On the Properties for Iteration of a Compact Operator with Unstructured Perturbation

Jarmo Malinen
Helsinki University of Technology
Institute of Mathematics
Research Reports A360 (1996)

Jarmo Malinen: On the Properties for Iteration of a Compact Operator with Unstructured Perturbation, Helsinki University of Technology, Institute of Mathematics, Research Reports A360 (1996)

Abstract. We consider certain speed estimates for Krylov subspace methods (such as GMRES) when applied upon systems consisting of a compact operator K with small unstructured perturbation B. Information about the decay of singular values of K is also assumed. Our main result is that the Krylov method will perform initially at superlinear speed when applied upon such preconditioned system. However, with large iteration numbers the effect of B is seen to be dominant. As a byproduct, we present several upper speed estimates with explicit constant that should be useful for numerical purposes.

AMS subject classification. 65F10, 65F40, 47A55, 47A30

Key words. Krylov subspace methods, preconditioning, compact operator, Schatten class, speed estimate, superlinear, linear

1 Introduction

Given a bounded linear operator $L \in \mathcal{L}(H)$, what kind of computational problem is the inversion of equation

$$x = Lx + g \tag{1.1}$$

by iterative means? How fast can it be done by Krylov subspace methods? How to make good speed estimates for the iterations? The answers to these (and related) questions give us much information about L.

In particular the properties for iteration of a compact operator are by now rather well known, see [**1**]. In the theory of preconditioning and iterative inversion of Toeplitz operators we meet another kind of problem, which is related to the case of the compact operator — but not quite, see [**2**]. We get linear operators consisting of a large compact part K and a small non-compact perturbation B to invert. In the Toeplitz context we have a complicated interaction between K and B that we cannot lay our hands on. However, we may always generalize and look at K with a small unstructured perturbation B, whose norm alone is known. The study of such systems is the subject of this paper. Another aspect on the properties for iteration of $K + B$ can be found in [**3**].

We shall see that an abstract Krylov subspace method will start with a superlinear convergence rate corresponding to the structure of the compact part K. When the compact part has been depleted, solver will "attack" the unstructured small perturbation, and convergence rate is now linear. In applications the small perturbation part could be a preconditioning error (see [**2**]) or a linear operator arising from noise. In both cases, the compact part is what matters and iteration should be stopped once it enters the linear phase - at least if the small perturbation is noise. So we may say, that the properties for iteration of K are in a sense preserved under small perturbations.

The notion of determinant is quite crucial in our approach. In our language, the determinant of a compact operator K is an entire function $\phi_K(\lambda)$, which has zeroes deep enough to remove the singularities of $(\lambda - K)^{-1}$ other that the origin. In the constructions of determinants we shall need convergence estimates for the singular values of K. The additional requirement that K lies in some Schatten class $S_p(H)$, as defined below, can serve as such.

Definition 1.1. Let H be Hilbert, $p \in (0, \infty)$ and σ_j be the singular values of $K \in \mathcal{LC}(H)$.

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

$$||K||_{S_p} := (\sum_{j=0}^{\infty} |\sigma_j|^p)^{\frac{1}{p}} \tag{1.2}$$

for each $K \in \mathcal{LC}(H)$.

(ii) By $S_p(H)$ denote the set of such $K \in \mathcal{LC}(H)$ that $||K||_{S_p} < \infty$. We call the space $S_p(H)$ the Schatten p-class.

At the end, in Lemma 6.5, we have analytic (but no longer entire) functions that we wish to call generalized determinants for the non-compact operators $K + B$, where $K \in S_p(H)$ and B is small.

The general organization of this paper is as follows: In Section 2 we give the basic definitions of the theory of polynomial acceleration of iterations, following the guidelines of [**1**]. In Section 3 we define a determinant for $K \in S_p(H), p \in (0, 1]$ and use it prove a so called Carleman inequality (Theorem 3.8) for such K. Section 4 is devoted to the study of determinants for $K \in S_p(H), p \in (1, \infty)$. The results of Section 4 are used in Section 5 to prove the Carleman inequality for $K \in S_p(H), p \in (1, \infty)$ (Theorem 5.4).

These Carleman inequalities (giving upper bounds for the resolvents of certain compact operators) are by now classical (see [**4**]), but here we give precise proofs with explicit, carefully analysed constants. These results are very useful for numerical purposes, and to our knowledge not previously published.

Finally in Section 6 we apply the results of all previous sections in order to produce our main results (Theorems 6.7 and 6.9) that show us what an iterative solver can do on a compact operator with small perturbation. As a byproduct we obtain a "generalization" of Carleman inequality for operators of form $K+B$ with $K \in S_p(H)$ and B of small norm (formula 6.12).

The idea of studying this problem, and here presented approaches to solve is originated from the author's Master's Thesis about certain properties for iteration of Toeplitz operators, and are to our knowledge new. The construction of the polynomial sequence in Theorem 6.7 has been separately published in [**3**].

2 Polynomial acceleration of iterations

By X denote a Banach space. Let $L \in \mathcal{L}[X]$ such that $1 \notin \sigma(L)$. Assume that the following fixed point problem is to be solved:

$$x = Lx + g \tag{2.1}$$

Given x_0, define the sequence $\{x_j\}_{j=0}^{\infty}$ of elements of X by the recursion:

$$x_{j+1} := Lx_j + g \tag{2.2}$$

It is clear that if $\{x_j\}$ converges in the topology of X, then the limit x is the solution on (2.1). This is how we have constructed a rudimentary iterative solver for equation (2.1), called the method of successive approximations. Given a sequence of iterates $\{x_j\}$, the convergence of iteration (2.2) (and any other iteration that we shall later study) can be studied in terms of the following two sequences, which we shall call residuals and errors, respectively.

$$d_j := Lx_j - x_j + g \tag{2.3}$$

$$e_j := x - x_j \tag{2.4}$$

After n steps of iteration (2.2), we have a $n+1$ elements $\{x_j\}_{j=0}^{n}$ of X in hand. Also we know n residuals $\{d_j\}_{j=0}^{n-1}$, but the error sequence is not known $\{e_j\}_{j=0}^{n-1}$, because we simply do not have x. What we have is some information about the problem (2.1) — the Krylov data. One might justifiably ask three questions:

(i) Assuming that we want to calculate Krylov data in n steps, is it not true that we can get a better approximation to the solution of (2.1) than the last in the sequence $\{x_j\}$ of the iterates of the simplest iterative solver of all?

(ii) If the above question can be answered in positive, how do we find a better approximation than x_n?

(iii) Assuming that we could use the Krylov data of the problem "optimally", how "good" an approximation to the solution x of (2.1) could we get?

In many cases, the first question can be answered positively, and further study would reveal details that at least partially answer the latter two questions.

We might wish to search for the solution of form:

$$x_n = x_0 + \sum_{j=0}^{n-1} \gamma_{jn} L^j \, d_0 \tag{2.5}$$

for some coefficients γ_{jn}. For notational convenience, define two polynomials p_n and q_{n-1} as follows:

$$q_{n-1}(\lambda) := \sum_{j=0}^{n-1} \gamma_{jn} \lambda^j \tag{2.6}$$

and

$$p_n(\lambda) := 1 - (1-\lambda) q_{n-1}(\lambda) \tag{2.7}$$

Note that p_n satisfies the normalization condition $p_n(1) = 1$. Each polynomial p_n normalized in this way corresponds to some (not necessarily very efficient) method of using Krylov data to produce an approximate solution to (2.1). Thus the polynomials normalized in this way will be important, and so we shall give the following definition:

Definition 2.1. By P_n denote the set of polynomials of degree $\le n$, with complex coefficient, satisfying $(\forall p_n \in P_n) : p_n(1) = 1$.

