A CONTRIBUTION TO THE THEORY OF TOEPLITZ MATRICES

BY

ULF GRENANDER

1. **Introduction.** Consider an abstract space X with a σ-algebra $\mathfrak{X}$ of subsets on which a measure m is defined (see e.g. P. R. Halmos, *Measure theory*). The elements in X are denoted by x. Let $\phi_\nu(x)$, $\nu = 0, 1, \cdots$, be a sequence of functions on X, orthonormal with respect to m,

$$\int_X \phi_\nu(x)\phi_\mu^*(x)dm(x) = \delta_{\nu\mu}(^1); \qquad\qquad \nu = 0, 1, \cdots,$$

and complete in $L_2(X)$. If $f(x)$ is a bounded, $(\mathfrak{X})$-measurable, and real-valued function on X, we introduce the infinite matrix

$$M(f) = \{m_{\nu\mu}(f); \nu, \mu = 0, 1, \cdots \},$$

with the elements

$$m_{\nu\mu}(f) = \int_X \phi_\nu(x)\phi_\mu^*(x)f(x)dx.$$

Note that for $f(x) \equiv 1$ we get the identity matrix

$$M(1) = \{\delta_{\nu\mu}; \nu, \mu = 0, 1, \cdots \} = I.$$

As $m_{\mu\nu}(f) = m_{\nu\mu}^*(f)$ the matrix is Hermitian.

If $\mathfrak{H}$ is the complex Hilbert space of vectors $z \in \mathfrak{H}$, $z = (z_0, z_1, \cdots)$ with $\sum_{\nu=0}^\infty |z_\nu|^2 < \infty$, the matrix $M(f)$ defines a linear Hermitian bounded operator in $\mathfrak{H}$ with values in the same space. The boundedness follows easily from

$$(z, Mz) = \sum_{\nu,\mu=0}^\infty z_\nu \bar{z}_\mu m_{\nu\mu}(f) = \int_X |s(x)|^2 f(x)dm(x),$$

where

$$s(x) = \sum_{\nu=0}^\infty z_\nu \phi_\nu(x),$$

which sum converges in the mean. As

$$\int_X |s(x)|^2 dm(x) = \sum_{\nu=0}^\infty |z_\nu|^2 = \|z\|^2$$

(¹) Complex conjugates will be denoted by asterisks.

we get

(1) $$m\|z\|^2 \leq (z, Mz) \leq M\|z\|^2,$$

where $m = \inf_{x \epsilon X}$ ess $f(x)$, $M = \sup_{x \epsilon X}$ ess $f(x)$ so that

$$\|M\| \leq \max (|m|, |M|).$$

The operator $M(f)$ has a well defined spectrum $S(f)$ on the real axis and it follows from (1) that it is contained in the interval (m, M). It is easy to find $S(f)$ by considering the isometric problem obtained by mapping $\mathfrak{H}$ onto $L_2(X)$ by

$$z = (z_0, z_1, \cdots) \leftrightarrow \sum_{\nu=0}^{\infty} z_\nu \phi_\nu(x).$$

The operator $N(f)$ in $L_2(X)$ corresponding to $M(f)$ is of a simple form. If h_μ is a unit vector along the μth coordinate axis in $\mathfrak{H}$ (so that $h_\mu \leftrightarrow \phi_\mu(x)$) then

$$M(f)h_\mu = \sum_{\nu=0}^{\infty} m_{\nu\mu}(f)h_\nu,$$

so that

$$N(f)\phi_\nu(x) = \sum_{\nu=0}^{\infty} m_{\nu\mu}(f)\phi_\nu(x) = \sum_{\nu=0}^{\infty} (\phi_\nu; \phi_\mu f)\phi_\nu(x) = f(x)\phi_\nu(x).$$

As this holds for every μ, $N(f)$ consists simply in multiplication by the function $f(x)$,

$$N(f)g = f(x)g(x), \qquad\qquad g \in L_2(X).$$

Now it is well known that the spectrum of $N(f)$, which coincides with $S(f)$, is the point set of all the values of $f(x)$, when x runs through X. $S(f)$ is the "Wertevorrat" of $f(x)$.

$M(f)$ is called an *infinite Toeplitz matrix* because of its similarity to a type of infinite matrices studied by Toeplitz [1]. He considered essentially the case where X is the interval $(0, 2\pi)$, m is Lebesgue measure, and

$$\phi_\nu(x) = \frac{1}{(2\pi)^{1/2}} e^{i\nu x}, \qquad \nu = \cdots, -1, 0, 1, \cdots.$$

In this case ν runs through all the integers instead of only the non-negative ones, but this difference is not important.

In this paper we shall study the *finite* Toeplitz matrices $M_n(f)$, which are sections of $M(f)$,

$$M_n(f) = \{m_{\nu\mu}(f); \nu, \mu = 0, 1, \cdots, n - 1\}.$$

As $M_n(f)$ is still a Hermitian matrix it has n real eigenvalues $\lambda_0^{(n)}, \lambda_1^{(n)}, \cdots,$

$\lambda_{n-1}^{(n)}$. It is in most cases not possible to get explicit expressions for $\lambda_{\nu}^{(n)}$, but as we shall show, one can in many cases describe their asymptotic distribution as n tends to infinity. To do this we regard the associated distribution function

$$D_n(t; f) = \frac{\text{number of } \lambda_{\nu}^{(n)} \leqq t}{n}.$$

If there is a distribution function $D(t; f)$ such that

$$\lim_{n \to \infty} D_n(t; f) = D(t; f)$$

for every value of t at which $D(t; f)$ is continuous, then we shall say that the Toeplitz matrix has an *asymptotic eigenvalue distribution* given by the distribution function $D(t; f)$.

