty$ (Theorem 4.2).

The remainder of the article is organized as follows. In Section 2 we recall some basic notions, present formulas for the finite prediction error $\sigma^2_n(f)$, and state some preliminary results. Section 3 is devoted to the extension of the Rosenblatt’s first theorem (Theorem A). In Section 4 we extend Rosenblatt’s second theorem (Theorem B).

2. PRELIMINARIES: FORMULAS FOR THE PREDICTION ERROR

We recall some basic notions, present formulas for the finite prediction error $\sigma^2_n(f)$, and state some preliminary results, which will be used in the sequel.

2.1. Preliminaries

Let $X(t)$ be a centered discrete-time stationary process defined on a probability space $(\Omega, F, \mathbb{P})$ with covariance function $r(t)$, $t \in \mathbb{Z}$. By the Herglotz theorem (see, e.g., Brockwell and Davis, 1991, p. 117–8), there is a finite measure $\mu$ on $\Lambda$ such that the covariance function $r(t)$ admits the following spectral representation:

$$r(t) = \int_{-\pi}^{\pi} e^{-it\lambda} d\mu(\lambda), \quad t \in \mathbb{Z}. \quad (2.1)$$

The measure $\mu$ in (2.1) is called the spectral measure of the process $X(t)$. If $\mu$ is absolutely continuous (with respect to the Lebesgue measure), then the function $f(\lambda) : = d\mu(\lambda)/d\lambda$ is called the spectral density of $X(t)$. In this case, the spectral measure we will denote by $\mu_f$. We assume that $X(t)$ is a non-degenerate process, that is, $\text{Var}[X(0)] = \mathbb{E}[X(0)^2] = r(0) > 0$ and, without loss of generality, we may take $r(0) = 1$. Also, to avoid the trivial cases, we assume that the spectral measure $\mu$ is non-trivial, that is, $\mu$ has infinite support.

Remark 2.1. The parametrization of the unit circle $\mathbb{T}$ by the formula $z = e^{it}$ establishes a bijection between $\mathbb{T}$ and the interval $[-\pi, \pi]$. By means of this bijection the measure $\mu$ on $\Lambda$ generates the corresponding measure on the unit circle $\mathbb{T}$, which we also denote by $\mu$. Thus, depending on the context, the measure $\mu$ will be supported either on $\Lambda$ or on $\mathbb{T}$. We use the standard Lebesgue decomposition of the measure $\mu$:

$$d\mu(\lambda) = d\mu_a(\lambda) + d\mu_s(\lambda) = f(\lambda) d\lambda + d\mu_s(\lambda), \quad (2.2)$$

where $\mu_a$ is the absolutely continuous part of $\mu$ (with respect to the Lebesgue measure) and $\mu_s$ is the singular part of $\mu$, which is the sum of the discrete and continuous singular components of $\mu$.

Let $L^2(\Omega, P)$ denote the $L^2$-space of centered random variables $\xi$ defined on $\Omega$. Denote by $H(X)$ the time-domain of the process $X(t)$, that is, the closed linear subspace of $L^2(\Omega, P)$ spanned by the random variables $X(t, \omega), \ t \in \mathbb{Z}$, and by $L^2(\mu) = L^2(\mathbb{T}, \mu)$ the frequency-domain of $X(t)$, that is, the space of complex-valued square-integrable functions with respect to the measure $\mu$ functions on $\mathbb{T}$.

Kolmogorov’s isometric isomorphism theorem. For any stationary process $X(t)$ with spectral measure $\mu$ there exists a unique isometric isomorphism $V$ between the time-domain $H(X)$ and the frequency-domain $L^2(\mu)$, such that $V[X(t)] = e^{it\lambda}$ for any $t \in \mathbb{Z}$.

The next result describes the asymptotic behavior of the prediction error $\sigma^2_n(\mu)$ for a stationary process $X(t)$ with spectral measure $\mu$ of the form (2.2) and gives a spectral characterization of deterministic and non-deterministic processes (see, e.g., Grenander and Szegö, 1958, p. 44).

Kolmogorov–Szegö Theorem. Let $X(t)$ be a non-degenerate stationary process with spectral measure $\mu$ of the form (2.2). The following relations hold.

$$\lim_{n \to \infty} \sigma^2_n(\mu) = \lim_{n \to \infty} \sigma^2_n(f) = \sigma^2(f) = 2\pi G(f), \quad (2.3)$$
where \( G(f) \) is the geometric mean of \( f \), namely

\[
G(f) := \begin{cases} 
\exp \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln f(\lambda) \, d\lambda \right\} & \text{if } \ln f \in L^1(\Lambda) \\
0, & \text{otherwise.}
\end{cases}
\]  

(2.4)

It is remarkable that (2.3) is independent of the singular part \( \mu \). The condition \( \ln f \in L^1(\Lambda) \) in (2.4) is equivalent to the Szegő condition:

\[
\int_{-\pi}^{\pi} \ln f(\lambda) \, d\lambda > -\infty
\]

(2.5)

(this equivalence follows because \( \ln f(\lambda) \leq f(\lambda) \)) The Szegő condition (2.5) is also called the non-determinism condition.

In this article, following Rosenblatt (1957), we consider the class of deterministic processes with absolutely continuous spectra.

We will say that the spectral density \( f(\lambda) \) has a very high order of contact with zero at a point \( \lambda_0 \) if \( f(\lambda) \) is positive everywhere except for the point \( \lambda_0 \), due to which the Szegő condition (2.5) is violated. Observe that the Szegő condition is related to the character of the singularities (zeroes) of the spectral density \( f \), and does not depend on the differential properties of \( f \). For example, for any \( a > 0 \), the function \( f_{a}(\lambda) = \exp(-|\lambda|^{-a}) \) is infinitely differentiable. In addition, for \( a < 1 \) Szegő’s condition is satisfied, and hence the corresponding process \( X(t) \) is non-deterministic, while for \( a \geq 1 \) Szegő’s condition is violated, and \( X(t) \) is deterministic. Thus, according to the above definition, for \( a \geq 1 \) this function has a very high order of contact with zero at the point \( \lambda = 0 \).

### 2.2. Formulas for the prediction error

Suppose we have observed the values \( X(-n), \ldots, X(-1) \) of a centered, real-valued stationary process \( X(t) \) with spectral measure \( \mu \) of the form (2.2). The one-step ahead linear prediction problem in predicting a random variable \( X(0) \) based on the observed values \( X(-n), \ldots, X(-1) \) involves finding constants \( \hat{c}_k := \hat{c}_{k,n}, k = 1, 2, \ldots, n \), that minimize the one-step ahead prediction error:

\[
\sigma^2_n(\mu) := \min_{\{c_k\}} \left\{ \mathbb{E} \left[ X(0) - \sum_{k=1}^{n} c_k X(-k) \right]^2 \right\} = \mathbb{E} \left[ X(0) - \sum_{k=1}^{n} \hat{c}_k X(-k) \right]^2.
\]  

(2.6)

Using Kolmogorov’s isometric isomorphism \( V : X(t) \leftrightarrow e^{it\lambda} \), in view of (2.6), for the prediction error \( \sigma^2_n(\mu) \) we can write

\[
\sigma^2_n(\mu) = \min_{\{c_k\}} \left\{ \mathbb{E} \left[ 1 - \sum_{k=1}^{n} c_k e^{-ik\lambda} \right]^2, \right\} = \min_{\{q_{n,\mu}\}} \left\{ \mathbb{E} \left[ q_{n,\mu} \right]^2 \right\},
\]  

(2.7)

where \( \| \cdot \|_{2,\mu} \) is the norm in \( L^2(\mathbb{T}, \mu) \), and

\[
Q_n := \{ q_n : q_n(z) = z^n + c_1 z^{n-1} + \ldots + c_n \}
\]  

(2.8)

is the class of monic polynomials (i.e., with \( c_0 = 1 \)) of degree \( n \). Thus, the problem of finding \( \sigma^2_n(\mu) \) becomes to the problem of finding the solution of the minimum problem (2.7)–(2.8).

The polynomial \( p_n(z) := p_n(z, \mu) \) which solves the minimum problem (2.7)–(2.8) is called the optimal polynomial for \( \mu \) in the class \( Q_n \). This minimum problem was solved by G. Szegő by showing that the optimal polynomial
ASYMPTOTIC BEHAVIOR OF THE PREDICTION ERROR

627

$p_n(z)$ exists, is unique and can be expressed in terms of orthogonal polynomials on the unit circle with respect to the measure $\mu$ (see Proposition 2.1).

To state Szegő’s solution of the minimum problem (2.7)–(2.8), we first recall some facts from the theory of orthogonal polynomials on the unit circle (OPUC).

The system of orthogonal polynomials on the unit circle associated with the measure $\mu$: $\varphi_n(z) := \varphi_n(z; \mu) (z = e^{i\lambda}, n \in \mathbb{Z}_+)$ is uniquely determined by the following two conditions:

(i) $\varphi_n(z) = \kappa_n z^n + \ldots + \lambda_n$ is a polynomial of degree $n$, in which the coefficient $\kappa_n$ is positive;

(ii) $(\varphi_k, \varphi_j)_{\mu} = \delta_{kj}$ for arbitrary $k, j \in \mathbb{Z}_+$, where $\delta_{kj}$ is the Kronecker delta.

Define the monic ($p_n(z)$) and the reciprocal ($p_n^*(z)$) polynomials (see, e.g., Simon, 2005, p. 2):

\[ p_n(z) := p_n(z, \mu) = \kappa_n^{-1} \varphi_n(z) = z^n + \ldots + \lambda_n \kappa_n^{-1}, \quad (2.9) \]

\[ p_n^*(z) := p_n^*(z, \mu) = z^n p_n(1/z) = \tilde{\kappa}_n z^n + \ldots + 1. \quad (2.10) \]

We have

\[ ||p_n||_{2,\mu} = ||p_n^*||_{2,\mu} = \kappa_n^{-2}. \quad (2.11) \]

The polynomials $p_n(z)$ and $p_n^*(z)$ satisfy Szegő’s recursion relation (see Simon, 2005, p. 56):

\[ p_{n+1}(z) = zp_n(z) - \nu_{n+1} p_n^*(z), \quad n \in \mathbb{Z}_+ \quad (2.12) \]

where

\[ \nu_{n+1} = -\overline{p_{n+1}(0)} = \tilde{\kappa}_n^{-1} \kappa_{n+1}^{-1}, \quad |\nu_{n+1}| < 1. \quad (2.13) \]

In view of (2.12) we have (see Simon, 2005, p. 56)

\[ ||p_n||_{2,\mu} = (1 - |\nu_n|^2)||p_{n-1}||_{2,\mu} = \prod_{j=1}^{n} (1 - |\nu_j|^2), \quad n \in \mathbb{N}. \quad (2.14) \]

From (2.11) and (2.14) we obtain

\[ \kappa_n^2 \kappa_{n+1}^{-2} = 1 - |\nu_{n+1}|^2. \quad (2.15) \]

The parameters $\nu_n := \nu_n(\mu) (n \in \mathbb{N})$, which play an important role in the theory of OPUC, are called Verblunsky’s coefficients (also known as the Szegő, Schur, and canonical moments; see Simon, 2005, Sect. 1.1, and Dette and Studden, 1997, Sect. 9.4).

Note. The term ‘Verblunsky coefficient’ is from Simon, 2005. Observe that we write $\nu_{n+1}$ for Simon’s $a_n$, and so one has $n \in \mathbb{N}$ for Simon’s $n \in \mathbb{Z}_+$. Our notational convention is already established in the time-series literature and is more convenient in our context of the PACF (defined below), where $n \in \mathbb{N}$ (see Bingham, 2012, Brockwell and Davis, 1991, §5.2, Inoue, 2002, Pourahmadi, 2001, §7.3).

The following result shows that Verblunsky’s coefficients provide a convenient way for the parametrization of probability measures on the unit circle $T$ (see, e.g., Simon, 2005, p. 2).

Verblunsky’s theorem. Let $\mathbb{D}^\infty$ be the set of complex sequences $\nu := (\nu_n, n \in \mathbb{N})$ with $\nu_n \in \mathbb{D}$. The map $\mathcal{S} : \mu \mapsto \nu$ is a bijection between the set of non-trivial probability measures on $T$ and $\mathbb{D}^\infty$. 

\[ J. \text{Time Ser. Anal.} \ 42: \ 622–652 (2021) \  © 2020 John Wiley & Sons Ltd \  wileyonlinelibrary.com/journal/jtsa \  DOI: 10.1111/jtsa.12572 \]
This result was established by Verblunsky in 1935, in connection with OPUC. It was re-discovered by Ramsey (1974) in 1974, in connection with parametrization of time-series models.

