ON THE ASYMPTOTIC SPECTRUM OF HERMITIAN BLOCK TOEPLITZ MATRICES WITH TOEPLITZ BLOCKS

PAOLO TILLI

In loving memory of Ennio de Georgi

ABSTRACT. We study the asymptotic behaviour of the eigenvalues of Hermitian $n \times n$ block Toeplitz matrices $A_{n,m}$, with $m \times m$ Toeplitz blocks. Such matrices are generated by the Fourier coefficients of an integrable bivariate function f, and we study their eigenvalues for large n and m, relating their behaviour to some properties of f as a function; in particular we show that, for any fixed k, the first k eigenvalues of $A_{n,m}$ tend to inf f, while the last k tend to sup f, so extending to the block case a well-known result due to Szegö. In the case the $A_{n,m}$'s are positive-definite, we study the asymptotic spectrum of $P_{n,m}^{-1} A_{n,m}$, where $P_{n,m}$ is a block Toeplitz preconditioner for the conjugate gradient method, applied to solve the system $A_{n,m} x = b$, obtaining strict estimates, when n and m are fixed, and exact limit values, when n and m tend to infinity, for both the condition number and the conjugate gradient convergence factor of the previous matrices. Extensions to the case of a deeper nesting level of the block structure are also discussed.

1. INTRODUCTION

In this paper we study the asymptotic behaviour of the eigenvalues of Hermitian block Toeplitz matrices with Toeplitz blocks: such matrices arise naturally in many applications ([9]), such as signal processing, trigonometric moment problems, integral equations and elliptic partial differential equations with boundary conditions, solved by means of finite differences. In such problems, it is often necessary to solve linear systems of the kind

$$(1) \qquad A_{n,m} x = b,$$

with $A_{n,m}$ having block Toeplitz structure with Toeplitz blocks, i.e.

$$(2) \qquad A_{n,m} = \begin{pmatrix} A_0 & A_1 & \dots & A_{n-1} \\ A_{-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & A_1 \\ A_{-n+1} & \dots & A_{-1} & A_0 \end{pmatrix},$$

with

$$(3) \qquad A_p = \begin{pmatrix} a_{p,0} & a_{p,1} & \cdots & a_{p,m-1} \\ a_{p,-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{p,1} \\ a_{p,-m+1} & \cdots & a_{p,-1} & a_{p,0} \end{pmatrix}.$$

In this paper we consider the case where the matrices $A_{n,m}$ are related to the Fourier coefficients of a function $f : Q \to \mathbf{R}$, which is integrable on $Q = (-\pi, \pi)^2$, i.e. $f \in L^1(Q)$ and

$$(4) \qquad a_{j,k} = \frac{1}{4\pi^2} \int_Q f(x,y) e^{-i(jx+ky)} \, dx dy, \qquad j, k = 0, \pm 1, \pm 2, \ldots .$$

Once the natural numbers n and m have been fixed, we can associate the finite block Toeplitz matrix (2) with the function $f(x,y)$: in this case, we will say that $A_{n,m}$ is *generated* by $f(x,y)$. The matrix $A_{n,m}$ has $n \times n$ block Toeplitz structure, and each block is an $m \times m$ Toeplitz matrix with complex entries; it is immediate to see that $A_{n,m}$ is Hermitian, since $f(x,y)$ takes only real values.

The idea of studying the relationship between the spectra of Toeplitz matrices and their generating functions dates back to Szegö ([9]) for what concerns the one-dimensional case (i.e. the generating function depends on one variable only, and the associated matrices have scalar Toeplitz structure), while the study of the block case is a recent matter, motivated mainly by the need of finding good preconditioners for conjugate gradient and other iterative methods ([15, 3, 4, 5]).

Our purpose here is to study the extremities of the spectra of the set of matrices $\{A_{n,m}\}$, $n, m \in \mathbf{N}$, generated by $f(x,y)$, when n and m tend to infinity, and to relate their properties to some properties of $f(x,y)$ as a function.

If A is an $n \times n$ Hermitian matrix, throughout this paper we indicate by A^t the transpose matrix (without conjugating the elements), and by $\lambda_k(A)$, $k = 1, \ldots, n$, all the eigenvalues of A, counted with their multiplicities, numbered in non-decreasing way: in particular, $\lambda_1(A) = \lambda_{\min}(A)$ and $\lambda_n(A) = \lambda_{\max}(A)$.

In [15] it was proved that, under the hypothesis that $f(x,y)$ is continuous on the closed square $\bar{Q} = [-\pi, \pi]^2$ and not identically constant, for any n and m all the eigenvalues of $A_{n,m}$ lie in the open interval $(\min_Q f, \max_Q f)$, and it holds

$$(5) \qquad \lim_{n,m \to \infty} \lambda_{\min}(A_{n,m}) = \min_{\bar{Q}} f, \qquad \lim_{n,m \to \infty} \lambda_{\max}(A_{n,m}) = \max_{\bar{Q}} f$$

(observe that the limits are finite since $\bar{Q}$ is compact and f was supposed to be continuous on $\bar{Q}$). In section 3 we prove a stronger statement (Theorem 3.1): requiring only the integrability of f over Q, we prove that all the eigenvalues of $A_{n,m}$ (for any n and m) lie in the open interval $(\inf_Q f, \sup_Q)$ (of course, here $\inf_Q f$ is meant to be the essential infimum of f, and $\sup_Q f$ the essential supremum, with respect to Lebesgue measure), unless f is constant almost everywhere on Q. Moreover, still under the only hypothesis that f is integrable over Q, we prove that for any natural number k, it holds

$$(6) \qquad \lim_{n,m \to \infty} \lambda_k(A_{n,m}) = \inf_Q f, \qquad \lim_{n,m \to \infty} \lambda_{nm+1-k}(A_{n,m}) = \sup_Q f.$$