In the language of the polynomials q_{n-1}, p_n, formula (2.5) takes now form:

$$x_n = x_0 + q_{n-1}(L) \, d_0 \tag{2.8}$$

Furthermore, the following proposition in [**1, Prop. 1.4.1**] will show how the residuals, iterates and errors (as defined in (2.3) and (2.4)) relate to polynomials q_{n-1}, p_n:

Proposition 2.2. *Define the Krylov subspace method for the solution of* (2.1) *method by* (2.8)*, with polynomials* q_{n-1}*,* p_n *satisfying* (2.7)*. By* $x_0 \in X$ *denote the initial vector of the iteration. Then the residuals, iterates and errors satisfy:*

$$d_n = p_n(L)d_0 \tag{2.9}$$

$$x_n = p_n(L)x_0 + q_{n-1}(L)g \tag{2.10}$$

$$e_n = p_n(L)e_0 \tag{2.11}$$

Proof. By definitions in formulae (2.3),(2.8) and (2.7) we get:

$$\begin{aligned} d_n &:= -(I-L)x_n + g = -(I-L)(x_0 + q_{n-1}(L)\,d_0) + g \\ &= [I-(I-L)q_{n-1}(L)]\,d_0 + [g-(I-L)x_0 - d_0] = p_n(L)\,d_0 \end{aligned} \tag{2.12}$$

and (2.9) is proved. To attack (2.10), we calculate by using (2.8), (2.3) and (2.7):

$$\begin{aligned} x_n &:= x_0 + q_{n-1}(L)\,d_0 = x_0 + q_{n-1}(L)\,[(L-I)x_0 + g] \\ &= [I+(L-I)q_{n-1}(L)]\,x_0 + q_{n-1}(L)\,g = p_n(L)x_0 + q_{n-1}(L)g \end{aligned} \tag{2.13}$$

Finally we use (2.4), (2.8) and (2.7):

$$\begin{aligned} e_n &:= x - x_n = [(I-L)^{-1} - q_{n-1}(L)]g - p_n(L)x_0 \\ &= (I-L)^{-1}[I-(I-L)q_{n-1}(L)]g - p_n(L)x_0 \\ &= (I-L)^{-1}p_n(L)g - p_n(L)x_0 \\ &= p_n(L)[(I-L)^{-1}g - x_0] = p_n(L)(x-x_0) = p_n(L)e_0 \end{aligned} \tag{2.14}$$

Formula (2.11) now follows. □

It is also true that the polynomials q_{n-1}, p_n related by (2.7) transform the original problem (2.1) into a problem with the same solution. The following proposition is from [**1, Prop. 1.4.2**]:

Proposition 2.3. *Let* q_{n-1} *be an arbitrary polynomial and set* $p_n(\lambda) := 1-(1-\lambda)q_{n-1}(\lambda)$*. If* x *is the solution of*

$$x = Lx + g \tag{2.15}$$

then it also solves the problem

$$x = p_n(L)x + q_{n-1}(L)g \tag{2.16}$$

Conversely, if additionally $N(q_{n-1}(L)) = 0$*, then* (2.15) *follows from* (2.16) *(*N *denotes the kernel of a linear operator).*

Proof. See [**1, p. 4**]

Proposition 2.3, in essence, can be regarded as the core of the theory of polynomial acceleration in the inversion of linear operators. By choosing polynomial

$p_n \in P_n$ in a clever way, we transform equation (2.15) into the equivalent equation (2.16) with more pleasant invertibility properties. It is of course desirable to choose $p_n(L)$ to be as small in norm as possible. Then, by (2.16), the solution x of (2.15) is about $q_{n-1}(L)g$, provided that $||p_n(L)x|| \ll ||x||$. If this is the case, then also $q_{n-1}(L)$ is close to $(I-L)^{-1}$, by (2.1).

So it is our task to search for polynomials p_n such that $p_n(L)$ is small. In practice we usually do not calculate with operators in order to minimize the operator norm $||p_n(L)||$. This follows from the fact that typically operators on large spaces are clumsy, if not quite impossible objects to handle. Even if that was possible, operator powers for polynomials might be expensive to calculate, or the calculation and minimization of the operator norms (that we need when finding the optimal use for the Krylov information) might be an undesirable task. Because of these reasons it is customary not to invert operators themselves, but only equations (2.1) with given fixed g. Proposition 2.2 gives us the hint to minimize the errors or residuals, instead. However, by formula (2.11), the minimization of errors is not feasible, because the initial error $e_0 := x - x_0$ is never known. However, the initial residual $d_0 := (L-I)x_0 + g$ is known, because it is determined by our initial guess x_0 that does not contain information. This is exactly what the GMRES (generalized minimal residual) algorithm does, giving of course better performance with better initial guess x_0.

3 The Carleman inequality for $p \in (0,1]$

Let H be a separable Hilbert space and $K \in \mathcal{LC}(H)$. Certain conclusions on the size of the resolvent can be made, providing we have some knowledge on the distribution of the singular values $\sigma_j(K)$ of K. The Schatten norm of Definition 1.1[1] is this kind of information — the corresponding inequalities are called Carleman inequalities. The natural starting point is a result concerning operators on a finite dimensional Hilbert space, i.e. matrices. However, before that let us define the determinant for operators in $S_p(H)$, $p \in (0,1]$.

Definition 3.1. Let $p \in (0,1]$ and $K \in S_p(H)$. Let $\lambda \in \mathbf{C} \setminus \{0\}$ be arbitrary. Then define the determinant of $(I - \frac{K}{\lambda})$ by:

$$\det(I - \frac{K}{\lambda}) := \prod_{j=1}^{\infty} (1 - \frac{\lambda_j(K)}{\lambda}) \tag{3.1}$$

where $\lambda_j(K)$ is the sequence of eigenvalues for K.

It is well known that we must require some speed of decay from the eigenvalues $\lambda_j(K)$ in order to get the infinite product in (3.1) converge. One way to deal with this is require that K lies in some Schatten class $S_p(H)$ for $p \in (0,1]$. Consider the following definition and lemma:

[1] Actually we may speak about **norm** only if $p \geq 1$. However, we shall not let this disturb ourselves. To make the precise distinction here would only be a trick of technical nature.

Definition 3.2. Let $K \in \mathcal{LC}(H)$ with eigenvalues $\lambda_j(K)$. Let $p \in (0,\infty)$. Define:

$$\Lambda_p(K) := \sum_{j=1}^{\infty} |\lambda_j(K)|^p \tag{3.2}$$

There is relation between the Schatten norm $||K||_{S_p}$ of a given compact operator K and $\Lambda_p(K)$, which encodes the decay of the eigenvalues of K. The following classical result is by H. Weyl:

Lemma 3.3. *Let $p \in (0,\infty)$ and $K \in S_p(H)$. Then the series $\sum_{j=1}^{\infty} |\lambda_j(K)|^p$ converges absolutely and*

$$\Lambda_p(K) \leq ||K||_{S_p}^p \tag{3.3}$$

Proof. This is [**4, Corollary 1, p. 1093**].

It can be proved by standard argument that the product (3.1) defining the determinant converges uniformly on compact subsets of $\mathbf{C} \setminus \{0\}$ with respect to $\lambda \in \mathbf{C}$ if $K \in S_p(H)$. In particular this implies that $\det(I - \frac{K}{\lambda})$ is a holomorphic function in $\mathbf{C} \setminus \{0\}$.

Proposition 3.4. *For all $x \in \mathbf{R}_+$ and $p \in (0,1]$):*

$$1 + x \leq e^{\frac{x^p}{p}} \tag{3.4}$$

Proof. Write $f(t) := e^{\frac{t}{p}} - 1 - t^{\frac{1}{p}}$. For $t > 0$,

$$f'(t) = \frac{1}{p}(e^{\frac{t}{p}} - t^{\frac{1-p}{p}}) \tag{3.5}$$

Quite easily $f'(t) > 0$ if $t > 0$, and consequently f is an increasing function if $t \in [0,\infty)$. Clearly $f(0) = 0$. It follows that f is a positive function in $[0,\infty)$ which is equivalent with the inequality:

$$e^{\frac{t}{p}} \geq 1 + t^{\frac{1}{p}} \tag{3.6}$$

By setting $t = x^p$ formula (3.4) follows. $\square$

The following lemma is the first in the series of Carleman inequalities. It is a finite dimensional statement that will be generalized later in Theorem 3.8.

