On the Approximation Numbers of Large Toeplitz Matrices

A. Böttcher*

Received: January 14, 1997

Communicated by Alfred K. Louis

Abstract. The kth approximation number $s_k^{(p)}(A_n)$ of a complex $n \times n$ matrix A_n is defined as the distance of A_n to the $n \times n$ matrices of rank at most $n-k$. The distance is measured in the matrix norm associated with the l^p norm $(1 < p < \infty)$ on $\mathbf{C}^n$. In the case $p = 2$, the approximation numbers coincide with the singular values.

We establish several properties of $s_k^{(p)}(A_n)$ provided A_n is the $n \times n$ truncation of an infinite Toeplitz matrix A and n is large. As $n \to \infty$, the behavior of $s_k^{(p)}(A_n)$ depends heavily on the Fredholm properties (and, in particular, on the index) of A on l^p.

This paper is also an introduction to the topic. It contains a concise history of the problem and alternative proofs of the theorem by G. Heinig and F. Hellinger as well as of the scalar-valued version of some recent results by S. Roch and B. Silbermann concerning block Toeplitz matrices on l^2.

1991 Mathematics Subject Classification: Primary 47B35; Secondary 15A09, 15A18, 15A60, 47A75, 47A58, 47N50, 65F35

1. Introduction

Throughout this paper we tacitly identify a complex $n \times n$ matrix with the operator it induces on $\mathbf{C}^n$. For $1 < p < \infty$, we denote by $\mathbf{C}_p^n$ the space $\mathbf{C}^n$ with the l^p norm,

$$\|x\|_p := \Big(|x_1|^p + \ldots + |x_n|^p\Big)^{1/p},$$

and given a complex $n \times n$ matrix A_n, we put

$$\|A_n\|_p := \sup_{x \neq 0} \Big(\|A_n x\|_p / \|x\|_p\Big). \tag{1}$$

*Research supported by the Alfried Krupp Förderpreis für junge Hochschullehrer of the Krupp Foundation

We let $\mathcal{B}(\mathbf{C}_p^n)$ stand for the Banach algebra of all complex $n \times n$ matrices with the norm (1). For $j \in \{0, 1, \ldots, n\}$, let $\mathcal{F}_j^{(n)}$ be the collection of all complex $n \times n$ matrices of rank at most j, i.e., let

$$\mathcal{F}_j^{(n)} := \Big\{ F \in \mathcal{B}(\mathbf{C}_p^n) : \dim \operatorname{Im} F \leq j \Big\}.$$

The kth approximation number ($k \in \{0, 1, \ldots, n\}$) of $A_n \in \mathcal{B}(\mathbf{C}_p^n)$ is defined as

$$s_k^{(p)}(A_n) := \operatorname{dist}\,(A_n, \mathcal{F}_{n-k}^{(n)}) := \min \Big\{ \|A_n - F_n\|_p : F_n \in \mathcal{F}_{n-k}^{(n)} \Big\}. \tag{2}$$

(note that $\mathcal{F}_j^{(n)}$ is a closed subset of $\mathcal{B}(\mathbf{C}_p^n)$). Clearly,

$$0 = s_0^{(p)}(A_n) \leq s_1^{(p)}(A_n) \leq \ldots \leq s_n^{(p)}(A_n) = \|A_n\|_p.$$

It is easy to show (see Proposition 9.2) that

$$s_1^{(p)}(A_n) = \begin{cases} 1/\|A_n^{-1}\|_p & \text{if} \quad A_n \quad \text{is invertible,} \\ 0 & \text{if} \quad A_n \quad \text{is not invertible.} \end{cases} \tag{3}$$

Notice also that in the case $p = 2$ the approximation numbers $s_1^{(2)}(A_n), \ldots, s_n^{(2)}(A_n)$ are just the singular values of A_n, i.e., the eigenvalues of $(A_n^* A_n)^{1/2}$.

Let $\mathbf{T}$ be the complex unit circle and let $a \in L^\infty := L^\infty(\mathbf{T})$. The $n \times n$ Toeplitz matrix $T_n(a)$ generated by a is the matrix

$$T_n(a) := (a_{j-k})_{j,k=1}^n \tag{4}$$

where a_l ($l \in \mathbf{Z}$) is the lth Fourier coefficient of a,

$$a_l := \frac{1}{2\pi} \int_0^{2\pi} a(e^{i\theta}) e^{-il\theta} d\theta.$$

This paper is devoted to the limiting behavior of the numbers $s_k^{(p)}(T_n(a))$ as n goes to infinity.

Of course, the study of properties of $T_n(a)$ as $n \to \infty$ leads to the consideration of the infinite Toeplitz matrix

$$T(a) := (a_{j-k})_{j,k=1}^\infty.$$

The latter matrix induces a bounded operator on $l^2 := l^2(\mathbf{N})$ if (and only if) $a \in L^\infty$. Acting with $T(a)$ on $l^p := l^p(\mathbf{N})$ is connected with a multiplier problem in case $p \neq 2$. We let M_p stand for the set of all $a \in L^\infty$ for which $T(a)$ generates a bounded operator on l^p. The norm of this operator is denoted by $\|T(a)\|_p$. The function a is usually referred to as the symbol of $T(a)$ and $T_n(a)$.

In this paper, we prove the following results.

Theorem 1.1. *If $a \in M_p$ then for each k,*

$$s_{n-k}^{(p)}\Big(T_n(a)\Big) \to \|T(a)\|_p \text{ as } n \to \infty.$$

Theorem 1.2. *If $a \in M_p$ and $T(a)$ is not normally solvable on l^p then for each k,*

$$s_k^{(p)}\Big(T_n(a)\Big) \to 0 \text{ as } n \to \infty$$

Let $M_{\langle 2\rangle} := L^\infty$. For $p \neq 2$, we define $M_{\langle p\rangle}$ as the set of all functions $a \in L^\infty$ which belong to $M_{\tilde{p}}$ for all $\tilde{p}$ in some open neighborhood of p (which may depend on a). A well known result by Stechkin says that $a \in M_p$ for all $p \in (1,\infty)$ whenever $a \in L^\infty$ and the total variation $V_1(a)$ of a is finite and that in this case

$$\|T(a)\|_p \leq C_p\Big(\|a\|_\infty + V_1(a)\Big) \tag{5}$$

with some constant $C_p < \infty$ (see, e.g., [5, Section 2.5(f)] for a proof). We denote by PC the closed subalgebra of L^∞ constituted by all piecewise continuous functions. Thus, $a \in PC$ if and only if $a \in L^\infty$ and the one-sided limits

$$a(t \pm 0) := \lim_{\varepsilon \to 0 \pm 0} a(e^{i(\theta+\varepsilon)})$$

exist for every $t = e^{i\theta} \in \mathbf{T}$. By virtue of (5), the intersection $PC \cap M_{\langle p\rangle}$ contains all piecewise continuous functions of finite total variation.

Throughout what follows we define $q \in (1,\infty)$ by $1/p + 1/q = 1$ and we put

$$[p,q] := \Big[\min\{p,q\}, \max\{p,q\}\Big].$$

One can show that if $a \in M_p$, then $a \in M_r$ for all $r \in [p,q]$ (see, e.g., [5, Section 2.5(c)]).

Here is the main result of this paper.

Theorem 1.3. *Let a be a function in $PC \cap M_{\langle p\rangle}$ and suppose $T(a)$ is Fredholm of the same index $-k$ ($\in \mathbf{Z}$) on l^r for all $r \in [p,q]$. Then*

$$\lim_{n\to\infty} s_{|k|}^{(p)}\Big(T_n(a)\Big) = 0 \quad \text{and} \quad \liminf_{n\to\infty} s_{|k|+1}^{(p)}\Big(T_n(a)\Big) > 0.$$

For $p = 2$, Theorems 1.2 and 1.3 are special cases of results by Roch and Silbermann [20], [21]. Since a Toeplitz operator on l^2 with a piecewise continuous symbol is either Fredholm (of some index) or not normally solvable, Theorems 1.2 and 1.3 completely identify the approximation numbers (= singular values) which go to zero in the case $p = 2$.

Now suppose $p \neq 2$. If $a \in C \cap M_{\langle p\rangle}$, then $T(a)$ is again either Fredholm or not normally solvable, and hence Theorems 1.2 and 1.3 are all we need to see which approximation numbers converge to zero. In the case where $a \in PC \cap M_{\langle p\rangle}$ we have three mutually excluding possibilities (see Section 3):

(i) $T(a)$ is Fredholm of the same index $-k$ on l^r for all $r \in [p,q]$;

(ii) $T(a)$ is not normally solvable on l^p or not normally solvable on l^q;

(iii) $T(a)$ is normally solvable on l^p and l^q but not normally solvable on l^r for some $r \in (p,q) := [p,q] \setminus \{p,q\}$.

In the case (i) we can apply Theorem 1.3. Since

$$s_k^{(p)}\Big(T_n(a)\Big) = s_k^{(q)}\Big(T_n(a)\Big) \tag{6}$$

(see (35)), Theorem 1.2 disposes of the case (ii). I have not been able to settle the case (iii). My conjecture is as follows.

Conjecture 1.4. *In the case* (iii) *we have*

$$s_k^{(p)}\Big(T_n(a)\Big) \to 0 \ \text{ as } \ n \to \infty$$

for every fixed k.

The paper is organized as follows. Section 2 is an attempt at presenting a short history of the topic. In Section 3 we assemble some results on Toeplitz operators on l^p which are needed to prove the three theorems stated above. Their proofs are given in Sections 4 to 6. The intention of Sections 7 and 8 is to illustrate how some simple constructions show a very easy way to understand the *nature* of the Heinig/Hellinger and Roch/Silbermann results. Notice, however, that the approach of Sections 7 and 8 cannot replace the methods of these authors. They developed some sort of high technology which enabled them to tackle the block case and more general approximation methods, while in these two sections it is merely demonstrated that in the scalar case (almost) all problems can be solved with the help of a few crowbars (Theorems 7.1, 7.2, 7.4). Nevertheless, beginners will perhaps appreciate reading Sections 7 and 8 before turning to the papers [13] and [25], [20].

2. Brief history

The history of the lowest approximation number $s_1^{(p)}(T_n(a))$ is the history of the finite section method for Toeplitz operators: by virtue of (3), we have

$$s_1^{(p)}\Big(T_n(a)\Big) \to 0 \Longleftrightarrow \|T_n^{-1}(a)\|_p \to \infty.$$

We denote by $\Phi_k(l^p)$ the collection of all Fredholm operators of index k on l^p. The equivalence

$$\limsup_{n\to\infty} \|T_n^{-1}(a)\|_p < \infty \Longleftrightarrow T(a) \in \Phi_0(l^p) \tag{7}$$

was proved by Gohberg and Feldman [7] in two cases: if $a \in C \cap M_{\langle p\rangle}$ (where C stands for the continuous functions on $\mathbf{T}$) or if $p = 2$ and $a \in PC$. For $a \in PC \cap M_{\langle p\rangle}$, the equivalence

$$\limsup_{n\to\infty} \|T_n^{-1}(a)\|_p < \infty \Longleftrightarrow T(a) \in \Phi_0(l^r) \ \text{ for all } \ r \in [p,q] \tag{8}$$

holds. This was shown by Verbitsky and Krupnik [30] in the case where a has a single jump, by Silbermann and the author [3] for symbols with finitely many jumps, and finally by Silbermann [23] for symbols with a countable number of jumps. In the work of many authors, including Ambartsumyan, Devinatz, Shinbrot, Widom, Silbermann, it was pointed out that (7) is also true if

$$p = 2 \ \text{ and } \ a \in (C + H^\infty) \cup (C + \overline{H^\infty}) \cup PQC$$

(see [4], [5]). Also notice that the implication "$\Longrightarrow$" of (8) is valid for every $a \in M_p$. Treil [26] proved that there exist symbols $a \in M_{\langle 2 \rangle} = L^\infty$ such that $T(a) \in \Phi_0(l^2)$ but $\|T_n^{-1}(a)\|_2$ is not uniformly bounded; concrete symbols with this property can be found in the recent article [2, Section 7.7].

