BIT 38(2) (1998), 356–371.

Properties of Iteration of Toeplitz Operators with Toeplitz Preconditioners *

J. Malinen †

[1]*Institute of Mathematics, Helsinki University of Technology, P.O.BOX 1100, FIN-02015 HUT, Finland. email: Jarmo.Malinen@hut.fi*

Abstract.

We consider the problems of preconditioning and iterative inversion of Toeplitz operators on sequences of complex numbers. We divide the preconditioned operator into two parts, of which one is compact and the other is regarded as a small perturbation. It will be shown that the Krylov subspace methods (such as GMRES) will perform initially at superlinear speed when applied to such preconditioned system. However, with large iteration numbers, the speed will settle down to linear order. Most of our results are stated in terms of the symbol of the Toeplitz operator in question.

AMS subject classification: 65T10, 65T20, 47B35.

Key words: Toeplitz operator, Krylov subspace methods, preconditioning, speed estimate, superlinear, linear.

1 Introduction

Toeplitz operators are a class of linear operators whose theory is by now fairly well known ([3], [6], [14]). They are classical objects of interest within function and operator theory, partly because of the unilateral shift of sufficiently high multiplicity gives an universal model for all bounded linear operator in a Hilbert space (see [16], [20]).

During the last decades, the development of systems and control theory has increased the number of interesting numerical applications for Toeplitz (and Hankel) operators. For example, the inverse of a certain self adjoint Toeplitz operator (the so called "Popov function", having generally a matrix or operator valued symbol) is a central notion in the linear control theory, Riccati equation theory and theory of spectral factorizations. Roughly, under certain conditions, the computation of an inverse for a self-adjoint Toeplitz operator is equivalent with solving of an associated spectral factorization problem as well as finding a special (stabilizing) solution for an associated algebraic Riccati equation ([7], [10], [18]). This fact alone is sufficient to make the numerical properties of Toeplitz operators an interesting research subject.

From the numerical analysts' point of view, Toeplitz operators are interesting because they represent a limit case of $n \times n$ Toepliz matrix systems as $n \to \infty$.

*Received June 1996. Revised September 1997. Communicated by Lars Eldén

†This work was partially supported by a grant from Emil Aaltosen säätiö.

An analogous treatment for Toeplitz matrices, following the lines of this paper, is given in the companion paper [11]. Several aspects of the matrix case have been extensively studied (see e.g. [1], [2], [19]).

In this paper we study certain properties of iteration of Toeplitz operators with complex (scalar) valued symbol in the Hilbert space of square summable sequences of complex numbers. The scalar valued symbol is the easiest case, related to the unilateral shift of multiplicity 1.

Our main interest lies on the following questions:

- How to precondition the operator once the symbol of the Toeplitz operator is known?
- What is the preconditioned operator like?
- What can be said about the spectrum of the preconditioned operator?
- How fast can an iterative solver perform in the above mentioned case?
- How does the smoothness of the symbol affect the performance of the iterative solver?

We are not so much interested in the questions related to explicit parameterization and discretization of the Toeplitz operators for numerical purposes. However, the reader is advised to regard the operator-vector product $\mathcal{T}[f]\{a_j\}$ to be something "easily computable" (of course only in an approximative sense). If this was not the case, then the iterative inversion of $\mathcal{T}[f]\{a_j\} = \{b_j\}$ would be expensive at each iteration step, because what we use are the powers $\mathcal{T}[f]^k\{a_j\}$. Fortunately, the truncated version of the product $\mathcal{T}[f]\{a_j\}$ (i.e. the Toeplitz matrix -vector product) is inexpensive when calculated by FFT.

We shall show that a Toeplitz operator can be preconditioned by approximation theoretic means applied upon the symbol (also called the generating function) of the operator. The preconditioned operator will consist of a large compact part and a small Toeplitz part. Moreover, if the symbol of the Toeplitz operator is smooth, we get information about the distribution of singular values of the compact part. This in turn gives us information about the spectrum of the preconditioned operator and its properties of iteration.

Of course, it is not the cost of a single iteration step alone that gives the price of the whole computation. We also need to consider how many steps we have to calculate in order to get the required precision. Our conclusion is that a Krylov subspace method (such as GMRES) applied upon the preconditioned system initially converges at increasing speed (or "superlinearly"), until the compact part has been "killed off" and the small Toeplitz part begins to dominate.

2 Definitions and basic theory

We use the following notations throughout the paper: $\mathbf{Z}$ is the set of integers. $\mathbf{Z}_+ := \{j \in \mathbf{Z} \mid j \geq 0\}$. $\mathbf{Z}_+ := \{j \in \mathbf{Z} \mid j < 0\}$. $\mathbf{N} := \{j \in \mathbf{Z} \mid j > 0\}$. $\mathbf{T}$ is the unit circle of the complex plane. $C(\mathbf{T})$ denotes the class of continuous functions on $\mathbf{T}$ equipped with sup-norm $||\cdot||_\infty$. Given $f \in C(\mathbf{T})$ and $\alpha > 0$, the number $||f||_{Lip_\alpha(\mathbf{T})}$ is defined by

$$||f||_{Lip_\alpha(\mathbf{T})} = ||f||_\infty + \sup_{t_1 \neq t_2} \frac{|f(e^{it_1}) - f(e^{it_2})|}{|t_1 - t_2|^\alpha} \tag{2.1}$$

is called the Lipschitz norm of f. $Lip_\alpha(\mathbf{T}) \subset C(\mathbf{T})$ is the set of such f for which the expression (2.1) is finite. For $r \in \mathbf{Z}_+$, $C^{r,\alpha}(\mathbf{T})$ are those functions whose r.th derivative is in $Lip_\alpha(\mathbf{T})$. If $\alpha = 0$, then $C^{r,\alpha}(\mathbf{T}) := C^r(\mathbf{T})$. Given a Hilbert space H, we define the sequence Hilbert spaces

$$\ell^2(\mathbf{I}; H) := \big\{\{z_i\}_{i\in I} \subset H : \sum_{i \in I} ||z_i||_H^2 < \infty\big\}, \quad \text{where} \quad \mathbf{I} = \mathbf{Z}, \mathbf{Z}_+, \text{ or } \mathbf{Z}_-.$$

If H is a Hilbert space, then $\mathcal{L}(H)$ denotes the bounded and $\mathcal{LC}(H)$ the compact linear operators in H. Other notations are introduced when they are needed.

The following definition is quite standard.

Definition 2.1. *The (unilateral) shift operator $S \in \mathcal{L}(\ell^2(\mathbf{Z}_+))$ is defined by*

$$S\{a_0, a_1, a_2, \ldots a_j, \ldots\} := \{0, a_0, a_1 \ldots a_{j-1} \ldots\}$$

for any sequence $\{a_j\}_{j=0}^\infty \in \ell^2(\mathbf{Z}_+)$.