Of course, here $\inf_Q f$ might equal $-\infty$, as well as $\sup_Q f$ might equal $+\infty$, since f is only supposed to be $L^1(Q)$; what is more remarkable, anyway, is that the result (5), proved in [15] only for the first (respectively, the last) eigenvalue of $A_{n,m}$, is

here extended to the first k (respectively, the last k) eigenvalues of $A_{n,m}$ for any fixed k, even under weaker hypotheses on f. This is remarkable since it shows that the matrices $A_{n,m}$ are always asymptotically ill conditioned for large n and m, in the case (which arises very often in the applications) where the $A_{n,m}$'s are positive-definite and generated by a function $f \geq 0$, with $\inf f = 0$ (see Corollary 3.4): in this situation, in fact, since the first k eigenvalues of the $A_{n,m}$'s tend to zero (for any k fixed), those matrices are ill conditioned and the conjugate gradient method ([12]), applied to (1) without preconditioning, will converge very slowly on any subspace of fixed dimension. In view of these facts, the necessity of finding good preconditioners for the system (1) becomes apparent.

A natural way to build a positive-definite preconditioner $P_{n,m}$ for the system (1) is that of letting $P_{n,m}$ be generated by a function $p(x,y) \geq 0$, in such a way that the eigenvalues of the matrices $\{P_{n,m}^{-1} A_{n,m}\}$ have nicer properties than those of the unpreconditioned $\{A_{n,m}\}$, for large values of n and m (see [3, 5, 15]). In particular, Serra ([15]) proved that if the sets of matrices $\{A_{n,m}\}$, $\{P_{n,m}\}$ are generated by two nonnegative functions f and p (not identically zero), both continuous on $\bar{Q}$, then all the eigenvalues of $P_{n,m}^{-1} A_{n,m}$ lie in the open interval (r, R), where r is the infimum and R is the supremum of the ratio f/p. As a consequence, the condition number of the matrices $P_{n,m}^{-1/2} A_{n,m} P_{n,m}^{-1/2}$ resulted bounded from above by the ratio R/r. In Theorem 3.5 we improve this result; more precisely we prove that, even under weaker hypotheses ($f, p \in L^1(Q)$ with p continuous on the open square Q, $p \geq 0$ not identically zero), for any fixed natural number k it holds

$$(7) \qquad \lim_{n,m \to \infty} \lambda_k(P_{n,m}^{-1} A_{n,m}) = \inf_{Q_+} \frac{f}{p}, \qquad \lim_{n,m \to \infty} \lambda_{nm+1-k}(P_{n,m}^{-1} A_{n,m}) = \sup_{Q_+} \frac{f}{p},$$

where $Q_+ \subseteq Q$ is the set where $p > 0$. We remark that, in general, the set of matrices generated by the function f/p (when it belongs to $L^1(Q)$) has nothing to do with the set $\{P_{n,m}^{-1} A_{n,m}\}$; for this reason, Theorem 3.5 comes by no means as a consequence of Theorem 3.1.

Equation (7) once more stresses that the choice of the preconditioner is crucial in view of the convergence rate of the conjugate gradient (or other iterative methods, see [17]) because, if the infimum of the ratio f/p is still too close to zero, for large n and m the preconditioned system will show up the same bad features as the original one. According to this, in Corollary 3.6 we obtain that, when f is nonnegative, the condition number of the matrices $M_{n,m} = P_{n,m}^{-1/2} A_{n,m} P_{n,m}^{-1/2}$ verifies

$$\lim_{n,m \to \infty} \mu(M_{n,m}) = \frac{r}{R}$$

where $r = \inf_{Q_+} f/p$ and $R = \sup_{Q_+} f/p$, while the convergence factor of the conjugate gradient method verifies

$$\lim_{n,m \to \infty} \frac{\lambda_{\max}(M_{n,m}) - \lambda_{\min}(M_{n,m})}{\lambda_{\max}(M_{n,m}) + \lambda_{\min}(M_{n,m})} = \frac{R - r}{R + r}.$$

It is possible to extend all the results so far examined to a deeper nesting level in the block Toeplitz structure of the matrices considered; for example, if $f(x,y,z) \in L^1([-\pi,\pi]^3)$, it is possible to build its Fourier coefficients with respect to the canonical tensor basis:

$$a_{j,k,l} = \frac{1}{(2\pi)^3} \int_{[-\pi,\pi]^3} f(x,y,z) e^{-i(jx+ky+lz)} \, dx \, dy \, dz, \qquad j, k, l = 0, \pm 1, \pm 2, \dots.$$

Obviously, a three-level block Toeplitz matrix $A_{n,m,h}$ can be associated to f, for each triplet of natural numbers n, m and h.

It is not difficult to convince oneself that all the proofs presented in this paper are independent of the dimension, and could be easily adapted to the case where f depends on any number of variables; the difficulty would come mainly from notation, and this is why the results are proved in detail only in the two-dimensional case. For this reason, all the results presented here are still valid when a larger number of variables is involved, and such extensions come in a straightforward way.

We remark that the techniques used by Szegö ([9]) for the scalar Toeplitz structure (one variable only) do not lend themselves to be easily adapted to dimension higher than one, since they strongly rely on the explicit computation of the minima of certain functionals involving univariate orthogonal polynomials. On the other hand, in the one-dimensional case all the results on the convergence of the eigenvalues to inf f or sup f come as a straight consequence of the Theorem of Szegö on the asymptotic distribution of the spectrum of Toeplitz matrices ([9]), which is not available in the block case under the only hypothesis that $f \in L^1(Q)$. For this reason, Theorem 3.1 can be seen as an extension to the block case of results known only in the scalar case and, together with Theorem 3.5, it is an attempt to introduce new and more general techniques suitable for studying the distribution of the eigenvalues of large block Toeplitz matrices with Toeplitz blocks.