This problem was solved by Szegö [5] for the finite matrices corresponding to the original case considered by Toeplitz. He showed that

$$D(t; f) = \frac{\text{Lebesgue measure of } \{ x \mid f(x) \leqq t \}}{2\pi}.$$

Typical for this case is not only the existence of an asymptotic eigenvalue distribution but the simple dependence of this on the function $f(x)$. Something similar to this happens in other cases, and it will be convenient to introduce a concept to describe this. If the limiting distribution $D(t; f)$ depends upon f in the following way

$$D(t; f) = \mu\{ x \mid f(x) \leqq t \},$$

where μ is a normed measure on $(\mathfrak{X})$, $\mu(X) = 1$, then μ is said to be a *canonical distribution*. This term is motivated since all the distributions $D(t; f)$ can be obtained as the distribution of $f(x)$, when x is distributed over X according to the measure μ. For the classical Toeplitz forms Szegö's result is equivalent to saying that the matrices have the canonical distribution $\mu = 1/2\pi$ Lebesgue measure. Results of this type are of importance in various branches of pure and applied mathematics.

2. **Approximation of Toeplitz matrices.** As the eigenvalues are uniformly bounded, $m \leqq \lambda_{\nu}^{(n)} \leqq M$; $\nu = 0, 1, \cdots, n-1$, $n = 1, 2, \cdots$, we can describe the eigenvalue distributions completely by their moments

$$\mu_p^{(n)}(f) = \frac{1}{n} \sum_{\nu=0}^{n-1} \lambda_{\nu}^{(n)p} = \int_m^M t^p dD_n(t; f).$$

If we can show that there is a distribution $D(t)$ with moments

$$\mu_p = \int_m^M t^p dD(t),$$

such that for all positive integers p

(2)
$$\lim_{n \to \infty} \mu_p^{(n)}(f) = \mu_p,$$

then it is clear that the eigenvalues of $M_n(f)$ have $D(t)$ as an asymptotic distribution.

It will be of considerable help that in a rather general case it is not necessary to verify (2) for every p but only for $p = 1$. Before we discuss this, however, we shall have to consider questions of approximation of Toeplitz matrices.

We want to introduce a metric in the linear space of all $n \times n$ matrices. One that is commonly used is

$$\|A\| = \max_{\|z\|=1} | z^* A z |.$$

For our purpose it is convenient to use the definition

$$| A | = \left\{ \frac{1}{n} \sum_{\nu,\mu=0}^{n-1} | a_{\nu\mu} |^2 \right\}^{1/2} = \left\{ \frac{1}{n} \operatorname{tr} A A^* \right\}^{1/2}.$$

In terms of the eigenvalues we have, if the matrices are Hermitian,

$$\|A\| = \max_{\nu} | \lambda_\nu |, \qquad | A |^2 = \frac{1}{n} \sum_{\nu=0}^{n-1} \lambda_\nu^2.$$

Let us first prove

LEMMA 1. *Let A, B, C be Hermitian $n \times n$ matrices. Then*

$$\left| (1/n)\operatorname{rt}(ABC) \right|^2 \leq \|A\|^2 | B |^2 \|C\|^2.$$

Proof. First it is clear that if G is any $n \times n$ matrix, Hermitian or not, with the eigenvalues λ_ν

$$\det (G - \lambda_\nu I) = 0, \qquad \nu = 0, 1, \cdots, n - 1,$$

then

(3)
$$\left| (1/n) \operatorname{tr} G \right| \leq | G |.$$

Indeed we can find a unitary matrix U so that the transformed matrix U^*GU is of superdiagonal form

$$U^*GU = \{ \gamma_{\nu\mu}; \nu, \mu = 0, 1, \cdots, n - 1 \},$$
$$\gamma_{\nu\mu} = 0 \text{ if } \mu < \nu.$$

As the trace of the matrix is left unchanged and as the eigenvalues of G are $\gamma_{00}, \gamma_{11}, \cdots, \gamma_{n-1\,n-1}$ we find by applying Schwarz inequality

$$\left| \frac{1}{n} \operatorname{tr} G \right|^2 = \left| \frac{1}{n} \sum_0^{n-1} \gamma_{\nu\nu} \right|^2 \leqq \frac{1}{n} \sum_0^{n-1} | \gamma_{\nu\nu} |^2.$$

But

$$|G|^2 = \frac{1}{n} \operatorname{tr} GG^* = \frac{1}{n} \operatorname{tr} UGU^*UG^*U^* = \frac{1}{n} \sum_{\nu,\mu=0}^{n-1} | \gamma_{\nu\mu} |^2 \geqq \frac{1}{n} \sum_0^{n-1} | \gamma_{\nu\nu} |^2$$

which proves (3).

If G is Hermitian then

$$|GH|^2 = \frac{1}{n} \sum_{\nu,\mu=0}^{n-1} \left| \sum_{k=0}^{n-1} g_{\nu k} h_{k\mu} \right|^2 = \frac{1}{n} \sum_{\nu,\mu,l,k} g_{\nu k} \bar{g}_{\nu l} h_{k\mu} \bar{h}_{l\mu} = \frac{1}{n} \sum_{\mu} \sum_{l,k} f_{lk} h_{k\mu} \bar{h}_{l\mu},$$

where

$$G^2 = \{ f_{lk}; l, k = 0, 1, \cdots, n-1 \}.$$

Hence

$$|GH|^2 \leqq \|G\|^2 \frac{1}{n} \sum_{k,\mu} | h_{k\mu} |^2 = \|G\|^2 | H |^2.$$

Repeated use of this together with (3) proves the lemma.