The Toeplitz matrices

$$T_n(\varphi_\gamma) = \left(\frac{1}{j-k+\gamma} \right)_{j,k=1}^{n} \qquad (\gamma \notin \mathbf{Z})$$

are the elementary building blocks of general Toeplitz matrices with piecewise continuous symbols and have therefore been studied for some decades. The symbol is given by

$$\varphi_\gamma(e^{i\theta}) = \frac{\pi}{\sin \pi\gamma} e^{i\pi\gamma} e^{-i\gamma\theta}, \ \ \theta \in [0, 2\pi).$$

This is a function in PC with a single jump at $e^{i\theta} = 1$. Tyrtyshnikov [27] focussed attention on the singular values of $T_n(\varphi_\gamma)$. He showed that

$$s_1^{(2)}\Big(T_n(\varphi_\gamma)\Big) = O(1/n^{|\gamma|-1/2}) \ \text{ if } \ \gamma \in \mathbf{R} \ \text{ and } \ |\gamma| > 1/2$$

and that there are constants $c_1, c_2 \in (0, \infty)$ such that

$$c_1 / \log n \le s_1^{(2)}\Big(T_n(\varphi_{1/2})\Big) \le c_2 / \log n.$$

Curiously, the case $|\gamma| < 1/2$ was left as an open problem in [27], although from the standard theory of Toeplitz operators with piecewise continuous symbols it is well known that

$$T(\varphi_\gamma) \in \Phi_0(l^2) \Longleftrightarrow |\mathrm{Re}\, \gamma| < 1/2$$

(see, e.g., [7, Theorem IV.2.1] or [5, Proposition 6.24]), which together with (7) (for $p = 2$ and $a \in PC$) implies that

$$\liminf_{n\to\infty} s_1^{(2)}\Big(T_n(\varphi_\gamma)\Big) = 0 \ \text{ if } \ |\mathrm{Re}\, \gamma| \ge 1/2 \tag{9}$$

and

$$\liminf_{n\to\infty} s_1^{(2)}\Big(T_n(\varphi_\gamma)\Big) > 0 \ \text{ if } \ |\mathrm{Re}\, \gamma| < 1/2$$

(see [20]). A simple and well known argument (see the end of Section 3) shows that in (9) the liminf can actually be replaced by lim.

Also notice that it was already in the seventies when Verbitsky and Krupnik [30] proved that

$$\lim_{n\to\infty} s_1^{(p)}\Big(T_n(\varphi_\gamma)\Big) = 0 \iff |\mathrm{Re}\, \gamma| \ge \min\{1/p, 1/q\}$$

(full proofs are also in [4, Proposition 3.11] and [5, Theorem 7.37; in part (iii) of that theorem there is a misprint: the $-1/p < \operatorname{Re}\beta < 1/q$ must be replaced by $-1/q < \operatorname{Re}\beta < 1/p$]).

As far as I know, collective phenomena of $s_1^{(p)}(T_n(a)), \ldots, s_n^{(p)}(T_n(a))$ have been studied only for $p = 2$, and throughout the rest of this section we abbreviate $s_k^{(2)}(T_n(a))$ to $s_k(T_n(a))$.

In 1920, Szegö showed that if $a \in L^\infty$ is real-valued and F is continuous on $\mathbf{R}$, then

$$\frac{1}{n}\sum_{k=1}^{n} F\Big(s_k(T_n(a))\Big) \to \frac{1}{2\pi}\int_0^{2\pi} F\Big(|a(e^{i\theta})|\Big)\,d\theta. \tag{10}$$

In the eighties, Parter [15] and Avram [1] extended this result to arbitrary (complex-valued) symbols $a \in L^\infty$. Formula (10) implies that

$$\Big\{s_k(T_n(a))\Big\}_{k=1}^{n} \ \text{ and } \ \Big\{|a(e^{2\pi i k/n})|\Big\}_{k=1}^{n} \tag{11}$$

are equally distributed (see [9] and [29]).

Research into the asymptotic distribution of the singular values of Toeplitz matrices was strongly motivated by a phenomenon discovered by C. Moler in the middle of the eighties. Moler observed that almost all singular values of $T_n(\varphi_{1/2})$ are concentrated in $[\pi - \varepsilon, \pi]$ where ε is very small. Formula (10) provides a way to understand this phenomenon: letting $F = 1$ on $[0, \pi - 2\varepsilon]$ and $F = 0$ on $[\pi - \varepsilon, \pi]$ and taking into account that $|\varphi_{1/2}| = 1$, one gets

$$\frac{1}{n}\sum_{k=1}^{n} F\Big(s_k(T_n(\varphi_{1/2}))\Big) \to \frac{1}{2\pi}\int_0^{2\pi} F(1)\,d\theta = F(1) = 0,$$

which shows that the percentage of the singular values of $T_n(\varphi_{1/2})$ which are located in $[0, \pi - 2\varepsilon]$ goes to zero as n increases to infinity.

Widom [32] was the first to establish a second order result on the asymptotics of singular values. Under the assumption that

$$a \in L^\infty \ \text{ and } \ \sum_{n\in\mathbf{Z}} |n|\,|a_n|^2 < \infty$$

and that $F \in C^3(\mathbf{R})$, he showed that

$$\sum_{k=1}^{n} F\Big(s_k^2(T_n(a))\Big) = \frac{n}{2\pi}\int_0^{2\pi} F\Big(|a(e^{i\theta})|^2\Big)\,d\theta + E_F(a) + o(1)$$

with some constant $E_F(a)$, and he gave an expression for $E_F(a)$. He also introduced two limiting sets of the sets

$$\Sigma(T_n(a)) := \Big\{s_1(T_n(a)), \ldots, s_n(T_n(a))\Big\},$$

which, following the terminology of [19], are defined by

$$\begin{aligned}\Lambda_{\text{part}}\left(\Sigma(T_n(a))\right) &:= \{\lambda\in\mathbf{R}:\lambda \text{ is partial limit of some sequence}\\ &\qquad \{\lambda_n\} \text{ with } \lambda_n\in\Sigma(T_n(a))\},\\ \Lambda_{\text{unif}}\left(\Sigma(T_n(a))\right) &:= \{\lambda\in\mathbf{R}:\lambda \text{ is the limit of some sequence}\\ &\qquad \{\lambda_n\} \text{ with } \lambda_n\in\Sigma(T_n(a))\}.\end{aligned}$$

It turned out that for large classes of symbols a we have

$$\Lambda_{\text{part}}\left(\Sigma(T_n(a))\right)=\Lambda_{\text{unif}}\left(\Sigma(T_n(a))\right)=\operatorname{sp}\left(T(\overline{a})T(a)\right)^{1/2} \tag{12}$$

where $\operatorname{sp} A := \{\lambda\in\mathbf{C} : A-\lambda I \text{ is not invertible}\}$ denotes the spectrum of A (on l^2) and $\overline{a}$ is defined by $\overline{a}(e^{i\theta}):=\overline{a(e^{i\theta})}$. Note that $T(\overline{a})$ is nothing but the adjoint $T^*(a)$ of $T(a)$. Widom [32] proved (12) under the hypothesis that $a\in PC$ or that a is locally self-adjoint, while Silbermann [24] derived (12) for locally normal symbols. Notice that symbols in PC or even in PQC are locally normal.

In the nineties, Tyrtyshnikov [28], [29] succeeded in proving that the sets (11) are equally distributed under the sole assumption that $a\in L^2:=L^2(\mathbf{T})$. His approach is based on the observation that if $\|A_n-B_n\|_F=o(n)$, where $\|\cdot\|_F$ stands for the Frobenius (or Hilbert-Schmidt) norm, then A_n and B_n have equally distributed singular values. The result mentioned can be shown by taking $A_n=T_n(a)$ and choosing appropriate circulants for B_n.

The development received a new impetus from Heinig and Hellinger's 1994 paper [13]. They considered normally solvable Toeplitz operators on l^2 and studied the problem whether the Moore-Penrose inverses of $T_n^+(a)$ of $T_n(a)$ converge strongly on l^2 to the Moore-Penrose inverse $T^+(a)$ of $T(a)$. Recall that the Moore-Penrose inverse of a normally solvable Hilbert space operator A is the (uniquely determined) operator A^+ satisfying

$$AA^+A=A,\;\; A^+AA^+=A^+,\;\; (A^+A)^*=A^+A,\;\; (AA^+)^*=AA^+.$$

If $a\in C$, then $T(a)$ is normally solvable on l^2 if and only if $a(t)\neq 0$ for all $t\in\mathbf{T}$. When writing $T_n^+(a)\to T^+(a)$, we actually mean that $T_n^+(a)P_n\to T^+(a)$, where P_n is the projection defined by

$$P_n:\{x_1,x_2,x_3,\ldots\}\mapsto\{x_1,x_2,\ldots,x_n,0,0,\ldots\}. \tag{13}$$

It is not difficult to verify that $T_n^+(a)\to T^+(a)$ strongly on l^2 if and only if $T(a)$ is normally solvable and

$$\limsup_{n\to\infty}\|T_n^+(a)\|_2<\infty. \tag{14}$$

Heinig and Hellinger investigated normally solvable Toeplitz operators $T(a)$ with symbols in the Wiener algebra W,

$$a\in W \iff \|a\|_W:=\sum_{n\in\mathbf{Z}}|a_n|<\infty,$$

and they showed that then (14) is satisfied if and only if there is an $n_0\geq 1$ such that

$$\operatorname{Ker} T(a)\subset \operatorname{Im} P_{n_0} \text{ and } \operatorname{Ker} T(\overline{a})\subset \operatorname{Im} P_{n_0}, \tag{15}$$

where $\operatorname{Ker} A := \{x \in l^2 : Ax = 0\}$ and $\operatorname{Im} A := \{Ax : x \in l^2\}$. (This formulation of the Heinig-Hellinger result is due to Silbermann [25].) Conditions (15) are obviously met if $T(a)$ is invertible, in which case even $\|T_n^{-1}(a)\|_2$ is uniformly bounded. The really interesting case is the one in which $T(a)$ is not invertible, and in that case (15) and thus (14) are highly instable. For example, if a is a rational function (without poles on $\mathbf{T}$) and $\lambda \in \operatorname{sp} T(a)$, then

$$\limsup_{n\to\infty} \|T_n^+(a-\lambda)\|_2 < \infty \tag{16}$$

can only hold if λ belongs to $\operatorname{sp} T_n(a)$ for all sufficiently large n. Consequently, (16) implies that λ lies in $\Lambda_{\text{unif}}(\operatorname{sp} T_n(a))$, and the latter set is extremely "thin": it is contained in a finite union of analytic arcs (see [22] and [6]).

What has Moore-Penrose invertibility to do with singular values ? The answer is as follows: if $A_n \in \mathcal{B}(\mathbf{C}_2^n)$ and $s_k(A_n)$ is the smallest nonzero singular value of A_n, then

$$\|A_n^+\|_2 = 1/s_k(A_n).$$

Thus, (14) holds exactly if there exists a $d > 0$ such that

$$\Sigma(T_n(a)) \subset \{0\} \cup [d,\infty) \tag{17}$$

for all sufficiently large n.