2. NOTATION AND PRELIMINARY CONSIDERATIONS

If n and m are non-negative integers, we indicate by $\mathbf{P}_{n,m}$ the vector space of all complex bivariate trigonometric polynomials of degree at most n in the first variable and at most m in the second, that is

(8)

$$u(x,y) \in \mathbf{P}_{n,m} \iff u(x,y) = \sum_{j=-n}^{n} \sum_{k=-m}^{m} u_{jk} e^{i(jx+ky)}, \quad (x,y) \in \bar{Q}, \quad u_{jk} \in \mathbf{C}.$$

Such polynomials can naturally be seen as functions defined on the square $Q = [-\pi, \pi]^2$; on the other hand, it is sometimes useful to identify $\mathbf{P}_{n,m}$ with the complex Euclidean space $\mathbf{C}^{n^* m^*}$, where $m^* = 2m+1$ and $n^* = 2n+1$: in fact, any $u(x,y) \in \mathbf{P}_{n,m}$ is uniquely identified by the complex vector u of its coefficients, which can be usefully partitioned into blocks as follows:

(9) $\quad u = \begin{pmatrix} u_{-n} \\ u_{-n+1} \\ \vdots \\ u_n \end{pmatrix}, \quad u_j \in \mathbf{C}^{m^*}, \quad u_j = \begin{pmatrix} u_{j,-m} \\ u_{j,-m+1} \\ \vdots \\ u_{j,m} \end{pmatrix}, \quad j = -n \ldots n.$

Throughout, we will systematically use the isomorphism of vector spaces between $\mathbf{P}_{n,m}$ and $\mathbf{C}^{n^* m^*}$, and to stress this we will often use the same letter, say u, to indicate both the polyomial $u(x,y) \in \mathbf{P}_{n,m}$ and the vector of its coefficients $u \in \mathbf{C}^{n^* m^*}$, partitioned into blocks according to (9): it will always be clear from the context which space we are referring to.

If $x, y \in \mathbf{C}^n$ are complex vectors, we indicate by $\langle y, x \rangle = \sum_{i=1}^{n} \bar{y}_i x_i$ the usual scalar product on $\mathbf{C}^n$.

We now derive some simple identities which will be widely used in the rest of this paper. If both $u(x,y)$ and $v(x,y)$ belong to $\mathbf{P}_{n,m}$, i.e. if

$$u(x,y) = \sum_{j=-n}^{n} \sum_{k=-m}^{m} u_{j,k} e^{ijx} e^{iky}, \quad v(x,y) = \sum_{j=-n}^{n} \sum_{k=-m}^{m} v_{j,k} e^{ijx} e^{iky},$$

and $f(x,y) \in L^1(Q)$, then we have

$$(10) \qquad \int_Q f \bar{u} v = \sum_{j,j'=-n}^{n} \sum_{k,k'=-m}^{m} \bar{u}_{j,k} v_{j',k'} \int_Q f e^{-i(j-j')x} e^{-i(k-k')y} \, dx dy$$

$$= 4\pi^2 \sum_{j,j'=-n}^{n} \sum_{k,k'=-m}^{m} \bar{u}_{j,k} v_{j',k'} a_{j-j',k-k'} = 4\pi^2 \sum_{j,j'=-n}^{n} \langle A^t_{j'-j} u_j, v_{j'} \rangle$$

$$= 4\pi^2 \langle A^t_{n^*,m^*} u, v \rangle,$$

where $A_{n^*,m^*} = A_{2n+1,2m+1}$, according to (2), (3) and (4). Observe that, in the particular case $f(x,y) \equiv 1$, we obtain

$$(11) \qquad\qquad \frac{1}{4\pi^2} \int_Q \bar{u} v = \sum_{j=-n}^{n} \sum_{k=-m}^{m} \bar{u}_{j,k} v_{j,k} = \langle u, v \rangle.$$

Recall that, if A is a positive-definite Hermitian matrix, the convergence rate of the conjugate gradient method, applied to a linear system $Ax = b$, is strictly related to the quantity $\rho(A) = (M - m)/(M + m)$, where M and m are, respectively, the largest and the smallest eigenvalue of A: it holds $0 \le \rho(A) < 1$, and it can be shown that the absolute error is reduced, at each step, by a factor at least $\rho(A)$. Throughout, we will refer to this quantity as the *convergence factor* for the conjugate gradient, with respect to the matrix A. For more details, the reader is referred to [6, 17, 12, 1].

In the following, we will have to do with integrable functions with respect to Lebesgue measure, which we indicate by $\mathbf{L}$; when not explicitly indicated, we will always refer to $\mathbf{L}$ as the Lebesgue measure on the open square $Q = (-\pi, \pi)^2$, and we will indicate by $L^1(Q)$ the Banach space of all measurable functions which are integrable on Q, that is

$$f \in L^1(Q) \quad \Longleftrightarrow \quad f \text{ is measurable and } \|f\|_{L^1(Q)} = \int_Q |f| \, d\mathbf{L} < +\infty.$$

In this paper we will simply write $\int_Q f$ instead of $\int_Q f \, d\mathbf{L}$, Lebesgue measure being implicitly assumed. We will also refer to the closure of Q as the *closed square*, and we will indicate it by $\bar{Q}$. When we state a pointwise property of a function $f \in L^1(Q)$ (such as $f \ge 0$), we will always mean that such property holds *almost everywhere* on Q, i.e. that it holds at every point of Q but for a set of zero Lebesgue measure.

For a measurable function $f \in L^1(Q)$ there is a notion of *essential infimum* (essinf$_Q f$), which plays essentially the same role as the usual infimum notion, but disregards zero measure sets:

$$\text{essinf}_Q f = \sup\{y \in \mathbf{R} : f > y \text{ almost everywhere on } Q\}.$$

In this paper we will simply write $\inf_Q f$ to indicate the essential infimum of f on Q, if f is only supposed to be measurable on Q; the same considerations are valid for the notion of *essential supremum* which we will simply indicate by $\sup_Q f$.