Clearly S is a left but not right invertible operator satisfying $S^*S = I$. It is well known that S as an isometry can be extended into a unitary operator on a larger Hilbert space. Such an extension will serve as a useful tool in the study of Toeplitz operators. The following Definition 2.2 gives us the necessary concepts.

Definition 2.2.

(i) The extension space H of $\ell^2(\mathbf{Z}_+)$ is the cartesian product Hilbert space $\ell^2(\mathbf{Z}_+) \times \ell^2(\mathbf{Z}_-)$ with the inner product

$$\left\langle \begin{pmatrix} x_+ \\ x_- \end{pmatrix}, \begin{pmatrix} x'_+ \\ x'_- \end{pmatrix} \right\rangle_H := \langle x_+, x'_+\rangle_{\ell^2(\mathbf{Z}_+)} + \langle x_-, x'_-\rangle_{\ell^2(\mathbf{Z}_+)},$$

for any $x_+, x'_+ \in \ell^2(\mathbf{Z}_+)$ and $x'_-, x_- \in \ell^2(\mathbf{Z}_-)$.

(ii) The unitary extension of S is the operator $U \in \mathcal{L}(H)$ defined by

$$U := \begin{pmatrix} S & I - SS^* \\ 0 & S^* \end{pmatrix}.$$

(iii) Define the projection operators $\pi_+ \in \mathcal{L}(H, \ell^2(\mathbf{Z}_+))$ and $\pi_- \in \mathcal{L}(H, \ell^2(\mathbf{Z}_-))$ by

$$\pi_+ \begin{pmatrix} x_+ \\ x_- \end{pmatrix} := x_+; \quad \pi_- \begin{pmatrix} x_+ \\ x_- \end{pmatrix} := x_-,$$

for any $x_+ \in \ell^2(\mathbf{Z}_+)$ and $x_- \in \ell^2(\mathbf{Z}_-)$.

It is not difficult to show that U is indeed unitary, and that its adjoint is given by $U^* = \left(\begin{smallmatrix} S^* & 0 \\ I-SS^* & S \end{smallmatrix} \right)$. In fact, with the obvious identification of the spaces H and the sequence space $\ell^2(\mathbf{Z})$, the operator U can be regarded as a bilateral shift operator on $\ell^2(\mathbf{Z})$ in an obvious way. It is well known that the operator polynomials of U and U^* form a commutative sub-algebra in $\mathcal{L}(H)$, whose operator norm closure has a particularly simple commutative C^*-algebra structure. This is the content of the following lemma.

LEMMA 2.1. *Let A be the closure in $\mathcal{L}(H)$ of the set of operators $p(U, U^*)$, where p ranges over all polynomials $p(x, y)$ with complex coefficient. Then A is a commutative C^*-algebra that is isometrically isomorphic to $C(\mathbf{T})$. Moreover, there is an isometric isomorphism $\Psi : C(\mathbf{T}) \to A$ which satisfies*

(i) $\Psi \bar{f} = (\Psi f)^*$,

(ii) $(\forall \xi \in \mathbf{T} : f(\xi) = \xi) \Rightarrow \Psi f = U$.

PROOF. First note that $\sigma(U) = \mathbf{T}$, because U is a bilateral shift. Because U and U^* commute, the claim follows from [17, Theorem 11.19]. □

In the light of previous lemma, let us denote the elements of A by $\mathcal{C}[f] := \Psi f$ — the convolution operator with symbol $f \in C(\mathbf{T})$. Now we are in position to define Toeplitz, causal and anti-causal Hankel operators.

DEFINITION 2.3. *Let U, π_+ and π_- be as in Definition 2.2. Let $f \in C(\mathbf{T})$ be arbitrary and $\mathcal{C}[f]$ be the convolution operator with symbol f.*

(i) The Toeplitz operator $\mathcal{T}[f]$ with symbol f is the operator in $\mathcal{L}(\ell^2(\mathbf{Z}_+))$ defined by

$$\mathcal{T}[f] := \pi_+ \mathcal{C}[f] \pi_+.$$

(ii) The causal Hankel operator $\mathcal{H}_+[f]$ with symbol f is the operator in $\mathcal{L}(\ell^2(\mathbf{Z}_-), \ell^2(\mathbf{Z}_+))$ defined by

$$\mathcal{H}_+[f] := \pi_+ \mathcal{C}[f] \pi_-.$$

(iii) The anti-causal Hankel operator $\mathcal{H}_-[f]$ with symbol f is the operator in $\mathcal{L}(\ell^2(\mathbf{Z}_-), \ell^2(\mathbf{Z}_+))$ defined by

$$\mathcal{H}_-[f] := \pi_- \mathcal{C}[f] \pi_+.$$

It is customary in the literature to define Toeplitz operators as limits of polynomials $p(S, S^*)$ of the unilateral shift operator S rather than truncations of convolution operators; see for example [5]. Both the approaches give the same Toeplitz operators associated to the same symbols, as reader may check. The study of the truncation effect (that is always present in the inversion of Toeplitz operators) seems to be more concise in the frame we have chosen.

The following two lemmas give fundamental properties of Toeplitz and Hankel operators.

LEMMA 2.2. *It is true that for any $f \in C(\mathbf{T})$*

(i) $||\mathcal{T}[f]|| = ||f||_\infty$,

(ii) $||\mathcal{H}_\pm[f]|| \leq ||f||_\infty$.

PROOF. For claim (i), see [5]. Claim (ii) follows because the norms of the orthogonal projections $\pi_\pm$ equal 1, and the norm of $\mathcal{C}[f]$ is $||f||_\infty$ by Lemma 2.1. ▯

Note that the estimate (ii) for the Hankel norm is nothing but optimal. The Nehari extension theorem characterizes the norm of Hankel operator $\mathcal{H}_+[f]$ as $\inf ||f + g||$ over all g with vanishing positive Fourier coefficients.

The index or winding number of a function $f \in C(\mathbf{T})$ is defined to be the number, how many times the curve $f(\mathbf{T})$ rotates around the origin. More formally

$$Ind\,(f) := \frac{1}{2\pi}\, arg\, f(e^{i\theta})|_0^{2\pi}.$$

The invertibility condition of a Toeplitz operator can be stated in terms of the index, as shown in [5]

LEMMA 2.3. *Let $f \in C(\mathbf{T})$ be arbitrary. Then $\mathcal{T}[f]$ is invertible if and only if $0 \notin f(\mathbf{T})$ and $Ind\,(f) = 0$.*

3 Preconditioning of Toeplitz operators

The study of the symbol $f \in C(\mathbf{T})$ of a Toeplitz operator $\mathcal{T}[f]$ is a central concept when the preconditioning of the operator is concerned. In particular, we may construct a multitude of preconditioners for a Toeplitz operator $\mathcal{T}[f]$ starting from the function f (see Theorem 3.1). Furthermore, the preconditioned operator will have such properties (see Theorem 3.3) so that the Krylov subspace method will give superlinear convergence, as shown in Section 4.