Now we can easily dominate the norms of Toeplitz matrices. Indeed we have

LEMMA 2.

$$\|M_n(f)\| \leqq \sup_x \operatorname{ess} | f(x) |,$$

$$| M_n(f) | \leqq \sup_x \operatorname{ess} | f(x) |.$$

Proof. The first relation has already been proved in 1. To prove the second one we use Parseval's relation

$$\sum_{\mu=0}^{n-1} | m_{\nu\mu}(f) |^2 \leqq \sum_{\mu=0}^{\infty} | (\phi_\nu f; \phi_\mu) |^2 = \|\phi_\nu f\|^2 \leqq \sup_x \operatorname{ess} | f(x) |^2$$

so that

$$| M_n(f) |^2 = \frac{1}{n} \sum_{\nu,\mu=0}^{n-1} | m_{\nu\mu}(f) |^2 \leqq \sup_x \operatorname{ess} | f(x) |^2$$

and the result follows directly. The Toeplitz matrices are continuous transforms of f in both the metrics used.

It may also be noted that these transforms are monotonic. Consider e.g. the finite case $M_n(f)$. Then if $f(x) \leqq g(x)$, $x \in X$, we have

$$z^* M_n(f) z - z^* M_n(g) z = z^* M_n(f - g) z$$

$$= \int_X \left| \sum z_\nu \phi_\nu(x) \right|^2 [f(x) - g(x)] dm(x) \leq 0$$

so that we have $M_n(f) \leq M_n(g)$.

We finally prove a third lemma which implies that the limiting eigenvalue distribution of a sequence of Hermitian matrices depends continuously upon the matrices if the second metric is used.

LEMMA 3. *Let M_n, $n = 1, 2, \cdots$, be a sequence of Hermitian matrices with eivenvalue distribution $D_n(t)$ belonging to a finite interval. If there is another such sequence N_n with an asymptotic eigenvalue distribution D and if*

$$\lim_{n \to \infty} \left| M_n - N_n \right| = 0,$$

then M_n has also an asymptotic eigenvalue distribution and this coincides with D.

Proof. Writing $M_n = N_n + \Delta_n$ we have for any positive integer p

$$M_n^p = N_n^p + R_n,$$

where R_n is a sum of a number (not depending on n) of terms of the type $A_n \Delta_n B_n$, where A_n and B_n are Hermitian uniformly bounded matrices. Using Lemma 1 we obtain

$$\mu_p^{(n)} = \frac{1}{n} \sum_{\nu=0}^{n-1} \lambda_\nu^{(n)p} = \frac{1}{n} \operatorname{tr} M_n^p = \frac{1}{n} \operatorname{tr} N_n^p + o(1)$$

if $\lambda_\nu^{(n)}$ are the eigenvalues of M_n. But the eigenvalues $l_\nu^{(n)}$ of N_n have moments $m_p^{(n)} = (1/n) \sum_{\nu=0}^{n-1} l_\nu^{(n)p}$ that converge, say to m_p, as n tends to infinity. But $\lim_{n \to \infty} \mu_p^{(n)} = m_p$ implies that $D_n \to D$, which proves the lemma.

3. **Trace complete families of Toeplitz matrices.** Let us suppose in this section that we have already proved for $M_n(f)$ in some way the existence of $\lim_{n \to \infty} \mu_1^{(n)}(f) = \mu_1(f)$ for every f in the class of functions C that we want to consider. We shall see below how this can be done in a fairly simple way in many situations To obtain a complete result we would then have to deal with the moments of higher order, which will involve some cumbersome calculations. Sometimes we can avoid this in the following way.

Let f and g be two functions in the set C, which will always be chosen as a subclass of all real valued bounded and ($\mathfrak{X}$)-measurable functions; we consider the associated Toeplitz matrices $M(f)$, $M(g)$. As Parseval's relation holds

$$\{ M(f) M(g) \}_{\nu\mu} = \sum_{K=0}^{\infty} m_{\nu K}(f) m_{K\mu}(g) = \sum_{K=0}^{\infty} (f\phi_\nu; \phi_K)(g\phi_\mu; \phi_K)^*$$

$$= (f\phi_\nu; g\phi_\mu) = \{ M(fg) \}_{\nu\mu},$$

or

(4) $$M(f)M(g) = M(fg).$$

This is a relation between infinite Toeplitz matrices. No such relation holds in general for the finite matrices $M_n(f)$. But for an important class something similar holds and it is convenient to introduce the following.

DEFINITION. A class of Toeplitz matrices $M_n(f)$, $f \in A$, where A is a given set of functions, closed under multiplication, is said to be trace complete if for any $f_1, f_2, \cdots, f_r \in A$, one has

$$\lim_{n \to \infty} \frac{1}{n} \operatorname{tr} \left\{ M_n(f_1) M_n(f_2) \cdots M_n(f_r) - M_n(f_1 f_2 \cdots f_r) \right\} = 0.$$

THEOREM 1. *Suppose that for all the matrices $M_n(f)$, $f \in A$, the first order eigenvalue moments exist*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{r=0}^{n-1} \lambda_r^{(n)} = \mu_1(f),$$

and that the said matrices form a trace complete class. Then these matrices have asymptotic eigenvalue distributions with moments m_p given by

$$m_p = \mu_1(f^p).$$

Proof. The proof is immediate as

$$\mu_p^{(n)}(f) = \frac{1}{n} \operatorname{tr} M_n^p(f) = \frac{1}{n} \operatorname{tr} M_n(f^p) + o(1) = \mu_1^{(n)}(f^p) + o(1)$$

and $\lim_{n \to \infty} \mu_p^{(n)}(f) = \mu_1(f^p)$. In spite of its simplicity this result reduces the problem dealt with considerably in most cases.

