Toeplitz Eigenvalues for Radon Measures

E. E. Tyrtyshnikov and N. L. Zamarashkin

Preprint No. 9
Saarbrücken 2000
Toeplitz Eigenvalues for Radon Measures

E. E. Tyrtyshnikov  
Institute of Numerical Mathematics  
Russian Academy of Sciences  
Gubkina 8  
Moscow 117333, Russia  
E-Mail: tee@inm.ras.ru

N. L. Zamarashkin  
Institute of Numerical Mathematics  
Russian Academy of Sciences  
Gubkina 8  
Moscow 117333, Russia  
E-Mail: kolya@bach.inm.ras.ru

submitted: February 15, 2000

Preprint No. 9  
Saarbrücken 2000
Abstract

It is well known that for Toeplitz matrices generated by a “sufficiently smooth” real-valued symbol, the eigenvalues behave asymptotically as the values of the symbol on uniform meshes while the singular values, even for complex-valued functions, do as those values in modulus. These facts are expressed analytically by the Szegö and Szegö-like formulas, and, as is proved recently, the “smoothness” assumptions are as mild as those of $L_1$. In this paper, it is shown that the Szegö-like formulas hold true even for Toeplitz matrices generated by the so-called Radon measures.

Key-words: Toeplitz matrices, eigenvalues, singular values, Szegö formulas, Radon measures.


Acknowledgement: This work was supported by the Russian Fund of Basic Research under Grant 99-01-00017 and completed while the first author was a guest professor at the University of Saarland, Germany, under support of DAAD.
1 Introduction

We consider a sequence of Toeplitz matrices

\[ A_n = [a_{kl}], \quad a_{kl} = a_{k-l}, \quad 0 \leq k, l \leq n - 1, \]

constructed from the coefficients of a formal Fourier series

\[ f(x) \sim \sum_{k=-\infty}^{\infty} a_k e^{ikx}, \]

and will be interested in the asymptotic behavior of their eigenvalues \( \lambda_i(A_n) \) (in the Hermitian case) and singular values \( \sigma_i(A_n) \) (in the non-Hermitian case) as \( n \to \infty \). Due to G. Szegö [5] and successive works [1, 6, 8, 9, 11] we enjoy the following beautiful formula:

\[ \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} F(\lambda_i(A_n)) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(f(x)) \, dx, \]

which is valid for any test function \( F(x) \) from a suitable set \( \mathcal{F} \).

G. Szegö proved (3) for a real-valued \( f \in L_\infty \) and \( \mathcal{F} \) comprising all continuous functions on the interval \( \text{ess inf} f, \text{ess sup} f \). For \( f \in L_\infty \) this interval contains all \( \lambda_i(A_n) \). Since this is not the case for \( f \in L_p \) with \( p < \infty \), it was proposed in [8] to take up as \( \mathcal{F} \) all functions uniformly bounded and uniformly continuous for \( -\infty < x < \infty \); a bit more restrictive choice for \( \mathcal{F} \) might be all continuous functions with bounded support [8]. For both cases, the same formula (3) holds true for \( f \in L_2 \) [8, 9] and even for \( f \in L_1 \) [11].

If \( f \) is not necessarily real-valued, under the same "smoothness" assumptions on \( f \) and the same \( \mathcal{F} \) we have quite a similar formula for the singular values:

\[ \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} F(\sigma_i(A_n)) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(|f(x)|) \, dx. \]

An important and somewhat expected difference is that the eigenvalues behave as the values of \( f(x) \) (when \( f \) is real-valued and in some special cases of complex-valued \( f \) [10]) while the singular values do as the same values in modulus. Formula (4) was proposed by S. Parter [6] and proved first for a specific subclass of \( L_\infty \); then it was extended to the whole of \( L_\infty \) [1] and further to \( L_2 \) [8, 9] and even to \( L_1 \) [11].