Let $f \in C(\mathbf{T})$ satisfying the invertibility conditions of Lemma 2.3 be arbitrary. By $\{a_j\}_{j=0}^\infty$ denote a sequence in $\ell^2(\mathbf{Z}_+)$. Our problem is to iteratively invert

$$\{a_j\}_{j=0}^\infty = \mathcal{T}[f]\{b_j\}_{j=0}^\infty. \tag{3.1}$$

Let us multiply the both sides of equation (3.1) by an invertible Toeplitz operator $\mathcal{T}[g] \in \mathcal{L}(\ell^2(\mathbf{Z}_+))$ for a suitably chosen symbol g. This gives

$$\mathcal{T}[g]\{a_j\}_{j=0}^\infty = \mathcal{T}[g]\mathcal{T}[f]\{b_j\}_{j=0}^\infty,$$

or equivalently in the form of a fixed point problem

$$\{b_j\}_{j=0}^\infty = (I - \mathcal{T}[g]\mathcal{T}[f])\{b_j\}_{j=0}^\infty + \mathcal{T}[g]\{a_j\}_{j=0}^\infty. \tag{3.2}$$

We say that the equation (3.2) has been Toeplitz preconditioned, at least if $\mathcal{T}[g]$ is in some sense close to $\mathcal{T}[f]^{-1}$ (see [1],[2]). Note that the linear operator $I - \mathcal{T}[g]\mathcal{T}[f]$ is no longer Toeplitz; yet, in a sense, it is not very far from being Toeplitz. The following decomposition theorem will make this point precise.

THEOREM 3.1. *Assume that $f, g \in C(\mathbf{T})$. Then $I - \mathcal{T}[g]\mathcal{T}[f]$ can be decomposed as*

$$I - \mathcal{T}[g]\mathcal{T}[f] = (\mathcal{T}[gf] - \mathcal{T}[g]\mathcal{T}[f]) + \mathcal{T}[1 - gf]) \\ =: K_{f,g} + B_{f,g}, \tag{3.3}$$

where $K_{f,g}$ is a product of the causal and anticausal Hankel operators

$$K_{f,g} = \mathcal{H}_+[g]\mathcal{H}_-[f], \tag{3.4}$$

and $B_{f,g}$ is Toeplitz.

PROOF. By definition $B_{f,g}$ is Toeplitz. For $K_{f,g}$ we may write by Definition 2.3 and the fact that the projections on H satisfy $\pi_+ + \pi_- = I$

$$\begin{aligned} K_{f,g} &= \mathcal{T}[gf] - \mathcal{T}[g]\mathcal{T}[f] = \pi_+\mathcal{C}[gf]\pi_+ - \pi_+\mathcal{C}[g]\pi_+\pi_+\mathcal{C}[f]\pi_+ \\ &= \pi_+\mathcal{C}[gf]\pi_+ - \pi_+\mathcal{C}[g](I - \pi_-)\mathcal{C}[f]\pi_+ \\ &= (\pi_+\mathcal{C}[gf]\pi_+ - \pi_+\mathcal{C}[g]\mathcal{C}[f]\pi_+) + \pi_+\mathcal{C}[g]\pi_-\pi_-\mathcal{C}[f]\pi_+ \\ &= \pi_+(\mathcal{C}[gf] - \mathcal{C}[g]\mathcal{C}[f])\pi_+ + \pi_+\mathcal{C}[g]\pi_-\pi_-\mathcal{C}[f]\pi_+ \\ &= \pi_+\mathcal{C}[g]\pi_-\pi_-\mathcal{C}[f]\pi_+ = \mathcal{H}_+[g]\mathcal{H}_-[f], \end{aligned} \tag{3.5}$$

where we have used the fact that $\mathcal{C}[gf] - \mathcal{C}[g]\mathcal{C}[f] = 0$, by Lemma 2.1. □

At this stage we give the operators $K_{f,g}$, $B_{f,g}$ names. We propose the following:

DEFINITION 3.1. *Let $K_{f,g}$, $B_{f,g}$ be as in Theorem 3.1. We call the operators $K_{f,g}$, $B_{f,g}$ the truncation effect operator and perturbation operator, respectively.*

We can say even more about the structure of the truncation effect $K_{f,g}$ that will give us insight into the properties of iteration of $\mathcal{T}[f]$. For this purpose we first need to recall the definitions of approximation numbers and Schatten classes.

DEFINITION 3.2. *Let $T \in \mathcal{L}(\ell^2(\mathbf{Z}_+))$ and $k \in \mathbf{N}$. The approximation numbers by finite dimensional operators are defined by*

$$\sigma_k(T) := \inf_{rank\, F \leq k-1} ||T - F||.$$

In a Hilbert space the approximation numbers $\sigma_k(T)$ equal the singular values of T; see for example [4, p. 1089]. The closed ideal of compact operators $\mathcal{LC}(\ell^2(\mathbf{Z}_+))$ can now be divided into smaller spaces, if we look at the decay of the approximation numbers $\sigma_k(T)$. Consider the following definition:

DEFINITION 3.3. *Let $p \in (0, \infty)$.*

(i) By $||.||_{S_p}$ denote the number in $[0, \infty]$ given by

$$||T||_{S_p} := (\sum_{k=1}^{\infty} |\sigma_k(T)|^p)^{\frac{1}{p}},$$

for each $T \in \mathcal{LC}(\ell^2(\mathbf{Z}_+))$.

(ii) By S_p denote the set of such $T \in \mathcal{LC}(\ell^2(\mathbf{Z}_+))$ that $||T||_{S_p} < \infty$. The space S_p is called the Schatten p-class.

Note that $||.||_{S_p}$ is not actually a norm if $p \in (0,1)$ — the triangle inequality fails. However, S_p is a Banach space for $p \in [1,\infty)$.

One more detail is needed for the proof of Theorem 3.3, namely the following lemma which is a quite simple combination of two Jackson's theorems.

