Influence of Matrix Operations on the Distribution of Eigenvalues and Singular Values of Toeplitz Matrices

Evgenij E. Tyrtyshnikov
Institute of Numerical Mathematics
Russian Academy of Science
Leninskij Prospekt, 32-A,
Moscow 117334, Russia

Submitted by Richard A. Brualdi

ABSTRACT

Suppose some Toeplitz matrix families $\{A_n(f_\alpha)\}$ are given, generated by the Fourier expansions for f_α, and a new family$\{A_n\}$ is constructed from $A_n(f_\alpha)$ via basic matrix operations. Theorems are proved that describe the singular-value distribution for A_n in the terms of f_α, as well as the eigenvalue distribution for $H(A_n) \equiv (A_n + A_n^*)/2$ and $K(A_n) \equiv (A_n - A_n^*)/2i$. In particular, if $f_\alpha \in L_\infty$ and only multiplication is used, then we show the singular values of A_n are distributed as $|f(x)|$, where $f(x) = \Pi f_\alpha(x)$. At the same time, the eigenvalues of $H(A_n)$ are distributed as $\operatorname{Re} f(x)$, while those of $K(A_n)$ are distributed as $\operatorname{Im} f(x)$. The extension to multilevel Toeplitz matrices is also suggested. Finally, an application to circulant preconditioning is discussed.

1. INTRODUCTION

In this paper we study the following problem. Suppose we are given Toeplitz matrix families $\{A_n\{f_\alpha)\}$ generated by functions f_α, and construct from them a new family $\{A_n\}$, using basic matrix operations such as addition, multiplication, and inversion. Can anything be said about the distribution of eigenvalues and singular values for the new family?

If matrices $A_n(f_\alpha)$ are Hermitian, then we know their eigenvalues are distributed as $f_\alpha(x)$. This is established by the Szegö theorem [5] if $f_\alpha \in L_\infty$; as is shown in [9], the same holds if $f_\alpha \in L_2$. Whether matrices $A_n(f_\alpha)$ are

655 Avenue of the Americas, New York, NY 10010 0024-3795/94/$7.00

Hermitian or not, their singular values are distributed as $|f_\alpha(x)|$. This is proved by Avram and Parter [1, 6] for $f_\alpha \in L_\infty$; as is shown in [9], the same is valid if $f_\alpha \in L_2$.

A difficulty is that matrices A_n constructed from $A_n(f_\alpha)$ with the help of basic matrix operations are no longer Toeplitz in general. Just the same, in a certain sense they are generated by a function $f(x)$ constructed from $f_\alpha(x)$ via the corresponding operations. Therefore, we may expect that, under certain hypotheses, A_n's eigenvalues are distributed as $f(x)$, and A_n's singular values are distributed as $|f(x)|$. The goal of this paper is the formulation and proof of the relevant theorems.

In Section 2, all necessary definitions and other preliminaries are given. We shall chiefly make use of the framework proposed in [9].

In Section 3, main theorems are formulated and proved. Section 4 shows how these theorems can be extended to multilevel matrices.

In Section 5, we discuss an application to the problem of preconditioning when solving linear algebraic systems. Apparently, the theorems proposed in Sections 3 and 4 lead to the simplest and clearest explanation of the existence of clusters among eigenvalues or singular values.

In Section 6, some open questions are listed.

2. PRELIMINARIES

Two sequences of real numbers, $\{\lambda_k^{(n)}\}_{k=1}^n$ and $\{\mu_k^{(n)}\}_{k=1}^n$, are called equally distributed if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \left[F\left(\lambda_k^{(n)}\right) - F\left(\mu_k^{(n)}\right) \right] = 0 \tag{2.1}$$

for any continuous function F with bounded support.

Let f be a real-valued Lebesgue-integrable function that is defined on the whole real axis and has 2π as its period. A sequence $\{\lambda_k^{(n)}\}$ is said to be distributed as $f(x)$ if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} F\left(\lambda_k^{(n)}\right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(f(x))\, dx \tag{2.2}$$

for any continuous F with bounded support. In a way, these definitions generalize those by H. Weyl [5]. They were put forward in [9].

THEOREM 2.1 (A modification of Theorem 3.1 from [9]). *Given two sequences of complex matrices, $\{A_n\}$ and $\{B_n\}$, where n is the matrix order, suppose that there are matrices Δ_n such that*

$$\lim_{n\to\infty} \frac{1}{n}\|A_n - B_n + \Delta_n\|_F^2 = 0, \tag{2.3}$$

$$\lim_{n\to\infty} \frac{1}{n} \operatorname{rank} \Delta_n = 0. \tag{2.4}$$

Then the singular values of A_n and B_n are equally distributed. If A_n, B_n are Hermitian, then their eigenvalues are equally distributed too.

A matrix $A = [a_{ij}]_{ij=0}^n$ is called Toeplitz if $a_{ij} \equiv a_{i-j}$, and circulant if $a_{ij} \equiv a_{i-j(\mathrm{mod}(n+1))}$. For circulant A we will use the notation $A = \operatorname{circ}(a_0, \dots, a_n)$.

Toeplitz matrices $A_n = [a_{i-j}]_{ij=0}^n$ are said to be generated by the function $f(x)$ if a_{i-j} are the Fourier coefficients for $f(x)$, that is,

$$f(x) = \sum_{k=-\infty}^{\infty} a_k e^{ikx}, \qquad x \in \mathbf{R}. \tag{2.5}$$

Several families of circulant matrices are associated with the series (2.5). In particular, we have the so-called simple circulants

$$S_n \equiv \begin{cases} \operatorname{circ}(a_0, a_1 \dots a_m, a_{-m}, \dots, a_{-1}), & n = 2m, \\ \operatorname{circ}(a_0, a_1, \dots a_{m-1}, 0, a_{-m+1}, \dots, a_{-1}), & n = 2m-1, \end{cases} \tag{2.6}$$

and the Cesàro circulants

$$C_n \equiv \operatorname{circ}\left(c_0^{(n)}, \dots c_n^{(n)}\right), \tag{2.7}$$

$$c_k^{(n)} = \frac{1}{n+1} \sum_{\substack{i,j=0 \\ i-j=k(\mathrm{mod}(n+1))}}^{n} a_{i-j}.$$

The eigenvalues of simple and Cesàro circulants are expressed as follows:

$$\lambda_k(S_n) = f_{[n/2]}\left(\frac{2\pi k}{n+1}\right), \qquad k = 0, 1, \dots, n; \tag{2.8}$$

$$\lambda_k(C_n) = \sigma_n\left(\frac{2\pi k}{n+1}\right), \qquad k = 0, 1, \dots, n, \tag{2.9}$$

where

$$f_m(x) \equiv \sum_{k=-m}^{m} a_k e^{ikx}, \tag{2.10}$$

$$\sigma_n(x) \equiv \frac{1}{n+1} \sum_{m=0}^{n} f_m(x). \tag{2.11}$$

Equation (2.8) is well known; (2.9) is given in [2] (the derivation of both can be also found in [9]).

THEOREM 2.2. *Suppose $f \in L_2$ and the Toeplitz matrices A_n and circulants S_n, C_n are generated by series* (2.5). *Then*

$$\|A_n - S_n\|_F^2 = o(n), \tag{2.12}$$

$$\|S_n - C_n\|_F^2 = o(n), \tag{2.13}$$

$$\|A_n - C_n\|_F^2 = o(n) \tag{2.14}$$

($o(n)$ *signifies, as usual, that* $o(n)/n \to 0$ *as* $n \to \infty$).