Now Silbermann enters the scene. He replaced the Heinig-Hellinger problem by another one. Namely, given $T(a)$, is there a sequence $\{B_n\}$ of operators $B_n \in \mathcal{B}(\mathbf{C}_2^n)$ with the following properties: there exists a bounded operator B on l^2 such that

$$B_n \to B \;\text{ and }\; B_n^* \to B^* \;\text{ strongly on }\; l^2$$

and

$$\begin{gathered}\|T_n(a)B_nT_n(a) - T_n(a)\|_2 \to 0, \;\; \|B_nT_n(a)B_n - B_n\|_2 \to 0, \\ \|(B_nT_n(a))^* - B_nT_n(a)\|_2 \to 0, \;\; \|(T_n(a)B_n)^* - T_n(a)B_n\|_2 \to 0 \;?\end{gathered}$$

Such a sequence $\{B_n\}$ is referred to as an asymptotic Moore-Penrose inverse of $T(a)$. In view of the (instable) conditions (15), the following result by Silbermann [25] is surprising: if $a \in PC$ and $T(a)$ is normally solvable, then $T(a)$ always has an asymptotic Moore-Penrose inverse. And what is the concern of this result with singular values ? One can easily show $T(a)$ has an asymptotic Moore-Penrose inverse if and only if there is a sequence $c_n \to 0$ and a number $d > 0$ such that

$$\Sigma(T_n(a)) \subset [0,c_n] \cup [d,\infty). \tag{18}$$

One says that $\Sigma(T_n(a))$ has the splitting property if (18) holds with $c_n \to 0$ and $d > 0$. Thus, Silbermann's result implies that if $a \in PC$ and $T(a)$ is normally solvable on l^2, then $\Sigma(T_n(a))$ has the splitting property.

Only recently, Roch and Silbermann [20], [21] were able to prove even much more. The sets $\Sigma(T_n(a))$ are said to have the k-splitting property, where $k \ge 0$ is an integer, if (18) is true for some sequence $c_n \to 0$ and some $d > 0$ and, in addition, exactly k singular values lie in $[0,c_n]$ and $n-k$ singular values are located in $[d,\infty)$

(here multiplicities are taken into account). Equivalently, $\Sigma(T_n(a))$ has the k-splitting property if and only if

$$\lim_{n\to\infty} s_k(T_n(a)) = 0 \quad \text{and} \quad \liminf_{n\to\infty} s_{k+1}(T_n(a)) > 0. \tag{19}$$

A normally solvable Toeplitz operator $T(a)$ on l^2 with a symbol $a \in PC$ is automatically Fredholm and therefore has some index $k \in \mathbf{Z}$. Roch and Silbermann [20], [21] discovered that then $\Sigma(T_n(a))$ has the $|k|$-splitting property. In other words, if $a \in PC$ and $T(a) \in \Phi_k(l^2)$ then (19) holds with k replaced by $|k|$. Notice that this Theorem 1.3 for $p = 2$.

In fact, it was the Roch and Silbermann papers [20], [21] which stimulated me to do some thinking about singular values. It was the feeling that the $|k|$-splitting property must have its root in the possibility of "ignoring $|k|$ dimensions" which led me to the observation that none of the works cited in this section makes use of the fact that $s_k(A_n)$ may alternatively be defined by (2), i.e. that singular values may also be viewed as approximation numbers. I then realized that some basic phenomena of [20] and [21] can be very easily understood by having recourse to (2) and that, moreover, using (2) is a good way to pass from l^2 and C^*-algebras to l^p and Banach algebras.

3. Toeplitz operators on l^p

We henceforth always assume that $1 < p < \infty$ and $1/p + 1/q = 1$.

Let M_p and $M_{\langle p\rangle}$ be as in Section 1. The set M_p can be shown to be a Banach algebra with pointwise algebraic operations and the norm $\|a\|_{M_p} := \|T(a)\|_p$. It is also well known that

$$M_p = M_q \subset M_2 = L^\infty$$

and

$$\|a\|_{M_p} = \|a\|_{M_q} \ge \|a\|_{M_2} = \|a\|_\infty \tag{20}$$

(see, e.g., [5, Section 2.5]). We remark that working with $M_{\langle p\rangle}$ instead of M_p is caused by the need of somehow reversing the estimate in (20). Suppose, for instance, $p > 2$ and $a \in M_{\langle p\rangle}$. Then $a \in M_{p+\varepsilon}$ for some $\varepsilon > 0$, and the Riesz-Thorin interpolation theorem gives

$$\|a\|_{M_p} \le \|a\|_{M_2}^{\gamma} \, \|a\|_{p+\varepsilon}^{1-\gamma} = \|a\|_\infty^{\gamma} \, \|a\|_{M_{p+\varepsilon}}^{1-\gamma} \tag{21}$$

with some $\gamma \in (0,1)$ depending only on p and ε. The $\|a\|_{M_{p+\varepsilon}}$ on the right of (21) may in turn be estimated by $C_p(\|a\|_\infty + V_1(a))$ (recall Stechkin's inequality (5)) provided a has bounded total variation.

A bounded linear operator A on l^p is said to be normally solvable if its range, $\operatorname{Im} A$, is a closed subset of l^p. The operator A is called Fredholm if it is normally solvable and the spaces

$$\operatorname{Ker} A := \{x \in l^p : Ax = 0\} \quad \text{and} \quad \operatorname{Coker} A := l^p/\operatorname{Im} A$$

have finite dimensions. In that case the index $\operatorname{Ind} A$ is defined as

$$\operatorname{Ind} A := \dim \operatorname{Ker} A - \dim \operatorname{Coker} A.$$

We denote by $\Phi(l^p)$ the collection of all Fredholm operators on l^p and by $\Phi_k(l^p)$ the operators in $\Phi(l^p)$ whose index is k. The following four theorems are well known. Comments are at the end of this section.

Theorem 3.1. *Let $a \in M_p$.*

(a) *If a does not vanish identically, then the kernel of $T(a)$ on l^p or the kernel of $T(\overline{a})$ on l^q is trivial.*

(b) *The operator $T(a)$ is invertible on l^p if and only if $T(a) \in \Phi_0(l^p)$.*

Of course, part (b) is a simple consequence of part (a).

Theorem 3.2. *Let $a \in C \cap M_{\langle p \rangle}$. Then $T(a)$ is normally solvable on l^p if and only if $a(t) \neq 0$ for all $t \in \mathbf{T}$. In that case $T(a) \in \Phi(l^p)$ and*

$$\operatorname{Ind} T(a) = -\operatorname{wind} a,$$

where wind a *is the winding number of a about the origin.*

Now let $a \in PC,\ t \in \mathbf{T}$, and suppose $a(t-0) \neq a(t+0)$. We denote by

$$\mathcal{A}_p(a(t-0), a(t+0))$$

the circular arc at the points of which the line segment $[a(t-0), a(t+0)]$ is seen at the angle $\max\{2\pi/p, 2\pi/q\}$ and which lies on the right of the straight line passing first $a(t-0)$ and then $a(t+0)$ if $1 < p < 2$ and on the left of this line if $2 < p < \infty$. For $p = 2$, $\mathcal{A}_p(a(t-0), a(t+0))$ is nothing but the line segment $[a(t-0), a(t+0)]$ itself. Let $a_p^\#$ denote the closed, continuous, and naturally oriented curve which results from the (essential) range $\mathcal{R}(a)$ of a by filling in the arcs $\mathcal{A}_p(a(t-0), a(t+0))$ for each jump. In case this curve does not pass through the origin, we let $\operatorname{wind} a_p^\#$ be its winding number.

Theorem 3.3. *Let $a \in PC \cap M_{\langle p \rangle}$. Then $T(a)$ is normally solvable on l^p if and only if $0 \notin a_p^\#$. In that case $T(a) \in \Phi(l^p)$ and*

$$\operatorname{Ind} T(a) = -\operatorname{wind} a_p^\#.$$

For $a \in PC$ and $t \in \mathbf{T}$, put

$$\mathcal{O}_p\Big(a(t-0),\ a(t+0)\Big) := \bigcup_{r \in [p,q]} \mathcal{A}_r\Big(a(t-0),\ a(t+0)\Big). \tag{22}$$

If $a(t-0) \neq a(t+0)$ and $p \neq 2$, then $\mathcal{O}_p(a(t-0), a(t+0))$ is a certain lentiform set. Also for $a \in PC$, let

$$a_{[p,q]}^\# := \bigcup_{r \in [p,q]} a_r^\#.$$

Thus, $a^{\#}_{[p,q]}$ results from $\mathcal{R}(a)$ by filling in the sets (22) between the endpoints of the jumps. If $0 \notin a^{\#}_{[p,q]}$, then necessarily $0 \notin a^{\#}_{2}$ and we define $\operatorname{wind} a^{\#}_{[p,q]}$ as $\operatorname{wind} a^{\#}_{2}$ in this case.

From Theorem 3.3 we deduce that the conditions (i) to (iii) of Section 1 are equivalent to the following:

(i') $0 \notin a^{\#}_{[p,q]}$ and $\operatorname{wind} a^{\#}_{[p,q]} = k$;

(ii') $0 \in a^{\#}_{p} \cup a^{\#}_{q}$;

(iii') $0 \in a^{\#}_{[p,q]} \setminus (a^{\#}_{p} \cup a^{\#}_{q})$.

For $a \in M_p$, let $T_n(a) \in \mathcal{B}(\mathbf{C}^n_p)$ be the operator given by the matrix (4). One says that the sequence $\{T_n(a)\} := \{T_n(a)\}_{n=1}^{\infty}$ is stable if

$$\limsup_{n\to\infty} \|T_n^{-1}(a)\|_p < \infty.$$

Here we follow the practice of putting

$$\|T_n^{-1}(a)\|_p = \infty \ \text{ if } \ T_n(a) \ \text{ is not invertible.}$$

In other words, $\{T_n(a)\}$ is stable if and only if $T_n(a)$ is invertible for all $n \ge n_0$ and there exists a constant $M < \infty$ such that $\|T_n^{-1}(a)\|_p \le M$ for all $n \ge n_0$. From (3) we infer that

$$\{T_n(a)\} \ \text{ is stable } \ \Longleftrightarrow \liminf_{n\to\infty} s_1^{(p)}(T_n(a)) > 0.$$

Theorem 3.4. (a) *If $a \in C \cap M_{\langle p\rangle}$ then*

$$\{T_n(a)\} \ \textit{is stable} \ \Longleftrightarrow \ 0 \notin a(\mathbf{T}) \ \textit{and} \ \operatorname{wind} a = 0.$$

(b) If $a \in PC \cap M_{\langle p\rangle}$ then

$$\{T_n(a)\} \ \textit{is stable} \ \Longleftrightarrow \ 0 \notin a^{\#}_{[p,q]} \ \textit{and} \ \operatorname{wind} a^{\#}_{[p,q]} = 0.$$

As already said, these theorems are well known. Theorem 3.1 is due to Coburn ($p = 2$) and Duduchava ($p \ne 2$), Theorem 3.2 is Gohberg and Feldman's, Theorem 3.3 is the result of many authors in the case $p = 2$ and was established by Duduchava for $p \ne 2$, Theorem 3.4 goes back to Gohberg and Feldman for $a \in C \cap M_{\langle p\rangle}$ (general p) and $a \in PC$ ($p = 2$), and it was obtained in the work of Verbitsky, Krupnik, Silbermann, and the author for $a \in PC \cap M_{\langle p\rangle}$ and $p \ne 2$. Precise historical remarks and full proofs are in [5].

Part (a) of Theorem 3.4 is clearly a special case of part (b). In fact, Theorem 3.4(b) may also be stated as follows: $\{T_n(a)\}$ contains a stable subsequence

$\{T_{n_j}(a)\}$ $(n_j \to \infty)$ if and only if $0 \notin a^{\#}_{[p,q]}$ and $\operatorname{wind} a^{\#}_{[p,q]} = 0$. Hence, we arrive at the conclusion that if $a \in PC \cap M_{\langle p \rangle}$, then

$$\begin{aligned} & s_1^{(p)}(T_n(a)) \to 0 \\ & \Longleftrightarrow \{T_n(a)\} \text{ is stable} \\ & \Longleftrightarrow 0 \in a^{\#}_{[p,q]} \text{ or } \Big(0 \notin a^{\#}_{[p,q]} \text{ and } \operatorname{wind} a^{\#}_{[p,q]} \neq 0\Big). \end{aligned}$$

At this point the question of whether the lowest approximation number of $T_n(a)$ goes to zero or not is completely disposed of for symbols $a \in PC \cap M_{\langle p \rangle}$.