For a stationary process \( X(t) \) with a non-trivial spectral measure \( \mu \) the partial autocorrelation function (PACF) of \( X(t) \), denoted by \( \pi_n = \pi_n(\mu) \) \((n \in \mathbb{N})\), is defined to be the correlation coefficient between the forward and backward residuals in the linear prediction of the variables \( X(n) \) and \( X(0) \) on the basis of the intermediate observations \( X(1), \ldots, X(n-1) \), that is,

\[
\pi_n := \text{corr}(X(n) - \hat{X}(n), X(0) - \hat{X}(0)).
\]

It turns out that the Verblunsky coefficients \( v_n \) and the PACF \( \pi_n \) coincide, that is, \( v_n = \pi_n \) for all \( n \in \mathbb{N} \) (see Dette and Studden, 1997, Sect. 9.6). Thus, the Verblunsky sequence \( \nu := (v_n, n \in \mathbb{N}) \) provides a link between OPUC and time-series analysis, and, in view of the equality \( v_n = \pi_n \), the Verblunsky bijection gives a simple and unconstrained parametrization of stationary processes, in contrast to using the covariance function, which has to be positive-definite.

The next result by Szegő solves the minimum problem (2.7)–(2.8) (see, e.g., Grenander and Szegő, 1958, p. 38).

**Proposition 2.1.** The unique solution of the minimum problem (2.7)–(2.8) is the monic polynomial \( p_n(\mu) \) \( := p_n(z, \mu) \) given by formula (2.9), and the minimum in (2.7) is equal to \( ||p_n||_{2, \mu} = \kappa_n^{-2} \).

Thus, for the prediction error \( \sigma_n^2(\mu) \) we have the following formula:

\[
\sigma_n^2(\mu) = \min_{|q_n| \in \mathcal{Q}_n} ||q_n||^2_{2, \mu} = ||p_n(\mu)||^2_{2, \mu} = \kappa_n^{-2}.
\]  

(2.16)

**Remark 2.2.** Define

\[
Q^*_n := \{ q_n : q_n(z) = c_0 z^n + c_1 z^{n-1} + \cdots + c_n, \ c_n = 1 \},
\]  

(2.17)

and observe that the classes of polynomials \( \mathcal{Q}_n \) and \( Q^*_n \) defined in (2.8) and (2.17), respectively, differ by normalization: in (2.17) we have \( c_n = 1 \), while in (2.8) we have \( c_0 = 1 \). Also, the optimal polynomial for \( \mu \) in the class \( Q^*_n \) is the reciprocal polynomial \( p_n^*(z) \) (see (2.10)). Taking into account (2.11), we have the following formula for the prediction error \( \sigma_n^2(\mu) \) in terms of the optimal polynomial \( p_n^*(z) \):

\[
\sigma_n^2(\mu) = \min_{|q_n| \in \mathcal{Q}^*_n} ||q_n||^2_{2, \mu} = ||p_n^*(\mu)||^2_{2, \mu}.
\]  

(2.18)

In view of (2.16), the formulas (2.14) and (2.15) can be written as follows

\[
\sigma_n^2(\mu) = \prod_{j=1}^n (1 - |v_j|^2) \quad \text{and} \quad \frac{\sigma_{n+1}(\mu)}{\sigma_n^2(\mu)} = 1 - |v_n|^2.
\]  

(2.19)

**Remark 2.3.** From the second formula in (2.19), it follows that the convergence of the sequences \( |v_n| \) and \( \sigma_{n+1}(\mu)/\sigma_n(\mu) \) are equivalent, and, the greater the limiting value of \( |v_n| \), the faster the rate of decrease of \( \sigma_n(\mu) \). In particular, under the conditions of Theorem A, we have \( \lim_{n \to \infty} \sigma_{n+1}^2(f) / \sigma_n^2(f) = \sin^2(\alpha/2) \) and \( \lim_{n \to \infty} |v_n| = \cos(\alpha/2) \). Similarly, under the conditions of Theorem B, we have \( \lim_{n \to \infty} \sigma_{n+1}(f_n) / \sigma_n(f_n) = 1 \) and \( \lim_{n \to \infty} v_n(f_n) = 0 \).

For a general measure \( \mu \) of the form (2.2) the asymptotic relation

\[
\lim_{n \to \infty} v_n(\mu) = 0
\]  

(2.20)

is of special interest in the theory of OPUC. In this respect the following question arises naturally: what is the ‘minimal’ sufficient condition on the measure \( \mu \) ensuring the relation (2.20)? The next result of Rakhmanov (1987)
shows that for (2.20) (or equivalently, for \( \lim_{n \to \infty} \sigma_{\varphi_n}(\mu) / \sigma_n(\mu) = 1 \)) it is enough only to have a.e. positiveness on \( \mathbb{T} \) of the spectral density \( f \) (see also Simon, 2005, p. 5).

Rakhmanov’s theorem. Let the measure \( \mu \) have the form (2.2) with \( f > 0 \) a.e. on \( \mathbb{T} \). Then the asymptotic relation (2.20) is satisfied.

Bello and López (1998) proved the following extension of Rakhmanov’s theorem: Let \( \Gamma_\delta \) be a closed arc of the unit circle of length \( 2\delta \) (0 < \( \delta \) \( \leq \) \( \pi \)), and let \( \mu \) and \( (v_n) \) be as in Rakhmanov’s theorem. Assume that the measure \( \mu \) is supported on the arc \( \Gamma_\delta \) with \( f > 0 \) a.e. on \( \Gamma_\delta \). Then \( \lim_{n \to \infty} |v_n| = \cos(\delta/2) \). The case \( \delta = \pi \) corresponds to Rakhmanov’s theorem.

In the next proposition we list some properties of the prediction error \( \sigma_n^2(f) \), which can easily be verified (see also Grenander and Rosenblatt, 1957, p. 66).

**Proposition 2.2.** The prediction error \( \sigma_n^2(f) \) possesses the following properties.

a) The sequence \( \{ \sigma_n^2(f), n \in \mathbb{N} \} \) is decreasing in \( n \) : \( \sigma_{n+1}^2(f) < \sigma_n^2(f) \).

b) \( \sigma_n^2(f) \) is a non-decreasing functional of \( f \): \( \sigma_n^2(f_1) \leq \sigma_n^2(f_2) \) when \( f_1(\lambda) \leq f_2(\lambda) \).

c) If \( f(\lambda) = g(\lambda) \) almost everywhere on \( [-\pi, \pi] \), then \( \sigma_n(f) = \sigma_n(g) \).

3. AN EXTENSION OF ROSENBLATT’S FIRST THEOREM

We extend Rosenblatt’s first theorem (Theorem A) to a broader class of deterministic processes, possessing spectral densities. More precisely, we extend the asymptotic relation (1.2) to the case of several arcs, without having to stipulate continuity of the spectral density \( f \). Besides, we obtain necessary as well as sufficient conditions for the exponential decay of the prediction error \( \sigma_n(f) \). Also, we calculate the transfinite diameter of some subsets of the unit circle, and thus, obtain explicit asymptotic relations for \( \sigma_n(f) \) similar to Rosenblatt’s relation (1.2).

To state the corresponding results we first introduce some metric characteristics of compact (bounded closed) sets in the plane, such as, the transfinite diameter, the Chebyshev constant and the capacity, and discuss some properties of these characteristics.

3.1. Some metric characteristics of bounded closed sets in the plane

One of the fundamental result of geometric complex analysis is the classical theorem by Fekete and Szegö, stating that for any compact set \( F \) in the complex plane \( \mathbb{C} \) the transfinite diameter, the Chebyshev constant and the capacity of \( F \) coincide, although they are defined from very different points of view. Namely, the transfinite diameter of the set \( F \) characterizes the asymptotic size of \( F \), the Chebyshev constant of \( F \) characterizes the minimal uniform deviation of a monic polynomial on \( F \), and the capacity of \( F \) describes the asymptotic behavior of the Green function at infinity. For the definitions and results stated in this subsection we refer the reader to the following references: Fekete (1930), Goluzin (1969), Chapter 7, Saff (2010), Szegö (1967), Chapter 16, and Tsuji (1975), Chapter III.

**Transfinite diameter.** Let \( F \) be a compact (bounded closed) set in the complex plane \( \mathbb{C} \). Given a natural number \( n \geq 2 \) and points \( z_1, \ldots, z_n \in F \), we define

\[
d_n(F) := \max_{z_1, \ldots, z_n \in F} \left[ \prod_{1 \leq j < k \leq n} |z_j - z_k| \right]^{2/(n(n-1))},
\]

which is the maximum of products of distances between the \( \binom{n}{2} = n(n - 1)/2 \) pairs of points \( z_k \), \( k = 1, \ldots, n \), as the points \( z_k \) range over the set \( F \). Note that \( d_1(F) \) is the diameter of \( F \). The quantity \( d_n(F) \) is called the \( n \)th transfinite diameter of the set \( F \). It can be shown (see, e.g., Goluzin, 1969, p. 294) that \( d_n(F) \) decreases and does not exceed the diameter \( d_n(F) \) of \( F \), implying that \( d_n(F) \) has a finite limit as \( n \to \infty \). This limit, denoted by \( d_\infty(F) \),
is called the transfinite diameter of \( F \). Thus, we have

\[
d_\infty(F) := \lim_{n \to \infty} d_n(F). \tag{3.2}
\]

**Chebyshev constant.** For a bounded closed set \( F \) in the complex plane \( \mathbb{C} \), we put \( m_n(F) := \inf_{q \in \mathcal{P}} \max_{z \in F} |q_n(z)| \), where the infimum is taken over all monic polynomials \( q_n(z) \) from the class \( \mathcal{Q}_n \), where \( \mathcal{Q}_n \) is as in (2.8). Then there exists a unique monic polynomial \( T_n(z, F) \) from the class \( \mathcal{Q}_n \), called the Chebyshev polynomial of \( F \) of order \( n \), such that

\[
m_n(F) = \max_{z \in F} |T_n(z, F)|. \tag{3.3}
\]

Fekete (1930) proved that \( \lim_{n \to \infty} (m_n(F))^{1/n} \) exists. This limit, denoted by \( \tau(F) \), is called the Chebyshev constant for the set \( F \). Thus,

\[
\tau(F) := \lim_{n \to \infty} (m_n(F))^{1/n}. \tag{3.4}
\]

**Capacity.** Let \( F \) be a closed bounded set in the complex plane \( \mathbb{C} \), and let \( D_F \) denote the complementary domain to \( F \), containing \( \infty \) as an interior point. If the boundary \( \Gamma := \partial D_F \) of the domain \( D_F \) consists of a finite number of rectifiable Jordan curves, then for the domain \( D_F \) can be constructed a Green function \( G_{D_F}(z, \infty) \) with a pole at infinity. This function is harmonic everywhere in \( D_F \), except at the point \( z = \infty \), is continuous including the boundary \( \Gamma \) and vanishes on \( \Gamma \). It is known that in a vicinity of the point \( z = \infty \) the function \( G_{D_F}(z, \infty) \) admits the representation (see, e.g., Goluzin, 1969, p. 309–10):

\[
G_F(z, \infty) = \ln |z| + \gamma + O(z^{-1}) \quad \text{as} \quad z \to \infty. \tag{3.5}
\]

The number \( \gamma \) in (3.5) is called the Robin’s constant of the domain \( D_F \), and the number

\[
C(F) := e^{-\gamma} \tag{3.6}
\]

is called the capacity (or the logarithmic capacity) of the set \( F \).

Now we are in position to state the above mentioned fundamental result of geometric complex analysis, due to M. Fekete and G. Szegő (see, e.g., Goluzin, 1969, p. 197 and Tsuji, 1975, p. 73).

**Proposition 3.1.** (Fekete–Szegő theorem). For any compact set \( F \subset \mathbb{C} \), the transfinite diameter \( d_\infty(F) \) defined by (3.2), the Chebyshev constant \( \tau(F) \) defined by (3.4), and the capacity \( C(F) \) defined by (3.6) coincide, that is,

\[
d_\infty(F) = C(F) = \tau(F). \tag{3.7}
\]

It what follows, we will use the term ‘transfinite diameter’ and the notation \( \tau(F) \) for (3.7). In only very few cases can the transfinite diameter (and hence, the capacity and the Chebyshev constant) be exactly calculated.

In the next proposition we list a number of properties of the transfinite diameter (and hence, of the capacity and the Chebyshev constant), which will be used later.

**Proposition 3.2.** The transfinite diameter possesses the following properties.

(a) The transfinite diameter is monotone, that is, for any closed sets \( F_1 \) and \( F_2 \) with \( F_1 \subset F_2 \), we have \( \tau(F_1) \leq \tau(F_2) \) (see, e.g., Saff, 2010, p. 169, Tsuji, 1975, p. 56).

(b) If a set \( F_1 \) is obtained from a compact set \( F \subset \mathbb{C} \) by a linear transformation, that is, \( F_1 := aF + b = \{az + b; z \in F \} \), then \( \tau(F_1) = |a| \tau(F) \). In particular, the transfinite diameter \( \tau(F) \) is invariant with respect to parallel translation and rotation of \( F \) (see, e.g., Goluzin, 1969, p. 298, Saff, 2010, p. 169, Tsuji, 1975, p. 56).
(c) (Fekete theorem). Let $F$ be a bounded closed set in the complex $w$-plane, and let $p_n(z) = z^n + c_1 z^{n-1} + \ldots + c_n$ be an arbitrary monic polynomial of degree $n$. Let $F_n^*$ be the preimage of $F$ under the mapping $w = p_n(z)$, that is, $F_n^*$ is the set of all points $z$ such that $w := p_n(z) \in F$. Then $\tau(F_n^*) = [\tau(F)]^{1/n}$ (see, e.g., Goluzin, 1969, p. 299, Saff, 2010, p. 186).