LEMMA 3.2. *Let $r \in \mathbf{Z}_+$, $\alpha \geq 0$ such that $r+\alpha > 0$. Let $f \in C^{r,\alpha}(\mathbf{T})$. Define $E_k(f) := \inf_{deg\, p_k \leq k} ||p_k - f||_\infty$, where p_k is a trigonometric polynomial of degree k. Then*

$$E_k(f) \leq \frac{\pi^{r+\alpha}}{2^r}\, ||f^{(r)}||_{Lip_\alpha(\mathbf{T})}\, \frac{1}{(k+1)^{r+\alpha}}. \tag{3.6}$$

PROOF. In the proof of [15, Theorem 16.4], we find in formula (16.33)

$$E_k(f) \leq \frac{\pi}{2k+2} E_k(f'),$$

where $f'(e^{it}) = \frac{d}{dt} f(e^{it})$ is the derivative of f relative to the arch length parameter t of $\mathbf{T}$. Using this r times proves (3.6) for $\alpha = 0$. For $\alpha > 0$, we use it again r times. Taking into consideration [15, Theorem 16.2], we have

$$E_k(f) \leq (\frac{\pi}{2k+2})^r E_k(f^{(r)}) \leq (\frac{\pi}{2k+2})^r\, \omega(\frac{\pi}{k+1}), \tag{3.7}$$

where ω is the modulus of continuity of $f^{(r)}$. Because $f^{(r)}$ is Lipschitz of index α, we get for all $h > 0$, $e^{it_1}, e^{it_2} \in \mathbf{T}$ satisfying $0 < |t_1 - t_2| < h$

$$\begin{aligned}
&|f^{(r)}(e^{it_1}) - f^{(r)}(e^{it_2})| \\
&\leq h^\alpha \frac{|f^{(r)}(e^{it_1}) - f^{(r)}(e^{it_2})|}{|t_1 - t_2|^\alpha} \leq h^\alpha \sup_{|t_1 - t_2| \leq h} \frac{|f^{(r)}(e^{it_1}) - f^{(r)}(e^{it_2})|}{|t_1 - t_2|^\alpha} \\
&\leq h^\alpha \left[||f^{(r)}||_\infty + \sup_{t_1 \neq t_2} \frac{|f^{(r)}(e^{it_1}) - f^{(r)}(e^{it_2})|}{|t_1 - t_2|^\alpha} \right] =: h^\alpha\, ||f^{(r)}||_{Lip_\alpha(\mathbf{T})},
\end{aligned}$$

and by the previous

$$\omega(h) := \sup_{0 < |t_1 - t_2| \leq h} |f^{(r)}(e^{it_1}) - f^{(r)}(e^{it_2})| \leq h^\alpha\, ||f^{(r)}||_{Lip_\alpha(\mathbf{T})}.$$

This, together with formula (3.7), proves the claim. □

Note that because the space of polynomials of degree $\leq k$ is finite dimensional, the infimum $E_k(f)$ is in fact attained by some polynomial p_k.

Now we are ready to present a result about the relation between the smoothness of the symbol f and the distribution of the singular values of $K_{f,g}$.

THEOREM 3.3.

(i) If $f, g \in C(\mathbf{T})$, then $K_{f,g}$ is compact.

(ii) Let $r \in \mathbf{Z}_+$, $\alpha \geq 0$ such that $r+\alpha > 0$. If $f \in C^{r,\alpha}(\mathbf{T})$, then the approximation numbers of $K_{f,g}$ satisfy

$$\sigma_k(K_{f,g}) \leq \frac{\pi^{r+\alpha}}{2^r} \, ||f^{(r)}||_{Lip_\alpha(\mathbf{T})} \, ||g||_\infty \, k^{-r-\alpha} \quad \text{for all} \quad k \in \mathbf{N}.$$

(iii) Let $p \in (0,\infty)$. The Schatten information about $K_{f,g}$ is

$$||K_{f,g}||_{S_p} \leq \frac{\pi^{r+\alpha}}{2^r} \, ||f^{(r)}||_{Lip_\alpha(\mathbf{T})} \, ||g||_\infty \Big[\sum_{k=1}^{\infty} (k+1)^{-p(r+\alpha)} \Big]^{\frac{1}{p}}.$$

In particular, if $p > \frac{1}{r+\alpha}$, then $K_{f,g} \in S_p$.

Proof. We start with proving claim (i). It is more than enough to prove that the Hankel operators $\mathcal{H}_\pm[f]$ are compact for all $f \in C(\mathbf{T})$. Choose a sequence of trigonometric polynomials in $\xi \in \mathbf{T}$ written as $\{p_k(\xi)\} := \{\sum_{j=-k}^{k} c_j^{(k)} \, \xi^j\}$ of degree k such that $||p_k - f||_\infty \to 0$ as $k \to \infty$. This can be done by the Weierstrass approximation theorem. It follows from Lemma 2.1 that

$$\begin{aligned}\mathcal{C}[p_k] \equiv \Psi(p_k(\xi)) &= \Psi(\sum_{j=-k}^{k} c_j^{(k)} \, \xi^j) = \sum_{j=-k}^{k} c_j^{(k)} \, \Psi(\xi^j) \\ &= \sum_{j=0}^{k} c_j^{(k)} \, (\Psi(\xi))^j + \sum_{j=-k}^{-1} c_j^{(k)} \, (\Psi(\xi^{-1}))^{-j}.\end{aligned}$$

Note that $\Psi(\xi) = U$ by claim (ii) of Lemma 2.1 and furthermore $\Psi(\xi^{-1}) = \Psi(\bar{\xi}) = U^*$ by claim (i) of Lemma 2.1. This allows us to continue

$$\begin{aligned}\mathcal{C}[p_k] &= \sum_{j=0}^{k} c_j^{(k)} \, U^j + \sum_{j=-k}^{-1} c_j^{(k)} \, U^{*-j} \\ &= \sum_{j=0}^{k} c_j^{(k)} \, U^j + \sum_{j=-k}^{-1} c_j^{(k)} \, U^j = p_k(U),\end{aligned} \tag{3.8}$$

where the second equality is written because U is unitary. The operator polynomial $p_k(U)$, however, can also be computed in terms of S, see part (ii) of Definition 2.2. A direct computation gives for $j \in \mathbf{N}$

$$U^j = \begin{pmatrix} S^j & q_j(S,S^*) \\ 0 & S^{*j} \end{pmatrix}, \tag{3.9}$$

where $q_j(S,S^*)$ satisfies the recursion

$$q_j(S,S^*) = \begin{cases} I - SS^* & j = 1 \\ S q_{j-1}(S,S^*) + (I - SS^*) S^{*\,(j-1)} & j > 1. \end{cases}$$

After removing the recursion from the previous formula, we get

$$\begin{aligned} q_j(S,S^*)\{a_0,a_1,a_2,\dots a_{j-1},a_j\dots\} \\ =\{a_{j-1},a_{j-2},a_{j-3},\dots a_0,0,0,0,\dots\}. \end{aligned} \tag{3.10}$$