3. MAIN RESULTS

Theorem 3.1. *Let $f(x,y): Q \to \mathbf{R}$ belong to the space $L^1(Q)$, and let us indicate by $\{A_{n,m}\}_{n,m \in \mathbf{N}}$ the set of Hermitian matrices it generates, according to (2), (3), and (4). Then, for any natural numbers n and m, if λ is an eigenvalue of $A_{n,m}$ it holds*

$$(12) \qquad \inf_Q f \le \lambda \le \sup_Q f$$

where the inequalities are always strict, unless f is constant on Q. Moreover, however we fix a positive integer k, we have

$$(13) \qquad \lim_{n,m \to \infty} \lambda_k(A_{n,m}) = \inf_Q f, \qquad \lim_{n,m \to \infty} \lambda_{nm+1-k}(A_{n,m}) = \sup_Q f,$$

that is, for k fixed, the first k eigenvalues of $A_{n,m}$ tend to the essential infimum of f while the last k tend to the essential supremum of f, as n and m approach infinity.

For the proof of this theorem we need two lemmas, where some simple measure theory is involved; in particular, in the second lemma we repeatedly use the fact that a bounded set S is Lebesgue-measurable if and only if, for every $\epsilon > 0$, there exist a compact set $K \subseteq S$ and an open set $A \supseteq S$ such that $\mathbf{L}(A \setminus K) < \epsilon$. This is nothing but the equivalence of inner and outer measure for Lebesgue-measurable sets (see [7, 11] for more details).

Lemma 3.2. *Let $Q = (-\pi, \pi)^2$; suppose $S \subseteq Q$ is a measurable set with $\mathbf{L}(S) > 0$, and n is a positive integer. Then there exist n open sets $A_i \subseteq Q$, pairwise disjoint, such that $\mathbf{L}(S \cap A_i) > 0$ for $i = 1, \dots, n$.*

Proof. Let $U_1^0 = Q$, then divide Q into four equal squares of side π, and let U_j^1, $j = 1, \dots, 4$, indicate their internal parts. Iterate this procedure so that, for any natural number p, the original square is divided into 4^p equal squares of side $2\pi/2^p$, and let U_j^p, $j = 1, \dots, 4^p$, indicate their internal parts. For any p, the open squares $\{U_j^p\}$ are pairwise disjoint and cover almost all of Q, and since $S \subseteq Q$, they cover almost all of S. We seek the requested A_i's among some collection $\{U_j^p\}$, for a suitable p such that $4^p \ge n$. Suppose that this cannot be done for any p, i.e. suppose that, for every p with $4^p \ge n$, $\mathbf{L}(S \cap U_j^p) = 0$ for all $j = 1, \dots, 4^p$, except for at most $n-1$ indices. Since $\mathbf{L}(U_j^p) = \pi^2/4^{p-1}$, we then have

$$0 < \mathbf{L}(S) = \mathbf{L}(S \cap \bigcup_{j=1}^{4^p} U_j^p) = \sum_{j=1}^{4^p} \mathbf{L}(S \cap U_j^p) \le (n-1)\frac{\pi^2}{4^{p-1}},$$

and letting $p \to \infty$ we obtain $\mathbf{L}(S) = 0$, which is absurd. $\qquad\square$

Lemma 3.3. *Let $f(x,y): Q \to \mathbf{R}$ and $p(x,y): Q \to \mathbf{R}$ be two functions belonging to the space $L^1(Q)$, with $p \ge 0$, p continuous on $Q = (-\pi, \pi)^2$, p not identically zero. Let us indicate by Q_+ the subset of Q where $p(x,y) > 0$; if we set*

$$a = \inf_{(x,y) \in Q_+} \frac{f(x,y)}{p(x,y)} \in \mathbf{R} \cup \{-\infty\},$$

then for each real number $a' > a$ and each integer $N \geq 0$ there exist natural numbers n and m so large that, for every N-tuple of polynomials $h_i \in \mathbf{P}_{n,m}$, $i = 1, \ldots, N$, there exists a polynomial $u \in \mathbf{P}_{n,m}$, $u \not\equiv 0$, such that

$$\text{(14)} \qquad \frac{\int_Q f\,|u|^2}{\int_Q p\,|u|^2} \leq a',$$

$$\text{(15)} \qquad \int_Q \bar{h}_i u\,dx = 0, \quad i = 1, \ldots, N.$$

The idea underlying the proof of this lemma is that the modulus of the polynomial $u(x,y)$ should be large on some set where f/p is close to a, and small elsewhere; for such $u(x,y)$ to be orthogonal to any N fixed polynomials, we build $N + 1$ linearly independent polynomials whose moduli have the above property, and then we try to express $u(x,y)$ as a suitable linear combination of them. As we will show, these $N + 1$ linearly independent polynomials can be chosen depending only on f, p, a' and N, while the coefficients of their linear combination yelding $u(x,y)$ will depend, of course, also on the N polynomials to which $u(x,y)$ must be orthogonal.

Proof. Choose any real number a'' such that $a < a'' < a'$. Since f/p is measurable on Q_+ and $a'' > a$, the subset $S \subseteq Q_+$ where $f/p < a''$ is measurable and has positive measure (this follows from the definition of the essential infimum). By Lemma 3.2 we can find $N + 1$ pairwise disjoint open sets $A_i \subseteq Q$ such that, letting $S_i = S \cap A_i$, it holds $\mathbf{L}(S_i) > 0$, for $i = 1, \ldots, N + 1$. Let us set $m = \min_{i=1}^{N+1}\{\mathbf{L}(S_i)\}/2$; since the sets S_i's are measurable, their Lebesgue measure equals their inner measure, and so we can find $N + 1$ compact sets $C_i \subseteq S_i$, such that $\mathbf{L}(C_i) \geq m$, $i = 1, \ldots, N + 1$. We set $C = \bigcup_{i=1}^{N+1} C_i$; since C is compact, $p > 0$ on C and p is continuous on C, p achieves a minimum $\gamma > 0$ on C. Let us set $D = 4(N+1)^2/(\gamma m)$. Using the absolute continuity of Lebesgue integral, there exists $\delta > 0$ such that, for every measurable set $\Omega \subset Q$ with $\mathbf{L}(\Omega) < 2(N + 1)\delta$, it holds

$$\text{(16)} \qquad 4D \int_\Omega |f| \leq \frac{a' - a''}{2}, \text{ and}$$

$$\text{(17)} \qquad \frac{a''}{1 + 4D \int_\Omega p} \leq a'' + \frac{a' - a''}{8}$$