We omit the proof, because it is explicitly incorporated in the proofs of Theorems 4.1 and 5.1 from [9]. The next theorem is a generalization of the theorems of Szegö [5] and Avram and Parter [1, 6].

THEOREM 2.3 (Theorems 5.4 and 5.5 from [9]). *If $f \in L_2$, then the singular values of the matrices A_n, S_n, C_n are distributed as $|f(x)|$. If f is real-valued, then the matrices A_n, S_n, C_n are Hermitian, and their eigenvalues are distributed as $f(x)$.*

Below we will also refer to the well-known Fejer representation of the Cesàro sums:

$$\sigma_n(x) = \int_{-\pi}^{\pi} K_n(x, t) f(t)\, dt, \tag{2.15}$$

where

$$K_n(x,t) \equiv \frac{1}{2\pi(n+1)} \frac{\sin^2 (n+1)\dfrac{x-t}{2}}{\sin^2 \dfrac{x-t}{2}}, \tag{2.16}$$

and it is noted that for all x

$$\int_{-\pi}^{\pi} K_n(x,t)\, dt = 1. \tag{2.17}$$

Sometimes we will deal with the notion of a cluster. Let $\{\lambda_k^{(n)}\}$ be a sequence of real numbers, and $\gamma_n(\varepsilon)$ denote the number of those among $\lambda_k^{(n)}$, $k = 1, \dots, n$, which lie outside the ε-ball centered at c. If $\gamma_n(\varepsilon) = o(n)$ for any $\varepsilon > 0$, the c is called a cluster.

A cluster is called proper if $\gamma_n(\varepsilon) \leqslant C(\varepsilon)$, where $C(\varepsilon)$ is independent of n. Proper clusters are customarily used to obtain theoretical estimates of the convergence rate for the preconditioned conjugate gradients. From the practical point of view the distinction between proper and general clusters is not crucial, since moderate growth of $\gamma_n(\varepsilon)$ may well lead to fast convergence in practice.

3. MAIN RESULTS

Consider 2π-periodic complex-valued functions

$$f_\alpha(x), \quad \alpha = 1, \dots, r, \qquad g_\beta(x), \quad \beta = 1, \dots, q,$$

and suppose that $g_\beta(x) \neq 0$ for all x. Set

$$f(x) \equiv \prod_{\alpha=1}^{r} f_\alpha(x) \Big/ \prod_{\beta=1}^{q} g_\beta(x). \tag{3.1}$$

Let $C_n(f_\alpha)$ and $C_n(g_\beta)$ denote the Cesàro circulants associated with the Fourier series for f_α and g_β. Under certain hypotheses, the circulants $C_n(g_\beta)$ are guaranteed to be nonsingular, and we thus may introduce a new family of circulants defined as

$$C_n \equiv \prod_{\alpha=1}^{r} C_n(f_\alpha) \prod_{\beta=1}^{q} C_n^{-1}(g_\beta). \tag{3.2}$$

Note that it does not matter in which way the matrices in (3.2) are ordered, for all circulants commute, and the inverse of a circulant is also a circulant.

LEMMA 3.1. *Suppose f_α and g_β are continuous. Then for sufficiently large n the circulants $C_n(g_\beta)$ are nonsingular, and the singular values of the circulants C_n of the form* (3.2) *are distributed as $|f(x)|$, where f is defined by* (3.1). *At the same time, the eigenvalues of $H(A_n) \equiv (A_n + A_n^*)/2$ are distributed as* $\operatorname{Re} f(x) = [f(x) + \bar{f}(x)]/2$, *and the eigenvalues of $K(A_n) \equiv (A_n - A_n^*)/2i$ are distributed as* $\operatorname{Im} f(x) = [f(x) - \bar{f}(x)]/2i$.

Proof. Cesàro sums arising from any continuous function are known to be uniformly convergent to that function. Hence, condition $g_\beta(x) \neq 0$ provides that all eigenvalues of $C_n(g_\beta)$ are separated from zero for all sufficiently large n, since by (2.9) they are values of Cesàro sums $\sigma_n(g_\beta; x)$ of $g_\beta(x)$. Further, C_n's eigenvalues equal

$$\lambda_k^{(n)} = \lambda^{(n)}\left(\frac{2\pi k}{n+1}\right), \qquad k = 0, 1, \ldots, n, \tag{3.3}$$

where

$$\lambda^{(n)}(x) = \prod_{\alpha=1}^{r} \sigma_n(f_\alpha; x) \Big/ \prod_{\beta=1}^{q} \sigma_n(g_\beta; x). \tag{3.4}$$

The singular values of C_n are equal to $|\lambda_k^{(n)}|$. Clearly, the sequence of functions $\lambda^{(n)}(x)$ will uniformly converge to $f(x)$ given by (3.1). So for any ε and n sufficiently large we have

$$\begin{gathered}
\left| |\lambda_k^{(n)}| - \left| f\left(\frac{2\pi k}{n+1}\right) \right| \right| \leqslant \left| \lambda_k^{(n)} - f\left(\frac{2\pi k}{n+1}\right) \right| \leqslant \varepsilon, \\
\left| \operatorname{Re} \lambda_k^{(n)} - \operatorname{Re} f\left(\frac{2\pi k}{n+1}\right) \right| \leqslant \left| \lambda_k^{(n)} - f\left(\frac{2\pi k}{n+1}\right) \right| \leqslant \varepsilon, \\
\left| \operatorname{Im} \lambda_k^{(n)} - \operatorname{Im} f\left(\frac{2\pi k}{n+1}\right) \right| \leqslant \left| \lambda_k^{(n)} - f\left(\frac{2\pi k}{n+1}\right) \right| \leqslant \varepsilon, \\
k = 0, , \ldots, n.
\end{gathered} \tag{3.5}$$

Applying Lemma 3.2 from [9], we thus achieve the desired result. ■

Now, we are going to weaken the demand that generating functions should be continuous. From now on we will assume that they belong to L_∞. This suggests that the previous condition $g_\beta(x) \neq 0$ should be somehow strengthened. One possible way to do this is to assume that

$$\delta_1 \equiv \min_\beta \inf_x |g_\beta(x)| > \delta > 0. \tag{3.6}$$

However, (3.6) can not guarantee that the circulants $C_n(g_\beta)$ are nonsingular. For instance, if

$$g_\beta(x) = \begin{cases} 1, & 0 \leqslant x \leqslant \pi, \\ -1, & -\pi < x < 0, \end{cases}$$

then (3.6) is fulfilled with $0 < \delta < 1$, but due to (2.15), (2.16) we have $\sigma_n(g_\beta; 0) = 0$ for all n, i.e., all circulants $C_n(g_\beta)$ have zero as an eigenvalue. Nonetheless, (3.6) will be quite suited for our purposes, because it ensures that there are not "too many" zero eigenvalues. As a matter of fact, from Theorem 2.3 it follows that the number of those among $C_n(g_\beta)$'s eigenvalues which lie inside the δ-ball centered at zero is $o(n)$.

In order to freely consider "inversion" of singular matrices, we modify the operation of inversion. For an arbitrary square matrix A, we will write

$$A^{(-1)} \equiv (A + \Delta)^{-1}, \tag{3.7}$$

where:

(1) matrix $A + \Delta$ is nonsingular;

(2) rank Δ does not exceed the algebraic multiplicity of A's zero eigenvalue, if any; $\Delta = 0$ if A is nonsingular;

(3) if A is circulant then Δ is circulant;

(4) if A is Hermitian then Δ is Hermitian.