We now have to find a way to determine whether a given class of Toeplitz matrices is trace complete or not. The condition in the following theorem is probably far from necessary, but the criterion will suffice, when we study the matrices associated with some of the classical orthonormal systems.

THEOREM 2. *If for every $f \in A$ we have*

$$m_{\nu\mu}(f) = O\left(\frac{1}{1 + (\nu - \mu)^2} \right),$$

then $\left\{ M_n(f); f \in A \right\}$ is a trace complete class.

Proof. We have to consider two sums of the type

(5) $$\sum m_{\nu_1\nu_2}(f_1) m_{\nu_2\nu_3}(f_2) \cdots m_{\nu_p\nu_1}(f_p).$$

In one sum, S_n, the summation is extended over all indices between 0 and $n-1$,

$$0 \leqq \nu_j \leqq n - 1; \qquad\qquad j = 1, 2, \cdots, p.$$

Of course $S_n = \mathrm{tr}\ M_n(f_1) M_n(f_2) \cdots M_n(f_p)$. The other sum, T_n, is obtained by summing over

$$0 \leqq \nu_1 \leqq n - 1, \qquad 0 \leqq \nu_j < \infty\ ; \qquad i = 2, 3, \cdots, p.$$

Using (4) we see that this is $T_n = \mathrm{tr}\ M_n(f_1 f_2 \cdots f_p)$.

Consider the difference $T_n - S_n$. For $p = 1$ we have $T_n = S_n$ so that we need only consider $p > 1$. It can be written as a finite and fixed number of terms, each of which is of the form (5); here the summation extends over $0 \leqq \nu_1 \leqq n - 1$, at least one of the other subscripts runs from n to infinity and the remaining subscripts run from 0 to infinity. E.g. in the case $p = 2$ we have

$$T_n - S_n = \sum_{0 \leqq \nu_1 \leqq n-1, n \leqq \nu_2 < \infty} m_{\nu_1 \nu_2}(f_1) m_{\nu_2 \nu_1}(f_2),$$

and for larger values of p we obtain more terms of the type described above.

The indices in the third category are easily dealt with. Since

$$\sum_{\mu=0}^{\infty} m_{K\mu}(f) m_{\mu l}(g) = m_{Kl}(fg), \qquad\qquad fg \in A,$$

$T_n - S_n$ consists of a number of terms of the type

$$\sum m_{\nu_1 \nu_2}(g_1) m_{\nu_2 \nu_3}(g_2) \cdots m_{\nu_q \nu_1}(g_q),$$

where

$$q \leqq p,$$
$$g_1, g_2, \cdots, g_q \in A,$$
$$0 \leqq \nu_1 \leqq n - 1,$$
$$n \leqq \nu_j < \infty, \qquad\qquad j = 2, 3, \cdots, q.$$

For $n \leqq 1 \leqq k$ we have

$$\left| \sum_{\mu=n}^{\infty} m_{k\mu}(g) m_{\mu l}(h) \right| \leqq \sum_{n \leqq \mu \leqq (k+1)/2} + \sum_{(k+1)/2 < \mu < \infty} \frac{c}{[1 + (\mu - k)^2][1 + (\mu - 1)^2]}$$

$$\leqq \frac{2c}{1 + ((l - k)/2)^2} \sum_{-\infty}^{\infty} \frac{1}{1 + \nu^2} = O\left(\frac{1}{1 + (k - 1)^2} \right),$$

so that we finally arrived at sums of the type

$$\sum a_{\nu k} b_{k\nu}, \qquad 0 \leqq \nu \leqq n - 1, n \leqq k < \infty,$$

with the a's and b's bounded by $O(1/(1 + (\nu - k)^2))$. Then Schwarz' inequality gives

$$\lim_{n\to\infty} \frac{1}{n}(T_n - S_n) = 0,$$

which is the relation that was to be proved.

Theorem 2 enables us to deal with situations where A is a sufficiently restricted subset of C How this is done will be seen later. At present we note only that it is often possible to carry over the result to the whole set C. We have namely

THEOREM 3. *Let A form an algebra over the field of real numbers and be everywhere dense in C, $C \subset \overline{A}$. Let $\mathfrak{A}$ be a basis in A. If we have proved for every $g \in \mathfrak{A}$ that*

(i) $\lim_{n\to\infty} \mu_1^{(n)}(g)$ *exists and is, say, $\mu_1(g)$,*

(ii) $m_{\nu\mu}(g) = O\left(\dfrac{1}{1 + (\nu - \mu)^2}\right),$

then every Toeplitz matrix $M_n(f)$, $f \in C$, has an asymptotic eigenvalue distribution with the moments $m_p = L(f^p)$. Here $L(f)$ is a linear functional in C coinciding with $\mu_1(f)$ for $f \in \mathfrak{A}$.