Lemma 3.5. *Let $0 < p \leq 1$ and $K \in \mathcal{L}(E_d)$, where E_d is a d-dimensional Hilbert space. Then for $\forall \lambda \notin \sigma(K)$:*

$$|\det(I - \frac{K}{\lambda})|\, ||(I - \frac{K}{\lambda})^{-1}|| \leq e^{\frac{||K||_{S_p}^p}{p|\lambda|^p}} \tag{3.7}$$

Proof. Let $S \in E_d$ be invertible. Let us first show that:

$$||\det(S)S^{-1}|| \leq \prod_{j=1}^{d-1} \sigma_j(S) \tag{3.8}$$

where $\sigma_j(S)$ are the singular values of S in non-increasing order.

We may diagonalize S by unitary matrices U,U' to get $S = UMU'$, where M is diagonal and contains the singular values of S in non-increasing order. Now because the determinant of a unitary matrix has modulus 1:

$$\begin{aligned} &||det(S)S^{-1}|| \\ \leq &|\det(U)|\,|\det(M)|\,|\det(U')|\,||U'^{-1}||\,||M^{-1}||\,||U^{-1}|| \\ \leq &|\det(M)|\,||M^{-1}|| \end{aligned} \tag{3.9}$$

Clearly it suffices to show (3.8) for diagonal invertible matrices M only. Because $\det(M)$ is the product of its singular values, $\det(M)M^{-1}$ is a diagonal matrix with nth diagonal element equal to $\prod_{j=1;j\neq n}^{d} \sigma_j(M)$. By the fact that the operator norm on finite dimensional Hilbert space equals the largest singular value:

$$|| \det(M)M^{-1}|| = \max_{n\in\{1,2\dots,d\}} \prod_{j=1;j\neq n}^{d} \sigma_j(M) = \prod_{j=1}^{d-1} \sigma_j(M) \tag{3.10}$$

and (3.8) has been proved for invertible diagonal matrices, and hence for all invertible matrices. To prove the claim of the lemma, set $S := I - \frac{K}{\lambda}$, $\lambda \notin \sigma(K)$. We get:

$$\begin{aligned} &|\det(I-\frac{K}{\lambda})|\,||(I-\frac{K}{\lambda})^{-1}|| \leq \prod_{j=1}^{d-1} \sigma_j(I-\frac{K}{\lambda}) \\ &\leq \prod_{j=1}^{d-1}(I-\frac{\sigma_j(K)}{\lambda}) \leq \prod_{j=1}^{d-1}(I+\frac{\sigma_j(K)}{|\lambda|}) \\ &\leq \prod_{j=1}^{d-1} e^{\frac{\sigma_j(K)^p}{p|\lambda|^p}} \leq e^{\frac{\sum_{j=1}^{d-1}\sigma_j(K)^p}{p|\lambda|^p}} \leq e^{\frac{||K||_p^p}{p|\lambda|^p}} \end{aligned} \tag{3.11}$$

where we have used Proposition 3.4. The lemma has been proved. □

It is our intention to generalize the previous lemma to any separable Hilbert space H. The following lemma and proposition are tools that we shall need.

LEMMA 3.6. *Let $p \in (0,1]$ and $K \in S_p(H)$, where H is a separable Hilbert space. Then:*

$$|\det(I-\frac{K}{\lambda})| \leq e^{\frac{||K||_{S_p}^p}{p|\lambda|^p}} \tag{3.12}$$

PROOF. This is a direct consequence of Lemma 3.3 and Proposition 3.4 □

Proposition 3.7. *Let X be a Banach space and assume $W_n, W \in \mathcal{L}(X)$ such that*

(i) $(\forall n \leq N): ||W_n|| \leq C_n$
(ii) $(\forall x \in X): \lim_{n\to\infty} W_n x \to Wx$
(i.e. $W_n \to W$ in strong operator topology)
(iii) $\lim_{n\to\infty} C_n = C$

Then $||W|| \leq C$.

Proof. For contradiction, assume that there is $x_0 \in X$; $||x_0|| = 1$ such that $||Wx_0|| > C + \epsilon$ for some $\epsilon > 0$. Then by (ii) $(\exists N_1 \in \mathbf{N})$:

$$(\forall n > N_1): ||W_n x_0|| > C + \frac{\epsilon}{2} \tag{3.13}$$

Furthermore, by (iii) $(\exists N_2 \in \mathbf{N})$:

$$(\forall n > N_2): C_n < C + \frac{\epsilon}{2} \tag{3.14}$$

Now, if $n > \max(N_1, N_2)$ we have:

$$||W_n x_0|| > C_n \tag{3.15}$$

The contradiction against (i) proves the proposition. □.

Now we are ready to state and prove the Carleman inequality for Schatten classes $S_p(H)$, $p \in (0,1]$. For parallel expositions, see [**1**] Section 5.6 and [**4**].

Theorem 3.8. *Let $p \in (0,1]$ and $K \in S_p(H)$. For each $\lambda \notin \sigma(K) \cup \{0\}$ we have:*

$$| \, det(I - \frac{K}{\lambda})| \, ||(I - \frac{K}{\lambda})^{-1}|| \leq e^{\frac{||K||^p_{S_p}}{p|\lambda|^p}} \tag{3.16}$$

Proof. Fix $\lambda \notin \sigma(K) \cup \{0\}$. Pick a sequence $S_p(H) \ni K_n \to K$ in $S_p(H)$ such that $\dim K_n(H) = n$. Define the Hilbert subspace H_n of H by:

$$H_n := K_n(H) + K_n^*(H) \tag{3.17}$$

Clearly $\dim H_n \leq 2n$. By P_n denote the orthogonal projector onto H_n. One easily checks that $K_n = P_n K_n P_n$ and

$$(I - \frac{K_n}{\lambda})^{-1} = P_n(I - \frac{K_n}{\lambda})^{-1} P_n + (I - P_n) \tag{3.18}$$

because H_n is the smallest reducing subspace of K_n, outside of which K_n vanishes. So we may regard K_n as a matrix of dimension at most $2n \times 2n$, and use

Lemma 3.5. By Lemmas 3.5, 3.6 and the definition of operator norm we have:

$$(3.19)\ |\,det\,(I - \frac{K_n}{\lambda})|\,||(I - \frac{K_n}{\lambda})^{-1}||$$
$$= max\{|\,det\,(I - \frac{K_n}{\lambda})|\,||P_n(I - \frac{K_n}{\lambda})^{-1}P_n||, ||\,det\,(I - \frac{K_n}{\lambda})(I - P_n)||\}$$
$$= max\{|\,det\,(I - \frac{K_n}{\lambda})|\,||P_n(I - \frac{K_n}{\lambda})^{-1}P_n||, |\,det\,(I - \frac{K_n}{\lambda})|\}$$
$$\le max\{e^{\frac{||K_n||^p_{S_p}}{p|\lambda|^p}}, e^{\frac{||K_n||^p_{S_p}}{p|\lambda|^p}}\} = e^{\frac{||K_n||^p_{S_p}}{p|\lambda|^p}}$$

Denote:

$$W_n(\lambda) := P_n(I - \frac{K_n}{\lambda})^{-1}P_n + (I - P_n) \tag{3.20}$$

As $n \to \infty$,

$$W_n(\lambda) \to W(\lambda) := det\,(I - \frac{K}{\lambda})(I - \frac{K}{\lambda})^{-1} \tag{3.21}$$

in strong operator topology for each fixed $\lambda \notin \sigma(K) \cup \{0\}$. Denote:

$$C_n := e^{\frac{||K_n||^p_{S_p}}{p|\lambda|^p}} \tag{3.22}$$

Because $K_n \to K$ in $S_p(H)$, we have:

$$\lim_{n\to\infty} C_n = C := e^{\frac{||K||^p_{S_p}}{p|\lambda|^p}} \tag{3.23}$$

Now the hypotheses of Proposition 3.7 have been fulfilled, and it follows:

$$||W|| = |\det(I - \frac{K}{\lambda})|\,||(I - \frac{K}{\lambda})^{-1}|| \le e^{\frac{||K||^p_{S_p}}{p|\lambda|^p}} = C \tag{3.24}$$

This proves the theorem. □.