It is evident that items (1)–(4) can be easily satisfied. If A is nonsingular, then, by (2), $\Delta = 0$ and $A^{(-1)} = A^{-1}$. If A is singular, then there are many ways to satisfy (1)–(4). Any one of them is acceptable, so by $A^{(-1)}$ will be meant any matrix of the form (3.7), provided that (1)–(4) hold. For our purpose, that is, the study of the eigenvalue and singular-value distributions for $A^{(-1)}$, we do not need any more definite way to choose $A^{(-1)}$.

Instead of (3.2), we thus set

$$C_n \equiv \sum_{\alpha=1}^{r} C_n(f_\alpha) \prod_{\beta=1}^{q} C_n^{(-1)}(g_\beta). \tag{3.8}$$

If $C_n(g_\beta)$ are nonsingular, then, of course, (3.8) and (3.2) will produce the same matrix.

LEMMA 3.2. *Suppose* $f_\alpha, g_\beta \in L_\infty$ *and* (3.6) *holds. Then the singular values of* C_n, *defined by* (3.8), *are distributed as* $|f(x)|$, *where* f *is given by* (3.1). *The eigenvalues of* $H(A_n) \equiv (A_n + A_n^*)/2$ *are distributed as* $\operatorname{Re} f(x) = [f(x) + \bar{f}(x)]/2$, *while the eigenvalues of* $K(A_n) \equiv (A_n - A_n^*)/2i$ *are distributed as* $\operatorname{Im} f(x) = [f(x) - \bar{f}(x)]/2i$.

Proof. Take any $\varepsilon > 0$, arbitrary but sufficiently small, and choose continuous functions $f_\alpha^{(\varepsilon)}, g_\beta^{(\varepsilon)}$ such that

$$\begin{aligned} &\int_{-\pi}^{\pi} |f_\alpha(x) - f_\alpha^{(\varepsilon)}(x)|\,dx \leqslant \varepsilon, \qquad \alpha = 1,\dots,r, \\ &\int_{-\pi}^{\pi} |g_\beta(x) - g_\beta^{(\varepsilon)}(x)|\,dx \leqslant \varepsilon, \qquad \beta = 1,\dots,q. \end{aligned} \tag{3.9}$$

Let $I_n^{(\varepsilon)}$ be the set of those indices $k \in \{0, 1, \dots, n\}$ for which

$$\begin{aligned} &\left|\sigma_n\left(g_\beta; \frac{2\pi k}{n+1}\right)\right| \geqslant \delta, \qquad \beta = 1,\dots,q, \\ &\left|\sigma_n\left(g_\beta^{(\varepsilon)}; \frac{2\pi k}{n+1}\right)\right| \geqslant \delta, \qquad \beta = 1,\dots,q. \end{aligned} \tag{3.10}$$

It is readily seen that the number of indices which fall outside $I_n^{(\varepsilon)}$ is upper-bounded by $c_1 n\varepsilon$, where c_1 is independent of n and ε.

Suppose that $k \in I_n^{(\varepsilon)}$ is fixed, and set $x_k = 2\pi k/(n+1)$. We want to estimate the difference between $\lambda_k^{(n)}$, defined by (3.3), (3.4), and $\lambda_k^{(\varepsilon;n)}$ defined in a similar way:

$$\lambda_k^{(\varepsilon;n)} = \prod_{\alpha=1}^{r} \sigma_n\left(f_\alpha^{(\varepsilon)}; x_k\right) \Big/ \prod_{\beta=1}^{q} \sigma_n\left(g_\beta^{(\varepsilon)}; x_k\right). \tag{3.11}$$

To this end, we write

$$|\lambda_k^{(n)} - \lambda_K^{(\varepsilon;n)}| \leqslant \prod_{\beta=1}^{q} |\sigma_n^{-1}(g_\beta^{(\varepsilon)}; x_k)| \sum_{i=1}^{r} |\sigma_n(f_i^{(\varepsilon)}; x_k) - \sigma_N(f_i; x_k)|$$

$$\times \prod_{\alpha=1}^{i-1} |\sigma_n(f_\alpha; x_k)| \prod_{\alpha=i+1}^{r} |\sigma_n(f_\alpha^{(\varepsilon)}; x_k)|$$

$$+ \prod_{\alpha=1}^{r} |\sigma_n(f_\alpha; x_k)| \sum_{i=1}^{q} |\sigma_n^{-1}(g_i^{(\varepsilon)}; x_k) - \sigma_n^{-1}(g_i; x_k)|$$

$$\times \prod_{\beta=1}^{i-1} |\sigma_n^{-1}(g_\beta; x_k)| \prod_{\beta=i+1}^{q} |\sigma_n^{-1}(g_\beta^{(\varepsilon)}; x_k)|,$$

and using (3.9) and (3.10), we see that

$$|\lambda_k^{(n)} - \lambda_k^{(\varepsilon;n)}| \leqslant c_2 \sum_{i=1}^{r} |\sigma_n(f_i^{(\varepsilon)}; x_k) - \sigma_n(f_i; x_k)|$$

$$+ c_2 \sum_{i=1}^{q} |\sigma_n(g_i^{(\varepsilon)}; x_k) - \sigma_n(g_i; x_k)|, \qquad k \in I_n,$$

where c_2 does not depend on n and ε, provided that ε is sufficiently small. This does not imply that the difference between $\lambda_k^{(n)}$ and $\lambda_k^{(\varepsilon;n)}$ is bound to be small, but fortunately that is not what we are after. The only thing we need is that $\{\lambda_k^{(n)}\}$ and $\{\lambda_k^{(\varepsilon;n)}\}$ are "almost" equally distributed. Owing to (2.15)–(2.17) we find

$$\sum_{k \in I_n^{(\varepsilon)}} |\lambda_k^{(n)} - \lambda_k^{(\varepsilon;n)}| \leqslant \alpha_n c_2 (r+q)\varepsilon, \tag{3.12}$$

where

$$\alpha_n = \max_{0 \leqslant t \leqslant 2\pi} \frac{1}{2\pi(n+1)} \sum_{k=0}^{n} \frac{\sin^2 (n+1)\left(\dfrac{\pi k}{n+1} - \dfrac{t}{2}\right)}{\sin^2\left(\dfrac{\pi k}{n+1} - \dfrac{t}{2}\right)}. \tag{3.13}$$

A way in which α_n can be estimated was shown in [9]. Specifically, from the properties of the sine we have

$$\alpha_n \leqslant \max_{0 \leqslant \theta \leqslant 1} \frac{1}{\pi(n+1)} \sum_{k=0}^{[n/2]+1} \frac{\sin^2 (n+1) \dfrac{\pi(k+\theta)}{(n+1)}}{\sin^2 \dfrac{\pi(k+\theta)}{n+1}}$$

For $0 \leqslant k \leqslant [n/2] + 1$,

$$\left| \sin \frac{\pi(k+\theta)}{n+1} \right| \geqslant c_3 \frac{\pi(k+\theta)}{n+1},$$

where $c_3 > 0$ does not depend on n. Noting that $\sin \pi\theta \leqslant \pi\theta$, we thence deduce

$$\alpha_n \leqslant \frac{n+1}{\pi c_3^2} + \frac{n+1}{\pi^3 c_3^2} \sum_{k=0}^{[n/2]+1} \frac{1}{k^2} \leqslant c_0(n+1), \tag{3.14}$$

where $c_0 > 0$ does not depend on n. Now, by (3.12) and (3.14),

$$\sum_{k \in I_n^{(\varepsilon)}} |\lambda_k^{(n)} - \lambda_k^{(\varepsilon;n)}| \leqslant c_4 n \varepsilon, \tag{3.15}$$

where $c_4 > 0$ is independent of n and ε.