4. Proof of Theorem 1.1.

Contrary to what we want, let us assume that there is a $c < \|T(a)\|_p$ such that $s_{n-k}^{(p)}(T_n(a)) \le c$ for all n in some infinite set $\mathcal{N}$. Since $s_{n-k}^{(p)}(T_n(a)) = \operatorname{dist}(T_n(a), \mathcal{F}_k^{(n)})$, we can find $F_n \in \mathcal{F}_k^{(n)}$ $(n \in \mathcal{N})$ so that $\|T_n(a) - F_n\|_p \le c$. For $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$, we define

$$(x, y) := x_1 y_1 + \ldots + x_n y_n. \tag{23}$$

By [16, Lemma B.4.11], there exist $e_j^{(n)} \in \mathbf{C}_p^n$, $f_j^{(n)} \in \mathbf{C}_p^n$, $\gamma_j^{(n)} \in \mathbf{C}$ such that

$$F_n x = \sum_{j=1}^{k} \gamma_j^{(n)} \Big(x, f_j^{(n)}\Big) e_j^{(n)} \quad (x \in \mathbf{C}_p^n),$$

$\|e_j^{(n)}\|_p = 1$, $\|f_j^{(n)}\|_q = 1$, and

$$|\gamma_j^{(n)}| \le \|F_n\|_p \le \|T_n(a)\|_p + \|F_n - T_n(a)\|_p \le \|T(a)\|_p + c \tag{24}$$

for all $j \in \{1, \dots, k\}$.

Fix $x \in \mathbf{C}_p^n, y \in \mathbf{C}_q^n$ and suppose $\|x\|_p = 1$, $\|y\|_q = 1$. We then have

$$\left| \Big(T_n(a) x, y\Big) - \sum_{j=1}^{k} \gamma_j^{(n)} \Big(x, f_j^{(n)}\Big) \Big(e_j^{(n)}, y\Big) \right| \le \|T_n(a) - F_n\|_p \le c. \tag{25}$$

Clearly, $(T_n(a)x, y) \to (T(a)x, y)$. From (24) and the Bolzano-Weierstrass theorem we infer that the sequence $\{(\gamma_1^{(n)}, \dots, \gamma_k^{(n)})\}_{n \in \mathcal{N}}$ has a converging subsequence. Without loss of generality suppose the sequence itself converges, i.e.

$$\Big(\gamma_1^{(n)}, \dots, \gamma_k^{(n)}\Big) \to (\gamma_1, \dots, \gamma_k) \in \mathbf{C}^k$$

as $n \in \mathcal{N}$ goes to infinity. The vectors $e_j^{(n)}$ and $f_j^{(n)}$ all belong to the unit sphere of l^p and l^q, respectively. Hence, by the Banach-Alaoglu theorem (see, e.g., [18, Theorem IV.21]), $\{e_j^{(n)}\}_{n \in \mathcal{N}}$ and $\{f_j^{(n)}\}_{n \in \mathcal{N}}$ have subsequences converging in the weak $*$-topology. Again we may without loss of generality assume that

$$e_j^{(n)} \to e_j \in l^p, \; f_j^{(n)} \to f_j \in l^q$$

in the weak $*$-topology as $n \in \mathcal{N}$ goes to infinity.

From (25) we now obtain that if $x \in l^p$ and $y \in l^q$ have finite support and $\|x\|_p = 1$, $\|y\|_q = 1$, then

$$\left| \Big(T(a)x, y \Big) - \sum_{j=1}^{k} \gamma_j (x, f_j)(e_j, y) \right| \le c.$$

This implies that

$$\|T(a) - F\|_p \le c \tag{26}$$

where F is the finite-rank operator given by

$$Fx := \sum_{j=1}^{k} \gamma_j (x, f_j) e_j \quad (x \in l^p). \tag{27}$$

Let $\|T(a)\|^{(\mathrm{ess})}$ denote the essential norm of $T(a)$ on l^p, i.e. the distance of $T(a)$ to the compact operators on l^p. By (26) and (27),

$$\|T(a)\|_p^{(\mathrm{ess})} \le \|T(a) - F\|_p \le c < \|T(a)\|_p.$$

However, one always has $\|T(a)\|_p^{(\mathrm{ess})} = \|T(a)\|_p$ (see, e.g., [5, Proposition 4.4(d)]). This contradiction completes the proof.

5. Proof of Theorem 1.2.

We will employ the following two results.

Theorem 5.1. *Let A be a bounded linear operator on l^p.*

(a) *The operator A is normally solvable on l^p if and only if*

$$k_A := \sup_{x \in l^p,\ \|x\|_p = 1} \operatorname{dist}(x, \operatorname{Ker} A) < \infty.$$

(b) *If M is a closed subspace of l^p and $\dim (l^p/M) < \infty$, then the normal solvability of $A|M : M \to l^p$ is equivalent to the normal solvability of $A : l^p \to l^p$.*

A proof is in [8, pp. 159–160]. □

Theorem 5.2. *If M is a k-dimensional subspace of $\mathbf{C}_p^n$, then there exists a projection $\Pi : \mathbf{C}_p^n \to \mathbf{C}_p^n$ such that $\operatorname{Im} \Pi = M$ and $\|\Pi\|_p \le k$.*

This is a special case of [16, Lemma B.4.9]. □

Theorem 1.2 is trivial in case a vanishes identically. So suppose $a \in M_p \setminus \{0\}$ and $T(a)$ is not normally solvable on l^p. Then the adjoint operator $T(\overline{a})$ is not normally solvable on l^q. By Theorem 3.1(a), $\operatorname{Ker} T(a) = \{0\}$ on l^p or $\operatorname{Ker} T(\overline{a}) = \{0\}$ on l^q.

Since $s_k^{(p)}(T_n(a)) = s_k^{(q)}(T_n(\overline{a}))$, we may a priori assume that $\operatorname{Ker} T(a) = \{0\}$ on l^p. Abbreviate $T(a)$ and $T_n(a)$ to A and A_n, respectively.

Define P_n on l^p by (13) and let

$$V := l^p \to l^p, \ \ \{x_1, x_2, x_3, \ldots\} \mapsto \{0, x_1, x_2, x_3, \ldots\}.$$

As $A|\operatorname{Im} V^n : \operatorname{Im} V^n \to l^p$ has the same matrix as $AV^n : l^p \to l^p$, we deduce from Theorem 5.1(b) that there is no $n \geq 0$ such that AV^n is normally solvable. Note that $\operatorname{Ker}(AV^n) = \{0\}$ for all $n \geq 0$.

Let $l^p(n_1, n_2]$ denote the sequences $\{x_j\}_{j=1}^\infty \in l^p$ which are supported in $(n_1, n_2]$, i.e., for which $x_j = 0$ whenever $j \leq n_1$ or $j > n_2$.

Lemma 5.3. *There are* $0 = N_0 < N_1 < N_2 < \ldots$ *and* $z_j \in l^p(N_{j-1}, N_j]$ $(j \geq 1)$ *such that*

$$\|z_j\|_p = 1 \ \ \textit{and} \ \ \|Az_j\|_p \to 0 \ \ \textit{as} \ \ j \to \infty.$$

Proof. By Theorem 5.1(a), there is a $y_1 \in l^p$ such that $\|y_1\|_p = 2$ and $\|Ay_1\| < 1/2$. If N_1 is large enough, then $\|P_{N_1} y_1\|_p \geq 1$ and $\|AP_{N_1} y_1\|_p < 1$. Letting $z_1 := P_{N_1 y_1}/\|P_{N_1} y_1\|_p$ we get

$$z_1 \in l^p(0, N_1], \ \ \|z_1\|_p = 1, \ \ \|Az_1\|_p < 1.$$

Applying Theorem 5.1(a) to the operator AV^{N_1}, we see that there is an $y_2 \in l^p$ such that $\|y_2\|_p = 2$ and $\|AV^{N_1} y_2\|_p < 1/4$. For sufficiently large $N_2 > N_1$ we have $\|P_{N_2} V^{N_1} y_2\|_p \geq 1$ and $\|AP_{N_2} V^{N_1} y_2\|_p < 1/2$. Setting

$$z_2 := P_{N_2} V^{N_1} y_2 / \|P_{N_2} V^{N_1} y_2\|_p,$$

we therefore obtain

$$z_2 \in l^p(N_1, N_2], \ \ \|z_2\|_p = 1, \ \ \|Az_2\|_p < 1/2.$$

Continuing in this way we find z_j satisfying

$$z_j \in l^p(N_{j-1}, N_j], \ \ \|z_j\|_p = 1, \ \ \|Az_j\|_p < 1/j. \qquad \square$$

Contrary to the assertion of Theorem 1.2, let us assume that there exist $k \geq 1$ and $d > 0$ such that $s_k^{(p)}(A_n) \geq d$ for infinitely many n. We may without loss of generality assume that

$$s_k^{(p)}(A_n) \geq d \ \text{ for all } \ n \geq n_0. \tag{28}$$

Let $\varepsilon > 0$ be any number such that

$$2\varepsilon k^2 < d. \tag{29}$$

Choose z_j as in Lemma 5.3. Obviously, there are sufficiently large j and N such that

$$\|P_N z_l\|_p \geq 1/2, \ \ \|AP_N z_l\|_p < \varepsilon \ \text{ for } \ l \in \{j+1, \ldots, j+k\}. \tag{30}$$

Since $P_N z_l \in l^p(N_{l-1,l}]$, it is clear that $P_N z_{j+1}, \ldots, P_N z_{j+k}$ are linearly independent. Now let $n \geq N$. By Theorem 5.2, there is a projection Π_n of $\mathbf{C}_p^n$ onto $\mathrm{span}\{P_N z_{j+1}, \ldots, P_N z_{j+k}\}$ for which $\|\Pi_n\|_p \leq k$. Let I_n stand for the identity operator on $\mathbf{C}_p^n$. The space $\mathrm{Im}\,(I_n - \Pi_n) = \mathrm{Ker}\,\Pi_n$ has the dimension $n-k$ and hence, $I_n - \Pi_n \in \mathcal{F}_{n-k}^{(n)}$. Every $x \in \mathbf{C}_p^n$ can be uniquely written in the form

$$x = \gamma_1 P_N z_{j+1} + \ldots + \gamma_k P_N z_{j+k} + w \quad \text{with} \quad w \in \mathrm{Ker}\,\Pi_n.$$

Thus,

$$\begin{aligned} &\|A_n x - A_n(I_n - \Pi_n)x\|_p = \|A_n \Pi_n x\|_p \\ &= \|\gamma_1 A_n(P_N z_{j+1}) + \ldots + \gamma_k A_n(P_N z_{j+k})\|_p \leq |\gamma_1|\varepsilon + \ldots + |\gamma_k|\varepsilon, \end{aligned} \tag{31}$$

the estimate resulting from (30). Taking into account that the sequences $P_N z_l$ have pairwise disjoint supports, we obtain from (30) that

$$\begin{aligned} &\|\Pi_n x\|_p^p = \|\gamma_1 P_N z_{j+1} + \ldots + \gamma_k P_N z_{j+k}\|_p^p \\ &= |\gamma_1|^p \|P_N z_{j+1}\|_p^p + \ldots + |\gamma_k|^p \|P_N z_{j+k}\|_p^p \\ &\geq (1/2)^p \Big(|\gamma_1|^p + \ldots + |\gamma_k|^p\Big) \geq (1/2)^p \max_{1 \leq m \leq k} |\gamma_m|^p. \end{aligned} \tag{32}$$

Combining (31) and (32) we get

$$\|A_n x - A_n(I_n - \Pi_n)x\|_p \leq \varepsilon k \max_{1 \leq m \leq k} |\gamma_m| \leq 2\varepsilon k \|\Pi_n x\|_p \leq 2\varepsilon k^2 \|x\|_p,$$

whence $s_k^{(p)}(A_n) = \mathrm{dist}\,(A_n, \mathcal{F}_{n-k}^{(n)}) \leq \|A_n - A_n(I - \Pi_n)\|_p \leq 2\varepsilon k^2$. By virtue of (29), this contradicts (28) and completes the proof.

