Helsinki University of Technology Institute of Mathematics Research Reports
Teknillisen korkeakoulun matematiikan laitoksen tutkimusraporttisarja
Espoo 1998 A403

ROBUST BOUNDS FOR KRYLOV METHODS

Saara Hyvönen Olavi Nevanlinna

TEKNILLINEN KORKEAKOULU
TEKNISKA HÖGSKOLAN
HELSINKI UNIVERSITY OF TECHNOLOGY

Saara Hyvönen and Olavi Nevanlinna: Robust bounds for Krylov methods; Helsinki University of Technology Institute of Mathematics Research Reports A403 (1998).

Abstract: *In this paper we give bounds for polynomials of operators. These bounds are robust in low rank perturbations. This problem is encountered in the study of the convergence of Krylov methods. The central idea here is to view the resolvent as a meromorphic function.*

AMS subject classifications: 47A10, 65F10.

Keywords: Krylov methods, meromorphic functions, resolvent operator.

ISBN 951-22-4319-9
ISSN 0784-3143
Edita, Espoo, 1998

Helsinki University of Technology
Department of Engineering Physics and Mathematics
Institute of Mathematics
P.O. Box 1100, 02015 HUT, Finland
email: *math@hut.fi*
downloadables: *http://www.math.hut.fi/*

1 Introduction

The usual convergence analysis of Krylov methods uses different mathematical techniques for normal and for nonnormal operators. Yet, rank-one perturbations to an operator can change it from self-adjoint to quasinilpotent, and such an analysis cannot explain why the basic convergence speed is still essentially the same.

Our aim in this paper is to present an approach where normal and nonnormal operators are treated in the same way. Notice, that the convergence discussion can be divided into two parts. The first question, that of the speed with which $\|q_n\|$ decays as n increases, where q_n are monic polynomials such that $\|q_n(A)b\|$ is as small as possible, is insensitive to low rank perturbations. On the other hand the second question, that of solving for the spectrum or for a solution of an equation, is in general sensitive to low rank perturbations. We emphasize that our discussion aims to create reasonable upper bounds for $\inf \|q_n(A)b\|$, or in fact for $\inf \|q_n(A)\|$, in the case of one cluster of spectrum. Multiple clusters can be dealt with by mapping the clusters to the origin by a polynomial p and considering (for analysis only) the fixed point problem for $p(A)$ instead. Some of this approach has already appeared in our earlier papers [7], [8], [9], [10], [15], [16], [17], [18], [19], and those of our colleagues [6], [12].

We shall discuss bounded operators in separable complex Hilbert spaces. The reader should have a *fixed point formulation* in mind as a model problem:

$$x = Ax + b. \tag{1}$$

This is, of course, mathematically equivalent to the more common formulation $Ax = b$. However, the fixed point formulation is a natural one after the original problem has been preconditioned and it suggests more mathematical tools as then it is natural that the spectrum is clustered at the origin and the operator can be scanned simply by looking at the growth of the resolvent function

$$z \mapsto (1 - zA)^{-1}$$

as z grows.

We wish to be able to estimate the growth of the resolvent in terms that are not sensitive to low rank perturbations. This means in particular, that the analysis must not depend too strongly on the spectrum, as it is sensitive to low rank perturbations.

As an analogy, suppose that f is a scalar valued meromorphic function. If we “perturb” f by considering

$$g := \frac{1}{\frac{1}{f} - a},$$

where $a \in \mathbb{C}$ is a constant, then f and g are usually not close when evaluated pointwise, but as meromorphic functions they are still essentially equally large.

The outline of this paper is as follows. In Section 2 we give a representation for the resolvent in the case that $(1 - zA)^{-1}$ is meromorphic in the disc $|z| < R(\leq \infty)$. This representation gives a natural sequence of monic polynomials q_n such that

$$q_n(\lambda) = \lambda q_{n-1}(\lambda) + a_n$$

and that $q_n(A)$ appear as Taylor coefficients of an associated function $\Phi_A(z)$. Now, the decay of $\|q_n(A)\|$ as n tends to infinity is connected with the growth of $\sup_{|z| \leq r} \|\Phi_A(z)\|$ as r grows, and therefore the focus in this paper is in estimating the latter.

Notice that we do not discuss the infimum over all monic polynomials (nor the dependence of these polynomials on b) but instead focus on defining a natural "reference" sequence $\{q_n\}$, which always majorizes that created by, say, the Arnoldi method.

For example, when the operator is quasinilpotent, that is, $\sigma(A) = \{0\}$, then the polynomials $q_n(\lambda) = \lambda^n$ and $\Phi_A(z) = (1 - zA)^{-1}$ is an entire function. In general, however, we do not know the function Φ_A but the idea is that its growth (and thus the decay of $\|q_n(A)\|$) can still be estimated from the growth of $(1 - zA)^{-1}$ if the growth of the resolvent is measured as a meromorphic function.