Further, let continuous functions $h_\beta^{(\varepsilon)}$ be introduced such that

$$\int_{-\pi}^{\pi} |g_\beta^{-1}(x) + h_\beta^{(\varepsilon)}(x)|\, dx \leqslant \varepsilon, \qquad \beta = 1, \dots, q, \tag{3.16}$$

and set

$$\overline{\lambda}_i^{(\varepsilon;n)} = \prod_{\alpha=1}^{r} \sigma_n\big(f_\alpha^{(\varepsilon)}; x_k\big) \prod_{\beta=1}^{q} \sigma_n\big(h_\beta^{(\varepsilon)}; x_k\big), \tag{3.17}$$

$$f^{(\varepsilon)}(x) = \prod_{\alpha=1}^{r} f_\alpha^{(\varepsilon)}(x) \prod_{\beta=1}^{q} h_\beta^{(\varepsilon)}(x). \tag{3.18}$$

Lemma 3.1 ensures that $\bar{\lambda}_k^{(\varepsilon; n)}$ are distributed as $f_\alpha^{(\varepsilon)}(x)$. On the other hand, $\{\bar{\lambda}_k^{(\varepsilon; n)}\}$ and $\{\lambda_k^{(\varepsilon; n)}\}$ are "almost" equally distributed. Again from Lemma 3.1, for any β the quantities $\sigma_n(g_\beta^{(\varepsilon)}; x_k)\sigma_n(h_\beta^{(\varepsilon)}; x_k)$ are distributed as $g_\beta^{(\varepsilon)}(x)h_\beta^{(\varepsilon)}(x)$. This means that, in particular,

$$\frac{1}{n+1}\sum_{k=0}^{n}\sigma_n\big(g_\beta^{(\varepsilon)}; x_k\big)\sigma_n\big(h_\beta^{(\varepsilon)}; x_k\big) \to \frac{1}{2\pi}\int_{-\pi}^{\pi} g_\beta^{(\varepsilon)}(x)h_\beta^{(\varepsilon)}(x)\,dx,$$

and consequently,

$$\frac{1}{n+1}\sum_{k=0}^{n}\Big[\sigma_n\big(g_\beta^{(\varepsilon)}; x_k\big)\sigma_n\big(h_\beta^{(\varepsilon)}; x_k\big) - 1\Big] \to \frac{1}{2\pi}\int_{-\pi}^{\pi}\Big[g_\beta^{(\varepsilon)}(x)h_\beta^{(\varepsilon)}(x) - 1\Big]\,dx.$$

Since

$$\|g_\beta^{(\varepsilon)}h_\beta^{(\varepsilon)} - 1\|_{L_1} \leqslant \|\big(g_\beta^{(\varepsilon)} - f\big)h_\beta^{(\varepsilon)}\|_{L_1} + \|f\big(h_\beta^{(\varepsilon)} - f^{-1}\big)\|_{L_1},$$

from (3.9) and (3.16) it follows that

$$\|g_\beta^{(\varepsilon)}h_\beta^{(\varepsilon)} - 1\|_{L_1} = O(\varepsilon).$$

Therefore, for some $c_5 > 0$,

$$\left|\frac{1}{n+1}\sum_{k=0}^{n}\Big[\sigma_n\big(g_\beta^{(\varepsilon)}; x_K\big)\sigma_n\big(h_\beta^{(\varepsilon)}; x_k\big) - 1\Big]\right| \leqslant c_5\varepsilon.$$

Denote by $\bar{I}_{n;\beta}^{(\varepsilon)}$ the set of those indices k for which

$$|\sigma_n\big(g_\beta^{(\varepsilon)}; x_k\big)\sigma_n\big(h_\beta^{(\varepsilon)}; x_k\big) - 1| \leqslant \varepsilon^{1/2}. \tag{3.19}$$

If $\nu_{n;\beta}^{(\varepsilon)}$ designates the number of those k which fall outside $\bar{I}_{n;\beta}^{(\varepsilon)}$, then combining the last two inequalities yields

$$\nu_{n;\beta}^{(\varepsilon)} \leqslant c_5\varepsilon^{1/2}(n+1).$$

Setting

$$\bar{I}_n^{(\varepsilon)} = I_n^{(\varepsilon)} \cap \left(\bigcap_{\beta=1}^{q} \bar{I}_{n;\beta}^{(\varepsilon)} \right),$$

we see that outside $\bar{I}_n^{(\varepsilon)}$ there are at most $c_6 \varepsilon^{1/2} n$ indices, where $c_6 > 0$ is independent of n and ε. At the same time, for any $k \in \bar{I}_n^{(\varepsilon)}$ (3.19) and (3.10) imply

$$|\sigma_n\big(h_\beta^{(\varepsilon)}; x_k\big) - \sigma_n^{-1}\big(g_\beta^{(\varepsilon)}; x_k\big)| \leqslant \delta^{-1}\varepsilon^{1/2}, \qquad \beta = 1, \dots, q.$$

Hence, in view of the definitions of $\lambda_k^{(\varepsilon;n)}$ and $\bar{\lambda}_k^{(\varepsilon;n)}$, we conclude that

$$\sum_{k \in \bar{I}_n^{(\varepsilon)}} |\lambda_k^{(\varepsilon;n)} - \bar{\lambda}_K^{(\varepsilon;n)}| \leqslant c_7 n \varepsilon^{1/2}, \qquad c_7 > 0,$$

and by virtue of (3.15)

$$\sum_{k \in \bar{I}_n^{(\varepsilon)}} |\lambda_k^{(n)} - \bar{\lambda}_k^{(\varepsilon;n)}| \leqslant c_8 n \varepsilon^{1/2}, \qquad c_8 > 0. \tag{3.20}$$

Let $\tilde{I}_n^{(\varepsilon)} \subset \bar{I}_n^{(\varepsilon)}$ be the subset of those k for which

$$|\lambda_k^{(n)} - \bar{\lambda}_k^{(\varepsilon;n)}| \leqslant \varepsilon^{1/4}.$$

Then it immediately follows from (3.20) that outside $\tilde{I}_n^{(\varepsilon)}$ there can be at most $c\varepsilon^{1/4} n$ indices, where $c > 0$ does not depend on n and ε.

Finally, take an arbitrary continuous function $F(x)$ with a bounded support, which, say, belongs to $[m, M]$. Let $\omega(d; F)$ signify the continuity modulus of F:

$$\omega(d; F) \equiv \max_{\substack{m \leqslant x,\, y \leqslant M \\ |x-y| \leqslant d}} |F(x) - F(y)|.$$

Using the fact that

$$\lambda_k(C_n) = \lambda_k^{(n)}, \qquad k \in \bar{I}_n^{(\varepsilon)},$$

and applying the above established facts, we find

$$\left| \frac{1}{n+1} \sum_{k=0}^{n} F(|\lambda_k^{(n)}(C_n)|) - \frac{1}{n+1} \sum_{k=0}^{n} F(|\bar{\lambda}_k^{(\varepsilon;n)}|) \right|$$

$$\leqslant \frac{1}{n+1} \sum_{k \in \bar{I}_n^{(\varepsilon)}} \left| F(|\lambda_k(C_n)|) - F(|\bar{\lambda}_k^{(\varepsilon;n)}|) \right| + 2 \max_{m \leqslant x \leqslant M} |F(x)| c\varepsilon^{1/4}$$

$$\leqslant \omega(\varepsilon^{1/4}; F) + 2 \max_{m \leqslant x \leqslant M} |F(x)| c\varepsilon^{1/4} \equiv \rho(\varepsilon; F).$$