Remark 3.9. Note that in Theorem 3.8 the Carleman inequality has been proved only for $\lambda \notin \sigma(K) \cup \{0\}$. The spectrum on K, however, consists of discrete points (with the exception of the origin), and by continuity we may expand the result to apply for all $\lambda \neq 0$.

4 Generalized determinants

We have stated in Section 3 that the requirement $K \in S_p(H)$ for $p \in (0, 1]$ enforces sufficient amount of decay on the eigenvalues so that product (3.1) converges. On the contrary, it is true that a weaker requirement $K \in S_p(H)$ for $p \in (1, \infty)$ alone does not guarantee the convergence of the classical determinant.

So it must be accepted that the classical determinant of Definition 3.1 simply will not do for spaces $S_p(H)$, $p > 1$. However, the concept of an analytic function whose zeroes kill the singularities in the resolvent of K seems very appealing to us. Fortunately, it is possible to construct generalized determinants that satisfy our requirements. The building blocks for these new determinants are the Weierstrass elementary factors whose definitions and certain growth estimates are given in the following:

DEFINITION 4.1. Let $m \in \mathbf{N}$. The Weierstrass elementary factor is the entire function defined by:

$$E_m(z) := (1-z)\, e^{\sum_{j=1}^{m} \frac{z^j}{j}} \tag{4.1}$$

The study convergence questions of the above products are to be found in almost any book of basic function theory, see for example [**5, Theorem 15.9**].

The following propositions will be used in the proof of Lemma

PROPOSITION 4.2. *For any $m \in \mathbf{N}$ we have:*

$$\left|\frac{\sin m\theta}{\sin\theta}\right| \le m \tag{4.2}$$

PROOF. We shall give an induction proof. Clearly (4.2) holds for $m = 1$. Assume that it holds for $m \in \mathbf{N}$. Then:

$$\begin{aligned} \left|\frac{\sin(m+1)\theta}{\sin\theta}\right| &= \left|\frac{\sin m\theta\,\cos\theta + \cos m\theta\,\sin\theta}{\sin\theta}\right| \\ &\le \left|\frac{\sin m\theta\,\cos\theta}{\sin\theta}\right| + |\cos m\theta| \le \left|\frac{\sin m\theta}{\sin\theta}\right| + 1 \le m+1 \end{aligned} \tag{4.3}$$

This proves the claim. □

Certain radial size estimates for the Weierstrass elementary factors will be needed in the study on the generalized determinants. The following lemma is a result of O. Blumenthal and E. Lindelöf, see [**6, p. 131 - 147**].

LEMMA 4.3. *Let $m \in \mathbf{N}$ and $\alpha \in [0,1]$. Then:*

$$|E_m(z)| \le e^{C_{m,\alpha}\,|z|^{m+\alpha}} \tag{4.4}$$

where

$$C_{m,\alpha} \le \frac{m^{\alpha}(m+1)^{1-\alpha}}{m+\alpha} \le 2 \tag{4.5}$$

PROOF. The modulus of the Weierstrass elementary factor can be written in form:

$$|E_m(r\,e^{i\theta})| = e^{\log|1-r\,e^{i\theta}| + \sum_{j=1}^{m} \frac{r^j}{j}\,\cos j\theta} \tag{4.6}$$

So our task is clearly equivalent with finding an upper bound for the expression:

$$G_{m,\alpha}(r,\theta) := \frac{\log|1-r\,e^{i\theta}| + \sum_{j=1}^{m} \frac{r^j}{j}\cos j\theta}{r^{m+\alpha}} = \frac{\frac{1}{2}\log(1-2r\,\cos\theta+r^2) + \sum_{j=1}^{m} \frac{r^j}{j}\cos j\theta}{r^{m+\alpha}} \tag{4.7}$$

In order to maximize this function $G_{m,\alpha}$, we calculate the partial derivatives. By a straightforward calculation we get the following identities:

$$\frac{\partial G_{m,\alpha}}{\partial r}(r,\theta) = -\frac{m+\alpha}{r}\,G_{m,\alpha}(r,\theta) + \frac{1}{r^\alpha}\,\frac{r\cos m\theta - \cos(m+1)\theta}{1-2r\,\cos\theta+r^2} \tag{4.8}$$

and

$$\frac{\partial G_{m,\alpha}}{\partial r}(r,\theta) = -r^{1-\alpha}\frac{r\sin m\theta + \sin(m+1)\theta}{1-2r\,\cos\theta+r^2} \tag{4.9}$$

where we have used the helpful sum formulae:

$$\sum_{j=1}^{m} r^{j-1}\cos j\theta = \frac{r^{m+1}\sin m\theta - r^p\cos(m+1)\theta - r + \cos\theta}{1-2r\,\cos\theta+r^2} \tag{4.10}$$

and

$$\sum_{j=1}^{m} r^j \sin j\theta = \frac{r^{m+2}\sin m\theta - r^{m+1}\sin(m+1)\theta + s\sin\theta}{1-2r\,\cos\theta+r^2} \tag{4.11}$$

Let (r_0,θ_0) satisfy $\frac{\partial G_{m,\alpha}}{\partial r}(r_0,\theta_0) = \frac{\partial G_{m,\alpha}}{\partial r}(r_0,\theta_0) = 0$. From (4.8) we get:

$$G_{max} = \frac{r_0^{1-\alpha}}{p+\alpha}\,\frac{r_0\cos m\theta_0 - \cos(m+1)\theta_0}{1-2r_0\,\cos\theta_0+r_0^2} \tag{4.12}$$

where $G_{max} := G_{m,\alpha}(r_0,\theta_0)$. Formula (4.9) implies

$$r_0 = \frac{\sin(m+1)\theta_0}{\sin(m)\theta_0} \tag{4.13}$$

(or $r_0 = 0$, which is not a valid maximum). By inserting this into (4.12) we find:

$$\begin{aligned} G_{max} &= \frac{r_0^{1-\alpha}}{m+\alpha}\,\frac{\sin(m+1)\theta_0\,\cos m\theta_0 - \cos(m+1)\theta_0\,\sin m\theta_0}{\sin m\theta_0(1-2r_0\,\cos\theta_0+r_0^2)} \\ &= \frac{1}{m+\alpha}\,\frac{\sin\theta_0\,\cos^{1-\alpha}(m+1)\theta_0}{\sin^{2-\alpha} m\theta_0(1-2r_0\,\cos\theta_0+r_0^2)} \\ &= \frac{1}{m+\alpha}\,\frac{\sin^\alpha m\theta_0\,\sin^{1-\alpha}(m+1)\theta_0}{\sin\theta_0} \end{aligned} \tag{4.14}$$

where the last equality follows from:

$$1-2r_0\,\cos\theta_0+r_0^2 = \frac{\sin^2\theta_0}{\sin^2 m\theta_0} \tag{4.15}$$

Now we have immediately

$$G_{max} = \frac{1}{m+\alpha}\,\big(\frac{\sin m\theta_0}{\sin\theta_0}\big)^{\alpha}\big(\frac{\sin (m+1)\theta_0}{\sin\theta_0}\big)^{1-\alpha} \le \frac{1}{m+\alpha}\, m^{\alpha}(m+1)^{1-\alpha} \tag{4.16}$$

where the last upper estimate is by Proposition 4.2. The upper bound 2 in formula (4.5) is easy to show. This proves the lemma. □

The following corollary is nothing but the result of the previous Lemma in a form that will be useful for our purposes.

