SLAC-PUB-422
June 1968

TOEPLITZ DETERMINANTS AS GROUP AVERAGES*

C. Itzykson †
Stanford Linear Accelerator Center, Stanford, California

ABSTRACT

We use the expression of a Toeplitz determinant as an average over the unitary group of the same dimension, to exhibit its asymptotic behaviour, recovering an expression first derived by Szegö and Kac. Further related problems are suggested.

(Presented at the Symposium on Analytic Methods in Mathematical Physics, University of Indiana, Bloomington, Indiana, June 3, 1968.)

*Work supported by the U. S. Atomic Energy Commission.

†On leave of absence from: Department de Physique Theoretique, CEN Saclay, BP No. 2, Gif-sur-Yvette, France.

I

Toeplitz determinants arise in various problems of physics, a famous example being the computation of the correlation function in the two-dimensional Ising model. It turns out that such determinants can be expressed as averages over the unitary group of the same dimension. We shall make use of this remark to derive a number of their properties. The asymptotic behaviour of these determinants was first established rigourously by Szegö[1] in the case of positive definiteness of the corresponding matrix, using the theory of orthogonal polynomials with which Toeplitz determinants are intimately connected. Alternatively Kac[2] used combinatorial analysis and probability theory to derive the result in the general case. We shall see that it emerges quite naturally in the present context.

II

Let $a_0, a_1, a_{-1}, a_2, a_{-2} \ldots,$ be an infinite sequence of complex numbers to which we attach the formal generating function:

$$f\,(z) = \sum_{-\infty}^{+\infty} a_n z^n \tag{1}$$

The m-th Toeplitz determinant of $f\,(z)$, which we denote by $D_m(f)$, is defined as:

$$D_m(f) = \det \left\{ a_{p-q} \; ; \; 0 \leq p, \; q \leq m-1 \right\} \quad . \tag{2}$$

It is clear that as long as we let m vary on a finite range, say $m \leq N$, we can limit the series (1) by dropping all terms with $|n|$ larger than N, in which case it reduces to a finite Laurent series.

III

We now recall the introduction of an invariant integration over the group U_m of unitary $m \times m$ matrices.[3] Each such unitary matrix U can be diagonalized in the form:

$$U = V\,\delta\,V^{-1} \quad , \quad V \text{ unitary}, \ \delta = \begin{bmatrix} e^{i\phi_o} & 0 & 0 \\ 0 & e^{i\phi_2} & \\ 0 & & e^{i\phi_{m-1}} \end{bmatrix} \quad , \tag{3}$$

the order of the eigenvalues being immaterial. In the general case, when the eigenvalues are all distinct, V is arbitrary to the extent of multiplication on the right by a permutation matrix (a discrete set of operations) and by a diagonal unitary matrix. These matrices form a group equal to the direct product of m unitary groups in one dimension $\overset{m}{\otimes}\, U_1$. The invariant measure on U_m can then be factorized as:

$$dU = dV\, d\delta \ . \tag{4}$$

Here dV is a measure on the quotient space $U_m / \overset{m}{\otimes}\, U_1$, and $d\delta$ is a measure on the classes of U_m . Up to a constant factor, $d\delta$ can be written as:

$$d\delta = \frac{1}{m!} \left| \Delta \left(e^{i\phi_K}\right)\right|^2 \prod_{0\le i\le m-1} \frac{d\phi_i}{2\pi} \quad . \tag{5}$$

The symbol $\Delta\left(e^{i\phi}\right)$ stands for the Vandermonde determinant:

$$\Delta\left(e^{i\phi_K}\right) = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ e^{i\phi_o} & e^{i\phi_1} & \cdots & e^{i\phi_{m-1}} \\ e^{i2\phi_o} & e^{i2\phi_1} & \cdots & e^{i2\phi_{m-1}} \\ \vdots & & & \\ e^{i(m-1)\phi_o} & e^{i(m-1)\phi_1} & \cdots & e^{i(m-1)\phi_{m-1}} \end{vmatrix} = \prod_{0\le q<p\le m-1} \left(e^{i\phi_p} - e^{i\phi_q}\right).$$