(d) The transfinite diameter of an arbitrary circle of radius $R$ is equal to its radius $R$. In particular, the transfinite diameter of the unit circle $\mathbb{T}$ is equal to 1 (Tsuji, 1975, p. 84).

(e) The transfinite diameter of an arc $\Gamma_\alpha$ of a circle of radius $R$ with central angle $\alpha$ is equal to $R \sin(\alpha/4)$. In particular, for the unit circle $\mathbb{T}$, we have $\tau(\Gamma_\alpha) = \sin(\alpha/4)$ (Tsuji, 1975, p. 84).

(f) The transfinite diameter of an arbitrary line segment $F$ is equal to one-fourth its length, that is, if $F := [a, b]$, then $\tau(F) = \tau([a, b]) = (b - a)/4$. (see, e.g., Tsuji, 1975, p. 84).

3.2. An extension of Rosenblatt’s first theorem

In what follows, by $E_f$ we denote the support of the process $X(t)$, that is,

$$E_f := \{ e^{it} : f(\lambda) > 0 \}. \quad (3.8)$$

Thus, the closure $\overline{E_f}$ of $E_f$ is the support of the spectral density $f$.

Our first theorem extends Rosenblatt’s first theorem (Theorem A). More precisely, the result that follows extends the asymptotic relation (1.2) to the case of several arcs, without having to stipulate continuity of the spectral density $f$.

Theorem 3.1. Let the support $\overline{E_f}$ of the spectral density $f$ of the process $X(t)$ consist of a finite number of closed arcs of the unit circle $\mathbb{T}$, and let $f > 0$ a.e. on $\overline{E_f}$. Then the sequence $\sqrt{n} \sigma_n(f)$ converges, and

$$\lim_{n \to \infty} \sqrt{n} \sigma_n(f) = \tau_f, \quad (3.9)$$

where $\tau_f := \tau(\overline{E_f})$ is the transfinite diameter of $\overline{E_f}$.

Remark 3.1. A version of Theorem 3.1 was first proved in Babayan (1984) (see also Babayan, 1985). Here we will give a simplified short proof of this result.

Remark 3.2. It can be shown that under some natural additional conditions, the asymptotic relation (3.9) remains valid in the case where the support $\overline{E_f}$ consists of a countable number of arcs of the unit circle $\mathbb{T}$ (see Babayan, 1985).

Remark 3.3. In Theorem A, $\overline{E_f} = \{ e^{it} : \lambda \in [\pi/2 - a, \pi/2 + a] \}$, which represents a closed arc of length $2a$, and, according to Proposition 3.2(e), we have $\tau(\overline{E_f}) = \sin(2a/4) = \sin(a/2)$. Thus, the asymptotic relation (1.2) is a special case of (3.9).

We will need the following definition, which characterizes the rate of variation of a sequence of non-negative numbers compared with a geometric progression (see also Simon, 2005, p. 91).

Definition 3.1. (a) A sequence $\{a_n \geq 0, \ n \in \mathbb{N}\}$ is called exponentially neutral if $\lim_{n \to \infty} \sqrt{n} a_n = 1$.
(b) A sequence $\{b_n \geq 0, \ n \in \mathbb{N}\}$ is called exponentially decreasing if $\limsup_{n \to \infty} \sqrt{n} b_n < 1$.

For instance, the sequence $\{a_n = n^\alpha, \ \alpha \in \mathbb{R}, \ n \in \mathbb{N}\}$ is exponentially neutral because $\log \sqrt{n^\alpha} = (\alpha/n) \log n \to 0$ as $n \to \infty$. The geometric progression $\{b_n = q^n, \ 0 < q < 1, \ n \in \mathbb{N}\}$ is exponentially decreasing because
\[ \sqrt[n]{b_n} = q^{n/n} = q < 1. \] The sequence \( \{b_n = n^\alpha q^n, \alpha \in \mathbb{R}, 0 < q < 1, n \in \mathbb{N}\} \) is also exponentially decreasing because \( \sqrt[n]{b_n} = n^{n/n}q \to q < 1. \)

**Remark 3.4.** In fact, it can easily be shown that a sequence \( \{b_n \geq 0, n \in \mathbb{N}\} \) is exponentially decreasing if and only if there exists a number \( q (0 < q < 1) \) such that \( b_n = O(q^n) \) as \( n \to \infty. \)

**Remark 3.5.** It follows from relation (3.9) that the question of exponential decay of the prediction error \( \sigma_n(f) \) as \( n \to \infty \) is determined solely by the value of the transfinite diameter of the support \( \widetilde{E}_f \) of the spectral density \( f \), and does not depend on the values of \( f \) on \( \overline{E}_f \). Denoting \( \tau_n := \sigma_n(f)/\tau_f^n \) from (3.9) infer that \( \lim_{n \to \infty} \sqrt[n]{\tau_n} = 1 \) and

\[ \sigma_n(f) = \tau_f^n \cdot \gamma_n. \tag{3.10} \]

Thus, in the case where \( \tau_f < 1 \), the prediction error \( \sigma_n(f) \) is decomposed into a product of two factors, one of which \( (\tau_f^n) \) is a geometric progression, and the second \( (\gamma_n) \) is an exponentially neutral sequence. Also, if \( g \) is another spectral density satisfying the conditions of Theorem 3.1, then in view of (3.10), we have

\[ \frac{\sigma_n(g)}{\sigma_n(f)} = \left( \frac{\tau_g}{\tau_f} \right)^n \cdot \gamma'_n, \]

where \( \gamma'_n \) is an exponentially neutral sequence. It is worth noting that the last relation does not depend on the structures of the supports \( \overline{E}_f \) and \( \overline{E}_g \) (viz., the number and the lengths of arcs constituting these sets, as well as, their location on the unit circle \( \mathbb{T} \)).

The following result provides a sufficient condition for the exponential decay of \( \sigma_n(f) \).

**Theorem 3.2.** If the spectral density \( f \) of the process \( X(t) \) vanishes on an arc, then the prediction error \( \sigma_n(f) \) decreases to zero exponentially. More precisely, if \( f \) vanishes on an arc \( \Gamma_g \subset \mathbb{T} \) of length \( 2\delta (0 < \delta < \pi) \), then

\[ \lim_{n \to \infty} \sup_n \sqrt[n]{\sigma_n(f)} \leq \cos(\delta/2) < 1. \tag{3.11} \]

The next result gives a necessary condition for the exponential decay of \( \sigma_n(f) \).

**Theorem 3.3.** A necessary condition for the prediction error \( \sigma_n(f) \) to tend to zero exponentially is that the spectral density \( f \) should vanish on a set of positive Lebesgue measure.

**Remark 3.6.** Theorem 3.3 shows that if the spectral density \( f \) is almost everywhere positive, then it is impossible to obtain exponential decay of the prediction error \( \sigma_n(f) \), no matter how high the orders of the zeros of \( f \).

In view of (2.19), as a consequence of Theorem 3.1, we obtain the following result.

**Theorem 3.4.** Let the support \( \overline{E}_f \) and the spectral density \( f \) satisfy the conditions of Theorem 3.1. If the sequence of Verblunsky coefficients \( v_n(f) \) converges in modulus, then

\[ \lim_{n \to \infty} |v_n(f)| = \sqrt{1 - \tau_f^2}. \tag{3.12} \]

**Remark 3.7.** It is well-known that for an arbitrary sequence of positive numbers \( a_n \), the convergence \( a_{n+1}/a_n \to a \) implies the convergence \( \sqrt[n]{a_n} \to a \). The converse, in general, is not true, that is, the sequence \( \sqrt[n]{a_n} \) can be convergent, while \( a_{n+1}/a_n \) divergent. Indeed, for the sequence \( a_n \):

\[ a_n := \begin{cases} 2^{-3k} & \text{if } n = 2k - 1 \\ 2^{-(3k+1)} & \text{if } n = 2k, \quad k \in \mathbb{N}, \end{cases} \]

we have \( \lim_{n \to \infty} \sqrt[n]{a_n} = 2^{-3/2} \), while the limit \( \lim_{n \to \infty} a_{n+1}/a_n \) does not exist.
In the context of the considered sequences, $|v_n(f)|$ and $\sqrt{\sigma_n(f)}$, in view of (2.19), we can assert that the convergence of $|v_n(f)|$ (or equivalently $\sigma_{n+1}(f)/\sigma_n(f)$) implies the convergence of $\sqrt{\sigma_n(f)}$, but not the converse. Hence, the condition of convergence (in modulus) of Verblunsky’s sequence in Theorem 3.4 is essential.

As a consequence of Theorem 3.1 we obtain the following result (see Geronimus, 1948), which is a partial converse of Rakhmanov’s theorem:

**Theorem 3.5.** If the sequence $\sigma_n(f)$ satisfies the following condition:

$$\limsup_{n \to \infty} \sqrt{\sigma_n(f)} = 1$$  \hspace{1cm} (3.13)

(in particular, if $\lim_{n \to \infty} v_n(f) = 0$), then $\overline{E}_f = \mathbb{T}$, that is, the spectrum of the process is dense in $\mathbb{T}$.

### 3.3. Proof of the results of Section 3.2

We prove here Theorems 3.1–3.5. The proof of Theorem 3.1 is based on Lemma 3.1, where we use the following notions. A *continuum* $\Gamma$ is a continuous rectifiable Jordan curve in the complex plane $\mathbb{C}$. The *linear measure* $\mu(E)$ of a set $E \subset \Gamma$ is, by definition, the Lebesgue measure generated by the length of an arc of a continuum.

The following lemma is an immediate consequence of a result by Mazurkiewicz (1945).

**Lemma 3.1.** Let $\Gamma$ be a bounded closed set consisting of a finite number of continua. Then for any $\epsilon > 0$ there is a number $\delta = \delta(\epsilon, \Gamma) > 0$ such that for any closed subset $F \subset \Gamma$ and an arbitrary polynomial $q_n(z)$ of degree $n$ the following inequality holds:

$$M_n := \max_{x \in \Gamma} |q_n(z)| \leq (1 + \epsilon)^n \max_{z \in F} |q_n(z)|,$$  \hspace{1cm} (3.14)

provided that $\mu(\Gamma \setminus F) < \delta$.

In our proof of Theorem 3.1 given below, the set $\Gamma$ will be either the unit circle $\mathbb{T}$ or the union of a finite number of closed arcs of $\mathbb{T}$.

**Proof of Theorem 3.1.** We first prove the inequality

$$\limsup_{n \to \infty} \sqrt{\sigma_n(f)} \leq \tau(\overline{E}_f) \leq 1.$$  \hspace{1cm} (3.15)

To this end, we define the spectral density:

$$\tilde{f}(\lambda) := \begin{cases} f(\lambda) & \text{if } e^{i\lambda} \in E_f \\ 1 & \text{if } e^{i\lambda} \in \overline{E}_f \setminus E_f \\ 0 & \text{if } e^{i\lambda} \notin \overline{E}_f, \end{cases}$$  \hspace{1cm} (3.16)

and observe that $f(\lambda) \leq \tilde{f}(\lambda)$ for all $\lambda \in \Lambda$. Let $T_n(z, \overline{E}_f) := T_n(z, \overline{E}_f)$ be the Chebyshev polynomial of order $n$ of the set $\overline{E}_f$, and let (see (3.3))

$$m_n := m_n(\overline{E}_f) = \max_{z \in \overline{E}_f} |T_n(z, \overline{E}_f)|.$$  \hspace{1cm} (3.17)

Then we can write

$$\sigma_n^2(f) \leq \sigma_n^2(\overline{f}) = \|p_n(\overline{f})\|_{L_2}^2 \leq \|T_n(\overline{E}_f)\|_{L_2}^2 \leq c \cdot m_n^2(\overline{E}_f),$$  \hspace{1cm} (3.18)
where \( c := r(0, \tilde{f}) > 0 \) and \( r(t, \tilde{f}) \) is the covariance function corresponding to the spectral density \( \tilde{f} \). The first relation in (3.18) follows from Proposition 2.2(b), the second from (2.16), the third from the definition of optimal polynomial \( p_n(z, \tilde{f}) \), and the fourth from (3.17).