A similar kind of calculation can be made for the powers of $U^* \equiv U^{-1}$. Combining this with formula (3.8) and Definition 2.3 we get

$$\mathcal{C}[p_k] := \begin{pmatrix} \mathcal{T}[p_k] & \mathcal{H}_+[p_k] \\ \mathcal{H}_-[p_k] & \mathcal{T}[p_k(\bar{\xi})] \end{pmatrix},$$

where the Hankel operators $\mathcal{H}_\pm[p_k]$ are finite sums of j-dimensional operators $q_j(S,S^*)$

$$\begin{cases} \mathcal{H}_+[p_k] := \sum_{j=1}^{k} c_j^{(k)}\, q_j(S,S^*) \\ \mathcal{H}_-[p_k] := \sum_{j=-k}^{-1} c_j^{(k)}\, q_j(S,S^*). \end{cases}$$

Consequently, both the operators $\mathcal{H}_\pm[p_k]$ are at most k-dimensional, because the ranges $q_j(S,S^*)$ form an increasing sequence with increasing j by formula (3.10). By claim (ii) of Lemma 2.2 it follows that

$$||\mathcal{H}_\pm[f]-\mathcal{H}_\pm[p_k]|| = ||\mathcal{H}_\pm[f-p_k]|| \le ||f-p_k||_\infty \to 0 \quad \text{as} \quad k\to\infty, \tag{3.11}$$

where the limit is taken by the choise of polynomial sequence $\{p_k\}$. Thus $\mathcal{H}_\pm[f]$ is a norm limit of finite dimensional operators, and thus compact. This proves the first part of the theorem.

The proof of claim (ii) goes as follows. From Lemma 3.2 we know that the approximation properties of $f \in C(\mathbf{T})$ by trigonometric polynomials are consistent with the smoothness of f. More precisely, if $f \in C^{r,\alpha}(\mathbf{T})$, then we have a sequence of trigonometric polynomials p_k just as in the first part of the proof with the additional speed estimate for formula (3.11)

$$||\mathcal{H}_-[f]-\mathcal{H}_-[p_{k-1}]|| \le ||f-p_{k-1}||_\infty \le \frac{\pi^{r+\alpha}}{2^r}\,||f^{(r)}||_{Lip_\alpha(\mathbf{T})} k^{-r-\alpha} \tag{3.12}$$

The polynomial p_{k-1} exists because the infimum $E_k(f)$ in Lemma 3.2 is actually attained by some trigonometric polynomial. We estimate by Theorem 3.1 and by the fact that $rank\, \mathcal{H}_+[g]F \le rank\, F$ for $k\in\mathbf{N}$

$$\begin{aligned} \sigma_k(K_{f,g}) &:= \inf_{rank\, F\le k-1} ||K_{f,g}-F|| = \inf_{rank\, F\le k-1} ||\mathcal{H}_+[g]\mathcal{H}_-[f]-F|| \\ &\le \inf_{rank\, F\le k-1} ||\mathcal{H}_+[g]\mathcal{H}_-[f]-\mathcal{H}_+[g]F|| \le ||\mathcal{H}_+[g]|| \inf_{rank\, F\le k-1} ||\mathcal{H}_-[f]-F||. \end{aligned}$$

Claim (ii) of Lemma 2.2 gives us an upper estimate for $||\mathcal{H}_+[g]||$, and formula (3.12) gives us an upper estimate for $\inf_{rank\, F\le k-1}||\mathcal{H}_-[f]-F||$, because $\mathcal{H}_-[p_{k-1}]$ is at most $(k-1)$-dimensional. This proves the second part of the theorem. The last claim (iii) is a trivial consequence of the second part of the theorem. $\square$

In the proof Theorem 3.3 we used the fact that a Hankel operator having a trigonometric polynomial symbol is compact. The theorem of Kronecker (see [14, Theorem 3.11]) gives a complete characterization of all finite dimensional Hankel operators — they are exactly those having symbol that can be extended into a rational function in the complex plane, with a finite number of poles outside the unit disk of the complex plane. Furthermore, the rank of such Hankel operator is exactly the number of these poles +1. By AAK - theorem (see [14, Theorem 6.12]), we would get a sharp approximation for the singular values of $\mathcal{H}[f]$, by looking the approximation properties of $f \in C(\mathbf{T})$ by such rational functions. However, our crude trigonometric polynomial approach gives us quantitative estimates (via familiar Jackson approximation theorems) for the decay of the singular values of $K_{f,g}$, stated in terms of the smoothness of the symbol f.

4 On the convergence of iterations

Consider the following theorem:

THEOREM 4.1. *Let $f \in C(\mathbf{T})$ satisfy the invertibility condition of Lemma 2.3. Then there is a sequence of trigonometric polynomials $\{g_k\}_{k=0}^{\infty}$ of degree k satisfying:*

(i) $||B_{f,g_k}|| \to 0$ *as* $k \to \infty$, *and from some k on, g_k satisfies the invertibility condition of Lemma 2.3.*

(ii) For such large k, $\mathcal{T}[g_k]$ is a nonsingular preconditioner for $\mathcal{T}[f]$ as in formula 3.2, so that what is left after preconditioning has a large compact part and a small Toeplitz perturbation.

(iii) Moreover, assume that $f \in C^{r,\alpha}(\mathbf{T})$ for $r \in \mathbf{Z}_+$, $0 \leq \alpha \leq 1$ such that $r + \alpha > 0$. Then we may require $\{g_k\}$ to satisfy

$$||B_{f,g_k}|| \leq \frac{\pi^{r+\alpha}}{2^n} ||f||_\infty ||(\frac{1}{f})^{(r)}||_{Lip_\alpha(\mathbf{T})} (k+1)^{-r-\alpha}. \tag{4.1}$$

Also, if $p > \frac{1}{r+\alpha}$, then the family $\{K_{f,g_k}\}_{k\geq 0}$ is an uniformly bounded set in S_p.

PROOF. To prove claim (i), first note that if f satisfies the invertibility condition of Lemma 2.3, so does $\frac{1}{f}$. Choose a sequence $\{g_k\}$ so that $||g_k - \frac{1}{f}||_\infty \to 0$. By the definition of B_{f,g_k} and Lemma 2.2 we have

$$||B_{f,g_k}|| = ||f(\frac{1}{f} - g_k)||_\infty \leq ||f||_\infty ||(\frac{1}{f} - g_k)||_\infty \to 0. \tag{4.2}$$

Furthermore, for such sequence $\{g_k\}$ claim (i) of Lemma 2.2 implies that $||\mathcal{T}[g_k] - \mathcal{T}[\frac{1}{f}]|| \to 0$. Because the set of invertible (Toeplitz) operators is open, the first part of the theorem follows. Claim (ii) is a consequence of claim (i) and Theorem 3.1.