(we observe that, if $a'' \geq 0$, the last condition is always fulfilled). Again approximating Lebesgue measure first with inner measure and then with outer measure, we can find $N + 1$ compact sets $K_i \subseteq S_i$ such that $\mathbf{L}(S_i \setminus K_i) < \delta$, and $N + 1$ open sets $U_i \supseteq S_i$ such that $\mathbf{L}(U_i \setminus S_i) < \delta$. By replacing U_i with $U_i \cap A_i$ if necessary, we can assume that the U_i's are pairwise disjoint; similarly, by replacing K_i with $K_i \cup C_i$ we can assume $C_i \subseteq K_i$. Now let $U = \bigcup_{j=1}^{N+1} U_j$, and $K = \bigcup_{j=1}^{N+1} K_j$; it trivially holds $\mathbf{L}(U \setminus K) < 2(N + 1)\delta$, and we will use this later.

We now choose a number $0 < \epsilon < 1/(2N + 2)$ such that the following inequality holds

$$\text{(18)} \qquad \epsilon^2 D \|f\|_{L^1(Q)} \leq \frac{a' - a''}{4},$$

and, if $a'' < 0$, we choose ϵ so small that, for every measurable set $\Omega \subseteq Q$ with $\mathbf{L}(\Omega) < 2(N+1)\delta$, it holds

$$(19) \qquad \frac{a''}{1 + 4D \int_\Omega p + \epsilon^2 D \|p\|_{L^1(Q)}} \le a'' + \frac{a' - a''}{4}$$

(it is possible to do this because δ has already been fixed, and it verifies (17)). Now we consider $N + 1$ continuous functions ϕ_i defined on $\bar{Q}$ such that $\phi_i \equiv 1$ on K_i, $\phi_i \equiv 0$ on $Q \setminus U_i$, and $0 \le \phi_i \le 1$ everywhere, $i = 1, \dots, N+1$ (for any i, such ϕ_i exists by virtue of Urysohn's lemma ([14]), since the two sets K_i and $Q \setminus U_i$ are closed and disjoint). Since each ϕ_i is zero on the boundary of Q, it can be approximated in the uniform norm by a suitable trigonometric polynomial g_i, such that

$$(20) \qquad \sup_Q |\phi_i - g_i| < \epsilon.$$

We seek the polynomials $u(x, y)$ in the form

$$(21) \qquad u(x, y) = \sum_{j=1}^{N+1} c_j g_j(x, y),$$

where the c_j are complex numbers to be determined. It is obvious that $u(x, y) \in \mathbf{P}_{n,m}$ for suitable n and m, independently of how the numbers c_j are chosen.

Consider any N-tuple of polynomials $h_i(x, y) \in \mathbf{P}_{n,m}$, $i = 1, \dots, N$: we want to show that the complex numbers c_j can be chosen in such a way that both (14) and (15) hold, and this will complete the proof of the lemma.

According to equation (21), the orthogonality relations (15) become

$$(22) \qquad 0 = \sum_{j=1}^{N+1} c_j \int_Q \bar{h}_i(x, y) g_j(x, y) \, dx \, dy, \quad i = 1, \dots, N.$$

This is a system of N linear equations in $N + 1$ unknowns, so it always has a nontrivial solution $\bar{c} = (c_1, \dots, c_{N+1})$, for any choice of the $h_i(x, y)$'s. There is no loss in generality if we suppose that $\bar{c}$ is normalized, i.e., if we suppose

$$(23) \qquad \max_{1 \le j \le N+1} |c_j| = 1 = |c_k|,$$

for a suitable choice of k.

Owing to (20), we have that the modulus of any g_j is between $1 - \epsilon$ and $1 + \epsilon$ on K_j, is not greater than ϵ on $Q \setminus U_j$, and is no greater than $1 + \epsilon$ everywhere in Q. For this reason, from (23) we immediately obtain the estimate

$$(24) \qquad |u(x, y)| \le (N+1)(1 + \epsilon) \le 2(N+1).$$

Let us estimate $|u(x, y)|$ for $(x, y) \in K_k$ (we remark that $|c_k| = 1$); we have

$$(25) \qquad |u(x, y)| \ge 1 - \epsilon - N\epsilon, \qquad (x, y) \in K_k, \text{ and therefore}$$

$$(26) \qquad \int_Q p \, |u|^2 \ge \int_K p \, |u|^2 \ge [1 - (N+1)\epsilon]^2 \int_{K_k} p \ge \frac{1}{4} \gamma m,$$

since $C_k \subseteq K_k$, $\mathbf{L}(C_k) \ge m$, and $(N+1)\epsilon \le 1/2$. In particular, $u(x, y)$ is not identically zero.

By similar arguments we immediately obtain

$$(27) \qquad |u(x, y)| \le \epsilon(N+1), \qquad (x, y) \in Q \setminus U.$$

Finally, we have

$$\int_Q f\,|u|^2 = \int_K f\,|u|^2 + \int_{U\setminus K} f\,|u|^2 + \int_{Q\setminus U} f\,|u|^2 .$$

Since $f < a''p$ inside K, from (24) and (27) we have

$$\int_Q f\,|u|^2 \le a'' \int_K p\,|u|^2 + 4(N+1)^2 \int_{U\setminus K} |f| + \epsilon^2(N+1)^2 \int_{Q\setminus U} |f| .$$

Dividing both sides by $\int_Q p\,|u|^2$ and using (26) we obtain

$$(28) \qquad \frac{\int_Q f\,|u|^2}{\int_Q p\,|u|^2} \le a'' \frac{\int_K p\,|u|^2}{\int_Q p\,|u|^2} + 4D \int_{U\setminus K} |f| + \epsilon^2 D\,\|f\|_{L^1(Q)} .$$

Now, since $\mathbf{L}(U \setminus K) < 2(N+1)\delta$, from (16) and (18) we have

$$(29) \qquad \frac{\int_Q f\,|u|^2}{\int_Q p\,|u|^2} \le a'' \frac{\int_K p\,|u|^2}{\int_Q p\,|u|^2} + \frac{3}{4}(a' - a'').$$

If $a'' \ge 0$ we immediately obtain (14), since the ratio of the two integrals on the right side is not greater than 1.