However, we have long suspected that \( L_1 \) is still not the ultimate extension. For example, let

\[ a_k = 1, \quad k = 0, \pm 1, \pm 2, \ldots. \]
It this case \( f(x) \) (usually called a symbol or generating function) is not a function in the classical sense (it is a multiple of the Dirac delta function). Despite this, the eigenvalues of \( A_n = A_n(f) \) are easy to find explicitly:

\[
\lambda_1 = n; \quad \lambda_k = 0, \quad k = 2, \ldots, n.
\]

Therefore, the Szegő formula (3) gives the true asymptotic distribution even for this case if only we set \( f(x) \) to zero in the integrand. Thus, we obtain

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} F(\lambda_i(A_n)) = F(0)
\]

for any \( F \in \mathcal{F} \). From now onwards, let \( \mathcal{F} \) be the set of all uniformly bounded and uniformly continuous functions.

If \( A_n \) is an arbitrary sequence of matrices satisfying (5), we say that the eigenvalues of \( A_n \) have a cluster at zero. An equivalent definition reads [8, 9]: zero is a cluster for \( \lambda_i(A_n) \) if for any \( \varepsilon > 0 \) the number \( \gamma_n(\varepsilon) \) of those \( i \) from 1 to \( n \) for which \( |\lambda_i(A_n)| > \varepsilon \) is \( o(n) \) (that is, \( \frac{\gamma_n(\varepsilon)}{n} \to 0 \)). To denote the fact, we write \( \lambda(A_n) \sim 0 \). If (5) is fulfilled for the singular values, we write \( \sigma(A_n) \sim 0 \).

The above observation might suggest that we could have a cluster at zero in all cases when \( f \) is not a function modulo a function (that is, after subtracting any function from an appropriate space). Of course, it gives just a flavour of where we should look for a rigorous formulation. The purpose of this paper is to propose one by making a step from functions to “non-functions”.

Let us assume that the Fourier coefficients are the values of a linear bounded functional \( \mathcal{T}(\phi) \) on the space of continuous functions \( \phi \) on the basic closed interval \( \Pi = [-\pi, \pi] \). Such a functional is called a Radon measure [4]. It is well-known that there exists a bounded-variation function \( \mu \) on \( \Pi \) such that

\[
\mathcal{T}(\phi) = \int_{-\pi}^{\pi} \phi(x) \, d\mu(x),
\]

where the integral is understood in the sense of Stiltjes. Thus, it is \( \mu \) that can be viewed now as a symbol.

We know that any function \( \mu \) of bounded variation is a sum of three functions (see, for example, [5])

\[
\mu = \mu_a + \mu_s + \mu_j,
\]

where
where $\mu_a$ is an absolutely continuous function, $\mu_s$ is the so-called singular function (a continuous function with zero derivative at almost every point), and $\mu_j$ is a function of jumps. All three components are of bounded variation as well. The derivative $f \equiv \mu'_a$ of $\mu_a$ exists almost everywhere in the Lebesgue sense and belongs to $L_1$. The derivatives of $\mu_j$ and $\mu_s$ are almost everywhere equal to zero. Consequently, $\mu' = \mu'_a$ almost everywhere. Recall that, by definition, $\mu_j$ is a sum of a countable number of jumps:

$$
\mu_j(x) = \sum_{x<s_k} h^-_k + \sum_{x>s_k} h^+_k,
$$

where

$$
\sum_{k=1}^{\infty} |h_k^\pm| < \infty.
$$

(The values at $x = s_k$ do not count.) Note that $f = \mu'_a$ is determined uniquely as a function from $L_1$.

Our main result is the following theorem.

**Theorem 1.1** Suppose that $\mu$ is a function of bounded variation on $\Pi$, and $f \equiv \mu' \in L_1$ is its derivative. Let $A_n$ be Toeplitz matrices of the form (1) where

$$
a_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ikx} d\mu(x). \quad (8)
$$

Then, for any $F \in \mathcal{F}$, the relation (3) holds true, provided that $\mu$ is real-valued, and (4) holds true in case $\mu$ might be complex-valued. The test-function set $\mathcal{F}$ consists of all uniformly bounded and uniformly continuous functions.