If lim means any partial limit, we have

$$\left| \lim \frac{1}{n+1} \sum_{k=0}^{n} F(|\lambda_k(C_n)|) - \frac{1}{2\pi} \int_{-\pi}^{\pi} F(|f(x)|)\, dx \right|$$

$$\leqslant \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| F(|f(x)|) - F(|f^{(\varepsilon)}(x)|) \right| dx + \rho(\varepsilon; F).$$

The right-hand side tends to 0 as $\varepsilon \to 0$. Consequently,

$$\lim_{n \to \infty} \frac{1}{n+1} \sum_{k=0}^{n} F(|\lambda_k(C_n)|) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(|f(x)|)\, dx. \tag{3.21}$$

In each occurrence of $\lambda_k(C_n)$, $\bar{\lambda}_k^{(\varepsilon;n)}$, $\lambda_k^{(n)}$, $\lambda_k^{(\varepsilon;n)}$, $f(x)$, $f^{(\varepsilon)}(x)$ the modulus may be replaced by the real (or imaginary) part, and that will complete the proof. ■

COROLLARY. *If f_α, $g_\beta \in L_\infty$ are real-valued, then the eigenvalues of C_n are distributed as $f(x)$.*

LEMMA 3.3. *Given two batches of matrix families, $A_n^{(i)}$, $i = 1, \ldots, t$, and $B_n^{(i)}$, $i = 1, \ldots, t$, suppose that*

$$\| A_n^{(i)} - B_n^{(i)} + \Delta_n^{(i)} \|_F^2 = o(n), \qquad i = 1, \ldots, t, \tag{3.22}$$

$$\operatorname{rank} \Delta_n^{(i)} = o(n), \tag{3.23}$$

and in addition, for some $\Gamma > 0$,

$$\| A_n^{(i)} \|_2 \leqslant \Gamma, \quad \| B_n^{(i)} \|_2 \leqslant \Gamma, \qquad i = 1, \ldots, t. \tag{3.24}$$

Then the singular values of the matrices

$$A_n \equiv A_n^{(1)} \cdots A_n^{(t)} \tag{3.25}$$

and

$$B_n \equiv B_n^{(1)} \cdots B_n^{(t)} \tag{3.26}$$

are equally distributed. The eigenvalues of $H(A_n) \equiv (A_n + A_n^*)/2$ *and* $H(B_n) \equiv (B_n + B_n^*)/2$ *are equally distributed, and so also are the eigenvalues of* $K(A_n) \equiv (A_n - A_n^*)/2i$ *and* $K(B_n) \equiv (B_n - B_n^*)/2i$.

Proof. Write

$$\begin{aligned} A_n - B_n &= \left(A_n^{(1)} - B_n^{(1)}\right) A_n^{(2)} \cdots A_n^{(t)} \\ &\quad + B_n^{(1)}\left(A_n^{(2)} - B_n^{(2)}\right) A_n^{(3)} \cdots A_n^{(t)} + \cdots \\ &\quad + B_n^{(1)} \cdots B_N^{(t-1)}\left(A_n^{(t)} - B_n^{(t)}\right) \\ &= \left(A_n^{(1)} - B_n^{(1)} + \Delta_n^{(1)}\right) A_n^{(2)} \cdots A_n^{(t)} \\ &\quad + B_n^{(1)}\left(A_n^{(2)} - B_n^{(2)} + \Delta_n^{(2)}\right) A_n^{(3)} \cdots A_n^{(t)} + \cdots \\ &\quad + B_n^{(1)} \cdots B_n^{(t-1)}\left(A_n^{(t)} - B_n^{(t)} + \Delta_n^{(t)}\right) - \Delta_n, \end{aligned}$$

where

$$\Delta_n \equiv \Delta_n^{(1)} A_n^{(2)} \cdots A_n^{(t)} + B_n^{(1)} \Delta_n^{(2)} A_n^{(3)} \cdots A_n^{(t)} + \cdots + B_n^{(1)} \cdots B_n^{(t-1)} \Delta_n^{(t)}.$$

By (3.23), rank $\Delta_n = o(n)$, and at the same time we see that

$$\| A_n - B_n + \Delta_n \|_F \leqslant \Gamma^{t-1} \sum_{i=1}^{t} \| A_n^{(i)} - B_n^{(i)} + \Delta_n^{(i)} \|_F .$$

Hence, due to (3.22), $\| A_n - B_n + \Delta_n \|_F^2 = o(n)$. It remains to apply Theorem 2.1 concerning A_n and B_n.

Setting $H(\Delta_n) = (\Delta_n + \Delta_n^*)/2$, $K(\Delta_n) = (\Delta_n - \Delta_n^*)/2i$, we obviously obtain

$$\|H(A_n) - H(B_n) + H(\Delta_n)\|_F^2 = o(n),$$

$$\|K(A_n) - K(B_n) + K(\Delta_n)\|_F^2 = o(n),$$

where rank $H(\Delta_n) = o(n)$, rank $K(\Delta_n) = o(n)$. The proof is thus over. ∎

COROLLARY. *If $\overline{B}_n \equiv (B_n^{(1)} + \Delta_n^{(1)}) \cdots (B_n^{(t)} + \Delta_n^{(t)})$, then the singular values of A_n and $\overline{B}_n$ are equally distributed and so are the eigenvalues of $H(A_n)$ and $H(\overline{B}_n)$ as well as those of $K(A_n)$ and $K(\overline{B}_n)$.*

All preparatory work is now completed, and we are in a position to enunciate our main theorems pertaining to Toeplitz matrices $A_n(f_\alpha)$ and $A_n(g_\beta)$ composed of Fourier coefficients for f_α and g_β. We are interested in the distribution of singular values (and eigenvalues, if they are real) of a product of matrices $A_n(f_\alpha)$ and $A_n^{(-1)}(g_\beta)$, taken in some order of our choice. Note that Toeplitz matrices do not commute, in general, and thus changing the ordering of matrices may lead to different products.

Set $t = r + q$, and consider a one-to-one mapping

$$\sigma : \{1, \ldots, t\} \to \{1, \ldots, r, -1, \ldots, -q\} \tag{3.27}$$

and also a marking mapping

$$\zeta : \{1, \ldots, t\} \to \{1, 0\}. \tag{3.28}$$

With the help of σ and ζ, we build up a new family $\{A_n\}$ in the following way:

$$A_n = B_n^{(1)} \ldots B_n^{(t)}, \tag{3.29}$$

where

$$B_n^{(k)} = \begin{cases} A_n(f_{\sigma(k)}) & \text{if} \quad \sigma(k) > 0, \quad \zeta(k) = 0, \\ C_n(f_{\sigma(k)}) & \text{if} \quad \sigma(k) > 0, \quad \zeta(k) = 1, \\ A_n^{(-1)}(g_{|\sigma(k)|}) & \text{if} \quad \sigma(k) < 0, \quad \zeta(k) = 0, \\ C_n^{(-1)}(g_{|\sigma(k)|}) & \text{if} \quad \sigma(k) < 0, \quad \zeta(k) = 1. \end{cases} \tag{3.30}$$

THEOREM 3.1. *Suppose f_α, $g_\beta \in L_\infty$ and* (3.6) *holds. Then the singular values of the matrices A_n, defined by* (3.29), (3.30) *are distributed as $|f(x)|$, where f is given by* (3.1).