The case $a'' < 0$ requires one more step. From (26) and (27) we have

$$\frac{\int_{Q\setminus K} p\,|u|^2}{\int_K p\,|u|^2} = \frac{\int_{U\setminus K} p\,|u|^2 + \int_{Q\setminus U} p\,|u|^2}{\int_K p\,|u|^2} \le 4D \int_{U\setminus K} p + \epsilon^2 D\,\|p\|_{L^1(Q)} ,$$

and so

$$\frac{\int_K p\,|u|^2}{\int_Q p\,|u|^2} = \frac{1}{1 + \frac{\int_{Q\setminus K} p|u|^2}{\int_K p|u|^2}} \ge \frac{1}{1 + 4D \int_{U\setminus K} p + \epsilon^2 D\,\|p\|_{L^1(Q)}};$$

since a'' is negative, putting the last estimate into (29) yields

$$(30) \qquad \frac{\int_Q f\,|u|^2}{\int_Q p\,|u|^2} \le \frac{a''}{1 + 4D \int_{U\setminus K} p + \epsilon^2 D\,\|p\|_{L^1(Q)}} + \frac{3}{4}(a' - a''),$$

and using condition (19) we conclude the proof. $\qquad\square$

Proof of Theorem 3.1. Since the spectrum of a matrix is unaffected by transposition, we can prove our theorem referring to the spectrum of the transposed matrices $\{A_{n,m}^t\}$, instead of $\{A_{n,m}\}$ (the reason for this will become clear throughout this proof).

We first show inequality (12); suppose that, for some n and m, λ is an eigenvalue of $A_{n,m}^t$ and

$$(31) \qquad\qquad\qquad \lambda \le \inf_Q f;$$

then it holds $\lambda_1(A_{n,m}^t) \le \inf_Q f$, which implies $\lambda_1(A_{n^*,m^*}^t) \le \inf_Q f$, since $A_{n,m}^t$ is a principal submatrix of A_{n^*,m^*}^t. Let $u \in \mathbf{C}^{n^*m^*}$ be an eigenvector of A_{n^*,m^*}^t relative to $\lambda_1(A_{n^*,m^*}^t)$, and let us indicate with the same letter u the polynomial

$u(x,y) \in \mathbf{P}_{n,m}$ which corresponds to the vector $u \in \mathbf{C}^{n^*m^*}$, according to (8), (9). Using (10), we obtain

$$(32) \qquad \lambda_1(A^t_{n^*,m^*}) = \frac{\langle A^t_{n^*,m^*}u, u\rangle}{\langle u, u\rangle} = \frac{\int_Q f\,|u|^2}{\int_Q |u|^2} \leq \inf_Q f,$$

from which it follows that

$$0 \leq \int_Q \left(f - \inf_Q f\right)|u|^2 \leq 0;$$

from this relation we obtain that $\left(f - \inf_Q f\right)|u|^2$ equals zero almost everywhere on Q and, since u is a nonidentically zero polynomial, $u \neq 0$ almost everywhere on Q implies $f = \inf_Q f$ almost everywhere. So, inequality (31) can never be strict, and can hold only as an equality if and only f is constant. This argument can be repeated considering $-f$ in place of f, and so we have proved (12), where the inequalities are always strict except when f is constant.

Now we can prove (13). Let the integer $k > 0$ be fixed; if we apply Lemma 3.3 to $f(x,y)$ (with $N = k - 1$ and $p \equiv 1$), we obtain that for every real number $a' > \inf_Q f$ there are n and m such that, however we choose $k - 1$ polynomials $h_i \in \mathbf{P}_{n,m}$, $i = 1, \ldots, k - 1$, there exists a polynomial $u \in \mathbf{P}_{n,m}$ such that

$$\frac{\int_Q f\,|u|^2}{\int_Q |u|^2} \leq a', \quad \text{and} \quad \int_Q \bar{h}_i u = 0, \quad i = 1, \ldots, k - 1.$$

Using the isomorphism between $\mathbf{P}_{n,m}$ and $\mathbf{C}^{n^*m^*}$ and the identities (10), (11), we can reformulate the above statement as follows: for each real number $a' > \inf_Q f$, there are integers n and m so large that, however we choose $k-1$ vectors $v_i \in \mathbf{C}^{n^*m^*}$, $i = 1, \ldots, k - 1$, there exists a vector $u \in \mathbf{C}^{n^*m^*}$ such that

$$\frac{\langle A^t_{n^*,m^*}u, u\rangle}{\langle u, u\rangle} \leq a', \quad \text{and} \quad \langle v_i, u\rangle = 0, \quad i = 1, \ldots, k - 1.$$

As the v_i's vary throughout $\mathbf{C}^{n^*m^*}$, they span all the subspaces $V \subseteq \mathbf{C}^{n^*m^*}$ whose dimension does not exceed k, so that we have

$$\sup_{\mathrm{codim}\,U=k} \inf_{u \in U} \frac{\langle A^t_{n^*,m^*}u, u\rangle}{\langle u, u\rangle} \leq a',$$

and according to the minimax theorem of Courant and Fisher we obtain that the k-th eigenvalue of the matrix $A^t_{n^*,m^*}$ is at most a'. Since $a' > \inf_Q f$ was arbitrary, with the aid of (12) we obtain

$$(33) \qquad \inf_{n,m} \lambda_k(A^t_{n,m}) = \inf_Q f, \quad \forall k \geq 1.$$