Taking the root of order \( 2n \) in (3.18), then passing to the limit as \( n \to \infty \), in view of (3.4), (3.7), Proposition 3.2(a) and (d), and the elementary relation \( \lim_{n \to \infty} \sqrt[n]{c} = 1 \), we obtain

\[
\limsup_{n \to \infty} \sqrt{n} \sigma_n(\tilde{f}) \leq \tau(\tilde{E}_f) \leq \tau(\mathcal{T}) = 1.
\]

Now we proceed to prove the inequality:

\[
\liminf_{n \to \infty} \sqrt{n} \sigma_n(\tilde{f}) \geq \tau(\tilde{E}_f). \tag{3.19}
\]

To this end, observe first that under the conditions of the theorem, by Proposition 2.2(c), we have

\[
\sigma_n^2(\tilde{f}) = \sigma_n^2(\tilde{f}) - \sigma_n^2(\tilde{f}) = \sigma_n^2(\tilde{f}).
\]

Consider a sequence of subsets \( \{E_n, n \in \mathbb{N}\} \) of \( \tilde{E}_f \), defined by

\[
E_n := \{z \in \tilde{E}_f : |p_n(z, \tilde{f})| > n\sigma_n(\tilde{f})\}. \tag{3.20}
\]

Then, in view of (2.16) and (3.20) we can write

\[
\sigma_n^2(\tilde{f}) = \int_{\tilde{E}_f} |p_n(z, \tilde{f})|^2 d\mu_{\tilde{f}} \geq \int_{E_n} |p_n(z, \tilde{f})|^2 d\mu_{\tilde{f}} > n^2 \sigma_n^2(\tilde{f}) \mu_{\tilde{f}}(E_n), \tag{3.21}
\]

where \( \mu_{\tilde{f}} \) is the measure on the unit circle \( \mathcal{T} \), corresponding to the spectral density \( \tilde{f} \). It follows from (3.21) that \( \mu_{\tilde{f}}(E_n) < n^{-2} \) and \( \lim_{n \to \infty} \mu_{\tilde{f}}(E_n) = 0 \). Next, since the spectral density \( \tilde{f} \) in (3.16) is strictly positive on \( \tilde{E}_f \), the Lebesgue measure \( \mu_\mathcal{T} \) is absolutely continuous with respect to the measure \( \mu_{\tilde{f}} \), and hence we have

\[
\lim_{n \to \infty} \mu_\mathcal{T}(E_n) = 0. \tag{3.22}
\]

Define the sets \( F_n := \tilde{E}_f \setminus E_n \), and observe that \( F_n \) are closed subsets of the set \( \tilde{E}_f \). Then, in view of (3.20), we have

\[
|p_n(z, \tilde{f})| \leq n\sigma_n(\tilde{f}), \quad z \in F_n. \tag{3.23}
\]

Given an arbitrary \( \varepsilon > 0 \) we choose \( \delta := \delta(\tilde{E}_f, \varepsilon) \) according to Lemma 3.1 with \( \Gamma = \tilde{E}_f \) and \( F = F_n \). Then, in view of (3.22), for large enough \( n \), we have \( \mu_\mathcal{T}(E_f \setminus F_n) = \mu_\mathcal{T}(E_n) < \delta \).

Therefore, we can write

\[
m_{\mathcal{T}}(\tilde{E}_f) = \max_{z \in \tilde{E}_f} |T_n(z, \tilde{f})| \leq \max_{z \in \tilde{E}_f} |p_n(z, \tilde{f})| \leq (1 + \varepsilon)^n \max_{z \in \tilde{E}_f} |p_n(z, \tilde{f})| \leq (1 + \varepsilon)^n n\sigma_n(\tilde{f}),
\]

Here the first and the second relations follow from the definition of Chebyshev polynomial (see (3.3)), the third from the relation (3.14), and the fourth from the inequality (3.23). The last relation implies that

\[
n\sigma_n(\tilde{f}) \geq (1 + \varepsilon)^{-n} m_{\mathcal{T}}(\tilde{E}_f).
\]
Taking the root of order $n$, and letting $n$ tend to infinity, in view of the relation $\lim_{n \to \infty} \sqrt[n]{n} = 1$ and the elementary inequality $(1 + \varepsilon)^{-1} > 1 - \varepsilon$, we obtain

$$\liminf_{n \to \infty} \sqrt[n]{\sigma_n(f)} \geq \tau(\overline{E}_f)(1 - \varepsilon).$$  \hfill (3.24)

Taking into account the arbitrariness of $\varepsilon$ and the equality $\sigma_n(f) = \sigma_n(\overline{f})$, from (3.24) we obtain (3.19). A combination of (3.15) and (3.19) implies (3.9), and thus completes the proof of Theorem 3.1.

**Proof of Theorem 3.2.** Denote by $\Gamma_{\alpha}$ the closure of the complement of $\Gamma_{\delta}$. Then we have $\overline{E}_f \subset \Gamma_{\alpha}$ and $\tau(\overline{E}_f) \leq \tau(\Gamma_{\alpha})$ (by Proposition 3.2(a)). Observing that the arc $\Gamma_{\alpha}$ is of length $2\alpha$ with $\alpha = \pi - \delta$, by Proposition 3.2(e) we have

$$\tau(\Gamma_{\alpha}) = \sin\left(\frac{2(\pi - \delta)}{4}\right) = \sin\left(\frac{\pi}{2} - \frac{\delta}{2}\right) = \cos\left(\frac{\delta}{2}\right).$$  \hfill (3.25)

Therefore, in view of (3.15) and (3.25), we obtain

$$\limsup_{n \to \infty} \sqrt[n]{\sigma_n(f)} \leq \tau(\overline{E}_f) \leq \tau(\Gamma_{\alpha}) = \cos\left(\frac{\delta}{2}\right),$$

and the relation (3.11) follows, completing the proof of Theorem 3.2.\hfill \blacksquare

**Proof of Theorem 3.3.** We argue by contradiction. According to the assumption of the theorem, we have a sequence of prediction errors $\sigma_n(f)$ that decreases to zero exponentially, that is,

$$\limsup_{n \to \infty} \sqrt[n]{\sigma_n(f)} < 1. \hfill (3.26)$$

Assuming that the spectral density $f$ is almost everywhere positive on $\mathbb{T}$, according to Theorem 3.1, we would have $\lim_{n \to \infty} \sqrt[n]{\sigma_n(f)} = \tau(\overline{E}_f) = \tau(\mathbb{T}) = 1$, which contradicts the inequality (3.26), completing the proof of Theorem 3.3. \hfill \blacksquare

**Proof of Theorem 3.4.** In view of the second formula in (2.19), we have

$$|v_n| = \sqrt{1 - \frac{\sigma_{n+1}^2(f)}{\sigma_n^2(f)}}. \hfill (3.27)$$

Since by assumption the sequence $|v_n|$ converges, the sequence $\sigma_{n+1}(f)/\sigma_n(f)$ also converges. Hence, in view of Remark 3.7 and (3.9), we obtain

$$\lim_{n \to \infty} \sigma_{n+1}(f)/\sigma_n(f) = \lim_{n \to \infty} \sqrt[n]{\sigma_n(f)} = \tau_f. \hfill (3.28)$$

Passing to the limit in (3.27), and taking into account (3.28), we obtain the relation (3.12). \hfill \blacksquare

**Proof of Theorem 3.5.** Observe first that in the proof of Theorem 3.1, the inequality (3.15) was established without using the assumptions on the structure of the set $E_{f'}$, that is, (3.15) is true for an arbitrary set $E_{f'}$. Then from (3.13) and (3.15) it follows that $\tau(\overline{E}_f) = 1$. From this we can conclude that $\overline{E}_f = \mathbb{T}$. Indeed, $\mathbb{T} \setminus \overline{E}_f$ is an open set, and hence if it contains at least one point, then it should contain an entire interval. Therefore, in this case, in view of Theorem 3.2 the inequality (3.11) should be satisfied, which contradicts the condition (3.13). The obtained contradiction completes the proof of Theorem 3.5. \hfill \blacksquare

### 3.4. Some consequences of Theorem 3.1

Motivated by Theorems A and 3.1 and Remark 3.3, the following question arises naturally: calculate the transfinite diameter $\tau(\overline{E}_f)$ of the set $\overline{E}_f$ consisting of several closed arcs of the unit circle $\mathbb{T}$, and thus, obtain an explicit
asymptotic relation for the prediction error $\sigma_p(f)$ similar to Rosenblatt’s relation (1.2). As was mentioned in Section 3.1, the calculation of the transfinite diameter (and hence, the capacity and the Chebyshev constant) is a challenging problem, and in only very few cases has the transfinite diameter been exactly calculated (see Proposition 3.2). One such example provides Theorem A, in which case the transfinite diameter of the set $E_f := \{e^{ik} : k \in [\pi/2 - \alpha, \pi/2 + \alpha]\}$ is $\sin(\alpha/2)$. Observe that in Rosenblatt (1957), Rosenblatt calculated the capacity of $E_f$. Below we give some other examples, where we can explicitly calculate the Chebyshev constant (and hence the transfinite diameter and the capacity) by using some properties of the transfinite diameter, stated in Proposition 3.2, and a result due to Robinson (1969) concerning the relation between the transfinite diameters of related sets.

In Robinson (1969), R. Robinson, extending the Fekete theorem (see Proposition 3.2(c)), proved the following important result about the transfinite diameters of related sets.

**Proposition 3.3.** Let $F$ be a bounded closed subset of the complex plane lying on the unit circle $\mathbb{T}$ and symmetric with respect to real axis, and let $F^\circ$ be the projection of $F$ onto the real axis. Then

$$\tau(F^\circ) = [2\tau(F)]^{1/2}.$$  \hfill (3.29)

The examples given below show that the formula (3.29) gives a simple way to calculate the transfinite diameters of some subsets of the circle, based only on the formula of the transfinite diameter of a line segment (see Proposition 3.2(f)).

We now give examples of calculation of transfinite diameters of some subsets of the unit circle. We will use the following notation: given $0 < \beta < 2\pi$ and $z_0 = e^{i\theta_0}, \theta_0 \in [-\pi, \pi]$, we denote by $\Gamma(\theta_0)$ an arc of the unit circle of length $\beta$ which is symmetric with respect to the point $z_0 = e^{i\theta_0}$:

$$\Gamma(\theta_0) := \{e^i\theta : |\theta - \theta_0| \leq \beta/2\} = \{e^i\theta : \theta \in [\theta_0 - \beta/2, \theta_0 + \beta/2]\}. \hfill (3.30)$$

**Example 3.1.** Let $\Gamma_{2\alpha} := \Gamma_{2\alpha}(0)$. Then the projection $\Gamma_{2\alpha}^\circ$ of $\Gamma_{2\alpha}$ onto the real axis is the segment $[\cos \alpha, 1]$ (see Figure 1(a)), and by Proposition 3.2(f) for the transfinite diameter $\tau(\Gamma_{2\alpha}^\circ)$ we have

$$\tau(\Gamma_{2\alpha}^\circ) = \frac{1 - \cos \alpha}{4} = \frac{\sin^2(\alpha/2)}{2}.$$  \hfill (3.31)

Hence, according to formula (3.29), we obtain

$$\tau(\Gamma_{2\alpha}(\theta_0)) = \sin(\alpha/2).$$  \hfill (3.32)

Taking into account that the transfinite diameter is invariant with respect to rotation (see Proposition 3.2(b)), from (3.31) for any $\theta_0 \in [-\pi, \pi]$ we have

$$\tau(\Gamma_{2\alpha}(\theta_0)) = \sin(\alpha/2).$$  \hfill (3.33)

Notice that formula in (3.31) was obtained by Rosenblatt (1957), where he calculated the capacity of the arc $\Gamma_{2\alpha}(\pi/2)$ by using the complex-analysis technique of conformal mappings and OPUC.