In other words, in the real-valued case the eigenvalues of $A_n$ are distributed as the values of $f(x)$, and in the complex-valued case the singular values of $A_n$ are distributed as the values of $|f(x)|$. Compared to the previous knowledge, a new message is that $f$ in the Szego-like formulas is not a generating function for $A_n$. It is the derivative of the Radon-measure symbol $\mu$, and it is $\mu$ that generates $A_n$. The Fourier series (2) is not associated with any function in the classical sense. However, at least in the Radon-measure case, it can be juxtaposed to some function from $L_1$ that describes the spectral distributions precisely by the Szego-like formulas.
2 Preliminaries

Given a matrix sequence $A_n$, we try to associate it with another sequence $B_n$ for which (3) or (4) is easier to prove and which is close, in a certain sense, to $A_n$. By definition, two sequences of $n$-tuples $\{\alpha_i^{(n)}\}_{i=1}^n$ and $\{\beta_i^{(n)}\}_{i=1}^n$ are equally distributed if, for any $F \in \mathcal{F}$,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n \left( F(\alpha_i^{(n)}) - F(\beta_i^{(n)}) \right) = 0. \quad (9)$$

We capitalize on the following lemma [9].

**Lemma 2.1** Let $G(x)$ be a continuous, nonnegative, and strictly increasing function for $x \geq 0$, and $G(0) = 0$. Let $c_1$ and $c_2$ be positive constants.

Given two matrix sequences $A_n$ and $B_n$, assume that for any $\varepsilon > 0$, there exists $N$ such that for all $n \geq N$, the difference between $A_n$ and $B_n$ can be split

$$A_n - B_n = E_n + R_n \quad (10)$$

so that

$$\sum_{i=1}^n G(\sigma_i(E_n)) \leq c_1 \varepsilon n \quad (11)$$

and

$$\text{rank}R_n \leq c_2 \varepsilon n. \quad (12)$$

Then the singular values of $A_n$ and $B_n$ are equally distributed.

If $A_n$ and $B_n$ are Hermitian, assume that $E_n$ and $R_n$ are Hermitian and, instead of (11), that

$$\sum_{i=1}^n G(|\lambda_i(E_n)|) \leq c_1 \varepsilon n. \quad (13)$$

Then, the eigenvalues of $A_n$ and $B_n$ are equally distributed as well.

An important example is $G(x) = x^2$; in this case (11) is equivalent to the Frobenius-norm (Schatten 2-norm) estimate

$$\|E_n\|_F^2 \leq c_1 \varepsilon n. \quad (14)$$
Another useful example is $G(x) = x$; in this case (11) is equivalent to the Schatten trace-norm estimate (see [2, 7])

$$\|E_n\|_{tr} \equiv \sum_{i=1}^{n} \sigma_i(E_n) \leq c_1 \varepsilon n. \quad (15)$$

Once having (14) or (15), from the Weyl inequalities we infer that (13) is also valid (for the respective $G(x)$).

The main vehicle to relate the eigenvalues with the symbol $\mu$ is the next observation. Consider the following one-to-one correspondence between vectors and polynomials:

$$p = \begin{bmatrix} p_0 \\ \vdots \\ p_{n-1} \end{bmatrix} \leftrightarrow p(x) = \sum_{i=0}^{n-1} p_i x^i. \quad (16)$$

If $A_n$ are Toeplitz matrices with the elements $a_k$ of the form (8), then

$$(A_n p, p) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |p(e^{ix})|^2 \, d\mu(x).$$

We take advantage of special probe vectors $p$ for which the “kernel” $|p(e^{ix})|^2$ can be expressed explicitly. As in [11], these are the columns of the Discrete Fourier Transform matrix:

$$p_k^{(n)} = \frac{1}{\sqrt{n}} \begin{bmatrix} e^{-i\frac{2\pi}{n} k 0} \\ \vdots \\ e^{-i\frac{2\pi}{n} k (n-1)} \end{bmatrix}, \quad k = 0, \ldots, n - 1. \quad (17)$$