6. Proof of Theorem 1.3.

The Hankel operator on l^p induced by a function $a \in M_p$ is given by the matrix

$$H(a) = (a_{j+k-1})_{j,k=1}^{\infty}.$$

For $a \in M_p$, define $\tilde{a} \in M_p$ by $\tilde{a}(e^{i\theta}) := a(e^{-i\theta})$. Clearly,

$$H(\tilde{a}) = (a_{-j-k+1})_{j,k=1}^{\infty}.$$

It is well known and easily seen that

$$T(ab) = T(a)T(b) + H(a)H(\tilde{b}) \tag{33}$$

for every $a, b \in M_p$. A finite section analogue of formula (33) reads

$$T_n(ab) = T_n(a)T_n(b) + P_n H(a)H(\tilde{b})P_n + W_n H(\tilde{a})H(b)W_n, \tag{34}$$

where P_n is as in (13) and W_n is defined by

$$W_n : \{x_1, x_2, x_3, \ldots\} \mapsto \{x_n, x_{n-1}, \ldots, x_1, 0, 0, \ldots\}.$$

The identity (34) first appeared in Widom's paper [31], a proof is also in [4, Proposition 3.6] and [5, Proposition 7.7].

We remark that $T_n(\tilde{a})$ is the transposed matrix of $T_n(a)$ and that the identity $T_n(\tilde{a}) = W_n T_n(a) W_n$ holds. In particular, we have

$$\begin{aligned}
s_k^{(q)}(T_n(a)) &= \min\left\{\|T_n(a) - F_{n-k}\|_q : F_{n-k} \in \mathcal{F}_{n-k}^{(n)}\right\} \\
&= \min\left\{\|T_n(\tilde{a}) - G_{n-k}\|_p : G_{n-k} \in \mathcal{F}_{n-k}^{(n)}\right\} \\
&= \min\left\{\|W_n(T_n(\tilde{a}) - G_{n-k})W_n\|_p : G_{n-k} \in \mathcal{F}_{n-k}^{(n)}\right\} \\
&= \min\left\{\|T_n(a) - H_{n-k}\|_p : H_{n-k} \in \mathcal{F}_{n-k}^{(n)}\right\} \\
&= s_k^{(p)}(T_n(a)) \qquad (35)
\end{aligned}$$

(note also that W_n is an invertible isometry on $\mathbf{C}_p^n$).

To prove Theorem 1.3, we need the following two (well known) lemmas.

Lemma 6.1. *If $A, B, C \in \mathcal{B}(\mathbf{C}_p^n)$ then*

$$s_k^{(p)}(ABC) \le \|A\|_p \, s_k^{(p)}(B) \|C\|_p \ \textit{ for all } k.$$

This follows easily from the definition of $s_k^{(p)}$. □

Lemma 6.2. *If $b \in M_p$ and $\{T_n(b)\}$ is stable on l^p, then $T(b)$ is invertible on l^p and $T_n^{-1}(b)\,(:= T_n^{-1}(b)P_n)$ converges strongly on l^p to $T^{-1}(b)$.*

This is obvious from the estimates

$$\begin{aligned}
&\|T_n^{-1}(b)P_n y - T^{-1}(b)y\|_p \\
&\le \|T_n^{-1}(b)\|_p \, \|P_n y - T_n(b)P_n T^{-1}(b)y\|_p + \|P_n T^{-1}(b)y - T^{-1}(b)y\|_p, \\
&\|x\|_p \le \liminf_{n\to\infty} \|T_n^{-1}(b)\|_p \, \|T(b)x\|_p, \quad \|\xi\|_q \le \liminf_{n\to\infty} \|T_n^{-1}(\tilde{b})\|_q \, \|T(\tilde{b})\xi\|_q.
\end{aligned}$$

□

We now establish two propositions which easily imply Theorem 1.3.

Define χ_k by $\chi_k(e^{i\theta}) = e^{ik\theta}$. Using Theorem 3.1(b) and formula (33) one can readily see that if $a \in M_p$, then $T(a) \in \Phi_{-k}(l^p)$ if and only if $a = b\chi_k$ and $T(b)$ is invertible on l^p.

Propostion 6.3. *If $b \in M_p$ and $\{T_n(b)\}$ is stable on l^p then for every $k \in \mathbf{Z}$,*

$$\liminf_{n\to\infty} s_{|k|+1}^{(p)}\Big(T_n(b\chi_k)\Big) > 0.$$

Proof. We can assume that $k \ge 0$, since otherwise we may pass to adjoints. Because $\|T_n(\chi_{-k})\|_p = 1$, we obtain from Lemma 6.1 that

$$\begin{aligned}
&s_{k+1}^{(p)}\Big(T_n(b\chi_k)\Big) = s_{k+1}^{(p)}\Big(T_n(b\chi_k)\Big)\|T_n(\chi_{-k})\|_p \\
&\ge s_{k+1}^{(p)}\Big(T_n(b\chi_k)T_n(\chi_{-k})\Big) = s_{k+1}^{(p)}\Big(T_n(b) - P_n H(b\chi_k)H(\chi_k)P_n\Big),
\end{aligned}$$

the latter equality resulting from (34) and the identities $H(\tilde{\chi}_{-k}) = H(\chi_k)$ and $H(\chi_{-k}) = 0$. As $\dim \operatorname{Im} H(\chi_k) = k$, we get that $F_k := P_n H(b\chi_k)H(\chi_k)P_n \in \mathcal{F}_k^{(n)}$, whence

$$\begin{aligned} s_{k+1}^{(p)}\Big(T_n(b) - F_k\Big) &= \inf\Big\{\|T_n(b) - F_k - G_{n-k-1}\|_p : G_{n-k-1} \in \mathcal{F}_{n-k-1}^{(n)}\Big\} \\ &\geq \inf\Big\{\|T_n(b) - H_{n-1}\|_p : H_{n-1} \in \mathcal{F}_{n-1}^{(n)}\Big\} = s_1^{(p)}(T_n(b)). \end{aligned}$$

Since $\{T_n(b)\}$ is stable, we infer from (3) that

$$\liminf_{n\to\infty} s_{k+1}^{(p)}(T_n(b\chi_k)) \geq \liminf_{n\to\infty} s_1^{(p)}(T_n(b)) > 0. \square$$

Proposition 6.4. *If $b \in M_p$ and $\{T_n(b)\}$ is stable on l^p then for every $k \in \mathbf{Z}$,*

$$\lim_{n\to\infty} s_{|k|}^{(p)}(T_n(b\chi_k)) = 0.$$

Proof. Again we may without loss of generality assume that $k \geq 0$. Using (34) and Lemma 6.1 we get

$$\begin{aligned} s_k^{(p)}(T_n(b\chi_k)) &= s_k^{(p)}\Big(T_n(\chi_k)T_n(b) + P_nH(\chi_k)H(\tilde{b})P_n\Big) \\ &\leq \|T_n(b)\|_p\, s_k^{(p)}\Big(T_n(\chi_k) + P_nH(\chi_k)H(\tilde{b})P_nT_n^{-1}(b)\Big). \end{aligned}$$

Put $A_n := T_n(\chi_k) + P_nH(\chi_k)H(\tilde{b})P_nT_n^{-1}(b)$. We have

$$A_n = \begin{pmatrix} * & C_n \\ I_{n-k} & 0 \end{pmatrix} = \begin{pmatrix} * & 0 \\ I_{n-k} & 0 \end{pmatrix} + \begin{pmatrix} 0 & C_n \\ 0 & 0 \end{pmatrix} =: B_n + D_n,$$

the blocks being of size $k \times (n-k)$, $k \times k$, $(n-k) \times (n-k)$, $(n-k) \times k$, respectively. Clearly, B_n has rank $n-k$ and thus $B_n \in \mathcal{F}_{n-k}^{(n)}$. It follows that

$$s_k^{(p)}(A_n) = s_k^{(p)}(A_n - B_n) = s_k^{(p)}(D_n) \leq \|D_n\|_p = \|C_n\|_p,$$

and we are left with showing that $\|C_n\|_p \to 0$.

Let b_n ($n \in \mathbf{Z}$) be the Fourier coefficients of b, let $e_j \in l^p$ be the sequence whose only nonzero entry is a unit at the jth position, and recall the notation (23). We have $C_n = (c_{jl}^{(n)})_{j,l=1}^k$, and it is easily seen that $c_{jl}^{(n)}$ equals $(b_{-k+j-1}, \ldots, b_{-k+j-n})$ times the $(n-k+l)$th column of $T_n^{-1}(b)$:

$$c_{jl}^{(n)} = (b_{-k+j-1} \ldots b_{-k+j-n})T_n^{-1}(b)P_ne_{n-k+l} = \Big(P_nf_{jk}, T_n^{-1}(b)P_ne_{n-k+l}\Big)$$

where

$$f_{jk} := \Big\{b_{-k+j-1},\ b_{-k+j-2},\ b_{-k+j-3},\ \ldots\Big\} = T(\chi_{-k+j-1})T(\tilde{b})e_1 \in l^q.$$

Consequently,

$$\begin{aligned} c_{jl}^{(n)} &= \Big(T_n^{-1}(\tilde{b})P_nf_{jk},\ e_{n-k+l}\Big) \\ &= \Big(T^{-1}(\tilde{b})f_{jk},\ e_{n-k+l}\Big) + \Big(T_n^{-1}(\tilde{b})P_nf_{jk} - T^{-1}(\tilde{b})f_{jk},\ e_{n-k+l}\Big). \qquad (36) \end{aligned}$$

The first term on the right of (36) obviously converges to zero as $n \to \infty$. The second term of (36) is at most

$$\|T_n^{-1}(\tilde{b})P_n f_{jk} - T^{-1}(\tilde{b})f_{jk}\|_q \tag{37}$$

(note that $\|e_{n-k+l}\|_p = 1$). Our assumptions imply that $\{T_n(\tilde{b})\}$ is stable on l^q. We so deduce from Lemma 6.2 that (37) tends to zero as $n \to \infty$.

Thus, each entry of the $k \times k$ matrix C_n approaches zero as $n \to \infty$. This implies that $\|C_n\|_p \to 0$. □

Now let a be as in Theorem 1.3. Since $T(a) \in \Phi_{-k}(l^r)$ for all $r \in [p,q]$, we have $a = b\chi_k$ where $T(b) \in \Phi_0(l^r)$ for all $r \in [p,q]$. From Theorems 3.3 and 3.4(b) we conclude that $\{T_n(b)\}$ is stable on l^p. The assertions of Theorem 1.3 therefore follows from Propositions 6.3 and 6.4.

We remark that Propositions 6.3 and 6.4 actually yield more than Theorem 1.3. Namely, let Π_p^0 denote the collection of all symbols $b \in M_p$ for which $\{T_n(b)\}$ is stable on l^p and let Π_p be the set of all symbols $a \in M_p$ such that $a\chi_{-k} \in \Pi_p^0$ for some $k \in \mathbf{Z}$. Notice that

$$\Pi_p = \Pi_q \subset \bigcup_{r\in[p,q]} \Pi_r$$

and

$$G(C+H^\infty) \cup G(C+\overline{H^\infty}) \cup G(PQC) \subset \Pi_2 \neq L^\infty,$$

where $G(B)$ stands for the invertible elements of a unital Banach algebra B. The following corollary is immediate from Propositions 6.3 and 6.4.

Corollary 6.5. *If $a \in \Pi_p$ and $T(a) \in \Phi_k(l^p)$ then*

$$\Sigma^{(p)}(T_n(a)) := \left\{ s_1^{(p)}(T_n(a)), \dots, s_n^{(p)}(T_n(a)) \right\}$$

has the $|k|$-splitting property. □

We also note that the proof of Proposition 6.4 gives estimates for the speed of convergence of $s_{|k|}^{(p)}(T_n(b\chi_k))$ to zero. For example, if $\sum_{n\in\mathbf{Z}} |n|^\mu |b_n| < \infty \quad (\mu > 0)$, then the finite section method is applicable to $T(b)$ on the space $l^{2,\mu}$ of all sequences $x = \{x_n\}_{n=1}^\infty$ such that

$$\|x\|_{2,\mu} := \left(\sum_{n=1}^\infty n^{2\mu} |x_n|^2 \right)^{1/2} < \infty$$

whenever $T(b)$ is invertible (see [17, pp. 106–107] or [5, Theorem 7.25]). Since

$$\|e_{n-k-l}\|_{2,-\mu} = (n-k+l)^{-\mu} = O(n^{-\mu}),$$

the proof of Proposition 6.4 implies the following result.