If $g(U)$ is a scalar function defined on U_m, its mean value $\langle g \rangle_m$ is defined to be:

$$\langle g \rangle_m = \left[\int_{U_m} dU \right]^{-1} \int_{U_m} g(U)\, dV \quad . \tag{6}$$

The index m reminds that the average is taken on the group U_m. In particular if g is a class function: $g(U) = g(U_1 U U_1^{-1})$, it is a symmetric function of the roots $e^{i\phi_k}$ of U and one has:

$$\langle g \rangle_m = \int g(\delta)\, d\delta = \frac{1}{m!} \int_0^{2\pi(m)} g(\delta) \left| \Delta\left(e^{i\phi_k}\right)\right|^2 \prod_{0\le i\le m-1} \frac{d\phi_i}{2\pi} \quad , \tag{7}$$

since:

$$\int_0^{2\pi(m)} \left| \Delta\left(e^{i\phi_k}\right)\right|^2 \prod_{0\le i\le m-1} \frac{d\phi_i}{2\pi} = m! \quad .$$

IV

We return to the series (1) and assume the coefficients a_n to be interpreted as the Fourier coefficients of $f(z)$ for $|z| = 1$:

$$a_n = \int_0^{2\pi} \frac{d\phi}{2\pi}\, e^{-in\phi}\, f(e^{i\phi}) \quad . \tag{8}$$

In the computation of $D_n(f)$ one can use that fact that a determinant is a linear function of each of its rows, so that:

$$D_m(f) = \int_0^{2\pi(m)} \prod_{0\le k\le m-1} \left[\frac{d\phi_k}{2\pi}\, f\left(e^{i\phi_k}\right)\right] \begin{vmatrix} 1 & e^{-i\phi_0} & \cdots\cdots & e^{-i(m-1)\phi_0} \\ e^{i\phi_1} & 1 & \cdots\cdots & e^{-i(m-1)\phi_1} \\ \vdots & \vdots & & \vdots \\ e^{i(m-1)\phi_{m-1}} & & \cdots\cdots & 1 \end{vmatrix}$$

The determinant in the integrand is equal to

$$e^{i \sum_{0\le n\le m-1} n\phi_n} \quad \overline{\Delta\left(e^{i\phi_k}\right)} \quad .$$

We can further replace this expression by its average over all permutations of the dummy variables ϕ_k. Since under such a permutation $\Delta_P\left(e^{i\phi_k}\right)$ is left invariant except for the signature factor $(-1)^P$ of the permutation we obtain :

$$\frac{1}{m!}\left(\sum_P (-1)^P e^{i \sum_{0\le n\le m-1} n\,\phi_{P_n}}\right) \overline{\Delta\left(e^{i\phi_k}\right)} = \left|\Delta\left(e^{i\phi_k}\right)\right|^2 \quad .$$

Finally:

$$D_m(f) = \frac{1}{m!}\int_0^{2\pi\,(m)} \prod_{0\le k\le m-1}\left[\frac{d\phi_k}{2\pi} \quad f\left(e^{i\phi_k}\right)\right] \left|\Delta\left(e^{i\phi_k}\right)\right|^2 \quad . \tag{9}$$

We can use the series (1) to define on U_m the matrix valued function:

$$F(U) = \sum_{-\infty}^{+\infty} a_n \, U^n \quad , \tag{10}$$

with the same properties of convergence as $f(z)$ had. It is then recognized that:

$$\prod_{0\le k\le m-1} f\left(e^{i\delta_k}\right) = \det F(\delta) = \det F(V\delta V^{-1}) \quad .$$

Comparing equations (7) and (9) we have:

$$D_m(f) = \langle \det F(U) \rangle_m \quad . \tag{11}$$

One observes some immediate consequences. If $f(z)$ for $|z| = 1$ is real positive,then $D_m(f) > 0$. Furthermore, if $a < f(e^{i\theta}) < A$ then $a < \sqrt[m]{D_m(f)} < A$, hence the quantities $\sqrt[m]{D_m(f)}$ have at least one limit point.