**Example 3.2.** Let $\Gamma_{2\alpha}(\alpha)$ be an arc of length $2\alpha$, defined by (3.30): $\Gamma_{2\alpha}(\alpha) = \{e^i\theta : \theta \in [0, 2\alpha]\}$, and let $\Gamma(2)$ be the preimage of the arc $\Gamma_{2\alpha}(\alpha)$ under the mapping $p(z) = z^2$. We show that the set $\Gamma(2)$ is the union of two closed arcs of equal lengths $\alpha$, symmetrically located with respect to the center of the unit circle (see Figure 1(b)):

$$\Gamma(2) = \{e^{i\omega} : \omega \in [-\pi, -\pi + \alpha] \cup [0, \alpha]\}. \hfill (3.33)$$
Indeed, the preimage $z = e^{i\omega}$ of an arbitrary point $e^{i\theta} \in \Gamma_{2\alpha}(a)$, $\theta \in [0, 2\alpha]$, under the mapping $p(z) = z^2$ satisfies the equality $z^2 = e^{2i\omega} = e^{i\theta}$. This, in view of the $2\pi$-periodicity of $e^{i\theta}$ implies that $2\omega = \theta - 2\pi k$, $k \in \mathbb{Z}$, and hence

$$\omega = \omega(k) = \theta/2 - \pi k, \quad k \in \mathbb{Z}. \quad (3.34)$$

Again using the $2\pi$-periodicity of $e^{i\theta}$, we conclude that from the countable set of values of $\omega(k)$ in (3.34) only two values $\omega(0) = \theta/2$ and $\omega(1) = \theta/2 - \pi$ correspond to distinct preimages $z = e^{i\omega}$ of the point $e^{i\theta} \in \Gamma_{2\alpha}(a)$, $\theta \in [0, 2\alpha]$. Thus, each point $e^{i\theta} \in \Gamma_{2\alpha}(a)$ has two distinct preimages $z_1 = e^{i\omega(0)} = e^{i\theta/2}$ and $z_2 = e^{i\omega(1)} = e^{i\theta/2 - \pi}$. Therefore, for the entire preimage $\Gamma(2)$ we have

$$\Gamma(2) = \{ e^{i\theta/2} : \theta \in [0, 2\alpha] \} \cup \{ e^{i\theta/2 - \pi} : \theta \in [0, 2\alpha] \}$$

$$= \{ e^{i\psi} : \psi \in [0, \alpha] \} \cup \{ e^{i\xi} : \xi \in [-\pi, -\pi + \alpha] \}$$

$$= \{ e^{i\omega} : \omega \in [-\pi, -\pi + \alpha] \cup [0, \alpha] \}.$$

and (3.33) follows. Then, by the Fekete theorem (see Proposition 3.2(c)) and formula (3.32), for the transfinite diameter $\tau(\Gamma(2))$ we have

$$\tau(\Gamma(2)) = [\tau(\Gamma_{2\alpha}(a))]^{1/2} = (\sin(\alpha/2))^{1/2}.$$

The above result can easily be extended to the case of $k (k \geq 2)$ arcs. Let $\Gamma(k)$ be the union of $k (k \in \mathbb{N}, \ k \geq 2)$ closed arcs of equal lengths $\alpha$, which are symmetrically located on the unit circle (the arcs are assumed to be equidistant). Arguments similar to those applied above can be used to show that the set $\Gamma(k)$ is the preimage (to within rotation) under the mapping $p(z) = z^k$ of the arc $\Gamma_{k\alpha}(ka/2)$ of length $ka$ defined by (3.30). Therefore, by the Fekete theorem and the invariance property of the transfinite diameter with respect to rotation (see Proposition 3.2(b)), for the transfinite diameter $\tau(\Gamma(k))$, we have

$$\tau(\Gamma(k)) = (\sin(ka/4))^{1/k}. \quad (3.35)$$

**Example 3.3.** Let $\alpha > 0$, $\delta \geq 0$ and $\alpha + \delta \leq \pi$. Let $\Gamma_{a,\delta}(\theta_0) := \Gamma_{a+\delta}(\theta_0) \setminus \Gamma_{\delta}(\theta_0)$ be the union of two arcs of the unit circle of lengths $\alpha$, the distance between which (over the circle) is equal to $2\delta$. Define (see Figure 2(a)):

$$\Gamma_{a,\delta} := \Gamma_{a,\delta}(0) = \{ e^{i\theta} : \theta \in [-(\delta + \alpha), -\delta] \cup [\delta, \delta + \alpha] \}. \quad (3.36)$$
Then the projection $\Gamma_{a,\delta}$ of $\Gamma_{a,\delta}$ onto the real axis is the segment $\Gamma_{a,\delta}^x = [\cos(a+\delta), \cos\delta]$, and by Proposition 3.2(f) for the transfinite diameter $\tau(\Gamma_{a,\delta})$ we have

$$\tau(\Gamma_{a,\delta}) = \frac{\cos\delta - \cos(a+\delta)}{4} = \frac{\sin(a/2)\sin(a/2+\delta)}{2}. \quad (3.37)$$

Hence, according to formula (3.29), for the transfinite diameter $\tau(\Gamma_{a,\delta})$, we obtain

$$\tau(\Gamma_{a,\delta}) = \left[2\tau(\Gamma_{a,\delta}^x)\right]^{1/2} = (\sin(a/2)\sin(a/2+\delta))^{1/2}. \quad (3.37)$$

In view of Proposition 3.2(b)), from (3.37) for any $\theta_0 \in [-\pi, \pi)$ we have

$$\tau(\Gamma_{a,\delta}(\theta_0)) = (\sin(a/2)\sin(a/2+\delta))^{1/2}. \quad (3.38)$$

Observe that for $\delta = 0$ we have $\Gamma_{a,\delta}(\theta_0) = \Gamma_a(\theta_0)$, and the formula (3.38) becomes (3.32).

**Example 3.4.** Let the arc $\Gamma_{a,\delta}$ be as in Example 3.3 (see (3.36)) with $a, \delta$ satisfying $a+\delta \leq \pi/2$, that is, $\Gamma_{a,\delta}$ is a subset of the right semicircle $T$. Denote by $\Gamma_{a,\delta}^* = \{\Gamma_{a,\delta} \cup \Gamma_{a,\delta}^x\}$ the symmetric to $\Gamma_{a,\delta}$ set with respect to both axes, that is,

$$\Gamma_{a,\delta}^* = \{e^{\theta} : \theta \in [-\pi+\delta, -\pi+\delta+\alpha] \cup [\pi-(\delta+\alpha), \pi-\delta]\}. \quad (3.39)$$

Define $\Delta_{a,\delta} := (\Gamma_{a,\delta}^* \cup \Gamma_{a,\delta}^x)$, and observe that the set $\Delta_{a,\delta}$ consists of four arcs of equal lengths $a$, which are symmetrically located with respect to both axes (see Figure 2(b)). Arguments similar to those applied in Example 3.2 can be used to show that the set $\Delta_{a,\delta}$ is the preimage (to within rotation) of the set $\Gamma_{2a,2\delta}$ under the mapping $p(z) = z^2$. Hence, according to the Fekete theorem (see Proposition 3.2(c)) and formula (3.37), for the transfinite diameter $\tau(\Delta_{a,\delta})$, we obtain

$$\tau(\Delta_{a,\delta}) = (\tau(\Gamma_{2a,2\delta})^{1/2} = (\sin a \sin(a+2\delta))^{1/4}. \quad (3.39)$$

Denote by $\Delta_{a,\delta}(\theta_0)$ the image of the set $\Delta_{a,\delta}$ under mapping $q(z) = e^{i\theta_0}z$, that is, under the rotation by the central angle $\theta_0$ around the origin. Then, in view of Proposition 3.2(b)), from (3.39) for any $\theta_0 \in [-\pi, \pi)$ we have

$$\tau(\Delta_{a,\delta}(\theta_0)) = (\sin a \sin(a+2\delta))^{1/4}. \quad (3.40)$$

Now we apply Theorem 3.1 to obtain the asymptotic behavior of the prediction error $\sigma_n(f)$ for some specific spectra. Putting together Theorem 3.1 and Examples 3.1–3.4, we can state the following result.
Theorem 3.6. Let $E_f$ be the support of the spectral density $f$ of a stationary process $X(t)$. Then

(a) If $E_f = \Gamma_2, \theta_0)$, where $\Gamma_2, \theta_0)$ is as in Example 3.1, then $\lim_{n \to \infty} \sqrt{n} \sigma_n(f) = \sin(a/2)$.

(b) If $E_f = \Gamma(k)$, where $\Gamma(k)$ is as in Example 3.2, then $\lim_{n \to \infty} \sqrt{n} \sigma_n(f) = (\sin(ka/4))^{1/k}$.

(c) If $E_f = \Gamma_{a,\delta}(\theta_0)$, where $\Gamma_{a,\delta}(\theta_0)$ is as in Example 3.3, then

$$\lim_{n \to \infty} \sqrt{n} \sigma_n(f) = (\sin(a/2) \sin(a/2 + \delta))^{1/2}.$$ 

(d) If $E_f = \Delta_{a,\delta}(\theta_0)$, where $\Delta_{a,\delta}(\theta_0)$ is as in Example 3.4, then

$$\lim_{n \to \infty} \sqrt{n} \sigma_n(f) = (\sin \alpha \sin(a + 2\delta))^{1/4}.$$ 

Remark 3.8. The assertion (a) is a slight extension of the Rosenblatt relation (1.2). The assertion (c) is an extension of assertion (a), which reduces to assertion (a) if $\delta = 0$.

4. AN EXTENSION OF ROSENBLATT’S SECOND THEOREM

4.1. Preliminaries

We analyze the asymptotic behavior of the prediction error in the case where the spectral density $f$ of the model has a very high order contact with zero at one or several points, so that the Szegő condition (2.5) is violated.

Based on Rosenblatt’s result for this case, namely Theorem B, we can expect that for any deterministic process with spectral density possessing a singularity of the type (1.4), the rate of the prediction error $\sigma_n^2(f)$ should be the same as in (1.5). However, the method applied in Rosenblatt (1957) does not allow us to prove this assertion. Here, using a different approach, we extend Rosenblatt’s second theorem to a broader class of spectral densities.

To state the corresponding results we need some definitions and preliminaries. We first introduce the notion of weakly varying sequences and state some of their properties.

Definition 4.1. A sequence of non-zero numbers $\{a_n, n \in \mathbb{N}\}$ is said to be weakly varying if $\lim_{n \to \infty} a_{n+1}/a_n = 1$.

For example, the sequence $\{n^a, a \in \mathbb{R}, n \in \mathbb{N}\}$ is weakly varying (for $a < 0$ it is weakly decreasing and for $a > 0$ it is weakly increasing), while the geometric progression $\{q^n, 0 < q < 1, n \in \mathbb{N}\}$ is not weakly varying.

In the next proposition we list some simple properties of the weakly varying sequences, which can easily be verified.

Proposition 4.1. The following assertions hold.

(a) If $a_n$ is a weakly varying sequence, then $\lim_{n \to \infty} a_{n+v}/a_n = 1$ for any $v \in \mathbb{N}$.

(b) If $a_n$ is such that $\lim_{n \to \infty} a_n = a \neq 0$, then $a_n$ is a weakly varying sequence.

(c) If $a_n$ and $b_n$ are weakly varying sequences, then $ca_n$ ($c \neq 0$), $a_n^\alpha$ ($\alpha \in \mathbb{R}$), $a_nb_n$ and $a_n/b_n$ also are weakly varying sequences.

(d) If $a_n$ is a weakly varying sequence, and $b_n$ is a sequence of non-zero numbers such that $\lim_{n \to \infty} b_n/a_n = c \neq 0$, then $b_n$ is also a weakly varying sequence.

(e) If $a_n$ is a weakly varying sequence of positive numbers, then it is exponentially neutral (see Definition 3.1(a) and Remark 3.7).

Properties of the geometric mean and trigonometric polynomials. Recall that a trigonometric polynomial $t(\lambda)$ of degree $\nu$ is a function of the form:

$$t(\lambda) = a_0 + \sum_{k=1}^{\nu} (a_k \cos k\lambda + b_k \sin k\lambda) = \sum_{k=-\nu}^{\nu} c_k e^{ik}, \quad \lambda \in \mathbb{R},$$

where $a_0, a_k, b_k \in \mathbb{R}$, $c_0 = a_0$, $c_k = 1/2(a_k - ib_k)$, $c_{-k} = 1/2(a_k + ib_k)$, $k = 1, 2, \ldots, \nu$. 

J. Time Ser. Anal. 42: 622–652 (2021) © 2020 John Wiley & Sons Ltd wileyonlinelibrary.com/journal/jtsa

DOI: 10.1111/jtsa.12572
In the next proposition we list some properties of the geometric mean \( G(h) \) of a function \( h \) (see formula (2.4)) and trigonometric polynomials.

**Proposition 4.2.** The following assertions hold.

(a) Let \( c > 0, \alpha \in \mathbb{R}, f \geq 0 \) and \( g \geq 0 \). Then

\[
G(c) = c; \quad G(fg) = G(f)G(g); \quad G(f^\alpha) = G^\alpha(f), \quad G(f) > 0.
\]

(4.1)

(b) If \( f(\lambda) \leq g(\lambda) \), then \( G(f) \leq G(g) \).

(c) (Fej’er–Riesz theorem). Let \( t \) be a non-negative trigonometric polynomial of degree \( \nu \). Then there exists an algebraic polynomial \( s_\nu(z) \) of the same degree \( \nu \), such that \( s_\nu(z) \neq 0 \) for \( |z| < 1 \), and

\[
t(\lambda) = |s_\nu(e^{i\lambda})|^2.
\]

(4.2)

Under the additional condition \( s_\nu(0) > 0 \) the polynomial \( s_\nu(z) \) is determined uniquely.

(d) Let \( t(\lambda) \) and \( s_\nu(z) \) be as in Assertion (b). Then \( G(t) = |s_\nu(0)|^2 > 0 \).