On having made this choice, we obtain

$$(A_n p_k^{(n)}, p_k^{(n)}) = \int_{-\pi}^{\pi} \Phi_n(k, x) \, d\mu(x), \quad \Phi_n(k, x) = \frac{1}{2\pi} |p_k^{(n)}(e^{ix})|^2. \quad (18)$$

A direct calculation yields [11]

$$\Phi_n(k, x) = \frac{\sin^2(H_n(k, x) n)}{2\pi n \sin^2 \theta_n(k, x)}, \quad (19)$$

where

$$H_n(k, x) = \frac{2\pi k + xn}{2n}.$$
Lemma 2.2 Let $0 < \delta < \pi$. Then, for any $n$,

$$\max_{-\delta \leq x \leq \delta} \Phi_n(k, x) \leq \frac{c_1(\delta)}{n}, \quad c_1(\delta) = \frac{1}{2\pi \sin^2 \frac{\delta}{2}},$$

(20)

for all $k \in \{0, \ldots, n-1\}$ except for at most $c_2\delta n + 1$ indices with $c_2 = 2/\pi$.

Proof. Denote by $\nu_n$ the number of $k \in \{0, \ldots, n-1\}$ for which (20) does not hold, and let $\tau_n$ be the number of those $k$ for which the denominator in (19) is strictly less than $n/c_1(\delta)$. That means that

$$\min_{-\delta < x < \delta} \sin H_n(k, x) < \sin \frac{\delta}{2}.$$  

(21)

Since (20) takes place whenever (21) does not, we conclude that $\nu_n \leq \tau_n$.

To estimate $\tau_n$, assume by the moment that $\delta \leq \pi/2$. Then (21) amounts to the claim that

$$\frac{\pi m - \delta}{2} < -\frac{\pi k}{n} + \frac{x}{2} < \frac{\delta}{2} + \pi m$$

for some integer $m$ and $x \in [-\delta, \delta]$. The latter implies that

$$\frac{\pi m - \delta}{2} < -\frac{\pi k}{n} < \delta + \pi m.$$  

Since $0 \leq k \leq n-1$, it is possible only when $m = 0$ or $m = -1$. Thus, we can estimate $\tau_n$ by counting how many indices $k$ satisfy

$$1 - \frac{\delta}{\pi} < \frac{k}{n} < 1 \quad \text{or} \quad 0 \leq \frac{k}{n} < \frac{\delta}{\pi}.$$  

Thus, $\tau_n < \frac{2\delta}{\pi} n + 1$. The same estimate stands also when $\pi/2 < \delta \leq \pi$. \hfill \(\square\)

3 Main results

We call a Radon measure nonnegative if the corresponding symbol $\mu$ is a monotone nondecreasing function. The general case can be reduced to those because an arbitrary function of bounded variation is a difference of two monotone nondecreasing functions.

For a Radon measure, a point is called essential if the full variation in any its neighbourhood is nonzero. The closure of the set of all essential points is said to be a support of this measure. We are going to show that a “small” support for a nonnegative measure means that the eigenvalues of the corresponding Toeplitz matrices are “almost clustered” at zero.
Lemma 3.1  Consider a nonnegative Radon measure with symbol $\mu$, and assume that it is supported on a closed interval of length $\delta$. Then the Toeplitz matrices $A_n = A_n(\mu)$ generated by $\mu$ can be split

$$A_n = A_{1n} + A_{2n}$$

so that

$$\sigma(A_{1n}) \sim 0$$

and, for some $c > 0$ independent of $\delta$ and $n$,

$$\text{rank} A_{2n} \leq c \delta n$$

for all sufficiently large $n$.