V

Another series of remarks stems from the fact that one knows on the unitary group U_m a complete set of class functions, namely the characters of its irreducible representations (complete means for instance complete in the L^2-sense). Equation (1) can be considered as the scalar product of det F(U) with the identity character. Let $X_{\{\ell\}}$, where $\ell_o < \ell_1 < \ldots < \ell_{m-1}$ is an increasing sequence of integers, be the character of the corresponding irreducible representation of U_m:[3]

$$\Delta\left(e^{i\phi_k}\right) X_{\{\ell\}}\left(e^{i\phi_o}, \ldots, e^{i\phi_{m-1}}\right) = \sum_P (-1)^P \; e^{i \sum_{0\leq n\leq m-1} \ell_n \phi_{P_n}} \tag{12}$$

In this expression P stands for a permutation of the integrers $\{0, 1, \ldots, m-1\}$.
We can generalize the Toeplitz determinants by defining:

$$D_m^{\{\ell\}}(f) = \langle \overline{X}_{\{\ell\}}(U) \det F(U) \rangle_m = \det \left\{ a_{\ell_p - q} \; ; \; 0 \leq p, q \leq m-1 \right\} \tag{13}$$

The last equality is obtained by essentially reversing the steps which led to Eq. (11).
If det F(U) is square integrable one then has the expansion:

$$\left\{ \begin{array}{l} \det F(U) = f\left(e^{i\phi_o}\right) \ldots f\left(e^{i\phi_{m-1}}\right) = \sum_{\{\ell\}} D_m^{\{\ell\}}(f)\, X_{\{\ell\}}\left(e^{i\phi_o}, \ldots, e^{i\phi_{m-1}}\right) \\ D_m(|f|^2) = \sum_{\{\ell\}} \left| D_m^{\{\ell\}}(f) \right|^2 , \end{array} \right. \tag{14}$$

which yields in particular:

$$\left| D_m(f) \right|^2 \leq D_m(|f|^2).$$

By using similar arguments one can easily show that given a non-negative function $f(e^{i\theta})$ on the unit circle, the m-th orthonormalized polynomial in z with respect to the measure $f(e^{i\theta}) \frac{d\theta}{2\pi}$, is equal to:

$$P_m(z) = \left[D_m(f)\, D_{m+1}(f)\right]^{-1/2} \langle \det(z-U)\, f(U)\rangle_m , \tag{15}$$

so that the study of these polynomials is very similar to the study of Toeplitz determinants.

VI

In this section we compute $D_m(f)$ explicitely in the case where $f(z)$ enjoys the following properties:

(i) $f(e^{i\theta})$ is the inverse of a finite trigonometric polynomial (we shall call the "degree" of such a polynomial $\sum x_n e^{in\theta}$ the largest $n \geq 0$ such that $|x_n|^2 + |x_{-n}|^2 \neq 0$).

(ii) Once a choice of branch has been made at a given point, log $f(e^{i\theta})$ is a bounded and periodic function of θ .

Property (ii) is crucial for the asymptotic property of $D_m(f)$ in general. If $f(e^{i\theta})$ has this property so does $\frac{1}{f(e^{i\theta})}$. We need the following three lemmas:

Lemma 1 :

Let $g(e^{i\theta})$ be a trigonometric polynomial $g(e^{i\theta}) = \sum_{-n}^{+n} g_k c^{ik\theta}$ satisfying property (ii). There exist two polynomials in z of degree at most equal to n, $g_1(z)$ and $g_2(z)$ taking the value 1 at $z = 0$, and nonvanishing inside the closed unit circle $|z| \leq 1$, such that

$$g(e^{i\theta}) = \exp\left[\int_0^{2\pi} \frac{d\phi}{2\pi}\, g(e^{i\theta})\right] g_1(e^{i\theta})\, g_2(e^{-i\theta}) \quad . \tag{16}$$

This is an easy extension of a classical result of Fejer and Riesz.