In order to prove claim (iii) note that if $f \in C^{r,\alpha}(\mathbf{T})$, so does $\frac{1}{f}$. By Lemma 3.2 we can choose the sequence $\{g_k\}$ to satisfy formula (4.1). This can be done because the infimum $E_k(f)$ of Lemma 3.2 is actually attained by some trigonometric polynomial.

To prove the remaining claim, note that $\{g_k\}$ is an uniformly bounded family in $C(\mathbf{T})$ because it converges uniformly to a limit in $C(\mathbf{T})$. It follows that $\{K_{f,g_k}\}_{k\geq 0}$ is an uniformly bounded family in S_p, by part (iii) of Theorem 3.3. ∎

Theorems 3.1 and 4.1 show that after Toeplitz preconditioning, the operator consists of a compact truncation effect $K_{f,g}$ perturbed by a small $B_{f,g}$. Smoothness of the symbol f has thus a two-fold impact on the properties of iteration of $\mathcal{T}[f]$: For smooth f it is easier to control the "preconditioning error" $B_{f,g}$. And then, by Theorem 3.3, smoothness specifies the approximation properties of $K_{f,g}$, which imply speed estimates for the Krylov subspace methods as will be presented below.

When applying the Krylov methods upon $K_{f,g} + B_{f,g}$, we must first "invest" into the amount of work and storage space used for the preconditioning and then collect the "profits" by having superlinear convergence in the first steps of iteration. Clearly the required accuracy is one component in the game, too.

The preconditioning sequence $\{g_k\}$ can be constructed in great many ways, still preserving the speed estimate 4.1 in an asymptotic sense; see for example [8, pp.21, Ex. 2] for Lipschitz continuous symbols. On the numerical construction schemes for $\{g_k\}$, we refer to the ideas presented in [2]. In the numerical experiments given in Section 5 we have used trigonometric polynomial interpolation.

The setting proposed above leads us to study an even more general problem: how does a Krylov subspace method perform if it is applied upon an operator consisting of a compact K perturbed by an unstructured small B, if Schatten information of K is available (see [9],[13]).

In the study of Krylov subspace methods applied upon $K + B$ it is customary to look at how the sequence $||p_k(K+B)||^{\frac{1}{k}}$ behaves as $k \to \infty$, where $\{p_k\}$ is a sequence of normalized ($p_k(1) = 1$) degree k polynomials associated to the Krylov subspace method in question (see [9], [12], [13]). A function theoretic argument proves the following theorem where the normalization of the polynomials $\{\tilde{p}_k\}$ is slightly different, but without effect on the asymptotics as $k \to \infty$:

Theorem 4.2. *Let $p \geq 1$. Let H be Hilbert and $S_p(H)$ be the Schatten p-class of compact operators in $\mathcal{LC}(H)$. Take $K \in S_p(H)$ and let $B \in \mathcal{L}(H)$ be a small perturbation such that $1 \notin \sigma(K + B)$. Then there exists a sequence of essentially monic polynomials $\{\tilde{p}_k\}_{k=1}^{\infty}$ such that for all parameter values $\beta \in (0, 1]$*

$$||\tilde{p}_k(K + B)||^{\frac{1}{k}} \leq p^{\frac{1}{k}} \big(||B|| + ||K||_{S_p}\, k^{-\frac{\beta}{p}}\big) \Big(\frac{||B||\, k^{\frac{\beta}{p}}}{||K||_{S_p}} + 1\Big)^{\frac{1}{k}} e^{\frac{C_p}{k^{1-\beta}}}, \tag{4.3}$$

where C_p is a finite constant (less than 3). Furthermore, $\lim_{k\to\infty} |\tilde{p}_k(1)| > 0$.

Proof. See [9, Theorem 6.7]. ∎

An analogous theorem can be proved for the Schatten classes $p \in (0, 1]$ (see [9, Theorem 6.9]).

The expression "essentially monic" means that the leading term of all $\tilde{p}_k$ is a same nonzero complex number. The fact that $\lim_{k\to\infty} |\tilde{p}_k(1)| > 0$ makes it possible to renormalize $\tilde{p}_k$ for large k, and define

$$p_k(\lambda) := \frac{\tilde{p}_k(\lambda)}{\tilde{p}_k(1)}.$$

Now the sequence $\{p_k\}$ has the correct normalization $p_k(1) = 1$. Furthermore, the speed estimate like (4.3) holds for $\{p_k\}$ for all k large enough with an additional sequence $|\tilde{p}_k(1)|^{-\frac{1}{k}}$ of multiplicative constants. Note that because $\lim_{k\to\infty} |\tilde{p}_k(1)|^{-\frac{1}{k}} = 1$, the effect of the incorrect normalization of $\tilde{p}_k$ does not change the nature of speed estimate (4.3) in the asymptotic sense.

Theorem 4.2 tells us that in the first stages of the iteration the convergence factor $||p_k(K+B)||^{\frac{1}{k}}$ of order $||B|| + ||K||_{s_p} k^{-\frac{\beta}{p}}$ decreases (the "superlinear" stage) and is asymptotically only of order $||B||$ (the "linear" stage). Moreover, the rate of decrease of the convergence factor is dictated by the Schatten class of K. Note that the notion "superlinear" is usually used to describe something that happens in the asymptotics of the speed estimates. Here we are a bit unorthodox and regard "superlinear" stage of an iteration as those iteration steps when "speed is being gained". By the "linear" stage we refer to the analogous phenomenon.

The GMRES method for the inversion of non-symmetric problems can be regarded as a minimization algorithm that (at least implicitly) generates polynomial sequences to approximate the values of resolvents at certain points; this is the minimization of residuals. If the GMRES generates the polynomial sequence s_k with $\deg(s_k) = k$ (corresponding to the normalized sequence p_k given after Theorem 4.2), then the residual d_k after k steps is of size $||s_k(K+B)d_0||$ (see [9, Proposition 2.2] or [12, Chapter 1], and we have

$$||s_k(K+B)d_0|| \leq ||p_k(K+B)d_0|| \leq ||p_k(K+B)||\,||d_0||. \tag{4.4}$$

The former inequality is true because s_k is optimal at d_0, and p_k is possibly worse than optimal for the initial residual d_0. This is to say that the upper estimates we have for $||p_k(K+B)||$ are as well upper estimates for the GMRES residuals. The same kind of result is true so as to the error sequences with quite obvious modifications for the reasoning — we again refer at [9] or [12, Chapter 1].

5 Numerical experiments

We conclude this paper by presenting some results of numerical experiment about the norms of $||B_{f,g_k}||$ and $||K_{f,g_k}||_{S_p}$, when the degree k of preconditioner becomes large. This is the data that we need for the speed estimate of Theorem 4.2. We use the Toeplitz matrix theory given in [11] as a technical tool.