Proof. Resorting to the singular-value decomposition of $A_n(g_\beta)$, that is,

$$A_n(g_\beta) = V^* M U, \tag{3.31}$$

where V, U are unitary, and

$$M = \operatorname{diag}(\mu_1, \dots, \mu_n), \tag{3.32}$$

we set

$$\hat{A}_n(g_\beta) \equiv V^* \operatorname{diag}(\mu_1 + \phi_1, \dots, \mu_n + \phi_n)\, U, \tag{3.33}$$

where

$$\phi_i = \begin{cases} 0 & \text{if} \quad |\mu_i| \geqslant \delta, \\ \delta - \mu_i & \text{if} \quad |\mu_i| < \delta. \end{cases} \tag{3.34}$$

Also, let $\hat{C}_n(g_\beta)$ denote a matrix derived from $C_n(g_\beta)$ in the same way as in (3.31)–(3.34). Set

$$C_n \equiv C_n^{(1)} \cdots C_n^{(t)}, \qquad \hat{C}_n \equiv \hat{C}_n^{(1)} \cdots \hat{C}_n^{(t)}, \tag{3.35}$$

where

$$C_n^{(k)} = \begin{cases} C_n(f_{\sigma(k)}) & \text{if} \quad \sigma(k) > 0, \\ C_n^{(-1)}(g_{|\sigma(k)|}) & \text{if} \quad \sigma(k) < 0, \end{cases} \tag{3.36}$$

$$\hat{C}_n^{(k)} = \begin{cases} C_n(f_{\sigma(k)}), & \text{if} \quad \sigma(k) > 0, \\ \hat{C}_n^{(-1)}(g_{|\sigma(k)|}), & \text{if} \quad \sigma(k) < 0, \end{cases} \tag{3.37}$$

and let $\hat{A}_n$ be obtained from A_n by replacing $A_n^{(-1)}(g_\beta)$ and $C_n^{(-1)}(g_\beta)$ with $\hat{A}_n^{-1}(g_\beta)$ and $\hat{C}_n^{-1}(g_\beta)$, respectively. In accord with the above definitions,

$$\operatorname{rank}\left(A_n^{(k)} - \hat{A}_n^{(k)}\right) = o(n), \qquad \operatorname{rank}\left(C_n^{(k)} - \hat{C}_n^{(k)}\right) = o(n). \tag{3.38}$$

Hence, setting

$$\Delta_k^{(k)} \equiv \left(A_n^{(k)} - C_n^{(k)}\right) - \left(\hat{A}_n^{(k)} - \hat{C}_n^{(k)}\right), \tag{3.39}$$

we see that

$$\operatorname{rank} \Delta_k^{(k)} = o(n). \tag{3.40}$$

At the same time, by Theorem 2.2 $\| A_n^{(k)} - C_n^{(k)} \|_F^2 = o(n)$ together with (3.39), we have

$$\| \hat{A}_n^{(k)} - \hat{C}_n^{(k)} + \Delta_k^{(k)} \|_F^2 = o(n). \tag{3.41}$$

Thus, the matrices $\hat{A}_n$ and $\hat{C}_n$ satisfy all the hypotheses of Lemma 3.3. Consequently, their singular values are equally distributed. Lemma 3.2 states that the singular values of both $\hat{C}_n$ and C_n are distributed as $|f(x)|$, and this completes the proof. ■

REMARK. Theorem 3.1 still stands if all or some of Cesàro circulants are replaced by simple circulants of the form (2.6).

To see this, we rely on (2.13) from Theorem 2.2 and Lemma 3.3. If $A_n^{(i)}$ is a Cesàro circulant and $B_n^{(i)}$ is the corresponding simple circulant, then $\| A_n^{(i)} - B_n^{(i)} \|_F^2 = o(n)$. Because the generating function, say f_i, is from L_∞, the $\| A_n^{(i)} \|_2$ are uniformly bounded with respect to n. That is not necessarily so for the $\| B_n^{(i)} \|_2$; but since the singular values of Cesàro and simple circulants are equally distributed (Theorem 2.3), only $o(n)$ of those for the latter can be greater than $\| f_i \|_{L_\infty}$. Therefore, there exist matrices $\Delta_n^{(i)}$ such that $\tilde{B}_n^{(i)} = B_n^{(i)} - \Delta_n^{(i)}$ are circulants, rank $\Delta_n^{(i)} = o(n)$, and $\| \tilde{B}_n^{(i)} \|_2$ are uniformly bounded with respect to n. Thus, we have

$$\| A_n^{(i)} - \tilde{B}_n^{(i)} + \Delta_n^{(i)} \| = o(n), \qquad \operatorname{rank} \Delta_n^{(i)} = o(n),$$

and the two matrix products A_n and $\tilde{B}_n$, one with $A_n^{(i)}$ and another with $\tilde{B}_n^{(i)}$, satisfy all the hypotheses of Lemma 3.3. At the same time, if B_n is the matrix product that involves $B_n^{(i)}$, then $\operatorname{rank}(B_n - \tilde{B}_n) = o(n)$, and hence the singular values of B_n and $\tilde{B}_n$ are equally distributed.

THEOREM 3.2. *Suppose* f_α, $g_\beta \in L_\infty$ *and* (3.6) *holds. Then the eigenvalues of the Hermitian matrices*

$$H(A_n) \equiv \frac{A_n + A_n^*}{2} \tag{3.42}$$

are distributed as Re $f(x)$, *where* f *is given by* (3.1), *and the eigenvalues of the Hermitian matrices*

$$K(A_n) \equiv \frac{A_n - A_n^*}{2i} \tag{3.43}$$

are distributed as Im $f(x)$.

Proof. Let $\hat{A}_n$ and $\hat{C}_n$ be the same as in the proof of Theorem 3.1. For $\hat{A}_n$ and $\hat{C}_n$ all hypotheses of Lemma 3.3 are fulfilled, and that concludes the proof. ■

COROLLARY. *If the functions* f_α, $g_\beta \in L_\infty$ *are real-valued, then the eigenvalues of* $H(A_n)$ *are distributed as* $f(x)$, *while the eigenvalues of* $K(A_n)$ *have a cluster at zero.*

Note that A_n being a product of Hermitian matrices no longer means that A_n itself is Hermitian. That would be so if we knew that matrices in the product commute. That is true of circulants, but not of Toeplitz matrices in general. Thus Im $f(x) = 0$ does not necessarily imply $K(A_n) = 0$.

4. THEOREMS FOR MULTILEVEL MATRICES

A function $f(x_1, \dots, x_p)$ expressed by a multidimensional Fourier series

$$f(x_1, \dots, x_p) = \sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_p=-\infty}^{\infty} a_{k_1;\dots;k_p} \exp i(k_1 x_1 + \cdots + k_p x_p) \tag{4.1}$$

is naturally associated with p-level Toeplitz matrices $A_{\bar{n}}$ and p-level Cesàro circulants $C_{\bar{n}}$, where $\bar{n} = (n_1, \dots, n_p)$. By definition,

$$A_{\bar{n}} \equiv [a_{i_1 - j_1; \dots; i_p - j_p}], \tag{4.2}$$

$$0 \leqslant i_l, j_l \leqslant n_l; \qquad l = 1, \dots, p.$$

A matrix $C_{\bar{n}}$ is called a p-level circulant if

$$C_{\bar{n}} \equiv \left[c^{(\bar{n})}_{i_1 - j_1 (\mathrm{mod}\,(n_1+1)); \dots ; i_p - j_p (\mathrm{mod}(n_p+1))} \right], \tag{4.3}$$

and is called a p-level Cesàro circulant generated by (4.1) if

$$c^{(\bar{n})}_{k_1; \dots ; k_p} = \frac{1}{(n_1+1) \cdots (n_p+1)}$$

$$\times \sum_{\substack{i_1, j_1 = 0 \\ i_1 - j_1 = k_1 (\mathrm{mod}(n_1+1))}}^{n_1} \cdots \sum_{\substack{i_p, j_p = 0 \\ i_p - j_p = k_p (\mathrm{mod}(n_p+1))}}^{n_p} a_{i_1 - j_1; \dots ; i_p - j_p}. \tag{4.4}$$

The matrices $A_{\bar{n}}$ and $C_{\bar{n}}$ can be viewed as block matrices assembled from $(n_1 + 1) \times (n_1 + 1)$ blocks, each block being a block matrix composed of $(n_2 + 1) \times (n_2 + 1)$ smaller blocks, and so on. The multiindex $\bar{n}$ describes the structure of the above nested partitionings. Such operations as summation, multiplication, and inversion applied to p-level matrices with common $\bar{n}$ obviously lead to a p-level matrix with the same $\bar{n}$. Above all, when these operations affect p-level circulants, the resulting matrix is still a p-level circulant. In contrast, both multiplication and inversion of p-level Toeplitz matrices usually lead to a p-level matrix which no longer is p-level Toeplitz. For more detailed information about p-level Toeplitz and circulant matrices we refer to [9]. That paper suggests, in particular, the analogs of Theorems 2.1–2.3 that have to do with multilevel matrices.