Corollary 6.6. *If $\sum_{n\in\mathbf{Z}} |n|^\mu |a_n| < \infty$ for some $\mu > 0$ and $T(a) \in \Phi_k(l^p)$ then*

$$s_{|k|}^{(p)}(T_n(a)) = O(n^{-\mu}) \quad \text{as} \quad n \to \infty. \ \square$$

7. Remarks on the Hilbert space case

Some aspects of the asymptotic behavior of the approximation numbers (= singular values) of matrices in $\mathcal{B}(\mathbf{C}_2^n)$ can be very easily understood by having recourse to the following well known fact (the "singular value decomposition").

Theorem 7.1. *If $A_n \in \mathcal{B}(\mathbf{C}_2^n)$ then there exist unitary matrices $U_n, V_n \in \mathcal{B}(\mathbf{C}_2^n)$ such that $A_n = U_n S_n V_n$ where*

$$S_n = \operatorname{diag}\Big(s_1(A_n), \ldots, s_n(A_n)\Big). \ \square$$

Here and throughout this section we abbreviate $s_k^{(2)}(A_n)$ to $s_k(A_n)$.

To illustrate the usefulness of Theorem 7.1, we give another proof of Theorem 1.2 for $p = 2$. We still need the following result.

Theorem 7.2. *A bounded linear Hilbert space operator A is normally solvable if and only if there is a $d > 0$ such that*

$$\operatorname{sp}(A^*A) \subset \{0\} \cup [d, \infty).$$

For a proof see [10], [11], [20]. $\square$

Theorem 7.3. *Let $a \in L^\infty$ and suppose $T(a)$ is not normally solvable on l^2. Then $s_k(T_n(a)) \to 0$ as $n \to \infty$ for each $k \geq 1$.*

Proof. Assume there is a $k \geq 1$ such that $s_k(T_n(a))$ does not converge to zero. Let k_0 be the smallest k with this property. Then there are $n_1 < n_2 < \ldots$ and $d > 0$ such that

$$s_{k_0}(T_{n_j}(a)) \geq d \ \text{ and } \ s_k(T_{n_j}(a)) \to 0 \ \text{ for } \ k < k_0. \tag{38}$$

To simplify notation, let us assume that $n_j = j$ for all j.

Write $T_n(a) = U_n S_n V_n$ as in Theorem 7.1. If $\lambda \notin \{0\} \cup [d^2, \infty)$, then (38) implies that $S_n^2 - \lambda I_n$ is invertible for all sufficiently large n, say for $n \geq n_0$, and that

$$\|(S_n^2 - \lambda I_n)^{-1}\|_2 \leq M(\lambda)$$

with some $M(\lambda) < \infty$ independent of n. Because

$$T_n^*(a) T_n(a) - \lambda I_n = V_n^*(S_n^2 - \lambda I_n) V_n,$$

it follows that $T_n^*(a)T_n(a) - \lambda I_n$ is invertible for $n \geq n_0$ and that

$$\|(T_n^*(a)T_n(a) - \lambda I_n)^{-1}\|_2 \leq M(\lambda).$$

Consequently, for every $x \in l^2$ we have

$$\|(T^*(a)T(a) - \lambda I)x\|_2 \geq \Big(1/M(\lambda)\Big)\|x\|_2,$$

which implies that $T^*(a)T(a) - \lambda I$ is invertible. Thus,

$$\mathrm{sp}\,\Big(T^*(a)T(a)\Big) \subset \{0\} \cup [d^2, \infty),$$

and Theorem 7.2 shows that $T(a)$ must be normally solvable. $\square$

Things are more transparent by invoking a few (harmless) C^*-algebras. Let $\mathcal{S}$ denote the C^*-algebra of all sequences $\{A_n\} := \{A_n\}_{n=1}^{\infty}$ of operators $A_n \in \mathcal{B}(\mathbf{C}_2^n)$ such that

$$\|\{A_n\}\| := \sup_{n \geq 1} \|A_n\|_2 < \infty,$$

and let $\mathcal{S}_c$ be the C^*-algebra of all $\{A_n\} \in \mathcal{S}$ for which there exists a bounded linear operator A on l^2 such that $A_n \to A$ and $A_n^* \to A^*$ strongly. Finally, let $\mathcal{C}$ stand for the sequences $\{A_n\} \in \mathcal{S}$ for which $\|A_n\|_2 \to 0$. Clearly, $\mathcal{C}$ is a closed two-sided ideal in both $\mathcal{S}$ and $\mathcal{S}_c$.

Obviously, a sequence $\{A_n\} \in \mathcal{S}$ is stable if and only if $\{A_n\} + \mathcal{C}$ is invertible in $\mathcal{S}/\mathcal{C}$. Following [25] and [20], we call a sequence $\{A_n\} \in \mathcal{S}$ a Moore-Penrose sequence if there exists a sequence $\{B_n\} \in \mathcal{S}$ such that

$$\{A_nB_nA_n - A_n\} \in \mathcal{C}, \;\; \{B_nA_nB_n - B_n\} \in \mathcal{C}, \tag{39}$$

$$\Big\{(B_nA_n)^* - B_nA_n\Big\} \in \mathcal{C}, \;\; \Big\{(A_nB_n)^* - A_nB_n\Big\} \in \mathcal{C}. \tag{40}$$

An element a of a unital C^*-algebra $\mathcal{A}$ is said to be Moore-Penrose invertible if there is an element $a^+ \in \mathcal{A}$ such that

$$aa^+a = a, \;\; a^+aa^+ = a^+, \;\; (a^+a)^* = a^+a, \;\; (aa^+)^* = aa^+.$$

Thus, $\{A_n\} \in \mathcal{S}$ is a Moore-Penrose sequence if and only if $\{A_n\} + \mathcal{C}$ is Moore-Penrose invertible in $\mathcal{S}/\mathcal{C}$.

The following result is again from [10], [11], [20].

Theorem 7.4. *Let $\mathcal{A}$ be a unital C^*-algebra. An element $a \in \mathcal{A}$ is Moore-Penrose invertible in $\mathcal{A}$ if and only if there is a $d > 0$ such that* $\mathrm{sp}\,(a^*a) \subset \{0\} \cup [d, \infty)$. $\square$

The next theorem is Roch and Silbermann's [20]. The proof given here is different from theirs.

Theorem 7.5. *A sequence $\{A_n\} \in \mathcal{S}$ is a Moore-Penrose sequence if and only if*

$$\Sigma(A_n) = \Big\{s_1(A_n), \ldots, s_n(A_n)\Big\}$$

has the splitting property.

Proof. Write $A_n = U_n S_n V_n$ as in Theorem 7.1. We have

$$\begin{aligned}&\|A_nB_nA_n - A_n\|_2 \to 0\\ \Longleftrightarrow\ &\|U_nS_nV_nB_nU_nS_nV_n - U_nS_nV_n\|_2 \to 0\\ \Longleftrightarrow\ &\|S_n(V_nB_nU_n)S_n - S_n\|_2 \to 0,\end{aligned}$$

and since analogous equivalences hold for the remaining three sequences in (39) and (40), we arrive at the conclusion that $\{A_n\}$ is a Moore-Penrose sequence if and only if $\{S_n\} + \mathcal{C}$ is Moore-Penrose invertible in $\mathcal{S}/\mathcal{C}$. By Theorem 7.4, this is equivalent to the condition

$$\mathrm{sp}_{\,\mathcal{S}/\mathcal{C}}\Big(\{S_n^2\} + \mathcal{C}\Big) \subset \{0\} \cup [d^2, \infty) \ \text{ for some } \ d > 0. \tag{41}$$

Let $\mathcal{D} \subset \mathcal{S}$ denote the sequences $\{A_n\}$ constituted by diagonal matrices A_n. From the elementary theory of C^*-algebras we get

$$\mathrm{sp}_{\,\mathcal{S}/\mathcal{C}}\Big(\{S_n^2\} + \mathcal{C}\Big) = \mathrm{sp}_{\mathcal{D}/(\mathcal{D}\cap\mathcal{C})}\Big(\{S_n^2\} + \mathcal{D} \cap \mathcal{C}\Big). \tag{42}$$

Consider the infinite diagonal matrix

$$\mathrm{diag}\,(S_1^2, S_2^2, \ldots) = \mathrm{diag}\,(\varrho_1, \varrho_2, \varrho_3, \ldots)$$

(here $\mathcal{S}_m \in \mathcal{B}(\mathbf{C}_2^m)$ and $\varrho_m \in \mathbf{C}$). Obviously, the spectrum on the right of (42) coincides with the set $\mathcal{P}\{\varrho_m\}$ of the partial limits of the sequence $\{\varrho_m\}$. Consequently, (41) holds if and only if

$$\mathcal{P}\{\varrho_m\} \subset \{0\} \cup [d^2, \infty) \ \text{ for some } \ d > 0,$$

which is easily seen to be equivalent to the splitting property of $\Sigma(A_n)$. □

Also as in [20], we call a sequence $\{A_n\} \in \mathcal{S}$ an exact Moore-Penrose sequence if $\{A_n^+\}$ belongs to $\mathcal{S}$; here $A_n^+ \in \mathcal{B}(\mathbf{C}_2^n)$ is the Moore-Penrose inverse of A_n.

PROPOSITION 7.6. *Let $\{A_n\}$ be a sequence in $\mathcal{S}_c$ and let A be the strong limit of A_n. Then the following are equivalent:*

(i) *A_n^+ is strongly convergent;*

(ii) *A is normally solvable and $A_n^+ \to A^+$ strongly;*

(iii) *A is normally solvable and $\{A_n\}$ is an exact Moore-Penrose sequence.*

The simple proof is omitted. □

The following theorem was by means of different methods established in [20].

THEOREM 7.7. *A sequence $\{A_n\} \in \mathcal{S}$ is an exact Moore-Penrose sequence if and only if there is a $d > 0$ such that*

$$\Sigma(A_n) \subset \{0\} \cup [d, \infty) \ \textit{ for all } \ n \geq 1. \tag{43}$$

Proof. As in the proof of Theorem 7.5 we see that $\{A_n\}$ is an exact Moore-Penrose sequence if and only if $\{S_n\}$ enjoys this property. Write

$$\mathrm{diag}\,(S_1, S_2, \ldots) = \mathrm{diag}\,(\mu_1, \mu_2, \mu_3, \ldots)$$

(where again $S_m \in \mathcal{B}(\mathbf{C}_2^m)$ and $\mu_n \in \mathbf{C}$) and define $f : [0, \infty) \to [0, \infty)$ by

$$f(x) := \begin{cases} x^{-1} & \text{if} \quad x > 0 \\ 0 & \text{if} \quad x = 0. \end{cases}$$

Since

$$\mathrm{diag}\,(S_1^+, S_2^+, \ldots) = \mathrm{diag}\,\Big(f(\mu_1), f(\mu_2), f(\mu_3), \ldots\Big),$$

we conclude that $\{S_n^+\} \in \mathcal{S}$ if and only if $\{f(\mu_m)\}$ is a bounded sequence, which is equivalent to (43). □

Now let $A_n = T_n(a)$ with $a \in L^\infty$. If $\{T_n(a)\}$ is a Moore-Penrose sequence, then $T(a)$ must obviously be normally solvable. Thus, from Theorem 3.3 (for $p = 2$) and Theorem 1.3 (for $p = 2$) we deduce that if $a \in PC$, then $\{T_n(a)\}$ is a Moore-Penrose sequence if and only if $T(a)$ is Fredholm.

The following result, which is also taken from [20], characterizes the exact Moore-Penrose sequences constituted by the truncations of an infinite Toeplitz matrix. Our proof is again different from the one of [20].