Lemma 2 : (Cauchy)

$$\det\left\{\frac{1}{1-x_p y_q}\quad,\quad 1\le p,\ q\le N\right\} = \frac{\Delta(x_1,\ldots,x_N)\,\Delta(y_1,\ldots,y_N)}{\prod_{1\le p,q\le N}(1-x_p\,y_q)}$$

$$= \Delta(x_1,\ldots,x_N)\,\Delta(y_1,\ldots,y_N)\sum_{\substack{\{\ell\}\\ 0\le\ell_o<\ell_1..<\ell_{N-1}}} X_{\{\ell\}}(x_1,\ldots,x_N)X_{\{\ell\}}(y_1,\ldots,y_N) \tag{17}$$

For a proof of this we refer for instance to Ref. (3).

The next lemma combines the previous two and gives $D_m(f)$ for an $f(e^{i\theta})$ satisfying conditions (i) and (ii) when m is larger than the degree of the trigonometric polynomial $g(e^{i\theta}) = \frac{1}{f(e^{i\theta})}$. In view of lemma 1, we can write:

$$\left\{\begin{array}{l} f(z) = e^{h_o}\prod_{0\le k\le n-1}\dfrac{1}{(1-x_k z)(1-y_k z^{-1})} \\ \\ |x_k|<1 \quad,\quad |y_k|<1 \\ \\ h_o = \exp\left\{\displaystyle\int_0^{2\pi}\frac{d\phi}{2\pi}\ \log\ f(e^{i\phi})\right\} \end{array}\right. \tag{18}$$

The quantities $\{x_k\}$ and $\{y_k\}$ are the inverse of the roots of the polynomials $g_1(z)$ and $g_2(z)$; some of them may vanish.

Lemma 3 With f given by (18) and $m\ge n$:

$$D_m(f) = e^{mh_o}\prod_{0\le p,q\le n-1}\frac{1}{(1-x_p y_q)}\ . \tag{19}$$

The proof is straightforward. First we dispose of the factor e^{mh_o} by observing that $D_m(\lambda f) = \lambda^m D_m(f)$. Next we can clearly assume $m = n$ by adding factors with x_k or $y_k = 0$ in (18). So that it is sufficient to study $D_m(f)$ for $f(z) = \prod_{0\le k\le m-1}\frac{1}{(1-x_k z)(1-y_k z^{-1})}$; $|x_k|, |y|_k < 1$. Let δ be a diagonal matrix as in (3). By applying lemma 2 :

$$\det F(\delta) = \prod_{0\le p,q\le m-1} \frac{1}{(1-x_p e^{i\phi_q})(1-y_p e^{-i\phi_q})}$$

$$= \sum_{\substack{0\le \ell_0<\ell_1\cdots<\ell_{m-1}\\ 0\le \ell'_0<\ell'_1\cdots<\ell'_{m-1}}} X_{\{\ell\}}(x)X_{\{\ell\}}(\delta)\, X_{\{\ell'\}}(y)\, X_{\{\ell'\}}(\delta^{-1}) \quad .$$

The characters are orthonormal by integration on U_m, so that we reach the desired result:

$$D_m(f) = \sum_{0\le \ell_0<\ell_1\cdots<\ell_{m-1}} X_{\{\ell\}}(x)\, X_{\{\ell\}}(y) = \prod_{0\le p,q\le m-1} \frac{1}{(1-x_p y_q)} \quad .$$

We are now in position to state the main identity of this section. Let $f(z)$ satisfy the hypothesis (i) and (ii). We write it in the form (18). We can set $h_o = 0$ since the multiplicative factor e^{h_o} is easily taken care of in $D_m(f)$.

One defines:

$$\phi(z;x,y) = \prod_{0\le p\le n-1} \left(\frac{z-y_p}{1-zx_p}\right) = \sum_{0\le \ell} \phi_\ell(x,y) z^\ell \qquad ; \tag{20}$$

$\phi_\ell(x,y)$ can be expressed in terms of the elementary symmetric functions of $\{x\}$ and $\{y\}$.

Theorem The following identity holds:

$$D_m(f) = \prod_{0\le p,q\le n-1} \frac{1}{(1-x_p y_q)} \det\left(I - \begin{bmatrix} \phi_o(x,y) & 0 & \cdots & 0 \\ \phi_1(x,y) & \phi_o(x,y) & \cdots & 0 \\ \vdots & & & \\ \phi_{n-m-1}(x,y) & \cdots & & \phi_o(x,y) \end{bmatrix} \begin{bmatrix} \phi_o(y,x) & \phi(y,x) & \cdots & \phi_{n-m-1}(y,x) \\ 0 & \phi_o(y,x) & \cdots & \cdot \\ \vdots & & & \\ 0 & \cdots & & \phi_o(y,x) \end{bmatrix} \right)$$

If $n \le m$ the determinant on the right hand side is replaced by 1.