We are going to elaborate the above theory, extending Theorems 3.1 and 3.2 to the multilevel case. Consider complex-valued functions

$$f_\alpha(x_1, \dots, x_p), \qquad \alpha = 1, \dots, r,$$

$$g_\beta(x_1, \dots, x_p), \qquad \beta = 1, \dots, q,$$

and assume that each is 2π-periodic with respect to every argument. Additionally, $g_\beta(x_1, \dots, x_p) \neq 0$ for all $x_1, \dots, x_p$. Let $C_{\bar{n}}(f_\alpha)$ and $C_{\bar{n}}(g_\beta)$ denote p-level Cesàro circulants which correspond to f_α and g_β. By $A^{(-1)}$ will be meant, as earlier, a matrix of the form (3.7), that is, $A^{(-1)} = (A + \Delta)^{-1}$, where Δ obeys (1)–(4) from Section 3, with (3) being enhanced as follows:

(3) If A is a p-level circulant, then so is Δ.

Let us define

$$C_{\bar{n}} \equiv \prod_{\alpha=1}^{r} C_{\bar{n}}(f_\alpha) \prod_{\beta=1}^{q} C_{\bar{n}}^{(-1)}(g_\beta). \tag{4.5}$$

$$f(x_1, \dots, x_p) \equiv \prod_{\alpha=1}^{r} f_\alpha(x_1, \dots, x_p) \Big/ \prod_{\beta=1}^{q} g_\beta(x_1, \dots, x_p). \tag{4.6}$$

Sufficient smoothness of f_α, g_β leads to conclusions similar to those in Lemma 3.1. In particular, for $\bar{n}$ sufficiently large, p-level Cesàro circulants $C_{\bar{n}}(g_\beta)$ are nonsingular, and the singular values of $C_{\bar{n}}$ are distributed as $|f(x_1, \dots, x_p)|$, where f is given by (4.6). If f_α, g_β are real-valued, then the eigenvalues of $C_{\bar{n}}$ are distributed as $f(x_1, \dots, x_p)$.

In a general case, assume that f_α, $g_\beta \in L_\infty$ and also

$$\delta_1 \equiv \min_{\beta} \inf_{x_1, \dots, x_p} |g_\beta(x_1, \dots, x_p)| > \delta > 0. \tag{4.7}$$

The above conclusions about singular values and eigenvalues are still valid.

Further, denote by $A_{\bar{n}}(f_\alpha)$, $A_{\bar{n}}(g_\beta)$ p-level Toeplitz matrices allied with the series (4.1). Let σ and ζ be some marking functions defined by (3.27) and (3.28), and set

$$A_{\bar{n}} \equiv B_{\bar{n}}^{(1)} \cdots B_{\bar{n}}^{(t)}, \qquad t = r + q, \tag{4.8}$$

where $B_{\bar{n}}^{k}$ are expressed by formulas like (3.42), where n is to be replaced by $\bar{n}$.

THEOREM 4.1. *Suppose* f_α, $g_\beta \in L_\infty$ *and* (4.7) *holds. Then the singular values of the matrices* $A_{\bar{n}}$, *defined by* (4.8), *are distributed as* $|f(x_1, \dots, x_p)|$, *where* f *is given by* (4.1).

THEOREM 4.2. *Suppose* f_α, $g_\beta \in L_\infty$ *and* (4.7) *holds. Then the eigenvalues of*

$$H(A_{\bar{n}}) \equiv \tfrac{1}{2}(A_{\bar{n}} + A_{\bar{n}}^*) \tag{4.9}$$

are distributed as $\operatorname{Re} f(x_1, \dots, x_p)$, *while the eigenvalues of*

$$K(A_{\bar{n}}) \equiv \frac{1}{2i}(A_{\bar{n}} - A_{\bar{n}}^*) \tag{4.10}$$

are distributed as $\operatorname{Im} f(x_1, \dots, x_p)$.

COROLLARY. *If f_α, $g_\beta \in L_\infty$ are real-valued, then eigenvalues of $K(A_n)$ have a cluster at zero.*

We omit the proofs, because they would almost entirely repeat those of Theorems 3.1 and 3.2.

5. APPLICATION TO PRECONDITIONING

Consider Toeplitz matrices A_n generated by a real-valued function $f(x)$ with Fourier expansion (2.5), and assume that the Cesàro circulants (2.7) serve as preconditioners. In fact the matrices C_n are the so-called optimal preconditioners proposed in [4]. It is shown in [7, 8] that if A_n is positive definite, then C_n also is.

Suppose $f \in L_\infty$ is such that

$$\inf_x f(x) \equiv \gamma > 0. \tag{5.1}$$

Then according to (2.9) and (2.15)–(2.17) we have

$$\lambda_k(C_n) \geqslant \gamma > 0, \tag{5.2}$$

that is, all eigenvalues of C_n are positive. Moreover, setting $x \equiv [z_0 \ \dots \ z_n]^T$, we find

$$z^* A_n z = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{l=0}^{n} z_l e^{ilx} \right|^2 f(x)\, dx \geqslant \gamma z^* z, \tag{5.3}$$

which means that all A_n's eigenvalues are positive, too. The restriction (5.1) thus implies that C_n and A_n are positive definite, and their minimum eigenvalues are greater than or equal to γ.

In order to discern whether C_n is a good preconditioner, we scrutinize the spectrum of $C_n^{-1} A_n$. The eigenvalues of interest are real, because

$$C_n^{-1} A_n = C_n^{-1/2} \left(C_n^{-1/2} A_n C_n^{-1/2} \right) C_n^{1/2}, \tag{5.4}$$

which means that $C_n^{-1} A_n$ is similar to the Hermitian matrix $H_n \equiv C_n^{-1/2} A_n C_n^{-1/2}$. By Lemma 3.3 all H_n's eigenvalues must be distributed like those of $C_n^{-1/2} C_n C_n^{-1/2} = I$. The circulant preconditioning thus leads to a cluster at 1.

Apparently the first explanation of the fact that there is a cluster at 1 was obtained in [3, 4], where f is assumed to belong to the Wiener class. In [9] the same is proved with a weaker assumption, that $f \in L_2$. In our exposition here, this fact emerges as a trivial inference from a general rule that connects the distribution for a product of matrices with the distributions of its factors.

We have very little to change concerning multilevel matrices. A cluster at 1 is a consequence of the next theorem.