Corollary 4.4. *Let $m \in \mathbf{N}$ and $p-1 \le m \le p$. Then*

$$|E_m(z)| \le e^{c_{m,p}|z|^p} \tag{4.17}$$

where

$$c_{m,p} := \frac{m^{p-m}(m+1)^{1-p+m}}{p} \le 2 \tag{4.18}$$

Proof. Trivial. □

Now we know enough about the nature of things to define the generalized determinants det_m for each $m \in \mathbf{N}$:

Definition 4.5. Let $m \in \mathbf{N}$ and $p-1 \le m < p$. Let $K \in S_p(H)$ and by $\lambda_j(K)$ denote the jth eigenvalue of K, ordered in the non-increasing order of absolute values, with multiplicities. Let $\lambda \neq 0$. Then

$$det_m\,(I - \frac{K}{\lambda}) := \prod_{j=1}^{\infty} E_m(\frac{\lambda_j(K)}{\lambda}) \tag{4.19}$$

is called the generalized determinant of order m.

The product defining the generalized determinant of order m converges uniformly on compact subsets of $\mathbf{C}\setminus\{0\}$. It follows that the limit is a holomorphic function in $\mathbf{C}\setminus\{0\}$.

It should not arrive as a surprise that the growth estimates on the Weierstrass elementary factors will imply corresponding growth estimates on the generalized determinants. The following lemma will state the immediate consequences of Corollary 4.4:

Lemma 4.6. *Let $m \in \mathbf{N}$ and $p-1 \le m < p$. Let $K \in S_p(H)$. Then*

$$|\,det_m\,(I - \frac{K}{\lambda})| \le e^{c_{m,p}\frac{||K||^p_{S_p}}{|\lambda|^p}} \tag{4.20}$$

where

$$c_{m,p} := \frac{m^{p-m}(m+1)^{1-p+m}}{p} \le 2 \tag{4.21}$$

PROOF. Use Corollary 4.4 as follows:

$$|\, det_m\,(I - \frac{K}{\lambda})| = \prod_{j=1}^{\infty} |E_m(\frac{\lambda_j(K)}{\lambda})| \le e^{\sum_{j=1}^{\infty} c_{m,p} |\frac{\lambda_j(K)}{\lambda}|^p} \tag{4.22}$$

$$= e^{c_{m,p} \frac{\Lambda_p(K)}{|\lambda|^p}} \le e^{c_{m,p} \frac{||K||_{S_p}^p}{|\lambda|^p}}$$

where the final conclusion is by Lemma 3.3. □

5 The Carleman inequality for $p \in (1, \infty)$

In this section we shall use the tools of Section 4 to prove Theorem 5.4; the Carleman inequality for compact operators in $S_p(H)$ for $p \in (1,\infty)$.

LEMMA 5.1. *Let $m \in \mathbf{N}$ and $K, B \in S_m(H)$. Let $\lambda \notin \sigma(K) \cup \{0\}$. Then the following determinant is analytic and satisfies:*

$$\frac{d}{dz}\, det_m\,(I - \frac{K}{\lambda} + zB)|_{z=0} \tag{5.1}$$

$$= det_m\,(I - \frac{K}{\lambda})\, Tr\,[\{(I - \frac{K}{\lambda})^{-1} - 1 - \ldots - (\frac{K}{\lambda})^{m-2}\}B]$$

PROOF. See [**4, Lemma 23, p.1110**]

LEMMA 5.2. *Let $m \in \mathbf{N}$ and $p - 1 \le m < p$. Then*

$$||\, det_m\,(I - \frac{K}{\lambda})\{(I - \frac{K}{\lambda})^{-1} - 1 - \ldots - (\frac{K}{\lambda})^{m-2}\}||_{S_q} \le e^{c_{m,p} \frac{||K||_{S_p}^p}{|\lambda|^p}} \tag{5.2}$$

where $c_{m,p}$ satisfies (4.21) *and $\frac{1}{p} + \frac{1}{q} = 1$.*

PROOF. Let $B \in S_p(H)$,$||B||_{S_p} = 1$ and $\frac{1}{p} + \frac{1}{q} = 1$. By Lemma 5.1 and the Residue Theorem:

$$det_m\,(I - \frac{K}{\lambda})\, Tr[\{(I - \frac{K}{\lambda})^{-1} - 1 - \ldots - (\frac{K}{\lambda})^{m-2}\}B] \tag{5.3}$$

$$= \quad -\frac{1}{2\pi i} \int_{|\xi|=\nu} \xi^{-2}\, det_m\,(I - \frac{K}{\lambda} + \xi B)\, d\xi$$

for all $\nu > 0$. By Lemma 4.6 (set $\lambda = -1$):

$$|\, det_m\,(I + S)| \le e^{c_{m,p}\, ||S||_{S_p}^p} \tag{5.4}$$

for any $S \in S_p$. Now take $S := -\frac{K}{\lambda} + \xi B$ and apply this upon (5.3):

$$|\, det_m\,(I - \frac{K}{\lambda})\; Tr\,[\{(I - \frac{K}{\lambda})^{-1} - 1 - \ldots - (\frac{K}{\lambda})^{m-2}\}B]| \tag{5.5}$$

$$\le e^{c_{m,p}\, \max_{|\xi|=\nu} ||-\frac{K}{\lambda} + \xi B||_{S_p}^p} \le e^{c_{m,p}\, [\frac{||K||_{S_p}}{|\lambda|} + \nu\, ||B||_{S_p}]^p}$$

Now let $\nu \to 0$. Then the previous formula gives:

$$(5.6) \quad |\, det_m\, (I - \frac{K}{\lambda})\ Tr\, [\{(I - \frac{K}{\lambda})^{-1} - 1 - \ldots - (\frac{K}{\lambda})^{m-2}\}B]| \le e^{c_{m,p}\, \frac{||K||^p_{S_p}}{|\lambda|^p}}$$

By the inverse Hölder inequality ([**4, Lemma 14, p. 1098**]) we have:

$$(5.7) \qquad ||A||_{S_q} = \sup_{||B||_{S_p} \le 1} |Tr(AB)|$$

This and (5.6) give:

$$(5.8) \quad ||\, det_m\, (I - \frac{K}{\lambda})\{(I - \frac{K}{\lambda})^{-1} - 1 - \ldots - (\frac{K}{\lambda})^{m-2}\}B||_{S_q} \le e^{c_{m,p}\, \frac{||K||^p_{S_p}}{|\lambda|^p}}$$

This proves the lemma. $\square$

One more constant will have to be calculated, before we are able to prove the Carleman inequality.

Proposition 5.3. *For all $x \ge 0$ we have:*

$$(5.9) \qquad x^j \le e^{\frac{j}{pe} x^p}$$

Proof. Trivial. $\square$

Theorem 5.4. *Let $m \in \mathbf{N}$ and $p - 1 \le m < p$. Let $K \in S_p(H)$. Then*

$$(5.10) \qquad |\, det_m\, (I - \frac{K}{\lambda})|\, ||(I - \frac{K}{\lambda})^{-1}|| \le m\, e^{(c_{m,p} + \frac{1}{e})\frac{||K||^p_{S_p}}{|\lambda|^p}}$$

where

$$(5.11) \qquad c_{m,p} = \frac{m^{p-m}(m+1)^{1-p+m}}{p} \le 2$$

Proof. By Lemma 5.2:

$$\begin{aligned}(5.12) \qquad & ||\, det_m\, (I - \frac{K}{\lambda})\{(I - \frac{K}{\lambda})^{-1} - 1 - \ldots - (\frac{K}{\lambda})^{m-2}\}|| \\ & \le ||\, det_m\, (I - \frac{K}{\lambda})\{(I - \frac{K}{\lambda})^{-1} - 1 - \ldots - (\frac{K}{\lambda})^{m-2}\}||_{S_q} \\ & \le e^{c_{m,p}\, \frac{||K||^p_{S_p}}{|\lambda|^p}}\end{aligned}$$

By the triangle inequality, we get immediately:

$$\begin{aligned}(5.13) \qquad & |\, det_m\, (I - \frac{K}{\lambda})|\, ||(I - \frac{K}{\lambda})^{-1}|| \\ & \le ||\, det_m\, (I - \frac{K}{\lambda})\{1 + \ldots + \big(\frac{K}{\lambda}\big)^{m-2}\}|| + e^{c_{m,p}\, \frac{||K||^p_{S_p}}{|\lambda|^p}} \\ & \le |\, det_m\, (I - \frac{K}{\lambda})| \sum_{j=0}^{m-2} \frac{||K^j||}{|\lambda|^j} + e^{c_{m,p}\, \frac{||K||^p_{S_p}}{|\lambda|^p}}\end{aligned}$$