Proof. The conditions (i), (ii) are clearly true also for every $g \in A$ and Theorems 2 and 1 show that $M_n(f)$, $f \in A$, has the asymptotic eigenvalue distribution determined by the moments $m_p = \mu_1(f^p)$. The first order moment $\mu_1(f)$ is a linear functional defined in A. It is bounded, because we have from Lemma 2

$$\left| \mu_1(f) \right| = \left| \lim_{n\to\infty} \frac{1}{n} \operatorname{tr} M_n(f) \right| \leq \limsup_{n\to\infty} \max_{\nu} \left| \lambda_\nu^{(n)} \right|$$

$$= \limsup_{n\to\infty} \left\| M_n(f) \right\| \leq \sup_{x} \operatorname{ess} \left| f(x) \right| = \|f\|.$$

There exists a linear bounded functional $L(f)$ defined for $f \in C$ and coinciding with $\mu_1(f)$ in A.

For a given $f \in C$ we choose functions $g_n \in A$ so that $\|f - g_n\| \leq 1/n$, $n = 1, 2, \cdots$. Then $M_n(g_n)$ is a sequence of Hermitian matrices of uniformly bounded norm $\|M_n(g_n)\| \leq \|f\| + 1$. But for any $p \geq 1$ we have

$$\frac{1}{n} \operatorname{tr} M_n^p(g_n) = \frac{1}{n} \operatorname{tr} M_n^p(g_m) + O(\|g_n - g_m\|).$$

Keeping m fixed we let n tend to infinity

$$\limsup_{n\to\infty} \left| \frac{1}{n} \operatorname{tr} M_n^p(g_n) - L(g_m^p) \right| = O(\|f - g_m\|).$$

so that

$$\lim_{n \to \infty} \frac{1}{n} \operatorname{tr} M_n^p(g_n) = L(f^p).$$

The assumptions of Lemma 3 are now satisfied if we choose $N_n = M_n(g_n)$ and the result follows.

4. Determination of the functional L. The general procedure dealing with problems of the type studied in this paper will be the following:

(i) Choose C, A, and $\mathfrak{A}$ in a way to simplify the formal manipulations.

(ii) Show that $M_n(g)$, $g \in A$, is a trace complete class of matrices, e.g., by applying Theorem 2.

(iii) Determine $\mu_1(g)$, $g \in A$, and extend the functional to C.

(iv) Calculate the limiting eigenvalue distribution with the moments $m_p = L(f^p)$. Find out whether these distributions are generated by a canonical distribution.

So far we have not considered the question how the existence of $\mu_1(g)$ should be proved or how this functional should be computed. Let us consider first two cases in which this can be done in a straightforward manner.

Consider the well-known complete and orthonormal system of Haar defined by (see e.g. Kaczmarz-Steinhaus, *Theorie der Orthogonalreihen*, p. 44)

$$\chi_0^{(0)}(x) = 1, \qquad\qquad 0 < x < 1,$$

$$\chi_0^{(1)}(x) = \begin{cases} 1, & 0 < x < 1/2, \\ -1, & 1/2 < x < 1, \end{cases}$$

$$\chi_1^{(1)}(x) = \begin{cases} 2^{1/2}, & 0 < x < 1/4, \\ -2^{1/2}, & 1/4 < x < 1/2, \\ 0, & 1/2 < x < 1, \end{cases}$$

$$\chi_1^{(2)}(x) = \begin{cases} 0, & 0 < x < 1/2, \\ 2^{1/2}, & 1/2 < x < 3/4, \\ -2^{1/2}, & 3/4 < x < 1, \end{cases}$$

$$\cdots \cdots \cdots,$$

$$\chi_m^{(k)}(x) = \begin{cases} 2^{m/2}, & \dfrac{2k-2}{2^{m+1}} < x < \dfrac{2k-1}{2^{m+1}}, \\ -2^{m/2}, & \dfrac{2k-1}{2^{m+1}} < x < \dfrac{2k}{2^{m+1}}, \\ 0, & \text{otherwise}, \end{cases}$$

$$k = 1, 2, 3, \cdots, 2^m,$$

$X = (0, 1)$ and $d\mu$ is Lebesgue measure.

It will be convenient to use a simple index ν for the enumeration of these indices, so that we have the functions $\phi_\nu(x)$, $\nu = 1, 2, \cdots, n$, where $n = 2 + 2 + 2^2 + \cdots + 2^m = 2^{m+1}$. Note that we have not completely specified the order for the $\chi_m^{(k)}(x)$ because in each group with fixed m the order has not been defined. We shall return to this point later. Let the midpoints $(s+1/2)/2^{m+1}$ of the intervals $(s/2^{m+1}, (s+1)/2^{m+1})$ be denoted by x_s, $s = 0, 1, \cdots, n-1$.

Let us choose C as the set of all real-valued continuous functions on the interval $(0, 1)$. Take $\mathfrak{A}$ as the set $\{\phi_\nu(x), \nu = 1, 2, \cdots\}$ and A as the corresponding linear hull. Of course $C \subset \overline{A}$.

It is easy to show that $\{M_n(g); g \in A\}$ is a trace complete class. Indeed if $g \in \mathfrak{A}$, that is,

$$g(x) = \chi_{m_0}^{(j)}(x),$$

then $\phi_\nu(x)g(x) = \text{const.} \cdot \phi_\nu(x)$ if $\nu > 2^{m_0+1}$. Hence

$$m_{\nu\mu}(g) = \int_0^1 \phi_\nu(x)\phi_\mu(x)g(x)dx = 0,$$

if $\nu, \mu > 2^{m_0+1}$ and $\nu \neq \mu$, and the condition of Theorem 2 is satisfied.