Until now there has been no need to take any position what kind of data structure or discretization could or should be used when calculating with objects such as Toeplitz operators and their truncation effects. Now that we are actually going to calculate something, a position will have to be taken. In this paper, we simply make all the calculations in the (Toeplitz) matrix algebra of sufficiently high dimension, and leave the general question how to implement an efficient iterative Toeplitz manipulator or solver as an open problem.

Even though the truncation effect $K_{f,g}$ generally contains an infinite amount of data, it is possible to find $K_{f,g}$ exactly by making a finite number of matrix algebra operations, provided that the symbols f, g are trigonometric polynomials. The general symbols $f, g \in C(\mathbf{T})$ must be handled by a limit argument. In [11] it is shown that if we define the Toeplitz matrix truncation effect (analogous to $K_{f,g}$) by

$$K^{(n)}_{f,g} := T_n[gf] - T_n[g]T_n[f],$$

then we may decompose $K^{(n)}_{f,g} = K^{(n)+}_{f,g} + K^{(n)-}_{f,g}$, where $K^{(n)+}_{f,g} \to K_{f,g}$ in the operator norm and $K^{(n)-}_{f,g} \to 0$ in strong operator topology. Here $T_n[f] := \pi_{[0,n-1]}\mathcal{T}[f]\pi_{[0,n-1]}$ is the $n \times n$ Toeplitz matrix with symbol $f \in C(\mathbf{T})$, and $\pi_{[0,n-1]}$ is the orthogonal projection of $\ell^2(\mathbf{Z}_+)$ onto the first n components of the sequence space. For notations and details, see [11, Theorem 7 and Lemma 10]. Our basic tool will be the following two facts whose proofs will be omitted.

(i) Let $f, g \in C(\mathbf{T})$. Let $\{f^{(k)}\}$, $\{g^{(k)}\}$ be sequences of trigonometric polynomials, $\deg(f^{(k)}) \leq k$, $\deg(g^{(k)}) \leq k$, such that $f^{(k)} \to f$, $g^{(k)} \to g$ uniformly. Then

$$(5.1) \qquad ||K_{f,g} - K^{(n)+}_{f^{(j)},g^{(k)}}|| \to 0 \quad \text{as} \quad \min(j,k,n) \to \infty.$$

(ii) If, in addition, $K_{f,g} \in S_p(H)$, then $||K^{(n)+}_{f^{(j)},g^{(k)}}||_{S_p} \to ||K_{f,g}||_{S_p}$, as $\min(j,k,n) \to \infty$.

These facts make it possible to approximate the S_p-norm of an infinite dimensional truncation effect $K_{f,g}$ by studying the objects $K^{(n)+}_{f^{(j)},g^{(k)}}$ that are $n \times n$ matrices. Their S_p-norms can be calculated by SVD. So it remains to figure out how to practically obtain the $K^{(n)+}_{f^{(j)},g^{(k)}}$'s. The problem is that we need to eliminate the interference of $K^{(n)-}_{f^{(j)},g^{(k)}}$ with $K^{(n)+}_{f^{(j)},g^{(k)}}$, to obtain the latter unperturbed.

The solution lies in the structure of the matrix truncation effect $K^{(n)}_{f,g}$. If g is a trigonometric polynomial of degree d and $n \geq 2d$, then the ranges of $K^{(n)+}_{f,g}$ and $K^{(n)-}_{f,g}$ are orthogonal (see [11, Lemma 15]). More precisely, $K^{(n)+}_{f,g} = \pi_{[0,d-1]}K^{(n)+}_{f,g}$ and $K^{(n)-}_{f,g} = \pi_{[n-d,n-1]}K^{(n)-}_{f,g}$, where the ranges of the orthogonal projections $\pi_{[0,d-1]}$, $\pi_{[n-d,n-1]}$ are othogonal. By requiring $j = n$ and keeping $k \leq n/2$ in formula (5.1), we can easily separate $K^{(n)+}_{f^{(j)},g^{(k)}}$ from $K^{(n)}_{f^{(j)},g^{(k)}}$.

Also a calculation of $||B_{f,g_k}||$ is problematic, because we would need $||g_k - \frac{1}{f}||_\infty$ which is not available for great many practical preconditioning sequences g_k. Here we calculate the norm of a n-dimensional Toeplitz matrix $B^{(n)}_{f,g_k}$ that approximates B_{f,g_k}.

Now that we know what to calculate and how, we present a result of a numerical experiment, realized as Matlab code. We generate a class of $Lip_\alpha(\mathbf{T})$ symbols u_α for $0 < \alpha < 1$, defined by

$$u_\alpha(e^{it}) = \sum_{j=0}^{\infty} \frac{\cos(3^j t)}{3^{\alpha j}} + (\frac{1}{1-3^{-\alpha}} + 1). \tag{5.2}$$

It can be shown that $u_\alpha \in Lip_\alpha(\mathbf{T})$ but $u_\alpha \notin Lip_{\alpha-\epsilon}(\mathbf{T})$ for any $\epsilon > 0$ (see [8, Theorem 4.6 and Ex. 8, p. 27]). Moreover, the additive constant in (5.2) is chosen so that the invertibility condition of Lemma 2.3 is satisfied.

Let $n = 2^{l-1}$, where $l \in \mathbf{N}$ is the fixed approximation level of the experiment. We define the trigonometric polynomials f_α, (g_α) of degree $\leq n$ by trigonometric polynomial interpolation of the function u_α ($1/u_\alpha$, respectively) at the $2^l = 2n$ interpolation points $\{e^{2\pi i \frac{j}{2n}}\}_{j=\{0,\ldots,2n-1\}} \subset \mathbf{T}$. We use the uniformly spaces interpolation mesh of 2^l nodes because it is convenient for the FFT techniques. The function f_α is the symbol of the Toeplitz operator for the experiment; it satisfies the invertibility condition of Lemma 2.3 for n large enough because so does u_α. Note that the trigonometric polynomial interpolation does not give us the optimal uniform convergence speed described in Lemma 3.2.

To obtain the preconditioning sequence, the interpolant g is truncated to a degree k ($k \leq n/2$) trigonometric polynomial by zeroing the extra Fourier coefficients of g_α. This gives the preconditioning sequence $\{(g_\alpha)_k\}$ for $2k \leq n = 2^{l-1}$, where k is the preconditioning parameter. The preconditioning parameter has to be kept below 2^{l-1} to ensure that the truncation effects $K^{(n)+}_{f_\alpha,(g_\alpha)_k}, K^{(n)-}_{f_\alpha,(g_\alpha)_k}$ do not mix, as discussed above.