In the case $n \leq m$ this is the result of lemma 3. The idea of the proof for $n > m$ is simply to reduce it to the previous case by using properties of symmetric functions. We have no space to give the argument here.

The whole point of the remarkable identity (21) is to establish the departure of $D_m(f)$ from the simple expression $\prod_{p,q} \frac{1}{(1-x_p y_q)}$. This is fully answered by the theorem as long as we deal with functions with properties (i) and (ii). However, the restriction that $f(e^{i\theta})$ be the inverse of a trigonometric polynomial is too severe. In the last part we shall indicate heuristically an avenue to a direct proof without providing, however, a careful analysis of the convergence of our process.

VII

We should have remarked, when dealing with the functions $f(e^{i\theta})$ of section VI, that writing:

$$\log f(e^{i\theta}) = h(e^{i\theta}) = \sum_{-\infty<k<+\infty} h_k e^{ik\theta} ,$$

one has:

$$h(e^{i\theta}) = h_o - \log \prod_{0\leq p\leq n-1} (1 - x_p e^{i\theta})(1-y_p e^{-ip\theta})$$

$$= h_o + \sum_{1\leq k} \left\{ e^{ik\theta} \left(\sum_{0\leq p\leq n-1} \frac{x_p^k}{k} \right) + e^{-ik\theta} \left(\sum_{0\leq p\leq n-1} \frac{y_p^k}{k} \right) \right. . \qquad (22)$$

So that, according to lemma 3, for $m \geq n$:

$$D_m(f) = e^{mh_o} \prod_{0\leq p,q\leq n-1} \frac{1}{(1-x_p y_q)} = e^{mh_o + \sum_{1\leq k} k h_k h_{-k}} \qquad (23)$$

The theorem of Szego and Kac states that this is in fact the correct asymptotic expression for an "arbitrary" function $f(e^{i\theta})$. We now present a direct hint of this result. Assuming again $f(e^{i\theta}) = e^{h(e^{i\theta})}$ where $h(e^{i\theta})$ is "sufficiently regular", we introduce the matrix function $H(U) = \sum_k h_k U^k$ on U_m and write, using

$\det\ e^{A} = e^{\mathrm{tr} A}$

$$D_m(f) = \langle e^{\mathrm{tr}\, H(U)} \rangle_m = \langle e^{\sum_{-\infty < p < +\infty} h_p \,\mathrm{tr}\, U^p} \rangle_m \quad .$$

Now it is very tempting to proceed as in statistical mechanics to define the "cumulants" $\mathscr{C}_p^{(m)}(r_1, \ldots, r_p)$ of the unitary group U_m through the following algorithm:

$$\langle e^{\sum_{-\infty<p<+\infty} h_p \,\mathrm{tr}\, U^p} \rangle_m = e^{\sum_{1 \leq p} \left[\sum_{-\infty<r_1 \leq \ldots \leq r_p < \infty} \mathscr{C}_p^{(m)}(r_1, \ldots, r_p) h_{r_1} \ldots h_{r_p} \right]} \tag{24}$$

This relation can be thought as a generating function for the cumulants, which can however be computed through finite algebraic manipulations. We set:

$$C_p^{(m)}(h) = \sum_{-\infty<r_1 \leq \ldots \leq r_p < \infty} \mathscr{C}_p^{(m)}(r_1, \ldots, r_p)\, h_{r_1} \ldots h_{r_p} \tag{25}$$

Then:

$$C_1^{(m)}(h) = \langle \mathrm{tr}\ H(U) \rangle_m$$

$$C_2^{(m)}(h) = \frac{1}{2} \langle [\mathrm{tr}\ H(U)]^2 \rangle_m - \frac{1}{2} [\langle \mathrm{tr}\ H(U) \rangle_m]^2$$

$$C_3^{(m)}(h) = \frac{1}{6} \langle [\mathrm{tr}\ H(U)]^3 \rangle_m - \frac{1}{2} \langle \mathrm{tr}\ H(U) \rangle_m \langle [\mathrm{tr}\ H(U)]^2 \rangle_m + \frac{1}{3} \langle \mathrm{tr}\ H(U) \rangle_m^3$$

. . . .