To determine L we note that

$$\mu_1^{(n)}(g) = \frac{1}{n} \sum_{\nu=1}^n \int_0^1 \phi_\nu^2(x)g(x)dx.$$

But if $\nu > 2^{m_0+1}$ then

$$\int_0^1 \phi_\nu^2(x)g(x)dx = g(x_s),$$

where x_s is the midpoint of the interval where $\phi_\nu(x) \neq 0$. Now it is clear that when $n = 2^m$ increases, then the points x_s will tend to be more and more uniformly distributed over $(0, 1)$. As $g(x)$ is a step-wise continuous function the Riemann sums will converge and we find

$$\lim_{n \to \infty} \mu_1^{(n)}(g) = \int_0^1 g(x)dx,$$

so that the completed functional becomes

$$L(f) = \int_0^1 f(x)dx.$$

Now we only have to find the distribution with moments

$$m_p = \int t^p dD(t) = L(f^p) = \int_0^1 f^p(x)dx.$$

It is evident that $D(t) = m\{x \mid f(x) \leq t\}$, in other words, there is a canonical distribution, viz. Lebesgue measure on $(0, 1)$.

The reader may have noted that we let n run through only the integers of the form 2^{m+1}, $m = 1, 2, \cdots$. This restriction can be removed easily, but one has then to specify the ordering of the $\chi_m^{(k)}$ in each group with subscript m. This can be done so that we still arrive at the same canonical distribution. It is not difficult to see that by ordering the ϕ_ν's differently we can obtain a situation where not only no canonical distribution exists, but the eigenvalue distributions diverge unless $g(x)$ is identically constant.

The Toeplitz matrices of Haar type show a behavior similar to that of the classical Toeplitz matrices mentioned in the first section. In both cases the canonical distributions are uniform over the intervals $(0, 1)$ and $(0, 2\pi)$ respectively.

The result is different when we start from the Legendre polynomials,

$$P_\nu(x) = \frac{1}{2^\nu \nu!} \frac{d^\nu}{dx^\nu} (x^2 - 1)^\nu,$$

or after norming

$$\phi_\nu(x) = \left(\nu + \frac{1}{2}\right)^{1/2} P_\nu(x).$$

Here $X = (-1, 1)$ and $d\mu$ is Lebesgue measure. Let us take for C all real continuous functions on $(-1, 1)$ and, for $\mathfrak{A}$, the set $\{x^\nu, \nu = 0, 1, \cdots\}$ so that A consists of all polynomials with real coefficients.

If $g \in \mathfrak{A}$, $g(x) = x^r$, then $g(x)\phi_\nu(x)$ is a polynomial of order $\nu + r$ and hence orthogonal to $\phi_\mu(x)$ if $\mu > \nu + r$. As

$$m_{\nu\mu}(g) = \int_{-1}^{1} \phi_\nu(x)\phi_\mu(x)g(x)dx = 0$$

if $|\nu - \mu| > r$ it is clear that Theorem 2 is immediately applicable.

The easiest way to find L is to use the relation

$$P_\nu^2(x) = \sum_{s=0}^{\nu} \frac{A_{\nu-s}^2 A_s}{A_{2\nu-s}} \frac{4\nu - 4s + 1}{4\nu - 2s + 1} P_{2\nu-2s}(x),$$

where

$$A_m = \frac{1 \cdot 3 \cdot 5 \cdots (2m - 1)}{m!},$$

(see Whittaker-Watson p. 331). If $g \in A$ it can be written as a linear combination of P_ν's and it will be sufficient to deal with $g(x) = P_l(x)$. If l is odd,

$m_{\nu\nu} = \int_{-1}^{2} \phi_\nu^1(x) P_1(x) dx = 0$. Otherwise, $l = 2k$, we have, using the above relation and the orthonormality of the ϕ's,

$$m_{\nu\nu}(g) = \frac{2\nu + 1}{2} A_k^2 \frac{A_{\nu-k}}{A_{\nu+k}} \frac{4k + 1}{2\nu + 2k + 1} \frac{2}{4k + 1}$$

$$\sim A_k^2 \frac{(\nu - k + 1)(\nu - k + 2) \cdots (\nu + k)}{(2\nu - 2k + 1)(2\nu + 2k + 3) \cdots (2\nu + 2k - 1)} \sim \frac{A_k^2}{2^{2k}}$$

as ν tends to infinity. On the other hand we have (see Erdélyi [1, p. 183])

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} P_{2k}(\cos \theta) d\theta = \frac{A_k^2}{2^{2k}},$$

so that

$$\lim_{\nu \to \infty} m_{\nu\nu}(g) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_{2k}(\cos \theta) d\theta.$$

Hence

$$\mu_1(g) = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\cos \theta) d\theta = \frac{1}{2} \int_{-1}^{1} g(x) \frac{dx}{(1 - x^2)^{1/2}}.$$

The functional L has the same form and the limiting eigenvalue distribution has the moments

$$m_p = L(f^p) = \frac{1}{2} \int_{-1}^{1} f^p(x) \frac{dx}{(1 - x^2)^{1/2}}; \qquad f \in C.$$

In other words there is a canonical distribution μ absolutely continuous with respect to Lebesgue measure on 1 with the density $\mu'(x) = 1/2(1 - x^2)^{1/2}$.