**Proof.** Assertions (a) and (b) immediately follow from the definition of the geometric mean (see formula (2.4)) and the properties of exponent and logarithm. The proof of Assertion (c) can be found, for example, in Grenander and Szegő (1958), p. 20–2. Assertion (d) follows from Assertion (c). Indeed, observing that \( \ln |s_\nu(z)|^2 \) is a harmonic function, by the well-known mean-value theorem for harmonic functions (see, e.g., Ahlfors, 1979, p. 165), we have

\[
\ln |s_\nu(0)|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln |s_\nu(e^{i\lambda})|^2 d\lambda = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln t(\lambda) d\lambda = \ln G(t),
\]

and the result follows. \( \square \)

In what follows we consider the class of deterministic processes for which the sequence of prediction errors \( \sigma_n(f) \) is weakly varying, that is,

\[
\lim_{n \to \infty} \frac{\sigma_{n+1}(f)}{\sigma_n(f)} = 1.
\]

(4.4)

**Remark 4.1.** According to the second relation in (2.19) the condition (4.4) is equivalent to \( \lim_{n \to \infty} v_n(f) = 0 \). Hence, by Rakhmanov’s theorem, if the spectral density \( f > 0 \) a.e. on \( \Lambda \), then (4.4) is satisfied, that is, the sequence \( \sigma_n(f) \) is weakly varying. Thus, the considered class of stationary processes includes all the deterministic processes possessing a.e. positive spectral densities. Also, from Theorem 3.5 and Remark 3.7 we infer that a necessary condition for (4.4) is that the spectrum \( E_f \) is dense in \( \Lambda \).

4.2. An extension of Rosenblatt’s second theorem

We first examine the asymptotic behavior as \( n \to \infty \) of the ratio

\[
\frac{\sigma_n^2(fg)}{\sigma_n^2(f)},
\]
where \( g \) is a non-negative function. We suppose that \( fg \) is a density'. By this we mean that \( fg \) can be normalized so that it can be viewed as a probability density with respect to the normalized Lebesgue measure on \([−π, π]\).

Supposing that \( fg \) is a density is a way to exclude cases like \( fg \equiv 0 \), which occurs for instance if \( f \) and \( g \) have disjoint supports.

To clarify the approach, we first assume that \( f \) is the spectral density of a non-deterministic process, in which case the geometric mean \( G(f) \) is positive (see (2.3) and (2.4)). We can then write

\[
\lim_{n→∞} \frac{σ_n^2(fg)}{σ_n^2(f)} = \frac{σ_n^2(fg)}{σ_n^2(f)} = \frac{\sigma_n^2(fg)}{\sigma_n^2(f)} = G(f)G(g) = G(g).
\]

It turns out that under some additional assumptions imposed on functions \( f \) and \( g \), the asymptotic relation (4.5) remains also valid in the case of deterministic processes, that is, when \( G(f) = 0 \).

We are now in position to state the main results of this section.

The following theorem describes the asymptotic behavior of the ratio \( \sigma_n^2(fg)/\sigma_n^2(f) \) for the class of processes described above, and essentially states that if the spectral density \( f \) is such that the sequence \( \sigma_n(f) \) is weakly varying, and \( g \) is the spectral density of a non-deterministic process satisfying some conditions, in particular that \( fg \) is a density' as defined above, then the sequence \( \sigma_n(fg) \) and \( \sigma_n(f) \) have the same asymptotic behavior.

**Theorem 4.1.** Suppose that \( f \) is the spectral density of a deterministic process such that the sequence \( \sigma_n(f) \) is weakly varying, that is, (4.4) is satisfied. Let \( g \) be a function of the form:

\[
g(\lambda) = h(\lambda) \cdot \frac{t_1(\lambda)}{t_2(\lambda)},
\]

where \( h \in B_0^+ \) and \( t_1, t_2 \) are non-negative trigonometric polynomials, such that \( fg \) is a density'. Then \( g \) is the spectral density of a non-deterministic process and the following relation holds:

\[
\lim_{n→∞} \frac{\sigma_n^2(fg)}{\sigma_n^2(f)} = G(g) > 0,
\]

where \( G(g) \) is the geometric mean of \( g \).

In view of Remark 4.1, as a consequence of Theorem 4.1 we obtain the following result.

**Corollary 4.1.** Let the spectral density \( f \) of a deterministic process \( X(t) \) be a.e. positive, and let \( g \) be as in Theorem 4.1. Then \( g \) is the spectral density of a non-deterministic process and the relation (4.7) holds.

As an immediate consequence of Theorem 4.1 and Proposition 4.1(d), we have the following result.

**Corollary 4.2.** Let the spectral densities \( f \) and \( g \) be as in Theorem 4.1. Then the sequence \( \sigma_n(fg) \) is also weakly varying.

Taking into account that the sequence \( \{n^{-α}, \ n \in \mathbb{N}, \ α > 0\} \) is weakly varying, as an immediate consequence of Theorem 4.1, we obtain the following result.

**Corollary 4.3.** Let the spectral densities \( f \) and \( g \) be as in Theorem 4.1, and let \( \sigma_n(f) \sim cn^{-α} \) \( (c > 0, α > 0) \) as \( n → ∞ \). Then \( \sigma_n(fg) \sim cG(g)n^{-α} \), where \( G(g) \) is the geometric mean of \( g \).

The next result, which immediately follows from Theorem B and Corollary 4.3, extends Rosenblatt’s second theorem (Theorem B).

**Theorem 4.2.** Let \( f = f_a g \), where \( f_a \) is defined by (1.3) and \( g \) satisfies the assumptions of Theorem 4.1. Then

\[
\sigma_n^2(f) ∼ \frac{Γ^2((α + 1)/2)G(g)}{π2^{−α}} n^{-α} \quad \text{as} \quad n → ∞,
\]

where \( G(g) \) is the geometric mean of \( g \).
We thus obtain the same limiting behavior for $\sigma_n^2(f)$ as in the Rosenblatt’s relation (1.5) up to an additional positive factor $G(g)$.

4.3. Auxiliary lemmas

To prove Theorem 4.1, we first establish a number of lemmas.

**Lemma 4.1.** Assume that the sequence $\sigma_n(f)$ is weakly varying. Then for any non-negative trigonometric polynomial $t$ we have

$$\lim \inf_{n \to \infty} \frac{\sigma_n^2(ft)}{\sigma_n^2(f)} \geq G(t) > 0. \tag{4.8}$$

**Proof.** Let the polynomial $t$ be of degree $\nu$, and let $s_\nu(z)$ be the algebraic polynomial of degree $\nu$ from the Fejér-Riesz representation (4.2).

Let $p^*_n(z,ft)$ be the optimal polynomial of degree $n$ for $ft$ in the class $Q^*_n$ (see formula (2.18)). We now introduce a new polynomial:

$$r_{n+\nu}(z) := p^*_n(z,ft) \frac{s_\nu(z)}{s_\nu(0)}, \tag{4.9}$$

and observe that $r_{n+\nu}(z) \in Q^*_{n+\nu}$. Therefore

$$\int_{-\pi}^{\pi} |r_{n+\nu}(e^{i\lambda})|^2 f(\lambda) d\lambda \geq \int_{-\pi}^{\pi} |p^*_n(e^{i\lambda},ft)s_\nu(e^{i\lambda})|^2 f(\lambda) d\lambda, \tag{4.10}$$

where $p^*_n(z,ft)$ is the optimal polynomial of degree $n + \nu$ for $f$ in the class $Q^*_n$.

Next, we can write

$$\sigma_n^2(ft) = \int_{-\pi}^{\pi} |p^*_n(e^{i\lambda},ft)|^2 f(\lambda) d\lambda = \int_{-\pi}^{\pi} |p^*_n(e^{i\lambda},ft)s_\nu(e^{i\lambda})|^2 f(\lambda) d\lambda$$

$$= |s_\nu(0)|^2 \int_{-\pi}^{\pi} |r_{n+\nu}(e^{i\lambda})|^2 f(\lambda) d\lambda \geq |s_\nu(0)|^2 \int_{-\pi}^{\pi} |p^*_n(e^{i\lambda},f)|^2 f(\lambda) d\lambda = |s_\nu(0)|^2 \sigma_n^2(f).$$

Here the first relation follows from formula (2.18), the second from Fejér-Riesz representation (4.2), the third from (4.9), the fourth from (4.10), and the fifth from (2.18). Therefore, in view of Proposition 4.2(d), we obtain

$$\lim \inf_{n \to \infty} \frac{\sigma_n^2(ft)}{\sigma_n^2(f)} \geq |s_\nu(0)|^2 = G(t). \tag{4.11}$$

Now, taking into account (4.4) and Proposition 4.1(a), from (4.11) we obtain (4.8).

**Lemma 4.2.** Let the sequence $\sigma_n(f)$ satisfy (4.4), and let $t$ be a non-negative trigonometric polynomial such that the function $ft \in B$. Then the following inequality holds:

$$\lim \sup_{n \to \infty} \frac{\sigma_n^2(f/t)}{\sigma_n^2(f)} \leq G(1/t), \quad G(1/t) > 0. \tag{4.12}$$
Proof. Let \( s_n(z) \) be the algebraic polynomial of degree \( n \) from the Fejér-Riesz representation (4.2) for polynomial \( t \), and let \( p_n^*(z,f/t) \) be the optimal polynomial of degree \( n \) for \( f/t \) in the class \( Q_n^* \) (see formula (2.18)). For \( n > n \) we set

\[
    r_n(z) := p_{n-1}^*(z,f) \frac{s_n(z)}{s_0(0)},
\]

and observe that \( r_n(z) \in Q_n^* \). Therefore, we have

\[
    \sigma_n^2(f/t) = \int_{-\pi}^{\pi} |p_n^*(e^{i\lambda}f/t)|^2 f(\lambda)/t(\lambda) d\lambda \leq \int_{-\pi}^{\pi} |r_n(e^{i\lambda})|^2 f(\lambda)/t(\lambda) d\lambda
\]

\[
    = \frac{1}{|s_0(0)|^2} \int_{-\pi}^{\pi} |p_{n-1}^*(e^{i\lambda}f)|^2 f(\lambda) d\lambda = \frac{1}{|s_0(0)|^2} \sigma_{n-1}^2(f),
\]

which, in view of Proposition 4.2(d) and (4.1), implies that

\[
    \limsup_{n \to \infty} \frac{\sigma_n^2(f/t)}{\sigma_{n-1}^2(f)} \leq \frac{1}{|s_0(0)|^2} = G(1/t).
\]

Finally, taking into account (4.4) and Proposition 4.1(a), from (4.13) we obtain (4.12). □

In the next lemma we approximate in the space \( L^1 \) a function from the class \( B_{\infty}^+ \) by a trigonometric polynomial with special features.

**Lemma 4.3.** Let \( h \) be a function from the class \( B_{\infty}^+ \). Then for any \( \varepsilon > 0 \) a trigonometric polynomial \( t \) can be found to satisfy the following condition:

\[
    \|h - t\|_1 := \int_{-\pi}^{\pi} |h(\lambda) - t(\lambda)| d\lambda \leq \varepsilon.
\]

Moreover, if \( m \) and \( M \) are the constants from the Definition 4.2 (see (4.3)), then the polynomial \( t \) can be chosen so that for all \( \lambda \in [-\pi, \pi] \) one of the following relations is satisfied:

\[
    m - \varepsilon < t(\lambda) < h(\lambda),
\]

\[
    h(\lambda) < t(\lambda) < M + \varepsilon.
\]

**Proof.** We first prove the combination of inequalities (4.14) and (4.15).

Let \( \{\lambda_i\} (-\pi = \lambda_0 < \lambda_1 < \ldots < \lambda_k = \pi) \) be a partition of the segment \([-\pi, \pi] \), and let \( s \) be the Darboux lower sum corresponding to this partition:

\[
    s = \sum_{i=1}^{k} m_i \Delta \lambda_i, \quad m_i = \inf_{\lambda \in \Delta_i} h(\lambda), \quad \Delta \lambda_i = [\lambda_{i-1}, \lambda_i], \quad \Delta \lambda_i = \lambda_i - \lambda_{i-1}, \quad i = 1, \ldots, k.
\]

On the segment \([-\pi, \pi] \) we define a step-function \( \varphi(\lambda) \) corresponding to given partition as follows (see Figure 3):

\[
    \varphi(\lambda) := \begin{cases} 
        m_i, & \text{if } \lambda \in (\lambda_{i-1}, \lambda_i), \quad i = 1, \ldots, k - 1, \\
        \min\{m_i, \lambda_{i+1}\}, & \text{if } \lambda = \lambda_i,
    \end{cases}
\]

\[
    \min\{m_i, \lambda_k\}, \quad \text{if } \lambda = \lambda_0 \text{ or } \lambda = \lambda_k.
\]
Observe that if \( m_1 = m_k \), then the steps of the function \( \varphi_\lambda(\lambda) \) are intervals, and the number of steps is equal to \( k \).

In the case where \( m_1 \neq m_k \), one more step (the fist or the last) is added, which consists of one point (in Figure 3, this is the first step, consisting of the point with coordinates \((\lambda_0, m_k))\).