Theorem 7.8. *Let $a \in PC$. Then $\{T_n(a)\}$ is an exact Moore-Penrose sequence if and only if $T(a)$ is Fredholm and*

$$\dim \mathrm{Ker}\, T_n(a) = |\mathrm{Ind}\, T(a)| \tag{44}$$

for all sufficiently large n.

Proof. If $\{T_n(a)\}$ is an exact Moore-Penrose sequence, then $T(a)$ is normally solvable and thus Fredholm. Let $T(a) \in \Phi_k(l^2)$. Then

$$s_{|k|}(T_n(a)) \to 0 \;\text{ and }\; s_{|k|+1}(T_n(a)) \geq d > 0$$

by virtue of Theorem 1.3 (for $p = 2$). Since

$$\mathrm{dist}\,\Big(T_n(a),\ \mathcal{F}_{n-|k|-1}^{(n)}\Big) > 0,$$

we see that

$$\mathrm{rank}\, T_n(a) \geq n - |k|. \tag{45}$$

From Theorem 7.7 we deduce that $\{T_n(a)\}$ is an exact Moore-Penrose sequence if and only if $s_{|k|}(T_n(a)) = 0$ for all $n \geq n_0$. Because

$$s_{|k|}(T_n(a)) = \mathrm{dist}\,\Big(T_n(a),\ \mathcal{F}_{n-|k|}^{(n)}\Big)$$

and $\mathcal{F}^{(n)}_{n-|k|}$ is a closed subset of $\mathcal{B}(\mathbf{C}^n_2)$, we have $s_{|k|}(T_n(a)) = 0$ if and only if

$$\operatorname{rank} T_n(a) \leq n - |k|. \tag{46}$$

Combining (45) and (46) we obtain that $\{T_n(a)\}$ is an exact Moore-Penrose sequence if and only if $T(a) \in \Phi_k(l^2)$ for some $k \in \mathbf{Z}$ and

$$\dim \operatorname{Ker} T_n(a) = n - \operatorname{rank} T_n(a) = |k|$$

for all $n \geq n_0$. □

8. The Heing-Hellinger theorem

Of course, condition (44) is difficult to check. In this section we give a new proof of the Heinig-Hellinger theorem, which provides a criterion (in terms of only the symbol a) for (44) to hold.

If $a \in PC$ and $T(a)$ is Fredholm of index zero and thus invertible, then the sequence $\{T_n(a)\}$ is stable (Theorems 3.3 and 3.4 for $p = 2$). In this case $\Sigma(T_n(a)) \subset [d, \infty)$ and $\dim \operatorname{Ker} T_n(a) = 0$ for all sufficiently large n and hence each of Theorems 7.7 and 7.8 yields that $\{T_n(a)\}$ is an exact Moore-Penrose sequence; however, we have $T_n^+(a) = T_n^{-1}(a)$ for all sufficiently large n and therefore consideration of Moore-Penrose inverses is not at all necessary in this situation.

The really interesting case is the one in which $T(a)$ is Fredholm of nonzero index. The rest of this section is devoted to the proof of the following result.

Theorem 8.1 (Heinig and Hellinger). *Let $a \in PC$. Suppose $T(a)$ is Fredholm on l^2 and $\operatorname{Ind} T(a) \neq 0$. If $\operatorname{Ind} T(a) < 0$, then the following are equivalent:*

(i) $\dim \operatorname{Ker} T_n(a) = |\operatorname{Ind} T(a)|$ *for all sufficiently large n;*

(ii) $\operatorname{Ker} T(\tilde{a}) \subset \operatorname{Im} P_{n_0}$ *for some $n_0 \geq 1$;*

(iii) *the Fourier coefficients $(a^{-1})_{-m}$ are zero for all sufficiently large m.*

If $\operatorname{Ind} T(a) > 0$, then the following are equivalent:

(i') $\dim \operatorname{Ker} T_n(a) = \operatorname{Ind} T(a)$ *for all sufficiently large n;*

(ii') $\operatorname{Ker} T(a) \subset \operatorname{Im} P_{n_0}$ *for some $n_0 \geq 1$;*

(iii') $(a^{-1})_m = 0$ *for all sufficiently large m.*

For the sake of definiteness, let us assume that $\operatorname{Ind} T(a) = -k < 0$. The proofs of the implications (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) are easy.

Proof of the implication (iii) $\Rightarrow$ (ii). Let $x \in \operatorname{Ker} T(\tilde{a})$. Then, by (33),

$$T(\tilde{a}^{-1})T(\tilde{a}) = I - H(\tilde{a}^{-1})H(a),$$

which shows that $x = H(\tilde{a}^{-1})H(a)x$, and since $H(\tilde{a}^{-1})$ has only a finite number of nonzero rows, it follows that $x_m = 0$ for all sufficiently large m. □

Proof of the implication (ii) $\Rightarrow$ (i). If n is large enough then $s_{k+1}(T_n(\tilde{a})) \geq d > 0$ by Theorem 1.3 (or Proposition 6.3), whence $\operatorname{rank} T_n(\tilde{a}) > n - k + 1$ and thus,

$$\dim \operatorname{Ker} T_n(\tilde{a}) < k + 1. \tag{47}$$

If $x \in \operatorname{Ker} T(\tilde{a}) \subset \operatorname{Im} P_{n_0}$ and $n \geq n_0$, then $T_n(\tilde{a})P_n x = P_n T(\tilde{a})x = 0$, which implies that

$$\dim \operatorname{Ker} T_n(\tilde{a}) \geq \dim \operatorname{Ker} T(\tilde{a}) = k \tag{48}$$

(recall Theorem 3.1(a) for the last equality). Clearly, equality (i) follows from (47) and (48). □

The proof of the implication (i) $\Rightarrow$ (iii) is less trivial and is based on the following deep theorem. Recall that χ_n is defined by $\chi_n(t) = t^n$ for $t \in \mathbf{T}$.

Theorem 8.2 (Heinig). *Let $a \in L^\infty$ and let $k > 0$ be an integer. Then*

$$\dim \operatorname{Ker} T_n(a) = k \text{ for all sufficiently large } n$$

if and only if a or $\tilde{a}$ is of the form $\chi_{p+k}(r + h)$ where h is a function in H^∞, r is a rational function in L^∞, r has exactly p poles in the open unit disk $\mathbf{D}$ (multiplicities taken into account), r has no pole at the origin, and $r(0) + h(0) \neq 0$.

A proof is in [12, Satz 6.2 and formula (8.4)]. Also see [14, Theorem 8.6]. □

Proof of the implication (i) $\Rightarrow$ (iii). Let $\chi_{p+k}(r + h)$ be the representation of a or $\tilde{a}$ ensured by Theorem 8.2 and put $b := \chi_{p+k}(r + h)$. Denote by $\alpha_1, \ldots, \alpha_p$ and $\beta_1, \ldots, \beta_q$ the poles of r inside and outside $\mathbf{T}$, respectively. For $t \in \mathbf{T}$,

$$\begin{aligned} r(t) &= \frac{u_+(t)}{(t - \alpha_1) \ldots (t - \alpha_p)(t - \beta_1) \ldots (t - \beta_q)} \\ &= \frac{t^{-p} v_+(t)}{(1 - \alpha_1/t) \ldots (1 - \alpha_p/t)(1 - t/\beta_1) \ldots (1 - t/\beta_q)} \end{aligned}$$

with polynomials $u_+, v_+ \in H^\infty$. Clearly,

$$s_+(t) := (1 - t/\beta_1)^{-1} \ldots (1 - t/\beta_q)^{-1} \in H^\infty.$$

Letting

$$c_+(t) := t^k v_+(t)s_+(t) + t^{p+k}(1 - \alpha_1/t) \ldots (1 - \alpha_p/t)h(t),$$

we get

$$b(t) = (1 - \alpha_1/t)^{-1} \ldots (1 - \alpha_p/t)^{-1} c_+(t).$$

The function c_+ lies in H^∞ and has a zero of order at least k at the origin. Obviously, $(1 - \alpha_1/t)^{-1} \ldots (1 - \alpha_p/t)^{-1}$ is a function which together with its inverse belongs to

$\overline{H^\infty}$. If c_+ would have infinitely many zeros in $\mathbf{D}$, then $T(c_+)$ and thus $T(b)$ were not Fredholm (see, e.g., [5, Theorem 2.64]). Hence, c_+ has only a finite number $\lambda \geq k$ of zeros in $\mathbf{D}$. It follows that $\operatorname{Ind} T(c_+) = -\lambda$ (again see, e.g., [5, Theorem 2.64]) and therefore $\operatorname{Ind} T(b) = \operatorname{Ind} T(c_+) = -\lambda$. If $b = a$, then λ must equal k. Consequently, $c_+(z) = z^k \varphi_+(z)$ with φ_+ and φ_+^{-1} in H^∞. This implies that

$$a^{-1}(t) = t^{-k}(1-\alpha_1/t)\dots(1-\alpha_p/t)\varphi_+^{-1}(t)$$

has only finitely many nonzero Fourier coefficients with negative index. If b would be equal to $\tilde{a}$, it would result that $\operatorname{Ind} T(\tilde{a})$ is negative, which is impossible due to the equality $\operatorname{Ind} T(\tilde{a}) = -\operatorname{Ind} T(a)$. □

Corollary 8.3. *If $a \in PC \setminus C$ then $\{T_n(a)\}$ is an exact Moore-Penrose sequence on l^2 if and only if $\{T_n(a)\}$ is stable on l^2.*

Proof. The "if part" is trivial. To prove the "only if" portion, suppose $\{T_n(a)\}$ is an exact Moore-Penrose sequence. Then $T(a)$ is Fredholm by Theorem 7.8. If $T(a)$ has index zero, then $\{T_n(a)\}$ is stable. If $\operatorname{Ind} T(a) \neq 0$, then Theorem 7.8 and the implication (i) $\Rightarrow$ (iii) of Theorem 8.1 tell us that a^{-1} is a polynomial times a function in H^∞ or $\overline{H^\infty}$. As functions in H^∞ or $\overline{H^\infty}$ cannot have jumps, this case is impossible. □

We remark that Heinig and Hellinger [13] proved the equivalence (i) $\Leftrightarrow$ (iii) of Theorem 8.2 for symbols in the Wiener algebra W. Corollary 8.3 was known to Silbermann and led him to the introduction of condition (ii). In the case of block Toeplitz matrices, (iii) and (ii) are no longer equivalent; Silbermann proved that then the validity of (15) for some $n_0 \geq 1$ implies that

$$\{T_n(a)\} \text{ is an exact Moore-Penrose sequence,} \tag{49}$$

and he conjectures that (49) is even equivalent to (15) for some $n_0 \geq 1$ (see [25]). The proofs of [13] and [25] differ from the proof given above.

9. l^p versus l^2

As shown in the previous section, many l^2 results can be derived with the help of Theorem 7.1, which reduces problems for $\{A_n\}$ to questions about the infinite diagonal operator

$$\operatorname{diag}\left(s_1^{(2)}(A_1),\, s_1^{(2)}(A_2),\, s_2^{(2)}(A_2),\, s_1^{(2)}(A_3),\, s_2^{(2)}(A_3),\, s_3^{(2)}(A_3),\, \dots\right).$$

It would therefore be very nice to have an analogous result for l^p. For example, one could ask the following: given $A_n \in \mathcal{B}(\mathbf{C}_p^n)$, are there invertible isometries $U_n, V_n \in \mathcal{B}(\mathbf{C}_p^n)$ and a diagonal matrix $S_n \in \mathcal{B}(\mathbf{C}_p^n)$ such that $A_n = U_n S_n V_n$? If the answer were "yes", we had

$$\Sigma^{(p)}(A_n) = \Sigma^{(p)}(S_n),$$

and Theorem 11.11.3 of [16] would tell us that $\Sigma^{(p)}(S_n)$ is the collection of the moduli of the diagonal elements of S_n.