   5. **Existence of limiting and canonical distributions.** In the last section we showed that $\mu_1^{(n)}(g)$ converges to $\mu_1(g)$ in two special cases and how L can be determined. We shall now consider a method of more general applicability. For this purpose we shall take advantage of the fact that the function

$$K(x; z) = \sum_{\nu=0}^{\infty} |\phi_\nu(x)|^2 z^\nu$$

has been explicitly evaluated for most of the classical orthogonal systems (see Watson [7] and Meixner [3]) inside the domain of convergence of the series. We can then compute the function

(6) $$G(g; z) = \sum_{\nu=0}^{\infty} z^\nu \int_X |\phi_\nu(x)|^2 g(x) dm(x) = \sum_{\nu=0}^{\infty} z^\nu m_{\nu\nu}(g), \qquad g \in \mathfrak{A}.$$

As g is bounded we can always assume that $0 < m \leq g(x) \leq M < \infty$ after adding

a constant to g which affects the eigenvalues only by adding the same constant to them. But then

$$m \leq m_{\nu\nu}(g) \leq M,$$

so that $G(g; z)$ is clearly regular inside the unit circle $|z| < 1$ and has a singularity at the point $z = 1$. Let us assume that this is a pole; of course this has to be investigated from case to case. As $G(g; z) \leq M/(1 - |z|)$; this pole is then of order one. We obtain

THEOREM 4. *Assume that for each $g \in \mathfrak{A}$ the function $G(g; z)$ has a pole at $z = 1$ and denote the residue by $R(g)$. If $\{M_n(g); g \in \mathfrak{A}\}$ is a trace complete class, then the Toeplitz matrices $M_n(f)$, $f \in C$, have limiting eigenvalue distributions with the moments $m_p = L(f^p)$, where L is the linear functional in C obtained by extending $R(g)$.*

Proof. As $\{M_n(g); g \in \mathfrak{A}\}$ was a trace complete class, Theorem 3 tells us almost immediately that we need only show the convergence of $\mu_1^{(n)}(g)$ to $R(g^p)$. But the coefficients of the power series (6) are positive and $G(g; z) \sim R(g)/(1 - z)$ when $z \to 1$, say taking real values. Then a famous theorem of Hardy-Littlewood (see [2]) implies that

$$\frac{1}{n} \mu_1^{(n)}(g) = \frac{1}{n} \operatorname{tr} M_n(g) = \frac{1}{n} \sum_{\nu=0}^{n-1} m_{\nu\nu}(g) \sim R(g),$$

and the proof is completed.

In a situation that is more special but still of some generality we can show more, viz. the existence of a canonical distribution.

THEOREM 5. *Let X be a compact set in a finite-dimensional Euclidean space and let the functions in C be continuous. The existence of limiting eigenvalue distributions for the trace complete Toeplitz matrices $M_n(g) \in \mathfrak{A}$ implies the existence of a canonical distribution.*

Proof. Just as before $\mu_1(g)$ can be extended to a bounded linear functional L defined on C. But then according to a theorem of F. Riesz [4] L can be written as

$$L(f) = \int_X f(x) dW(x),$$

where W is a set function of bounded variation on X. But as $L(f)$ is a monotonic functional (see §1) it is clear that W is also monotonic. On the other hand for $f(x) \equiv 1$ we have $M_n(1) = I$ and $L(1) = 1$ so that the total variation of W is 1. Hence W is a measure normed to one on X, and as the limiting eigenvalue moments are given by

$$m_p = L(f^p) = \int_X f^p(x) dW(x), \qquad p = 1, 2, \cdots,$$

it is clear that the corresponding limiting distribution is

$$D(t) = W\{x \mid f(x) \leqq t\},$$

or in other words, W is a canonical distribution.

6. **Toeplitz forms of Hermite type.** Let us now consider the case when $X = (-\infty, \infty)$, $d\mu(x) = e^{-x^2}dx$, and where the ϕ's are the normalized Hermite polynomials

$$\phi_\nu(x) = \frac{e^{x^2}}{\pi^{1/4}2^{\nu/2}(\nu!)^{1/2}} \frac{d^\nu}{dx^\nu} e^{-x^2}.$$

If we choose C as $L(-\infty, \infty)$ or some other set of functions tending to zero or some other definite value on the average as $x \to \pm\infty$, we can show that the eigenvalue distributions converge but the result will be rather trivial. Instead we shall choose C as the set of all real-valued functions which are almost periodic (in Bohr's sense) on $(-\infty, \infty)$. We take for $\mathfrak{A}$ the set $\{ce^{i\lambda x} + \overline{c}e^{i\lambda x};$ c complex, λ real$\}$ so that A will consist of all finite and real trigonometric "polynomials."

The author has not been able to prove that this is a trace complete class. Still if we assume that this has been done, we could proceed to the derivation of $\mu_1(g)$. One knows that for $|z| < 1$

$$K(x, z) = \sum_{\nu=0}^\infty \phi_\nu^2(x)z^\nu = \frac{1^\cdot}{(\pi(1-z^2))^{1/2}} \exp\left\{\frac{2x^2z}{1+z}\right\},$$

(see Erdélyi [1, p. 194]). Then using Lebesgue's theorem on dominated convergence we find

$$G(e^{i\lambda x}; z) = \int_{-\infty}^\infty K(x; z)\bar{e}^{x^2+i\lambda x}dx$$

$$= \frac{1}{(\pi(1-z^2))^{1/2}}\int_{-\infty}^\infty \exp\left\{-\frac{x^2(1-z)}{1+z} + i\lambda x\right\}dx$$

$$= \frac{1}{1-z}e^{-\lambda^2\frac{1+z}{1-z}}.$$

If $\lambda = 0$ then this has the residue 1 at $z = 1$, otherwise $R(e^{i\lambda x}) = 0$. Hence.

$$R(e^{i\lambda x}) = \delta_{\lambda 0},$$

so that for any $g \in \mathfrak{A}$ we see that $R(g)$ is simply the coefficient of the constant term of the trigonometric polynomial. This can be written

$$R(g) = \lim_{A\to\infty} \frac{1}{2A}\int_{-A}^A g(x)dx,$$

and the completed functional L has exactly the same form. Hence the limiting eigenvalue moments are given by

$$m_p = L(f^p) = \lim_{A \to \infty} \int_{-A}^{A} f^p(x)dx,$$

so that the asymptotic eigenvalue distribution D is

$$D(t) = \lim_{A \to \infty} \frac{\text{Lebesgue measure of } \{x \mid f(x) \leq t, \mid x \mid < A\}}{2A}.$$

These distributions are not generated by any canonical distribution.