It is clear that the function \( \varphi_\lambda(\lambda) \) satisfies the following conditions:

\[
\varphi_\lambda(\lambda) \leq h(\lambda), \quad \lambda \in [-\pi, \pi] \quad \text{and} \quad \int_{-\pi}^{\pi} \varphi_\lambda(\lambda) d\lambda = s. \tag{4.17}
\]

Since the function \( h \) is integrable, for an arbitrary given \( \epsilon > 0 \) a partition of the segment \([-\pi, \pi]\) can be found so that the corresponding Darboux lower sum satisfies the condition:

\[
\int_{-\pi}^{\pi} h(\lambda) d\lambda - s = \int_{-\pi}^{\pi} [h(\lambda) - \varphi_\lambda(\lambda)] d\lambda = ||h - \varphi_\lambda||_1 < \frac{\epsilon}{3}. \tag{4.18}
\]

Now using the function \( \varphi_\lambda(\lambda) \) we construct a new function that is continuous on \([-\pi, \pi]\). To this end, we connect all the adjacent steps of the graph of \( \varphi_\lambda(\lambda) \) by slanting line segments as follows: for each partition point \( \lambda_i, i = 1, \ldots, k-1 \), at which the function \( \varphi_\lambda(\lambda) \) is discontinuous, the endpoint of the lower step of the graph with abscissa \( \lambda_i \), we connect by a line segment with some interior point of the adjacent upper step, with the abscissa \( \lambda_i^* \) satisfying the condition (see Figure 4):

\[
|\lambda_i - \lambda_i^*| < \frac{\epsilon}{3} (3M). \tag{4.19}
\]

Then, we remove the part of the upper step lying under the constructed slanting segment. The obtained polygonal line is a graph of some continuous piecewise linear function, which we denote by \( h_\lambda \). According to the construction and (4.17), this function satisfies the condition:

\[
h_\lambda(\lambda) \leq \varphi_\lambda(\lambda) \leq h(\lambda) \leq M, \quad \lambda \in [-\pi, \pi]. \tag{4.20}
\]

Taking into account that the functions \( h_\lambda(\lambda) \) and \( \varphi_\lambda(\lambda) \) coincide outside the segments \([\lambda_i, \lambda_i^*] \) (or \([\lambda_i^*, \lambda_i] \)), in view of (4.19) and (4.20), we can write

\[
||\varphi_\lambda - h_\lambda||_1 = \int_{-\pi}^{\pi} [\varphi_\lambda(\lambda) - h_\lambda(\lambda)] d\lambda = \sum_{i=1}^{k-1} \int_{\lambda_i}^{\lambda_i^*} [\varphi_\lambda(\lambda) - h_\lambda(\lambda)] d\lambda < \frac{\epsilon}{3}. \tag{4.21}
\]

Notice that the function \( h_\lambda(\lambda) \) is continuous on the segment \([-\pi, \pi]\) and satisfies the condition \( h_\lambda(-\pi) = h_\lambda(\pi) \).

Hence, according to the Weierstrass theorem (see, e.g., Grenander and Szegö, 1958, p. 15), a trigonometric polynomial \( t \) can be found so that uniformly for all \( \lambda \in [-\pi, \pi] \),

\[
-\frac{\epsilon}{12\pi} < h_\lambda(\lambda) - t(\lambda) < \frac{\epsilon}{12\pi}. \tag{4.22}
\]

Setting \( t(\lambda) := \tilde{t}(\lambda) - \frac{\epsilon}{12\pi} \), from (4.22) we get

\[
0 < h_\lambda(\lambda) - t(\lambda) < \frac{\epsilon}{6\pi}. \tag{4.23}
\]

Therefore

\[
||h_\lambda - t||_1 = \int_{-\pi}^{\pi} [h_\lambda(\lambda) - t(\lambda)] d\lambda < \frac{\epsilon}{3}. \tag{4.24}
\]
Combining the inequalities (4.18), (4.21) and (4.24), we obtain

$$
\|h - t\|_1 \leq \|h - \varphi_k\|_1 + \|\varphi_k - h_k\|_1 + \|h_k - t\|_1 \leq \epsilon,
$$

and the inequality (4.14) follows.

Now we proceed to prove the inequality (4.15). Observe first that the second inequality in (4.15) follows from the first inequality in (4.23) and (4.20). To prove the first inequality in (4.15), observe that by the construction of the function $h_k$, we have

$$
h_k(\lambda) \geq \min\{m_1, \ldots, m_k\} \geq m. \quad (4.25)
$$

Next, in view of the second inequality in (4.23), we get

$$
t(\lambda) \geq h_k(\lambda) - \frac{\epsilon}{6\pi} > h_k(\lambda) - \epsilon. \quad (4.26)
$$

Combining (4.25) and (4.26), we obtain the first inequality in (4.15).

The combination of inequalities (4.14) and (4.16) can be proved similarly with the following changes: instead of the Darboux lower sum the upper sum should be used, in the definition of function $\varphi_k$ instead of minima should be taken maxima, and in the construction of function $h_k$, the endpoint of the upper step of the function $\varphi_k$ should be connected with an interior point of the adjacent lower step.

**Lemma 4.4.** Let $h \in B_+^-$ and the sequence $\sigma_n(f)$ satisfy (4.4). Then the following relation holds:

$$
\lim_{n \to \infty} \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} = G(h) > 0.
$$

**Proof.** Observe first that together with $h$ the function $1/h$ also belongs to the class $B^-_+$:

$$
m \leq h(\lambda) \leq M \quad \text{and} \quad 1/M \leq 1/h(\lambda) \leq 1/m. \quad (4.28)
$$

By Lemma 4.3, for a given small enough $\epsilon > 0$, we can find two trigonometric polynomials $t_1$ and $t_2$ to satisfy the following conditions:

$$
\|h - t_1\|_1 < \epsilon, \quad \|1/h - t_2\|_1 < \epsilon, \quad (4.29)
$$
\[ m/2 < t_1(\lambda) < h(\lambda), \quad 1/(2M) < t_2(\lambda) < 1/h(\lambda), \tag{4.30} \]

and hence
\[ m/2 < t_1(\lambda) < h(\lambda) < 1/t_2(\lambda) < 2M. \tag{4.31} \]

Now in view of (4.31), Proposition 2.2(b), and Lemmas 4.1 and 4.2, we obtain
\[
\liminf_{n \to \infty} \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} \geq \liminf_{n \to \infty} \frac{\sigma_n^2(f_{t_1})}{\sigma_n^2(f)} \geq G(t_1),
\]

and
\[
\limsup_{n \to \infty} \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} \leq \limsup_{n \to \infty} \frac{\sigma_n^2(f_{t_2})}{\sigma_n^2(f)} \leq G(1/t_2).
\]

Therefore
\[
G(t_1) \leq \liminf_{n \to \infty} \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} \leq \limsup_{n \to \infty} \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} \leq G(1/t_2). \tag{4.32}
\]

Next, in view of the first inequality in (4.28), the last inequality in (4.31) and the second inequality in (4.29), we can write
\[
\|h - 1/t_2\|_1 = \|h/t_2(t_2 - 1/h)\|_1 \leq 2M^2\epsilon. \tag{4.33}
\]

From the first inequality in (4.29) and (4.33) we get
\[
\|t_1 - 1/t_2\|_1 \leq \|t_1 - h\|_1 + \|h - 1/t_2\|_1 \leq c(1 + 2M^2). \tag{4.34}
\]
We now can write

\[ 0 < \ln \frac{G(1/t_2)}{G(t_1)} = \ln G \left( \frac{1}{t_1 t_2} \right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln \frac{1}{t_1(\lambda) t_2(\lambda)} d\lambda \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \left( \frac{1}{t_1(\lambda) t_2(\lambda)} - 1 \right) d\lambda \]

\[ = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{t_1(\lambda)} \left( \frac{1}{t_2(\lambda)} - t_1(\lambda) \right) d\lambda \leq \frac{1}{\pi m} ||t_1 - 1/t_2||_1 \leq \frac{e}{\pi m} (1 + 2M^2). \quad (4.35) \]

Here the first relation follows from the inequality $1/t_2 > t_1$ (see (4.31)), the second from (4.1), the third from (2.4), the fourth from the inequality $\ln x \leq x - 1$ ($x > 0$), the sixth from the first inequality in (4.30), and the seventh from (4.34).

Thus, the quantities $G(t_1)$ and $G(1/t_2)$ can be made arbitrarily close. On the other hand, in view of Proposition 4.2(b) and (4.31), we have

\[ G(t_1) \leq G(h) \leq G(1/t_2). \quad (4.36) \]

Finally, from (4.32), (4.35) and (4.36) we obtain (4.27). The inequality $G(h) > 0$ follows from (4.31).

Taking into account Proposition 4.1(d), from Lemma 4.4 we obtain the following result.

**Corollary 4.4.** If the sequence $\sigma_n(f)$ is weakly varying and $h \in B^+$, then the sequence $\sigma_n(fh)$ is also weakly varying.

**Lemma 4.5.** Let the sequence $\sigma_n(f)$ be weakly varying, and let $h \in B^-$. Then

\[ \limsup_{n \to \infty} \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} \leq G(h). \quad (4.37) \]

**Proof.** Observe that the function $h_\varepsilon = h + \varepsilon$ belongs to the class $B^+_\varepsilon$, and $h_\varepsilon \to h$ as $\varepsilon \to 0$. Then we have the asymptotic relation (see, Grenander and Szegö, 1958, p. 46):

\[ \lim_{\varepsilon \to 0} G(h_\varepsilon) = G(h). \quad (4.38) \]

Hence, using Proposition 2.2(b) and Lemma 4.4, we obtain

\[ \lim_{n \to \infty} \sup \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} \leq \lim_{n \to \infty} \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} = G(h_\varepsilon). \]

Passing to the limit as $\varepsilon \to 0$, and taking into account (4.38), we obtain (4.37).

As an immediate consequence of Lemma 4.5, we have the following result.

**Corollary 4.5.** Let the sequence $\sigma_n(f)$ be weakly varying, and let $g \in B^-$ with $G(g) = 0$. Then $\sigma_n(fg) = o(\sigma_n(f))$ as $n \to \infty$.

Thus, multiplying singular spectral densities we obtain a spectral density with higher ‘order of singularity’.

**Lemma 4.6.** Let the sequence $\sigma_n(f)$ be weakly varying, and let $h \in B_+$. Then

\[ \liminf_{n \to \infty} \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} \geq G(h). \quad (4.39) \]
Proof. Let \( h_l \) denote the truncation of \( h \) at the level \( l \in \mathbb{N} \):

\[
 h_l(\lambda) = \begin{cases} 
 h(\lambda), & h(\lambda) \leq l \\
 l, & h(\lambda) > l.
\end{cases}
\]

Then by the monotone convergence theorem of Beppo Levi (see, e.g., Bogachev, 2007, Theorem 2.8.2), we have

\[
 \lim_{l \to \infty} G(h_l) = G(h). \tag{4.40}
\]

Next, since \( h_k \leq h \), in view of Proposition 2.2(b) and Lemma 4.4, we get

\[
 \liminf_{n \to \infty} \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} \geq \lim_{n \to \infty} \frac{\sigma_n^2(fh_l)}{\sigma_n^2(f)} = G(h_l).
\]

Hence passing to the limit as \( l \to \infty \), and taking into account (4.40) we obtain (4.39). \( \blacksquare \)

As an immediate consequence of Lemma 4.6, we have the following result.

**Corollary 4.6.** Let the sequence \( \sigma_n(f) \) be weakly varying, \( g \in B^+ \) with \( G(g) = \infty \), and let \( fg \in B \). Then \( \sigma_n(f) = o(\sigma_n(fg)) \) as \( n \to \infty \).

### 4.4. Proof of Theorem 4.1

In this subsection we prove the main result of this section – Theorem 4.1.

**Proof of Theorem 4.1.** We have

\[
 \frac{\sigma_n^2(fg)}{\sigma_n^2(f)} = \frac{\sigma_n^2(fh/t_1)}{\sigma_n^2(f)} = \frac{\sigma_n^2(fh/t_1)}{\sigma_n^2(fh)} \cdot \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} \cdot \frac{\sigma_n^2(f)}{\sigma_n^2(fg)}.{ \tag{4.41}}
\]

Next, by Lemma 4.4 we have

\[
 \lim_{n \to \infty} \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} = G(h) > 0.
\]

This, in view of Corollary 4.4, implies that the sequence \( \sigma_n^2(fh) \) is also weakly varying. Therefore, by Lemma 4.1, we have

\[
 \liminf_{n \to \infty} \frac{\sigma_n^2(fh_1)}{\sigma_n^2(fh)} \geq G(t_1).
\]

On the other hand, since \( t_1 \in B^- \), then according to Lemma 4.5, we get

\[
 \limsup_{n \to \infty} \frac{\sigma_n^2(fh_1)}{\sigma_n^2(fh)} \leq G(t_1).
\]

Therefore

\[
 \lim_{n \to \infty} \frac{\sigma_n^2(fh_1)}{\sigma_n^2(fh)} = G(t_1) > 0. \tag{4.43}
\]
This implies that the sequence $\sigma_n^2(fht_1)$ is also weakly varying. Hence we can apply Lemma 4.2, to obtain

$$\limsup_{n \to \infty} \frac{\sigma_n^2(fht_1/t_2)}{\sigma_n^2(fht_1)} \leq G(1/t_2).$$

Next, it is easy to see that $1/t_2 \in B_+$. Hence, according to Lemma 4.6, we obtain

$$\liminf_{n \to \infty} \frac{\sigma_n^2(fht_1/t_2)}{\sigma_n^2(fht_1)} \geq G(1/t_2).$$

Therefore

$$\lim_{n \to \infty} \frac{\sigma_n^2(fht_1/t_2)}{\sigma_n^2(fht_1)} = G(1/t_2). \quad (4.44)$$

Finally, combining the relations (4.41) - (4.44), we obtain

$$\lim_{n \to \infty} \frac{\sigma_n^2(fg)}{\sigma_n^2(f)} = G(1/t_2)G(t_1)G(h) = G(ht_1/t_2) = G(g) > 0.$$  

Theorem 4.1 is proved.