Proof. Assume, first, that the interval of length $\delta$ is inside $[-\delta, \delta]$. Set $P_n = [P_{1n}, P_{2n}]$, where $P_{1n}$ contains all the columns $p_k^{(n)}$ for which (20) is fulfilled, all other $p_k^{(n)}$ being relegated to $P_{2n}$. Then

$$A_{1n} = P_n \begin{bmatrix} P_{1n}^* A_{1n} P_{1n} & 0 \\ 0 & 0 \end{bmatrix} P_n^*, \quad A_{2n} = P_n \begin{bmatrix} 0 & * \\ * & * \end{bmatrix} P_n^*.$$

From (16) and thanks to the nonnegativeness of the Radon measure, $A_n$ are Hermitian nonnegative matrices. Obviously, $A_{1n}$ is also a Hermitian nonnegative matrix. Hence,

$$\sum_{k=1}^{n} \sigma_k(A_{1n}) = \text{trace } A_{1n} = \text{trace } P_{1n}^* A_{1n} P_{1n},$$

and by Lemma 2.2,

$$\text{trace } P_{1n}^* A_{1n} P_{1n} \leq c_1(\delta) \int_{-\pi}^{\pi} d\mu = o(n).$$

Consequently, $\sigma(A_{1n}) \sim 0$ and, from Lemma 2.2, the rank of $A_{2n}$ does not exceed $(c_2 + 1)\delta n$ for all sufficiently large $n$.

If $I$ is an arbitrarily located interval of length $\delta$, then we choose a shift $s$ so that $s + I \subset [-\delta, \delta]$. Thus, the said-above splitting is taken for granted for Toeplitz matrices $A_n$ generated by $\mu(s + x)$. As is readily seen from (8),

$$\tilde{A}_n = D_n^* A_n D_n, \quad \text{where } D_n = \begin{bmatrix} e^{i\sigma 0} \\ & \ddots \\ & & e^{i\sigma (n-1)} \end{bmatrix}$$

8
is a unitary diagonal matrix. Having had \( \tilde{A}_n = \tilde{A}_{1n} + \tilde{A}_{2n} \), now we set

\[
A_{1n} = D_n \tilde{A}_{1n} D_n^*, \quad A_{2n} = D_n \tilde{A}_{2n} D_n^*,
\]

which completes the proof. \( \square \)

**Lemma 3.2** Assume that Toeplitz matrices \( A_n \) are generated by a nonnegative Radon measure with a compact support of the Lebesgue measure \( \delta \). Then \( A_n = A_{1n} + A_{2n} \) so that (23) and (24) are valid.

**Proof.** Since the support of the Radon measure is a compact set, it can be covered by finitely many (say, \( m \)) open intervals \( (a_i, b_i) \) so that

\[
\sum_{i=1}^{m} (b_i - a_i) < 2\delta.
\]

Let \( \mu_i = \mu \) on \( [a_i, b_i] \) and an appropriate constant elsewhere so that \( \mu = \sum_{i=1}^{m} \mu_i \).

Now we obtain

\[
A_n(\mu) = \sum_{i=1}^{n} A_n(\mu_i)
\]

and apply Lemma 3.1 to every \( A_n(\mu_i) \). The claim follows immediately. \( \square \)

Denote by \( \text{var} \mu \) the full variation of \( \mu \). By \( \text{meas} \ \text{supp} \mu \), it is meant the Lebesgue measure of the support of \( \mu \). The next lemma is a rather well-known assertion [5] (we give a bit more straightforward proof).

**Lemma 3.3** Let \( \mu \) be a singular function or function of jumps coupled with a nonnegative Radon measure. Then for any \( \varepsilon > 0 \), \( \mu \) can be split

\[
\mu = \mu_1 + \mu_2
\]

so that

\[
\text{meas} \ \text{supp} \mu_1 \leq \varepsilon
\]

and

\[
\text{var} \mu_2 \leq \varepsilon.
\]