THEOREM 5.1. *Given two real-valued functions*

$$f(x_1, \ldots, x_p),\ g(x_1, \ldots, x_p) \in L_\infty,$$

suppose that

$$\inf_{x_1, \ldots, x_p} g(x_1, \ldots, x_p) \equiv \gamma > 0 \tag{5.5}$$

and consider p-level Toeplitz matrices $A_{\bar{n}}(f)$, $A_{\bar{n}}(g)$ *and p-level Cesàro circulants* $C_{\bar{n}}(f), C_{\bar{n}}(g)$. *Then for all* $\bar{n}$ *the matrices* $A_{\bar{n}}(g)$ *and* $C_{\bar{n}}(g)$ *are positive definite, and the eigenvalues for each of the six products*

$$\begin{array}{ccc} A_{\bar{n}}(f)\, A_{\bar{n}}(g), & A_{\bar{n}}(f) C_{\bar{n}}(g), & C_{\bar{n}}(f)\, A_{\bar{n}}(g), \\ A_{\bar{n}}(g)\, A_{\bar{n}}(f), & C_{\bar{n}}(g)\, A_{\bar{n}}(f), & A_{\bar{n}}(g) C_{\bar{n}}(f) \end{array}$$

are real and distributed as $f(x_1, \ldots, x_p) g(x_1, \ldots, x_p)$, *while the eigenvalues for each of the six "quotients"*

$$\begin{array}{ccc} A_{\bar{n}}(f)\, A_{\bar{n}}^{(-1)}(g), & A_{\bar{n}}(f) C_{\bar{n}}^{(-1)}(g), & C_{\bar{n}}(f)\, A_{\bar{n}}^{(-1)}(g), \\ A_{\bar{n}}^{(-1)}(g)\, A_{\bar{n}}(f), & C_{\bar{n}}^{(-1)}(g)\, A_{\bar{n}}(f), & A_{\bar{n}}^{(-1)}(g) C_{\bar{n}}(f) \end{array}$$

are real and distributed as $f(x_1, \ldots, x_p)/g(x_1, \ldots, x_p)$.

Proof. We first of all observe that eigenvalues of the square roots $A_{\bar{n}}^{1/2}(g)$ and $C_{\bar{n}}^{1/2}(g)$ are equally distributed. To this end, we refer to an obvious statement: if real sequences $\{\lambda_k^{(n)}\}$ and $\{\mu_k^n\}$ are equally distributed, the $\{|\lambda_k|^{1/2}\}$ and $\{|\mu_k^{(n)}|^{1/2}\}$ are as well. Further, the eigenvalues of $A_{\bar{n}}(f)\, A_{\bar{n}}(g)$ coincide with those of $A_{\bar{n}}^{1/2}(g)\, A_{\bar{n}}(f) A_{\bar{n}}^{1/2}(g)$, which, in turn, are distributed in the same way as the eigenvalues of $C_{\bar{n}}^{1/2}(g)\, C_{\bar{n}}(f) C_{\bar{n}}^{1/2}(g)$, as follows from Lemma 3.3. The rest of the proof needs nothing besides repetition of similar arguments. ∎

6. DISCUSSION

Let us consider an example. If we choose

$$f(x) = 1 + e^{ix}, \qquad g(x) = 1 + e^{-ix},$$

then

$$A_n(f) = \begin{bmatrix} 1 & & & & 0 & \\ 1 & 1 & & & & \\ & 1 & 1 & & & \\ & & \ddots & \ddots & & \\ & 0 & & 1 & 1 & \\ & & & & 1 & 1 \end{bmatrix},$$

$$A_n(g) = \begin{bmatrix} 1 & 1 & & & & \\ & 1 & 1 & & 0 & \\ & & \ddots & \ddots & & \\ & & & 1 & 1 & \\ & 0 & & & 1 & 1 \\ & & & & & 1 \end{bmatrix},$$

and setting

$$A_n = A_n(f)A_n(g) = \begin{bmatrix} 1 & 1 & & & & \\ 1 & 2 & 1 & & 0 & \\ & 1 & 2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & 0 & & 1 & 2 & 1 \\ & & & & 1 & 2 \end{bmatrix}_{(n+1)\times(n+1)},$$

we can easily calculate that

$$\lambda_k(A_n) = 2 + 2\cos\left(\frac{2\pi k}{n+1}\right), \qquad k = 0, 1, \dots, n+1. \tag{6.1}$$

At the same time, Theorem 3.2 states that the eigenvalues of matrices $A_n = (A_n + A_n^*)/2$ are to be distributed as

$$\mathrm{Re}[f(x)g(x)] = f(x)g(x) = 2 + e^{ix} + e^{-ix} = 2 + 2\cos x. \quad (6.2)$$

Obviously, (6.2) agrees entirely with (6.1).

This example can also caution against a possible misunderstanding or wrong extension. We never stated that if the real eigenvalues of Toeplitz matrices $\{U_n\}$ and $\{V_n\}$ are distributed as $u(x)$ and $v(x)$, then the real parts of the eigenvalues of $\{U_n V_n\}$ ought to be distributed as $u(x)v(x)$. In fact, if $U_n \equiv A_n(f)$, $V_n \equiv A_n(g)$ as above, then we can take $u(x) = v(x) \equiv 1$, and unity as a distribution function does not fit (6.1).

Nevertheless, we may look for some true generalizations, and here are some open questions of interest.

(1) What should we demand of generating functions f and g so as to be guaranteed that real parts of the eigenvalues of matrices $\{A_n(f)A_n(g)\}$ are distributed as $\mathrm{Re}[f(x)g(x)]$? (If so then the corresponding imaginary parts will be distributed as $\mathrm{Im}[f(x)g(x)]$.)

(2) Is there an extension of Theorem 5.1 to products of more than two factors?

(3) Suppose we are given two matrix families, $\{U_n\}$ and $\{V_n\}$, with singular values distributed as $u(x)$ and $v(x)$. In which cases will the singular values of $\{U_n V_n\}$ be distributed as $u(x)v(x)$? [Note that one such case was discussed in this paper: the required property obtains when U_n and V_n are Toeplitz matrices generated by $f(x)$ and $g(x)$: $u(x) = |f(x)|$ and $v(x) = |g(x)|$.]

I would like to thank Raymond Chan for his truly pertinent remarks on the manuscript.

REFERENCES

1 F. Avram, On bilinear forms on Gaussian variables and Toeplitz matrices, *Probab. Theory Related Fields* 79:37–45 (1988).

2 R. Chan and M.-C. Yeung, Circulant Preconditioners for Complex Toeplitz Matrices, Research Rep. HKUM-91-6, Dept. of Mathematics, Hong Kong Univ., 1991.

3 R. H. Chan and G. Strang, Toeplitz equations constructed by conjugate gradients with circulant preconditioners, *SIAM J. Sci. Statist. Comput.* 10:104–119 (1989).

4 T. F. Chan, An optimal circulant preconditioner for Toeplitz systems, *SIAM J. Sci. Statist. Comput.* 9:766–771 (1988).

5 U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*, Univ. of California Press, Berkeley, 1958.
6 S. V. Parter, On the distribution of the singular values of Toeplitz matrices, *Linear Algebra Appl.* 80:115–130 (1986).
7 E. E. Tyrtyshnikov, Toeplitz Matrices: Some of Their Analogs and Applications (in Russian), Dept. of Numerical Mathematics, USSR Acad. of Sciences, Moscow, 1989.
8 E. E. Tyrtyshnikov, Optimal and superoptimal circulant preconditioners, *SIAM J. Matrix Anal. Appl.* 13(2):459–473 (1992).
9 E. E. Tyrtyshnikov, A unifying approach to some old and new theorems on distribution and clustering, to appear.

Received 9 June 1992; final manuscript accepted 5 February 1993