### 4.5. Examples

We discuss examples demonstrating the result obtained in Theorem 4.1. In these examples we assume that \{X(t), t \in \mathbb{Z}\} is a stationary deterministic process with a spectral density $f$ satisfying the conditions of Theorem 4.1, and the function $g$ is given by formula (4.6). To compute the geometric means we use the properties stated in Proposition 4.2(a).

**Example 4.1.** Let the function $g(\lambda)$ be as in (4.6) with $h(\lambda) = c > 0$ and $t_1(\lambda) = t_2(\lambda) = 1$, that is, $g(\lambda) = c > 0$. Then for the geometric mean $G(g)$ we have

$$G(g) = G(c) = c,$$  

and in view of (4.7), we get

$$\lim_{n \to \infty} \frac{\sigma_n^2(fg)}{\sigma_n^2(f)} = G(g) = c.$$

Thus, multiplying the spectral density $f$ by a constant $c > 0$ multiplies the prediction error by $c$.

**Example 4.2.** Let the function $g$ be as in (4.6) with $h(\lambda) = e^{\phi(\lambda)}$, where $\phi(\lambda)$ is an arbitrary odd function, and let $t_1(\lambda) = t_2(\lambda) = 1$, that is, $g(\lambda) = e^{\phi(\lambda)}$. Then for the geometric mean $G(g)$ we have

$$G(g) = G(e^{\phi(\lambda)}) = \exp \left\{ \frac{1}{2\pi} \int_{-\pi}^\pi \ln g(\lambda) \, d\lambda \right\} = \exp \left\{ \frac{1}{2\pi} \int_{-\pi}^\pi \phi(\lambda) \, d\lambda \right\} = e^0 = 1,$$  

and in view of (4.7), we get

$$\lim_{n \to \infty} \frac{\sigma_n^2(fg)}{\sigma_n^2(f)} = G(g) = 1.$$
Thus, multiplying the spectral density $f$ by the function $e^{\varphi(\lambda)}$ with odd $\varphi(\lambda)$ does not change the asymptotic behavior of the prediction error.

**Example 4.3.** Let the function $g$ be as in (4.6) with $h(\lambda) = \lambda^2 + 1$ and $t_1(\lambda) = t_2(\lambda) = 1$, that is, $g(\lambda) = \lambda^2 + 1$. Then for the geometric mean $G(g)$ by direct calculation we obtain

$$G(g) = \exp\left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln(\lambda^2 + 1) \, d\lambda \right\} = \exp\{\ln(1 + \pi^2) - 2 + \frac{2}{\pi} \arctan\pi\} \approx 3.3,$$

and in view of (4.7), we get

$$\lim_{n \to \infty} \frac{\sigma_n^2(fg)}{\sigma_n^2(f)} = G(g) = \exp\{\ln(1 + \pi^2) - 2 + \frac{2}{\pi} \arctan\pi\} \approx 3.3.$$  

Thus, multiplying the spectral density $f$ by the function $\lambda^2 + 1$ multiplies the prediction error approximately by 3.3.

**Example 4.4.** Let the function $g$ be as in (4.6) with $h(\lambda) = t_2(\lambda) = 1$, and $t_1(\lambda) = \sin^{2k}(\lambda - \lambda_0)$, where $k \in \mathbb{N}$ and $\lambda_0$ is an arbitrary point from $[-\pi, \pi]$, that is, $g(\lambda) = \sin^{2k}(\lambda - \lambda_0)$. To compute the geometric mean $G(g)$, we first find the algebraic polynomial $s_2(z)$ in the Fejér-Riesz representation (4.2) of the non-negative trigonometric polynomial $\sin^2(\lambda - \lambda_0)$ of degree 2. For any $\lambda_0 \in [-\pi, \pi]$ we have

$$\sin^2(\lambda - \lambda_0) = |\sin(\lambda - \lambda_0)|^2 = \left| \frac{1}{2} (e^{2i(\lambda - \lambda_0)} - 1) \right|^2 = |s_2(e^{i\lambda})|^2,$$

where

$$s_2(z) = \frac{1}{2} (e^{-2i\lambda_0} - z^2).$$

Therefore, by Proposition 4.2(d) and (4.48) we have

$$G(\sin^2(\lambda - \lambda_0)) = |s_2(0)|^2 = \left( \frac{1}{2} \right)^2 = \frac{1}{4}.$$  

(4.49)

Now, in view of Proposition 4.2(a) and (4.49), for the geometric mean of $g(\lambda) = t_1(\lambda) = \sin^{2k}(\lambda - \lambda_0)$ ($k \in \mathbb{N}$), we obtain

$$G(g) = G(\sin^{2k}(\lambda - \lambda_0)) = G^k(\sin^2(\lambda - \lambda_0)) = 4^{-k},$$

(4.50)

and in view of (4.7), we get

$$\lim_{n \to \infty} \frac{\sigma_n^2(fg)}{\sigma_n^2(f)} = G(g) = \frac{1}{4^k}.$$  

Thus, multiplying the spectral density $f$ by the non-negative trigonometric polynomial $\sin^{2k}(\lambda - \lambda_0)$ of degree $2k$ ($k \in \mathbb{N}$), yields a $4^k$-fold asymptotic reduction of the prediction error.

**Example 4.5.** Let the function $g$ be as in (4.6) with $h(\lambda) = t_1(\lambda) = 1$, and $t_2(\lambda) = \sin^{2l}(\lambda - \lambda_0)$, where $l \in \mathbb{N}$ and $\lambda_0$ is an arbitrary point from $[-\pi, \pi]$, that is, $g(\lambda) = \sin^{2l}(\lambda - \lambda_0)$. Then, in view of the third equality in (4.1) and (4.50) for the geometric mean $G(g)$ we have

$$G(g) = G(\sin^{2l}(\lambda - \lambda_0)) = G^{-1}(\sin^{2l}(\lambda - \lambda_0)) = 4^l.$$  

(4.51)
and in view of (4.7), we get

\[
\lim_{n\to\infty} \frac{\sigma_n^2(fg)}{\sigma_n^2(f)} = G(g) = 4^l.
\]

Thus, dividing the spectral density \( f \) by the non-negative trigonometric polynomial \( \sin^{2l}(\lambda - \lambda_0) \) of degree \( 2l \) \((l \in \mathbb{N})\), yields a \( 4^l \)-fold asymptotic increase of the prediction error.

Notice that the values of the geometric mean \( G(g) \) obtained in (4.50) and (4.51) do not depend on the choice of the point \( \lambda_0 \in [-\pi, \pi] \).

Putting together Examples 4.1–4.5 and using Proposition 4.2(a) we have the following summary example.

**Example 4.6.** Let \( \{X(t), t \in \mathbb{Z}\} \) be a stationary deterministic process with a spectral density \( f \) satisfying the conditions of Theorem 4.1. Let \( h(\lambda) = ce^{\varphi(\lambda)}(\lambda^2 + 1) \), \( t_1(\lambda) = \sin^{2k}(\lambda - \lambda_1) \) and \( t_2(\lambda) = \sin^{2l}(\lambda - \lambda_2) \), where \( c \) is an arbitrary positive constant, \( \varphi(\lambda) \) is an arbitrary odd function and \( \lambda_1, \lambda_2 \) are arbitrary points from \( [-\pi, \pi] \). Let the function \( g \) be defined as in (4.6), that is,

\[
g(\lambda) = h(\lambda) \cdot \frac{t_1(\lambda)}{t_2(\lambda)} = ce^{\varphi(\lambda)}(\lambda^2 + 1) \frac{\sin^{2k}(\lambda - \lambda_1)}{\sin^{2l}(\lambda - \lambda_2)}. \tag{4.52}
\]

Then, in view of Proposition 4.2(a) and relations (4.45)–(4.47) and (4.50)–(4.52), we have

\[
G(g) = G(h) \frac{G(t_1)}{G(t_2)} = G(c)G(e^{\varphi})G(\lambda^2 + 1)G(\sin^{2k}(\lambda - \lambda_1))G(\sin^{-2l}(\lambda - \lambda_2))
\]

\[
= (c)(1)\exp\{\ln(1 + \pi^2) - 2 + \frac{2}{\pi}\arctan \pi\}(4^{-k})(4^l) \approx 3.3c4^{l-k}, \tag{4.53}
\]

and in view of (4.7) and (4.53), we get

\[
\lim_{n\to\infty} \frac{\sigma_n^2(fg)}{\sigma_n^2(f)} = G(g) \approx 3.3c4^{l-k}.
\]

**ACKNOWLEDGEMENTS**

The authors would like to thank the referees for their constructive comments and suggestions. The first two authors are grateful to their supervisor Professor I.A. Ibragimov for introducing them to this research area. Murad S. Taqqu was supported in part by a Simons Foundation grant #569118 at Boston University.

**DATA AVAILABILITY STATEMENT**

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

**REFERENCES**

Ahlfors LV. 1979. *Complex Analysis: An Introduction to the Theory of Analytic Functions of One Complex Variable*, 3rd ed., McGraw-Hill: New York.

Baxter G. 1962. An asymptotic result for the finite predictor. *Mathematica Scandinavica* 10:137–144.

Babayan NM. 1984. On the asymptotic behavior of prediction error. *Journal of Soviet Mathematics* 27(6):3170–3181.

Babayan NM. 1985. On asymptotic behavior of the prediction error in the singular case. *Theory of Probability and its Applications* 29(1):147–150.

Bello MH, López GL. 1998. Ratio and relative asymptotics of polynomials orthogonal on an arc of the unit circle. *Journal of Approximation Theory* 92:216–244.
Bingham NH. 2012. Szegő’s theorem and its probabilistic descendants. *Probability Surveys* 9:287–324.

Bogachev V. 2007. *Measure Theory*, Vol. I, Springer: Berlin.

Brockwell PJ, Davis RA. 1991. *Time Series: Theory and Methods*, 2nd ed., Springer-Verlag: New York.

Davison LD. 1965. Prediction of time series from finite past. *Journal of the Society for Industrial and Applied Mathematics* 13(3):819–826.

Dette H, Studden WJ. 1997. *The Theory of Canonical Moments with Applications in Statistics, Probability, and Analysis*, John Wiley & Sons: New York.

Devinaiz A. 1964. Asymptotic estimates for the finite predictor. *Mathematica Scandinavica* 15:111–120.

Fekete M. 1930. Über den transfiniten durchmesser ebener punktmengen. *Zweite Mitteilung, Mathematische Zeitschrift* 32:215–221.

Fortus MI. 1990. Prediction of a stationary time series with the spectrum vanishing on an interval. *Akademiia Nauk SSSR, Izvestiiia, Fizika Atmosfery i Okeana* 26:1267–1274.

Geronimus YL. 1948. Polynomials orthogonal on circle and their applications. *Zapiski Nauchno-Issledovatelskogo Instituta Matematiki i Mechaniki, Harkovskogo Matematicheskogo Obshhestva* 4(19):35–120.

Ginovian MS. 1999. Asymptotic behavior of the prediction error for stationary random sequences. *Journal of Contemporary Mathematical Analysis* 4(1):14–33.

Golinskii BL. 1974. On asymptotic behavior of the prediction error. *Theory of Probability & Its Applications* 19(4):724–739.

Golinskii BL, Ibragimov IA. 1971. On G. Szegő limit theorem. *Izvestiya Akademii Nauk Armianskoy SSR, Matematika* 35:408–427.

Goluzin GM. 1969. *Geometric Theory of Functions of a Complex Variable*, American Mathematical Society: Providence, RI.

Grenander U. 1941a. Interpolation and Extrapolation of stationary random sequences. *Annali di Matematica Pura ed Applicata* 51:107–138.

Helson H, Szegő G. 1960. A problem in prediction theory. *Annali di Matematica Pura ed Applicata* 51:107–138.

Ibragimov IA. 1964. On asymptotic behavior of the prediction error. *Theory of Probability & Its Applications* 9(4):695–703.

Ibragimov IA, Solev VN. 1968. The asymptotic behavior of the prediction error of a stationary sequence with the spectral density function of a special form. *Theory of Probability & Its Applications* 13(4):746–750.

Inoue A. 2002. Asymptotic behavior for partial autocorrelation functions of fractional ARIMA processes. *The Annals of Applied Probability* 12(4):1471–1491.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.

Kolmogorov AN. 1941b. Interpolation and Extrapolation of stationary random sequences. *Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya* 5:3–14.