Moreover, the support of \( \mu_1 \) is a union of finitely many closed intervals.
Proof. We know that $\mu' = 0$ almost everywhere. Therefore, the set of those $x$ where $\mu'(x) > \varepsilon/2$ or does not exist is of zero Lebesgue measure. Thus, for any $\delta > 0$, it can be covered by a union of countably many non-intersecting open intervals $(a_i, b_i)$ such that

$$\sum_{i=1}^{\infty} (b_i - a_i) < \delta.$$ 

Denote by $\text{var} (\mu; a_i, b_i)$ the full variation on the interval $[a_i, b_i]$. Since

$$\sum_{i=1}^{\infty} \text{var} (\mu; a_i, b_i) \leq \text{var} \mu < +\infty,$$

for a sufficiently large $m = m(\varepsilon)$ we obtain $\sum_{i=m+1}^{\infty} \text{var} (\mu; a_i, b_i) \leq \varepsilon/2$. Set $E = \bigcup_{i=1}^{m} [a_i, b_i]$ and write $\mu = \mu_1 + \mu_2$ so that $\mu_1$ is supported within $E$ and $\mu_1 = \mu$ on $E$. It is clear that $\text{meas \ supp} \mu_1 \leq \delta$ and, also,

$$\text{var} \mu_2 \leq \sum_{i=m+1}^{\infty} \text{var} (\mu; a_i, b_i) + \frac{1}{2\pi} \int_{[-\pi, \pi] \setminus E} \mu'(x) \, dx \leq \varepsilon.$$ 

The choice $\delta = \varepsilon$ completes the proof. □

Lemma 3.4 Let $\mu$ be a symbol of a nonnegative Radon measure. Then

$$\frac{1}{n} \sum_{k=1}^{n} \sigma_k(A_n) \leq \frac{1}{2\pi} \text{var} \mu.$$ 

(28)

Proof. We take into account that $A_n = A_n^* \geq 0$. Hence, the singular values coincide with the eigenvalues, and their sum is equal to trace $A_n$. Since $A_n$ is a Toeplitz matrix, it is sufficient to show that $\alpha_0 \leq \frac{1}{2\pi} \text{var} \mu$. This trivially emanates from (8). □

Proof of Theorem 1.1. Assume, first, that $\mu$ is a monotone non-decreasing function. Then $\mu = \mu_a + \mu_s + \mu_j$, where $\mu_a$ is an absolutely continuous function, $\mu_s$ is a singular function, $\mu_j$ is a function of jumps, and all three are also monotone non-decreasing functions. Apart from $A_n = A_n(\mu)$, consider Toeplitz matrices $B_n$ generated by $\mu_a$. We intend to show that $A_n$ and $B_n$ enjoy the premises of Lemma 2.1.
Take an arbitrary \( \varepsilon > 0 \). Using Lemma 3.3, we can write \( \mu_\varepsilon + \mu_j = \mu_1 + \mu_2 \) so that (26) and (27) are fulfilled. Denote by \( T_n \) and \( U_n \) the Toeplitz matrices generated by \( \mu_1 \) and \( \mu_2 \), respectively.

Due to Lemma 3.2, we have \( T_n = T_{1n} + T_{2n} \) with trace \( T_{1n} = o(n) \) and \( \text{rank} T_{2n} \leq c_2 \varepsilon n \). By Lemma 3.4, trace \( U_n \leq \frac{1}{c_2} \varepsilon n \). Thus, setting up \( E_n = U_n + T_{1n} \) and \( R_n = T_{2n} \), we obtain, for some \( c > 0 \),

\[
||E_n||_F \leq c \varepsilon n \quad \text{and} \quad \text{rank} R_n \leq c \varepsilon n
\]

for all sufficiently large \( n \). As Lemma 2.1 states, \( A_n \) and \( B_n \) are bound to have equally distributed singular values (and eigenvalues).

In the general case, we write \( \mu = \mu_+ - \mu_- \), where \( \mu_+ \) and \( \mu_- \) are monotone non-decreasing functions. Then, we consider the above splittings and make use of the triangular inequality for the trace norm and that the rank of a sum does not exceed the sum of ranks. The Szegö-like formulas for Toeplitz matrices generated by the absolutely continuous component of \( \mu \) were proved in [11]. \( \square \)

References