Let us study more closely the sum in the previous expression. We have by Proposition 5.3:

$$\frac{||K^j||}{|\lambda|^j} \leq \frac{||K||^j}{|\lambda|^j} \leq e^{\frac{j}{pe}\frac{||K||^p}{|\lambda|^p}} \leq e^{\frac{j}{pe}\frac{||K||^p_{S_p}}{|\lambda|^p}} \tag{5.14}$$

and moreover:

$$\sum_{j=1}^{m-2} \frac{||K^j||}{|\lambda|^j} \leq (m-1)\, e^{\frac{m-2}{pe}\frac{||K||^p_{S_p}}{|\lambda|^p}} \leq (m-1)\, e^{\frac{1}{e}\frac{||K||^p_{S_p}}{|\lambda|^p}} \tag{5.15}$$

From Lemma 4.6 we remember:

$$|\, det_m\,(I - \frac{K}{\lambda})| \leq e^{c_{m,p}\frac{||K||^p_{S_p}}{|\lambda|^p}} \tag{5.16}$$

To conclude the proof, let us combine (5.13), (5.15) and (5.16) in the following manner:

$$\begin{aligned} &|\, det_m\,(I - \frac{K}{\lambda})|\, ||(I - \frac{K}{\lambda})^{-1}|| \\ &\leq e^{c_{m,p}\frac{||K||^p_{S_p}}{|\lambda|^p}} + e^{c_{m,p}\frac{||K||^p_{S_p}}{|\lambda|^p}}\big((m-1)\, e^{\frac{1}{e}\frac{||K||^p_{S_p}}{|\lambda|^p}}\big) \\ &= e^{c_{m,p}\frac{||K||^p_{S_p}}{|\lambda|^p}}\big(1 + (m-1)\, e^{\frac{1}{e}\frac{||K||^p_{S_p}}{|\lambda|^p}}\big) \end{aligned} \tag{5.17}$$

or a more beautiful upper bound:

$$\leq e^{c_{m,p}\frac{||K||^p_{S_p}}{|\lambda|^p}}\big(m\, e^{\frac{1}{e}\frac{||K||^p_{S_p}}{|\lambda|^j}}\big) = m\, e^{(c_{m,p}+\frac{1}{e})\frac{||K||^p_{S_p}}{|\lambda|^p}}$$

This proves the theorem. □

6 On the polynomial acceleration of a compact operator with a small perturbation

By availing the Carleman inequalities and certain points of vector valued function theory, conclusions on the polynomial acceleration properties of operator $L := K + B$ (as presented in Section 2) can be drawn, where $K \in \mathcal{LC}(H)$ for a separable Hilbert space H, $B \in \mathcal{L}(H)$ is small when compared to K and $1 \notin \sigma(K+B)$. Everything before Theorem 6.7 will be just technical preparations.

Definition 6.1. Let $K \in \mathcal{LC}(H)$ and $B \in \mathcal{L}(H)$. Denote:

$$K_\lambda := (\lambda - B)^{-1} K \tag{6.1}$$

for all $\lambda \notin \sigma(B)$.

Remark 6.2. In this section we are keeping the small perturbation B always the same. That is the reason why we have chosen not to write the dependency of K_λ on B explicitly.

The ideal structure of the set of finite dimensional operators and the fact that the singular values of a given compact operator $K \in \mathcal{LC}(H)$ are equal to the approximation numbers makes it possible to prove that the information about the Schatten class of K is preserved under the mapping $K \mapsto K_\lambda$. The following Proposition makes this point precise.

PROPOSITION 6.3. *Let $p \in (0, \infty)$, $K \in S_p(H)$, $B \in \mathcal{L}(H)$ and $\lambda \notin \sigma(B)$. Then $K_\lambda \in S_p(H)$ and it applies*

$$||K_\lambda||_{S_p} \leq ||(\lambda - B)^{-1}|| \, ||K||_{S_p} \tag{6.2}$$

PROOF. Let $\lambda \notin \sigma(B)$. The core of the proof is in the following calculation:

$$\begin{aligned} \sigma_j(K_\lambda) &= \inf_{rankF \leq n} ||K_\lambda - F|| \\ &= \inf_{rankF \leq j} ||(\lambda - B)^{-1}(K - (\lambda - B)F)|| \\ &\leq ||(\lambda - B)^{-1}|| \inf_{rankF \leq j} ||K - (\lambda - B)F|| \\ &\leq ||(\lambda - B)^{-1}|| \, \sigma_j(K) \end{aligned} \tag{6.3}$$

because $rank\,(\lambda - B)F \leq rank\,F$. Raising to power p and summing up the singular values give formula (6.2). □

LEMMA 6.4. *Let $m \in \mathbf{N}$. Let $K_n \in \mathcal{LC}(H)$ be an operator with n-dimensional range. Then*

$$\phi_m^{(n)}(\lambda) := det_m\,(I - K_{n\,\lambda}) \tag{6.4}$$

is a holomorphic function of $\lambda \in \mathbf{C} \setminus \sigma(B)$.

PROOF. For brevity, denote $D := \mathbf{C} \setminus \sigma(B)$. Because $\dim K_n(H) = n$, we have $\dim (K_n)_\lambda(H) \leq n$ for all $\lambda \in D$. Moreover, there is an r: $1 \leq r \leq n+1$ such that $r := \max_{\lambda \in D} card\,\sigma((K_n)_\lambda)$.

By $E \subset D$ denote the set of such λ's that $card\,\sigma((K_n)_\lambda) < r$. By [**7, Theorem 6.25**] we know that E is a discrete closed set. Pick any $\lambda_0 \in D \setminus E$. Because E is closed, we have a neighborhood $N_1(\lambda_0)$ of λ_0 such that

$$(\forall \lambda \in N_1(\lambda_0)) : \; card\,\sigma((K_n)_\lambda) = r \tag{6.5}$$

It follows that in $N_1(\lambda_0)$ the operator $(K_n)_\lambda$ has exactly $r-1$ non-zero distinct eigenvalues $\{\lambda_j((K_n)_\lambda))\}_{j=1}^n$. Define $\epsilon > 0$ to be the minimum distance between any two eigenvalues in $\{\lambda_j((K_n)_\lambda))\}_{j=1}^n$. By [**7, Theorem 6.4**] there exists a $\delta > 0$ such that

$$|\lambda - \lambda_0| < \delta \implies \{\lambda_j((K_n)_\lambda))\}_{j=1}^n \subset \cup_{j=1}^n B(\lambda_j(K_{n\,\lambda_0}), \frac{\epsilon}{3}) \tag{6.6}$$

Now by [**7, Theorem 6.20**] we see that each mapping $\lambda \mapsto \lambda_j((K_n)_\lambda)$ is holomorphic in $N(\lambda_0) := \{\lambda | \, |\lambda - \lambda_0| < \delta\}$. From the holomorphicity of Weierstrass elementary factors and the definition of det_m we conclude that $\phi_m^{(n)}(\lambda)$ is indeed holomorphic in $N(\lambda_0)$ and consequently in $D \setminus E$.

Because $\phi_m^{(n)}(\lambda)$ is a continuous function on D, the Rado extension theorem ([**5, Theorem 12.14**]) implies that $\phi_m^{(n)}(\lambda)$ is in fact holomorphic in whole D. This proves the lemma. $\square$

In the following lemma we present an analytic function ϕ_m that can be regarded as a generalized determinant of the non-compact operator $K + B$.

Lemma 6.5. *Let* $m \in \mathbf{N}$ *and* $p - 1 \leq m < p$. *Take* $K \in S_p(H)$. *Then* $\phi_m(\lambda) := det_m\,(I - K_\lambda)$ *is a holomorphic function of* $\lambda \in \mathbf{C} \setminus \sigma(B)$.