However, the answer to the above question is "no". The reason is the dramatic loss of symmetry when passing from l^2 to l^p. Looking at the (real) unit spheres

$$\mathbf{S}_1^{(p)} := \Big\{(x,y) \in \mathbf{R}^2 : |x|^p + |y|^p = 1\Big\},$$

we see that $\mathbf{S}_1^{(2)}$ has the symmetry group $O(2)$, while the symmetry group of $\mathbf{S}_1^{(p)}$ $(p \neq 2)$ is the dieder group D_4, which contains only 8 elements. Equivalently, the invertible isometries in $\mathcal{B}(\mathbf{C}_2^2)$ are the 2×2 unitary matrices, whereas a matrix $U_2 \in \mathcal{B}(\mathbf{C}_p^2)$ $(p \neq 2)$ is an invertible isometry if and only if

$$U_2 = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \text{ or } U_2 = \begin{pmatrix} 0 & \lambda \\ \mu & 0 \end{pmatrix} \text{ with } (\lambda, \mu) \in \mathbf{T}^2.$$

Thus, a matrix $A_2 \in \mathcal{B}(\mathbf{C}_p^2)$ $(p \neq 2)$ is of the form $A_2 = U_2 S_2 V_2$ with invertible isometries U_2, V_2 and a diagonal matrix S_2 if and only if

$$A_2 = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \text{ or } A_2 = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix} \text{ with } (a, b) \in \mathbf{C}^2.$$

I even suspect that relaxing the above question will not be successful.

Conjecture 9.1. *Fix $p \neq 2$ and let $1/p + 1/q = 1$. There is no number $M \in (1, \infty)$ with the following property: given any sequence $\{A_n\}$ of matrices $A_n \in \mathcal{B}(\mathbf{C}_p^n)$ such that $\sup \|A_n\|_p < \infty$ and $\sup \|A_n\|_q < \infty$, there are invertible matrices $U_n, V_n \in \mathcal{B}(\mathbf{C}_p^n)$ and diagonal matrices $S_n \in \mathcal{B}(\mathbf{C}_p^n)$ such that $A_n = U_n S_n V_n$ and*

$$\|U_n\|_p \leq M, \;\; \|U_n^{-1}\|_p \leq M, \;\; \|V_n\|_p \leq M, \;\; \|V_n^{-1}\|_p \leq M$$

for all n.

Finally, for the reader's convenience, we add a proof of (3).

Proposition 9.2. *If $A \in \mathcal{B}(\mathbf{C}_p^n)$, then $s_1^{(p)}(A) = 1/\|A^{-1}\|_p$ if A is invertible and $s_1^{(p)}(A) = 0$ if A is not invertible.*

Proof. Suppose A is not invertible. Then $\operatorname{Ker} A \neq \{0\}$. Let Z be any direct complement of $\operatorname{Ker} A$ in $\mathbf{C}_p^n$ and let $P : \mathbf{C}_p^n \to Z$ be the projection onto Z parallel to $\operatorname{Ker} A$. Clearly, $P \in \mathcal{F}_{n-1}^{(n)}$ and thus $F := AP \in \mathcal{F}_{n-1}^{(n)}$. If $x \in \mathbf{C}^n$, then $x = x_0 + x_1$ with $x_0 \in \operatorname{Ker} A$ and $x_1 = Px \in Z$. Therefore

$$(A - F)x = Ax - APx = A(x_0 + Px) - APx = 0,$$

which implies that $A - F = 0$ and hence $\operatorname{dist}(A, \mathcal{F}_{n-1}^{(n)}) = 0$.

Now suppose A is invertible. We then have

$$\|A^{-1}\|_p = \sup_{x \neq 0} \frac{\|A^{-1}x\|_p}{\|x\|_p} = \sup_{z \neq 0} \frac{\|z\|_p}{\|Az\|_p} = \left(\inf_{z \neq 0} \frac{\|Az\|_p}{\|z\|_p} \right)^{-1},$$

whence

$$1/\|A^{-1}\|_p = \inf_{z\neq 0} \frac{\|Az\|_p}{\|z\|_p} = \min_{\|z\|_p=1} \|Az\|_p =: \|Ae_0\|_p \tag{50}$$

with some $e_0 \in \mathbf{C}_p^n$ of norm 1. Put $\operatorname{span}\{e_0\} = \{\lambda e_0 : \lambda \in \mathbf{C}\}$ and let X be any direct complement of $\operatorname{span}\{e_0\}$ in $\mathbf{C}_p^n$. The functional

$$\varphi : \operatorname{span}\{e_0\} \to \mathbf{C}, \quad \lambda e_0 \mapsto \lambda$$

clearly has the norm 1. By the Hahn-Banach theorem, there is a functional $\Phi : \mathbf{C}_p^n \to \mathbf{C}$ such that $\Phi(\lambda e_0) = \lambda$ and $\|\Phi\| = 1$. Define $F \in \mathcal{B}(\mathbf{C}_p^n)$ by $Fx := Ax - \Phi(x)Ae_0$. Since

$$F(\lambda e_0) = \lambda Ae_0 - \lambda Ae_0 = 0,$$

we see that $F \in \mathcal{F}_{n-1}^{(n)}$. Because

$$\|Ax - Fx\|_p = \|\Phi(x)Ae_0\|_p = |\Phi(x)|\,\|Ae_0\|_p \leq \|x\|_p\,\|Ae_0\|_p,$$

it results that $\|A - F\|_p \leq \|Ae_0\|_p$. From (50) we therefore deduce that $s_1^{(p)}(A) \leq 1/\|A^{-1}\|_p$.

To prove that $s_1^{(p)}(A) \geq 1/\|A^{-1}\|_p$, let G be any matrix in $\mathcal{F}_{n-1}^{(n)}$. If $\|I - A^{-1}G\|_p$ were less than 1, then $A^{-1}G$ and thus G were invertible, which is impossible. Thus $\|I - A^{-1}G\|_p \geq 1$. We therefore have

$$1 \leq \|I - A^{-1}G\|_p = \|A^{-1}(A - G)\|_p \leq \|A^{-1}\|_p\,\|A - G\|_p,$$

which implies that $1/\|A^{-1}\|_p \leq \|A - G\|_p$. As $G \in \mathcal{F}_{n-1}^{(n)}$ was arbitrary, it follows that $1/\|A^{-1}\|_p \leq s_1^{(p)}(A)$. □

References

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

[2] A. Böttcher, S. Grudsky: Toeplitz operators with discontinuous symbols: phenomena beyond piecewise continuity. *Operator Theory: Advances and Applications* 90 (1996) 55–118.

[3] A. Böttcher, S. Silbermann: Über das Reduktionsverfahren für diskrete Wiener-Hopf Gleichungen mit unstetigem Symbol. *Z. Analysis Anw.* 1 (1982), 1–5.

[4] A. Böttcher, B. Silbermann: *Invertibility and Asymptotics of Toeplitz Matrices.* Akademie-Verlag, Berlin 1983.

[5] A. Böttcher, B. Silbermann: *Analysis of Toeplitz Operators.* Akademie-Verlag, Berlin 1989 and Springer-Verlag, Berlin, Heidelberg, New York 1990.

[6] K.M. Day: Measures associated with Toeplitz matrices generated by Laurent expansions of an arbitrary rational function. *Trans. Amer. Math. Soc.* 209 (1975), 175–183.

[7] I. Gohberg, I.A. Feldman: *Convolution Equations and Projection Methods for Their Solution.* Amer. Math. Soc., Providence, RI, 1974 [Russian original: Nauka, Moscow 1971; German transl.: Akademie-Verlag, Berlin 1974].

[8] I. Gohberg, N. Krupnik: *One-Dimensional Linear Singular Integral Equations.* Vol. I, Birkhäuser Verlag, Basel, Boston, Berlin 1992 [Russian original: Shtiintsa, Kishinev 1973; German transl.: Birkhäuser Verlag, Basel 1979].

[9] U. Grenander, G. Szegö: *Toeplitz Forms and Their Applications.* Univ. of California Press, Berkeley 1958 [Russian transl.: Izd. Inostr. Lit., Moscow 1961].

[10] R. Harte, M. Mbekhta: On generalized inverses in C^*-algebras. *Studia Math.* 103 (1992), 71–77.

[11] R. Harte, M. Mbekhta: Generalized inverses in C^*-algebras. II. *Studia Math.* 106 (1993), 129–138.

[12] G. Heinig: Endliche Toeplitzmatrizen und zweidimensionale diskrete Wiener-Hopf-Operatoren mit homogenem Symbol. *Math. Nachr.* 82 (1978), 29–68.

[13] G. Heinig, F. Hellinger: The finite section method for Moore-Penrose inversion of Toeplitz operators. *Integral Equations and Operator Theory* 19 (1994), 419–446.

[14] G. Heinig, K. Rost: *Algebraic Methods for Toeplitz-Like Matrices and Operators.* Akademie-Verlag, Berlin 1984 and Birkhäuser Verlag, Basel, Boston, Stuttgart 1984.

[15] S.V. Parter: On the distribution of the singular values of Toeplitz matrices. *Lin. Algebra Appl.* 80 (1986), 115–130.

[16] A. Pietsch: *Operator Ideals.* Deutscher Verlag d. Wiss., Berlin 1978 [Russian transl.: Mir, Moscow 1982].

[17] S. Prössdorf, B. Silbermann: *Projektionsverfahren und die näherungsweise Lösung singulärer Gleichungen.* Teubner, Leipzig 1977.

[18] M. Reed, B. Simon: *Methods of Modern Mathematical Physics. Vol. 1: Functional Analysis.* Academic Press, New York, London 1972 [Russian transl.: Mir, Moscow 1977].

[19] S. Roch, B. Silbermann: Limiting sets of eigenvalues and singular values of Toeplitz matrices. *Asymptotic Analysis* 8 (1994), 293–309.

[20] S. Roch, B. Silbermann: Index calculus for approximation methods and singular value decomposition. Preprint, TU Chemnitz-Zwickau 1996.

[21] S. Roch, B. Silbermann: A note on singular values of Cauchy-Toeplitz matrices. Preprint TU Chemnitz-Zwickau 1996.

[22] P. Schmidt, F. Spitzer: The Toeplitz matrices of an arbitrary Laurent polynomial. *Math. Scand.* 8 (1960), 15–38.

[23] B. Silbermann: Lokale Theorie des Reduktionsverfahrens für Toeplitzoperatoren. *Math. Nachr.* 104 (1981), 137–146.

[24] B. Silbermann: On the limiting set of singular values of Toeplitz matrices. *Lin. Algebra Appl.* 182 (1993), 35–43.

[25] B. Silbermann: Asymptotic Moore-Penrose inversion of Toeplitz operators. *Lin. Algebra Appl.*, to appear 1996.

[26] S.Treil: Invertibility of Toeplitz operators does not imply applicability of the finite section method. *Dokl. Akad. Nauk SSSR* 292 (1987), 563–567 [Russian].

[27] E.E. Tyrtyshnikov: Singular values of Cauchy-Toeplitz matrices. *Lin. Algebra Appl.* 161 (1992), 99–116.

[28] E.E. Tyrtyshnikov: New theorems on the distribution of eigenvalues and singular values of multilevel Toeplitz matrices. *Dokl. Adad. Nauk* 333 (1993), 300–303 [Russian].

[29] E.E. Tyrtyshnikov: A unifying appraoch to some old and new theorems on distribution and clustering. *Lin. Algebra Appl.* 232 (1996), 1–43.

[30] I.E. Verbitsky, N. Krupnik: On the applicability of the finite section method to discrete Wiener-Hopf equations with piecewise continuous symbol. *Matem. Issled.* 45 (1977), 17–28 [Russian].

[31] H. Widom: Asymptotic behavior of block Toeplitz matrices and determinants. II. *Adv. Math* 21 (1976), 1–29.

[32] H. Widom: On the singular values of Toeplitz matrices. *Z. Analysis Anw.* 8 (1989), 221–229.

Albrecht Böttcher
Fakultät für Mathematik
TU Chemnitz-Zwickau
09107 Chemnitz
Germany