In Figures 1, 2 and 3 we give numerical results for the symbols f_α for $\alpha = 0.5, 0.25$ and 0.125, respectively. We use the S_4 -norm (i.e. $p = 8$) to measure the decay of singular values for $K_{f_\alpha,(g_\alpha)_k}$ (see claim (iii) of Theorem 3.3). The norms $||B_{f_\alpha,(g_\alpha)_k}||$ and $||K_{f_\alpha,g_{\alpha k}}||_{S_4}$ for different preconditioning levels $k = 1 \ldots n/2$ with $n = 2^{l-1}, l = 9$.

In Figure 4 we present a typical example for a smooth symbol $f \in C^\infty(\mathbf{T})$. We use

$$u(e^{it}) = \left| \frac{\Pi_{j=1}^{50}(e^{it} - z_j)}{\Pi_{j=1}^{30}(e^{it} - p_j} \right| + 1,$$

with $z_j = e^{\frac{2\pi i + 100}{10000}}$ and $p_j = e^{\frac{-2\pi i + 100}{10000}}$ instead of u_α of formula (5.2). Then with obtain f in the same way as f_α above. The approximation level is set $l = 7$ and S_1-norm is used for K_{f,g_k}.

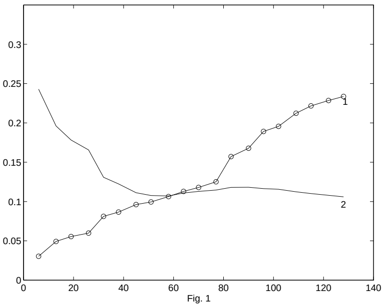

Fig. 1. $Lip_{\frac{1}{2}}$ **symbol**

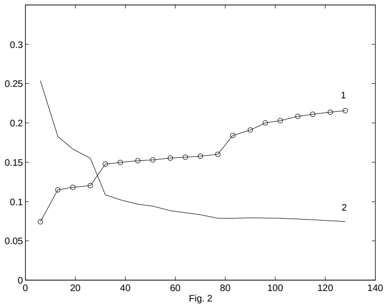

Fig. 2. $Lip_{\frac{1}{4}}$ **symbol**

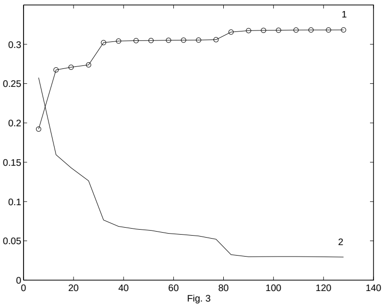

Fig. 3. $Lip_{\frac{1}{8}}$ **symbol**

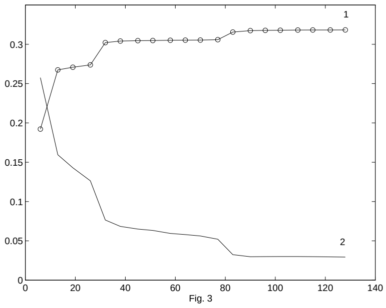

Fig. 4. C^{∞} **symbol**

In figures 1, 2 and 3, curve 1 is $||B_{f,g_k}||$ and curve 2 is $20 \cdot ||K_{f,g_k}||_{S_p}$ plotted against the preconditioning parameter k. Figure 4 is the same, except that we have plotted $10 \cdot ||K_{f,g_k}||_{S_p}$. In all these cases, $||B_{f,g_k}||$ seems to approach zero, and $||K_{f,g_k}||_{S_p}$ seems to remain bounded. We conclude that with better preconditioners g_k (i.e. larger k) the perturbation $||B_{f,g_k}||$ decreases, as the singular value norm of K_{f,g_k} remains bounded. Also the effect of the smoothness of the symbol is visible. This verifies the result of Theorem 4.1 in these special cases. Using the speed estimate (4.3), we conclude that increasing the preconditioning level k does not essentially increase the burden of the iterative solver, but only gives us more iteration steps of superlinear of convergence.

REFERENCES

1. R. Chan, *Toeplitz preconditioners for Toeplitz systems with nonnegative generating functions*, IMA J. Numer. Anal., 11 (1991), pp. 333–345.
2. R. Chan and K.-P. Ng, *Toeplitz preconditioners for hermitian Toeplitz systems*, Linear Algebra and its Applications, 190 (1993), pp. 181–205.
3. R. G. Douglas, *Banach algebra techniques in operator theory*, Academic Press, New York San Francisco London, 1972.
4. N. Dunford and J. Schwarz, *Linear Operators; Part II: Spectral Theory*, Interscience Publishers, Inc. (J. Wiley & Sons), New York, London, 1963.

5. I. Gohberg and I. Feldman, *Convolution Equations and Projection Methods for Their Solution*, vol. 41 of AMS Translations of Mathematical Monographs, American Mathematical Society, 1974.
6. U. Grenander and G. Szegö, *Toeplitz forms and their application*, Chelsea Publishing Company, New York, 1984.
7. V. Ionescu and M. Weiss, *Continuous and discrete-time Riccati theory: a Popov-function approach*, Linear Algebra and Applications, 193 (1993), pp. 173–209.
8. Y. Katznelson, *Introduction to Harmonic Analysis*, Dover Publications, Inc., New York, 2 ed., 1976.
9. J. Malinen, *On the properties for iteration of a compact operator with unstructured perturbation*, Helsinki University of Technology, Institute of mathematics, Research Report, A360 (1996).
10. ———, *Minimax control of distributed discrete time systems through spectral factorization*, Proceedings of EEC97, Brussels, Belgium, (1997).
11. ———, *Toeplitz preconditioning of Toeplitz matrices - an operator theoretic approach*, Preprint, (1997).
12. O. Nevanlinna, *Convergence of Iterations for Linear Equations*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, Boston, Berlin, 1993.
13. ———, *Convergence of Krylov-methods for sums of two operators*, BIT Numerical Mathematics, 36 (1996), pp. 775–785.
14. J. R. Partington, *An introduction to Hankel operators*, vol. 13 of Student texts, London Mathematical Society, London, 1988.
15. M. Powell, *Approximation Theory and Methods*, Cambridge University Press, Cambridge, 1981.
16. M. Rosenblum and J. Rovnyak, *Hardy classes and operator theory*, Oxford university press, New York, 1985.
17. W. Rudin, *Functional Analysis*, McGraw-Hill Book Company, New York, TMH ed., 1990.
18. O. J. Staffans, *Quadratic optimal control of stable well-posed linear systems*, To appear in Transactions of American Mathematical Society, (1997).
19. G. Strang, *A proposal for Toeplitz matrix calculations*, Studies in Applied Mathematics, 74 (1986), pp. 171–176.
20. B. Sz.-Nagy and C. Foias, *Harmonic Analysis of Operators on Hilbert space*, North-Holland Publishing Company, Amsterdam, London, 1970.