Proof. Pick a sequence of n-dimensional operators $K_n \to K$ in $S_p(H)$. Then by a continuity argument for the determinant we may show that for all $\lambda \in \mathbf{C} \setminus \sigma(B)$:

$$\phi_m(\lambda) = \lim_{n\to\infty} \phi_m^{(n)}(\lambda) \tag{6.7}$$

If we are able to show that $\{\phi_m^{(n)}\}$ is a normal family of holomorphic functions, then the limit $\phi_m(\lambda)$ itself is holomorphic by [**5, Definition 14.5 and Theorem 10.28**], thus proving the lemma. By Lemma 6.4 we already know that $\{\phi_m^{(n)}\}$ is a family of holomorphic functions. By [**5, Theorem 14.6**] it suffices to show that $\{\phi_m^{(n)}\}$ is uniformly bounded on the compact subsets of $\mathbf{C} \setminus \sigma(B)$.

The following calculation makes the trick:

$$\begin{aligned} |\phi_m^{(n)}(\lambda)| &= |\,det_m\,(I + (\lambda - B)^{-1}K_n)| \\ &= \prod_{j=1}^{\infty} |E_m(-\lambda_j((\lambda - B)^{-1}K_n))| \leq \prod_{j=1}^{\infty} e^{c_{m,p}|\lambda_j((\lambda-B)^{-1}K_n)|^p} \\ &= e^{c_{m,p}\,\Lambda_p((\lambda-B)^{-1}K_n)} \end{aligned} \tag{6.8}$$

where the inequality has been written by Corollary 4.4. We may continue the estimation by Lemma 3.3 as follows:

$$\begin{aligned} e^{c_{m,p}\,\Lambda_p((\lambda-B)^{-1}K_n)} &\leq e^{c_{m,p}\,||(\lambda-B)^{-1}K_n||_{S_p}^p} \\ &\leq e^{c_{m,p}\,||(\lambda-B)^{-1}||\,||K_n||_{S_p}^p} \end{aligned} \tag{6.9}$$

for all $\lambda \notin \sigma(B)$; the last inequality is by Proposition 6.3. Because $K_n \to K$ in $S_p(H)$, it follows that the family $\{K_n\}$ is uniformly bounded in the norm $||.||_{S_p}$. Now the upper bound for each $\phi_m^{(n)}$ given by the right side of (6.9) is uniformly bounded on compact subsets of $\mathbf{C} \setminus \sigma(B)$ because the resolvent of B is. This proves the lemma. $\square$

In the similar way we may proceed to prove also the following analogous lemma to Lemma 6.5:

Lemma 6.6. *Let* $0 < p \le 1$*. Take* $K \in S_p(H)$*. Then* $\phi(\lambda) := det\,(I - K_\lambda)$ *is a holomorphic function of* $\lambda \in \mathbf{C} \setminus \sigma(B)$*.*

Proof. Omitted.

It is time to present a lemma that has something to do with the polynomial acceleration of the operator $K + B$. We shall construct a sequence of non-normalized polynomials $\{\tilde{p}_k\}_{k=0}^{\infty}$ such that these polynomials "see" the operator $K + B$ small with increasing k. The existence of such non-normalized sequence itself is naturally a triviality, but later (in Lemma 6.8) we shall see that almost all of these polynomials in fact are almost normalized in the sense of Definition 2.1; i.e. $\{\tilde{p}_k(1)\}_{k=0}^{\infty}$ remains bounded from some k on. The approach presented here is due to O. Nevanlinna ([[]3]):

Theorem 6.7. *Let* $m \in \mathbf{N}$ *and* $p - 1 \le m < p$*. Take* $K \in S_p(H)$ *and let* $B \in \mathcal{L}(H)$ *be small 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* $\beta \in (0, 1]$*:*

$$||\tilde{p}_k(K + B)||^{\frac{1}{k}} \le m^{\frac{1}{k}}\big(||B|| + ||K||_{S_p}\, k^{-\frac{\beta}{p}}\big)\Big(\frac{||B||\, k^{\frac{\beta}{p}}}{||K||_{S_p}} + 1\Big)^{\frac{1}{k}}\, e^{\frac{c_{m,p}+\frac{1}{e}}{k^{1-\beta}}} \tag{6.10}$$

where $c_{m,p}$ *is defined in* (4.21).

Proof. Let us start with the easily proved fact for $|\lambda| > ||B||$:

$$||(\lambda - (K + B))^{-1}|| \le \frac{1}{|\lambda| - ||B||}||(I - K_\lambda)^{-1}|| \tag{6.11}$$

We also have by Theorem 5.4 and Proposition 6.3:

$$|\phi_m(\lambda)|\,||(I - K_\lambda)^{-1}|| \le me^{(c_{m,p}+\frac{1}{e})||K_\lambda||_{S_p}^p} \le me^{(c_{m,p}+\frac{1}{e})(\frac{||K||_{S_p}^p}{(|\lambda|-||B||)^p}} \tag{6.12}$$

On the other hand we can show that $(I - K_\lambda)^{-1}$ is a holomorphic operator valued function outside $\sigma(K+B)$ by studying the identity $\lambda-(K+B) = (\lambda-B)(I-K_\lambda)$. Also note that $\sigma(K + B) \subset \sigma(B) \cup \sigma_p(K + B)$ — for details see [**1, Theorem 2.2.15**]. Estimate (6.12) together with the fact that ϕ_m is holomorphic outside $\sigma(B)$ (Lemma 6.5) allow us to conclude that $\phi_m(\lambda)\,(I - K_\lambda)^{-1}$ is holomorphic in $\mathbf{C} \setminus \sigma(B)$ and same applies also for:

$$\phi_m(\lambda)\,(\lambda - (K + B))^{-1} \equiv \phi_m(\lambda)\,(I - K_\lambda)^{-1}\,(\lambda - B)^{-1} \tag{6.13}$$

Now the following estimate is a direct consequence of (6.12) and (6.13) for $|\lambda| > ||B||$:

$$|\phi_m(\lambda)|\,||(\lambda - (K + B))^{-1}|| \le \frac{m}{|\lambda| - ||B||}e^{(c_{m,p}+\frac{1}{e})(\frac{||K||_{S_p}^p}{(|\lambda|-||B||)^p}} \tag{6.14}$$

Because ϕ_m is holomorphic in $\mathbf{C}\backslash\sigma(B)$, we may write for $\lambda > ||B||$ the convergent Laurent series:

$$\phi_m(\lambda) = \sum_{j=0}^{\infty} \frac{a_j}{\lambda^j} \tag{6.15}$$

Moreover, if $\lambda > ||K + B||$, we may write the resolvent of $K + B$ in form:

$$(\lambda - (K + B))^{-1} = \sum_{j=0}^{\infty} \frac{(K + B)^j}{\lambda^{j+1}} \tag{6.16}$$

By multiplying the two previous formulae we get:

$$\phi_m(\lambda)\,(\lambda - (K + B))^{-1} = \sum_{k=0}^{\infty} \tilde{p}_k(K + B)\,\lambda^{-k-1} \tag{6.17}$$

where the polynomials $\{\tilde{p}_k\}_{k=0}^{\infty}$ are defined by:

$$\tilde{p}_k(\xi) := \sum_{j=0}^{k} a_{k-j}\xi^j \tag{6.18}$$

Formula (6.17) represents a convergent series for $\lambda > ||K + B||$. On the other hand, we know by the argument preceding formula (6.13) that the left side of (6.17) is holomorphic for all $\lambda > ||B||$; not just for those λ's that satisfy $\lambda > ||K + B||$. It follows that the right side of (6.17) converges for all $\lambda > ||B||$.