Now, since $m \geq m'$ and $n \geq n'$ imply $\lambda_k(A^t_{n,m}) \leq \lambda_k(A^t_{n',m'})$ (because, in this case, $A_{n',m'}$ is a principal submatrix of $A^t_{n,m}$), equation (33) becomes

$$\lim_{n,m \to \infty} \lambda_k(A_{n,m}) = \inf_Q f, \quad \forall k \geq 1.$$

By replacing f with $-f$ and repeating this argument, we immediately obtain the other limit in (13), the one relative to the k-th from the last eigenvalue. $\square$

Corollary 3.4. *Suppose $f \in L^1(Q)$ and $f \geq 0$ almost everywhere on Q, with $f > 0$ on a set of positive measure. Then the matrices $\{A_{n,m}\}$ generated by f are all positive-definite.*

Proof. In fact, by Lemma 3.3 all their eigenvalues are strictly positive. $\quad\square$

Theorem 3.5. *Let* $f(x,y) : Q \to \mathbf{R}$ *and* $p(x,y) : Q \to \mathbf{R}$ *be as in Lemma 3.3, and let* $\{A_{n,m}\}$, $\{P_{n,m}\}$, $n, m \in \mathbf{N}$, *be the sets of block Toeplitz matrices generated, respectively, by* $f(x,y)$ *and* $p(x,y)$. *Let us set*

$$r = \inf_{Q_+} \frac{f}{p}, \quad R = \sup_{Q_+} \frac{f}{p}, \quad where \quad Q_+ = \{(x,y) : p(x,y) > 0\},$$

and suppose $r \neq R$, $\mathbf{L}(Q \setminus Q_+) = 0$; *then, for any n and m, all the eigenvalues of* $P_{n,m}^{-1}A_{n,m}$ *lie in the open set* (r, R), *and for every fixed positive integer k it holds*

$$(34) \qquad \lim_{n,m\to\infty} \lambda_k(P_{n,m}^{-1}A_{n,m}) = r, \quad \lim_{n,m\to\infty} \lambda_{nm+1-k}(P_{n,m}^{-1}A_{n,m}) = R,$$

that is, for k fixed, the first k eigenvalues of $P_{n,m}^{-1}A_{n,m}$ *tend to the infimum of* f/p, *while the last k tend to the supremum of* f/p, *as n, m approach infinity.*

Proof. Throughout this proof, we will systematically use the fact that the eigenvalues of $P_{n,m}^{-1}A_{n,m}$ coincide with those of $(P_{n,m}^t)^{-1}A_{n,m}^t$. From Corollary 3.4, $P_{n,m}$ is positive-definite (and hence non-singular) for any value of n, m. Let k and $a'' > r$ be fixed; using the isomorphysm between $\mathbf{P}_{n,m}$ and $\mathbf{C}^{n^*m^*}$ and the identities (10), (11), from Lemma 3.3 we get the following: there are integers n and m so large that, however we choose $k - 1$ vectors $v_i \in \mathbf{C}^{n^*m^*}$, $i = 1, \dots, k - 1$, there exists a vector $u \in \mathbf{C}^{n^*m^*}$ such that

$$\frac{\langle A_{n^*,m^*}^t u, u\rangle}{\langle P_{n^*,m^*}^t u, u\rangle} \leq a'', \quad \text{and} \quad \langle v_i, u\rangle = 0, \quad i = 1, \dots, k - 1.$$

As the v_i's vary throughout $\mathbf{C}^{n^*m^*}$, they span all the subspaces $V \subseteq \mathbf{C}^{n^*m^*}$ whose dimension does not exceed $k - 1$ so that, according to the minimax theorem of Courant and Fisher for the generalized eigenvalue problem, we obtain that the k-th eigenvalue of the matrix $P_{n^*,m^*}^{-1}A_{n^*,m^*}$ is at most a'', i.e. $\lambda_k(A_{n^*,m^*}) \leq a''$. Since $a'' > r$ was arbitrary, we obtain

$$(35) \qquad \inf_{n,m} \lambda_k(P_{n,m}^{-1}A_{n,m}) \leq r, \quad \forall k \geq 1.$$

We want to show also the opposite inequality (it will suffice to prove it for $k = 1$). Suppose that there exist n and m such that

$$(36) \qquad \lambda_1(P_{n,m}^{-1}A_{n,m}) \leq r;$$

since $P_{n,m}$ is a principal submatrix of P_{n^*,m^*}, and $A_{n,m}$ is the corresponding submatrix of A_{n^*,m^*}, again from the minimax theory we have $\lambda_1(P_{n^*,m^*}^{-1}A_{n^*,m^*}) \leq r$; reasoning as in the proof of Theorem 3.1, i.e. considering an eigenvector relative to $\lambda_1(P_{n^*,m^*}^{-1}A_{n^*,m^*})$, there exists a a polynomial $u(x,y) \in \mathbf{P}_{n,m}$ (not identically zero) such that

$$\lambda_1(P_{n^*,m^*}^{-1}A_{n^*,m^*}) = \frac{\int_Q f\,|u|^2}{\int_Q p\,|u|^2} \leq r,$$

from which we obtain $0 \leq \int_{Q_+}(f - rp)\,|u|^2 \leq 0$ (we remark that $\mathbf{L}(Q \setminus Q_+) = 0$). Since $u \neq 0$ almost everywhere, we obtain $f = rp$ almost everywhere from which it follows that $r = R$, which is absurd since we had supposed $r < R$. So, inequality (36) cannot hold for any n and m, and we obtain that, for any value of n and m,

all the eigenvalues of $P_{n,m}^{-1}A_{n,m}$ are strictly greater than r; in particular, equality must hold in (35), so we have