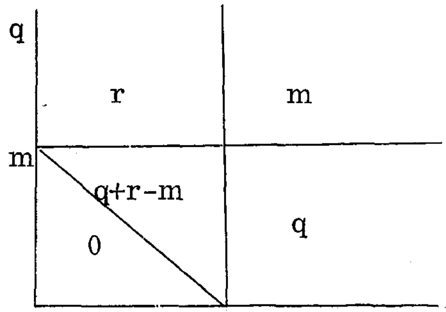
and so on.

One readily finds that:

$$\left\{ \begin{array}{l} C_1^{(m)}(h) = m\, h_o \\ C_2^{(m)}(h) = \sum_{1 \leq p \leq m} p\, h_p\, h_{-p} + \sum_{m+1 \leq q} m\, h_q\, h_{-q} \\ C_3^{(m)}(h) = 2^{-1} \sum_{0<q,\, r<\infty} \mathrm{Min} \{0, q, r, q+r-m, m\} \left(h_q h_r h_{-q-r} + h_{-q} h_{-r} h_{q+r} \right) \\ \ldots . \end{array} \right. \tag{26}$$

Clearly $\mathscr{C}_p^{(m)}(r_1, r_2, \ldots r_p)$ vanishes unless $r_1 + \ldots + r_p = 0$. The function $\psi(q, r) = \mathrm{Min}\,\{0, q, r, q+r-m, m\}$ enjoys nice symmetry properties and is represented in the following diagram:

$\psi(q, r)$:



It vanishes for $q+r < m$. This is a general property, since it follows from the previous section that for $p > 2$:

$$\mathscr{C}_p^{(m)}(r_1, \ldots r_p) = 0 \qquad \text{if} \qquad |r_1| + \ldots + |r_p| < 2m \tag{26}$$

Hence, if h_n decreases fast enough for $n \longrightarrow \infty$ one has for $p > 2$ $\lim\limits_{m \to \infty} C_p^{(m)}(h) = 0$. If it can be established that this limit is uniform in p for $p > 2$ one will have the Szegö-Kac result:

$$\lim_{m \to \infty} \frac{D_m(F)}{e^{mh_0}} = e^{\sum\limits_{1 \leq k} k h_k h_{-k}} \qquad . \tag{27}$$

VIII

The approach that has been sketched above raises a number of interesting questions and suggests a few other problems: (i) to obtain a closed form for the cumulants of the unitary group and study their properties; (ii) to provide under suitable regularity conditions on the function $h(e^{i\theta})$ the uniformity on p of $\lim\limits_{m \to \infty} C_p^{(m)}(h)$; (iii) to generalize the Toeplitz determinants to averages over other symmetric functions of $F(U)$ apart from the determinant. (iv) to study similar problems when the sequence of unitary groups is replaced by another sequence of groups, the orthogonal ones for instance.

There are various possible ramifications, including contact with the theory of continuous integrals, which show how much interest can stem from such an apparently simple looking object as a Toeplitz determinant. In the opening remarks of his treatise on orthogonal polynomials[4] Szegö seems to imply that the determinantal form of orthogonal polynomials is of little practical use. The preceding note was an attempt to challenge this point of view.

In conclusion it is a pleasure to pay a friendly tribute to J. F. Renardy with whom some part of this work has been done.

BIBLIOGRAPHY

1. G. Szegö, Communications du Seminaire Mathematique de l'Université de Lünd, tome supplémentaire dédié a M. Riesz, 228 (1952).
2. M. Kac, Duke Mathematical Journal 21, 501 (1954).
3. H. Weyl, The Classical Groups, Princeton University Press (1946).
4. G. Szegö, Orthogonal Polynomials, American Mathematical Society, New York (1959).