Now formula (6.17) and the standard theorem of residues gives:

$$\tilde{p}_k(K + B) = \frac{1}{2\pi i} \int_{|\lambda|=r} \lambda^k \phi_m(\lambda)\,(\lambda - (K + B))^{-1}\,d\lambda \tag{6.19}$$

for any $r > ||B||$. This gives immediately:

$$\begin{aligned} &||\tilde{p}_k(K + B)|| \\ &\le \frac{1}{2\pi} \int_{|\lambda|=r} |\lambda|^k |\phi_m(\lambda)|\,||(\lambda - (K + B))^{-1}||\,d|\lambda| \\ &\le \frac{1}{2\pi}\, r^k \frac{m}{r - ||B||} e^{(c_{m,p}+\frac{1}{e})(\frac{||K||_{S_p}^p}{r-||B||)^p}}\, 2\pi r \end{aligned} \tag{6.20}$$

where the first inequality is justified by formula (6.14).

By setting $r := ||B|| + ||K||_{S_p}\, k^{-\frac{\beta}{p}}$ and taking the kth root from the both ends on the previous formula gives (6.10). This proves the theorem. □

In Section 2 we found out that the polynomials $p_k(\lambda)$ satisfying the normalization condition $p_k(1) = 1$ are of special interest when studying the polynomial acceleration properties of a given linear system. On the other hand, polynomials $\tilde{p}_k$ are clearly not normalized in this manner - one might consider as a poor consolation that the coefficient of the highest degree term of $\tilde{p}_k$ is always a_0 of formula (6.15). In this sense these polynomials are monic.

It follows from formula (6.18) that

$$\tilde{p}_k(1) = \sum_{j=0}^{k} a_j \tag{6.21}$$

where a_j's have the same meaning as in formula (6.15). It also follows that

$$\lim_{k\to\infty} \tilde{p}_k(1) = \phi_m(1) \tag{6.22}$$

From the assumed nonsingularity of the problem (i.e. $1 \notin \sigma(K+B)$) and the general properties of the generalized determinant ϕ_m it follows that $\phi_m(1) \neq 0$. For large enough k's, $\tilde{p}_k(1)$ is bounded away from the origin, and we may define the correctly normalized polynomial sequence:

$$p_k(\lambda) := \frac{\tilde{p}_k(\lambda)}{\tilde{p}_k(1)} \tag{6.23}$$

This is exactly what we mean by stating before Theorem 6.7 that "almost all of the polynomials $\tilde{p}_k$ are almost almost normalized".

It would be very satisfactory indeed to be able to give powerful results on the convergence properties of the sequence $|\tilde{p}_k(1)|$ with only relatively general conditions imposed on the operators K and B. In particular, we would like to know how small the nonzero entity $|\phi_m(1)|$ actually is. In order to study this question we should make further assumptions about K and B; this is no longer under the subject of this paper. We are satisfied with giving a lemma that collects the asymptotic result of the above discussion and Theorem 6.7:

Lemma 6.8. *Let $m \in \mathbf{N}$ and $p-1 \leq m < p$. Take $K \in S_p(H)$ and let $B \in \mathcal{L}(H)$ be small such that $1 \notin \sigma(K+B)$. Then there exist an integer m and a sequence of polynomials $\{p_k\}_{k=m}^{\infty}$ satisfying $p_k(1) = 1$ such that for all $\beta \in (0,1]$:*

$$\begin{aligned} &||p_k(K+B)||^{\frac{1}{k}} \\ &\leq (m\, C_k)^{\frac{1}{k}} \left(||B|| + ||K||_{S_p}\, k^{-\frac{\beta}{p}}\right)\left(\frac{||B||\, k^{\frac{\beta}{p}}}{||K||_{S_p}} + 1\right)^{\frac{1}{k}} e^{\frac{c_{m,p}+\frac{1}{e}}{k^{1-\beta}}} \end{aligned} \tag{6.24}$$

where $c_{m,p}$ is defined in (4.21) and C_k satisfies:

$$\lim_{k\to\infty} C_k = C < \infty \tag{6.25}$$

Proof. Define $p_k(\lambda) := \frac{\tilde{p}_k(\lambda)}{\tilde{p}_k(1)}$ as in (6.23). Theorem 6.7 gives now formula (6.24), where $C_k := \frac{1}{|\tilde{p}_k(1)|}$. The existence of the finite C in formula (6.24) follows from the assumed nonsingularity of the problem as proposed just after formula (6.23). This proves the lemma. □

We could prove a theorem quite similar to Theorem 6.7 by using the determinant and Schatten classes $p \in (0,1]$. Nothing essential would change in the proof, we would just use Theorem 3.8 instead of Theorem 5.4. The result is stated without proof.

THEOREM 6.9. *Let $p \in (0,1]$. Take $K \in S_p(H)$ and let $B \in \mathcal{L}(H)$ be small 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 $\beta \in (0,1]$:*

$$||\tilde{p}_k(K+B)||^{\frac{1}{k}} \leq (||B|| + ||K||_{S_p} k^{-\frac{\beta}{p}})(\frac{||B||k^{\frac{\beta}{p}}}{||K||_{S_p}} + 1)^{\frac{1}{k}} e^{\frac{1}{p(k^{1-\beta})}} \tag{6.26}$$

7 Concluding remarks

Given a bounded linear operator L, the optimal reduction factor $\eta(L)$ defined by

$$\eta(l) := \inf_{k \geq 0; p_k \in p_k} ||p_k(L)||^{\frac{1}{k}} \tag{7.1}$$

limits the attainable asymptotic speed of the all Krylov subspace methods applied upon L. It can be proved that $\eta(L)$ is a function of the set $\sigma(L)$, and furthermore that it does not depend on the isolated spectral values of L (see [**1;Theorem 3.3.4**]). Now, if $L := K + B$, then $\eta(K+B) = \eta(B)$, because a compact perturbation can only add isolated points to the spectrum of the operator B. So it is true, that even a large compact part dies, because it cannot be seen in the optimal reduction factor.

There is certain interest to look at the asymptotics of $||p_k(K+B)||^{\frac{1}{k}}$ of Lemma 6.8 as $k \to \infty$. It is true that our sequence is asymptotically optimal in the sense that one cannot construct another sequence $p'_k(K+B)$ whose limit is always smaller *for all* B of fixed size. For such B whose spectrum fills the disk of radius $||B||$, the optimal reduction factor would equal $\rho(B) = ||B||$, by [**1;Theorem 3.3.4(v)**] and [**1;Theorem 3.6.3(iii)**].

So the asymptotic effect of K is inessential, but on the other hand we are not so interested in the asymptotics — it is only the small number of iterations that should ever get calculated in the real life. Lemma 6.8 should tell us that in the first stages 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 concept "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 of course refer at 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 given points; this is the minimization of residuals. If the GMRES generates the polynomial sequence s_k (corresponding to our sequence $\tilde{p}_k$), then the residual d_k after k steps is proportional to $||s_k(K+B)d_0||$ (see Proposition 2.2), and we have:

$$||s_k(K+B)d_0|| \leq ||\tilde{p}_k(K+B)d_0|| \leq ||\tilde{p}_k(K+B)||\,||d_0|| \tag{7.2}$$

The former inequality is true because s_k is optimal at d_0, and $\tilde{p}_k$ is superoptimal for the initial residual d_0. This is to say that the upper estimates we have for $\tilde{p}_k$ 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 Proposition 2.2.

References

1. Nevanlinna, O., *Convergence of Iterations for Linear Equations*, Monograph in series Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, Boston, Berlin, 1993.
2. Malinen J., *On Certain Properties or Iteration of Toeplitz Operators*, Helsinki University of Technology, Institute of Mathematics, Research Reports A361 (1996).
3. Nevanlinna O., *Convergence of Krylov-Methods for Sums of Two Operators*, Helsinki University of Technology, Institute of Mathematics, Research Report A353 (1995).
4. Dunford, N & Schwarz, J., *Linear Operators; Part II: Spectral Theory*, Interscience Publishers, Inc. (J. Wiley & Sons), New York, London, 1963.
5. Rudin, W., *Real and Complex Analysis*, 3rd edition, McGraw-Hill Book Company, New York, 1986.
6. Blumenthal, O., *Principes de la thèorie des fonctions entieres d'ordre infini*, Paris, 1910.
7. Aupetit, B., *A Primer on Spectral Theory*, Springer Verlag, Berlin, 1991.