(37) $$\inf_{n,m} \lambda_k(P_{n,m}^{-1}A_{n,m}) = r, \quad \forall k \geq 1, \text{ and}$$

(38) $$\lambda_k(P_{n,m}^{-1}A_{n,m}) > r, \quad \forall n,m \geq 1, \forall k, 1 \leq k \leq nm.$$

Moreover, since $\lambda_k(P_{n,m}^{-1}A_{n,m})$ is a non-increasing function of n and m (with respect to the usual partial ordering on the pairs (n,m)), the infimum in equation (37) is actually a limit, so we obtain

(39) $$\lim_{n,m\to\infty} \lambda_k(P_{n,m}^{-1}A_{n,m}) = r, \quad \forall k \geq 1.$$

Replacing f with $-f$, and repeating this reasoning, we obtain

(40) $$\lim_{n,m\to\infty} \lambda_{nm+1-k}(P_{n,m}^{-1}A_{n,m}) = R,$$

and any eigenvalue of $P_{n,m}^{-1}A_{n,m}$, for any value of n and m, is strictly smaller than R. So we have proved that the spectrum of any matrix $P_{n,m}^{-1}A_{n,m}$ lies in the open set (r,R), i.e.

(41) $$r < \lambda_k(P_{n,m}^{-1}A_{n,m}) < R, \qquad \forall n,m \geq 1, \quad \forall k, 1 \leq k \leq nm. \quad \square$$

Corollary 3.6. *Under the same hypotheses and notations as in Theorem 3.5, suppose that $f \geq 0$ almost everywhere, with $f > 0$ on a set of positive measure. Then, letting $M_{n,m} = P_{n,m}^{-\frac{1}{2}}A_{n,m}P_{n,m}^{-\frac{1}{2}}$, it holds*

(42) $$\mu(M_{n,m}) < \frac{R}{r}, \quad \forall n,m \geq 1,$$

(43) $$\lim_{n,m\to\infty} \mu(M_{n,m}) = \frac{R}{r},$$

where $\mu(A)$ indicates the condition number, i.e. $\mu(A) = \|A\|\,\|A^{-1}\|$, with respect to the Euclidean norm (it is understood that $R/r = +\infty$ if $r = 0$), and the limit is monotone increasing with respect to the usual partial ordering over the pairs (n,m); moreover, the convergence factor of the conjugate gradient method verifies

(44) $$\lim_{n,m\to\infty} \frac{\lambda_{\max}(M_{n,m}) - \lambda_{\min}(M_{n,m})}{\lambda_{\max}(M_{n,m}) + \lambda_{\min}(M_{n,m})} = \frac{R-r}{R+r}.$$

Proof. For any n and m, the matrix $M_{n,m}$ is Hermitian and is positive-definite, because it has the same eigenvalues as $P_{n,m}^{-1}A_{n,m}$, which by Theorem 3.5 lie in the open set (r,R), with $r \geq 0$. So we have

$$\mu(M_{n,m}) = \frac{\lambda_{\max}(P_{n,m}^{-1}A_{n,m})}{\lambda_{\min}(P_{n,m}^{-1}A_{n,m})}.$$

Then estimate (42) follows immediately from (41), while (43) follows from (39) and (40), passing to the limit in the above equation. The limit is monotone, being non-increasing in the limit (39) and non-decreasing in the limit (40). Finally, the limit relation (44) is proved in the same way. $\square$

REFERENCES

1. O. Axelsson and G. Lindskog, *On the rate of convergence of the preconditioned conjugate gradient method*, Numer. Math. **48** (1988), pp. 499–523. MR **88a:**65037b
2. J. R. Bunch, *Stability of methods for solving Toeplitz systems of equations*, SIAM J. Sci. Stat. Comp. **6** (1985), pp. 349–364. MR **87a:**65073
3. R. H. Chan, *Toeplitz preconditioniers for Toeplitz systems wit nonnegative generating functions*, IMA J. Numer. Anal. **11** (1991), pp. 333–345. MR **92f:**65041
4. R. H. Chan and X. Q. Jin, *A family of block preconditioners for block systems*, SIAM J. Sci. Stat. Comp. **13** (1992), pp. 1218–1235. MR **93e:**65050
5. F. Di Benedetto, G. Fiorentino and S. Serra, *C.G. preconditioning for Toeplitz matrices*, Comput. Math. Appl. **25** (1993), pp. 33–45. MR **93h:**65063
6. D. K. Faddeev and V. N. Faddeeva, *Computational Methods of Linear Algebra*, Freeman and Co., San Francisco, 1963. MR **28:**1742
7. H. Federer, *Geometric Measure Theory*, Springer-Verlag, New York, 1969. MR **41:**1976
8. G. H. Golub and C. F. Van Loan, *Matrix Computation*, The Johns Hopkins University Press, 1983. MR **85h:**65063
9. U. Grenander and G. Szegö, *Toeplitz forms and their applications*, Chelsea, New York, second edition, 1984. MR **88b:**42031
10. W. Hackbusch, *Multigrid Methods and Applications*, Springer Verlag, 1985. MR **87e:**65082
11. P. Halmos, *Measure Theory*, Springer Verlag, New York, 1974.
12. M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, Nat. Bur. Standards, J. of Res. **49** (1952), pp. 409–436. MR **15:**651a
13. I. S. Iokhvidov, *Hankel and Toeplitz forms: Algebraic Theory*, Birkhäuser, Boston, 1982. MR **83k:**15021
14. S. Lang, *Real and Functional Analysis*, Springer-Verlag, third edition, 1993. MR **94b:**00005
15. S. Serra, *Preconditioning strategies for asymptotically ill-conditioned block Toeplitz systems*, BIT **34** (1994), pp. 579–594.
16. G. Szegö, *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications, 1939. MR **1:**14b
17. R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, 1962. MR **28:**1725

SCUOLA NORMALE SUPERIORE, PIAZZA CAVALIERI 7, 56100 PISA, ITALY
E-mail address: tilli@cibs.sns.it