7. **A case of harmonic polynomials.** An example of a different character is obtained by choosing for X the unit circle $|x| \leq 1$, for dm the Lebesgue area, and for the orthonormal system:

$$\phi_0(x) = (1/\pi)^{1/2},$$

$$\phi_{2\nu}(x) = ((2\nu + 1)/\pi)^{1/2} r^\nu \cos \nu\theta, \qquad \nu = 1, 2, \cdots,$$

$$\phi_{2\nu+1}(x) = ((2\nu + 1)/\pi)^{1/2} r^\nu \sin \nu\theta, \qquad \nu = 1, 2, \cdots,$$

where $x = re^{i\theta}$. Let us first take $\mathfrak{A}$ as the set of functions $\{r^\nu \cos \nu\theta, r^\nu \sin \nu\theta; \nu = 0, 1, \cdots\}$, so that C consists of all real-valued functions harmonic in the closed unit circle. For $g \in \mathfrak{A}$, say $g(x) = r^p \sin p\theta$, it is clear that $m_{\nu\mu}(g) = 0$ for $|\nu - \mu| > p$ and similarly for $g(x) = r^p \cos p\theta$, so that Theorem 1 is applicable, showing that $\{M_n(g); g \in \mathfrak{A}\}$ is a trace complete class of Toeplitz matrices.

It is clear that for $g(x) = r^p \cos p\theta$

$$m_{2\nu,2\nu}(g) = \frac{2\nu + 1}{\pi} \int_{r=0}^{1} \int_{\theta=0}^{2\pi} r^{2\nu+1+p} \cos^2 \nu\theta \cos p\theta dr d\theta$$

$$= \frac{2\nu + 1}{(2\nu + 2 + p)\pi} \int_{0}^{2\pi} \cos^2 \nu\theta \cos p\theta d\theta \to \frac{1}{2\pi} \int_{0}^{2\pi} \cos p\theta d\theta$$

$$= \frac{1}{2\pi} \int_{0}^{\pi} g(e^{i\theta})d\theta,$$

as $\nu \to \infty$. Then

$$\lim_{n \to \infty} \mu_1^{(n)}(g) = \frac{1}{2\pi} \int_{0}^{\pi} g(e^{i\theta})d\theta$$

and according to Theorem 3 the eigenvalue distributions converge to the one with the moments $m_p = L(f^p)$ where the functional L is given by

$$(7) \qquad\qquad L(f) = \frac{1}{2\pi} \int_{0}^{2\pi} f(e^{i\theta})d\theta = f(0)$$

as the functions in C are harmonic in $|x| \leqq 1$. This means that there exists a canonical distribution (as we already know from Theorem 5) with all its variation concentrated in the single point $x = 0$.

8. Summary. We have seen that the eigenvalue distributions of Toeplitz matrices $M_n(f)$ can behave in various ways for large values of n. While in some cases (associated with Haar functions enumerated in a certain order) the distributions diverge, one may conjecture that these are also cases when the distributions converge but no canonical distribution exists. Finally there are cases when the canonical distribution exists and is absolutely continuous (for the classical Toeplitz matrices, those of Legendre type and others) or discrete (the case in 7).

For the important family of trace complete Toeplitz matrices we have shown that it is necessary to consider only the first order eigenvalue moments $\mu_1^{(n)}(f)$ as n tends to infinity. If this limit exists in a sufficiently wide class of functions $f \in A$, then the limiting distribution exists altogether. If moreover X and C have the properties of Theorem 5 there is a canonical distribution.

One direct and one indirect method of determining L have been described, but the possibilities are far from exhausted, and in more complicated situations other ways have probably to be devised.

BIBLIOGRAPHY

1. A. Erdélyi, ed., *Higher transcendental functions*, New York, 1953.
2. G. H. Hardy and J. E. Littlewood, *Tauberian theorems*, Proc. London Math. Soc. vol. 13 1914).
3. J. Meixner, *Erzeugende Funktionen der Charlierschen Polynome*, Math. Zeit. vol. 44 (1939).
4. F. Riesz and B. Sz. Nagy, *Leçons d'analyse fonctionelle*, Budapest, 1952.
5. G. Szegö, *Beiträge zur Theorie der Toeplitzschen Formen*, Math. Zeit. vol. 6 (1920) and vol. 9 (1921).
6. O. Toeplitz, *Zur Theorie der quadratischen und bilinearen Formen von unendlichvielen Veränderlichen*, Math. Ann. vol. 70 (1911).
7. G. N. Watson, *Notes on generating functions of polynomials*. J. London Math. Soc. vol. 8 (1933) and vol. 9 (1934).
8. E. T. Whittaker and G. N. Watson, *A course of modern analysis*, Cambridge, 1935.

UNIVERSITY OF STOCKHOLM,
    STOCKHOLM, SWEDEN.