In Section 3 we present the connections between the growth estimates for the analytic functions Φ_A (and of χ_A) and the meromorphic function $(1 - zA)^{-1}$. The growth of $(1 - zA)^{-1}$ is measured by a function

$$T_\infty(r, (1 - zA)^{-1}),$$

which is the straight forward generalization of the Nevanlinna characteristic function [20], used to measure the growth of scalar valued meromorphic functions. The growth of $T_\infty(r, (1 - zA)^{-1})$ can be used to divide operators into several classes. When

$$T_\infty(r, (1 - zA)^{-1}) < \infty \quad \text{for all } r > 0$$

the operator is *almost algebraic*, while

$$T_\infty(r, (1 - zA)^{-1}) = \mathcal{O}(\log r) \quad \text{as } r \to \infty$$

holds when the operator is *algebraic*. These classes are discussed in Sections 4 and 5 respectively. In Section 6 we show how perturbations of A by a small rank operator can be estimated. Section 7 is devoted to the decay of monic polynomials q_n. We show that the decay of $\|q_n(A)\|$ as $n \to \infty$ can be traced back to the growth of the resolvent, which is insensitive to low rank perturbations.

In Section 8 we return to the connection between estimating the decay of the monic polynomials $q_n(A)$ and the convergence of Krylov methods. We formulate a convergence result for the polynomials related to solving the fixed point equation. This gives an upper bound for the decay of the error of

GMRES. This bound is naturally divided into two parts. One depends solely on the spectrum, and consequently is sensitive to low rank perturbations. The other one depends on the growth of the resolvent, which, as stated above, is insensitive to low rank perturbations.

2 Representations for the resolvent

Let A be a bounded linear operator in a separable complex Hilbert space H. Unless otherwise stated we assume that $z \mapsto (1 - zA)^{-1}$ is meromorphic for $|z| < R(\leq \infty)$.

Definition 1 *A function $F(z)$ is called meromorphic for $|z| < R \leq \infty$ if around each z_0 in $|z_0| < R \leq \infty$ it has a representation of the form*

$$F(z) = \sum_{k=-h}^{\infty} F_k (z - z_0)^k. \tag{2}$$

Here F_k is a bounded linear operator in H and F_{-h} is nontrivial.

In order to discuss properties of $(1 - zA)^{-1}$ as a meromorphic function we need to have a natural representation for it and means to estimate its growth. The simplest representation

$$(1 - zA)^{-1} = \sum_{k=0}^{\infty} z^k A^k$$

only works for $|z| < \frac{1}{\rho(A)}$ and gives the function as an analytic function.

Theorem 1 *Let $z \mapsto (1 - zA)^{-1}$ be meromorphic for $|z| < R(\leq \infty)$. Then there exists a scalar valued function χ_A, analytic for $|z| < R$, $\chi_A(0) = 1$ and $\chi_A(z) \neq 0$ for $\frac{1}{z} \notin \sigma(A)$, such that*

$$z \mapsto \chi_A(z)(1 - zA)^{-1} =: \Phi_A(z),$$

is analytic for $|z| < R$.

Proof. Let $(1 - zA)^{-1}$ be meromorphic in a domain $D = \{z \,|\, |z| < R\}$ and let $\{\lambda_j\}$ be the nonzero points of $\sigma(A)$. Furthermore, let m_j denote the order of the pole of the resolvent at $z_j = \frac{1}{\lambda_j}$. (Note that $\{z_j\}$ may only have a limit point on the boundary of D.) Then by a generalization of Weierstrass' theorem (Theorem 3.3 in volume 3 of [13]) there exists a function $f(z)$ analytic in D such that $f(0) = 1$ and f has a zero at z_j of multiplicity m_j for every j, and is nonzero elsewhere. Now $\chi_A(z) := f(z)$ is a scalar valued function, analytic for $|z| < R$. Moreover, as we have locally around each z_j

$$(1 - zA)^{-1} = \sum_{k=-m_j}^{\infty} A_k (z - z_j)^k$$

and since χ_A has a zero of multiplicity m_j at z_j, the singularities of $\chi_A(z)(1-zA)^{-1}$ get removed, so $\Phi_A(z)$ is analytic for $|z| < R$. □

Note that the function χ_A (and thus Φ_A) are not uniquely defined, as χ_A can be multiplied by any normalized, analytic, nonvanishing function. However, it provides a natural representation for the resolvent:

Theorem 2 *Let χ_A and Φ_A be functions as in Theorem 1. If*

$$\chi_A(z) = 1 + a_1 z + a_2 z^2 + \dots$$

then define

$$q_j(\lambda) = \lambda^j + a_1 \lambda^{j-1} + \dots + a_j.$$

Then for $|z| < R$ we have

$$\Phi_A(z) = \sum_{j=0}^{\infty} q_j(A) z^j \tag{3}$$

and thus

$$(1 - zA)^{-1} = \frac{1}{\chi_A(z)} \sum_{j=0}^{\infty} q_j(A) z^j. \tag{4}$$

Conversely, if $\{a_j\} \subset \mathbb{C}$ is such that

$$\limsup \|q_j(A)\|^{1/j} =: \frac{1}{R}, \tag{5}$$

then χ_A and Φ_A are analytic for $|z| < R$.

Proof.

Since χ_A and Φ_A are analytic for $|z| < R$ they have convergent series expansions for $|z| < R$. Since

$$(1 - zA)^{-1} = \sum_{k=0}^{\infty} z^k A^k$$

converges for $|z| < \frac{1}{\rho(A)}$ we have (as necessarily $\frac{1}{\rho(A)} \leq R$) for $|z| < \frac{1}{\rho(A)}$ $(a_0 = 1)$

$$\chi_A(z)(1 - zA)^{-1} = (\sum_{k=0}^{\infty} a_k z^k)(\sum_{k=0}^{\infty} A^k z^k) = \sum_{j=0}^{\infty} \sum_{k=0}^{j} a_k A^{j-k} z^j = \sum_{j=0}^{\infty} q_j(A) z^j.$$

Thus, (3) holds for $|z| < \frac{1}{\rho(A)}$ and as the expansion of Φ_A converges for $|z| < R$, (3) must indeed hold for $|z| < R$.

Conversely, if (5) holds, then (3) converges for $|z| < R$. Moreover,

$$\chi_A(z) = (1 - zA)\Phi_A(z)$$

and as the right side converges for $|z| < R$ we can look at the coefficients which satisfy

$$q_j(A)z^j - zAq_{j-1}(A)z^{j-1} = a_j z^j$$

and thus we obtain the power series for χ_A. □

The representation (4) gives a natural sequence of monic polynomials $\{q_j\}$ the behavior of which is our primary interest in this paper. We view them as Taylor coefficients of the analytic function Φ_A and exploit the well known fact that the decay of coefficients of an analytic function is intimately related to the growth of the function. To that end, let us introduce the following notation. The maximum modulus of an analytic (vector valued) function F

$$M_\infty(r, F) := \sup_{|z|\leq r} \|F(z)\|.$$

Then we have the following.

Theorem 3 *Let Φ_A and q_j be as in Theorem 2. Then for $r < R$*

$$\|q_j(A)\| \leq M_\infty(r, \Phi_A) r^{-j}. \tag{6}$$

Conversely,

$$M_\infty(r, \Phi_A) \leq \sum_{j=0}^{\infty} \|q_j(A)\| r^j. \tag{7}$$

Proof. The first inequality (6) follows from the fact that

$$q_j(A) = -\frac{1}{2\pi \mathrm{i}} \int_{|z|=r} z^{-j-1} \Phi_A(z) \mathrm{d}z, \quad r < R,$$

while (7) follows from (3). □

The crucial step then is to estimate the growth of $M_\infty(r, \Phi_A)$ as $r \to R$. In our approach we base this on estimating $(1 - zA)^{-1}$ as a meromorphic function, because then it is insensitive to the actual location of the spectrum.

Corollary 1 *Assume $R = \infty$ and*

$$M_\infty(r, \Phi_A(z)) = M_\infty(r, \chi_A(z)(1 - zA)^{-1}) \leq C\mathrm{e}^{\tau r^\omega}$$

for $r \geq 0$. Then for $j \geq 1$

$$\|q_j(A)\| \leq C(\frac{\tau \mathrm{e} \omega}{j})^{j/\omega}. \tag{8}$$

Proof. The inequality (8) follows from (6) by choosing $r^\omega = \frac{j}{\tau\omega}$. □

Remark 1 In Section 7 we shall use the results presented in this section in such a way that for each $\eta < R$ we choose a $\chi_A(z) = \chi_{A,\eta}(z)$, valid for $r \leq \eta$. This leads to sharper estimates than choosing a single $\chi_A(z)$ valid for $|z| < R$.

3 Growth of meromorphic resolvents

As Theorem 3 shows, the polynomial sequence $\{q_j\}$ evaluated at A is in a natural way associated with the growth of $M_\infty(r, \Phi_A)$. In practice we do not know χ_A, and hence not Φ_A either, and estimating $M_\infty(r, \Phi_A)$ has to be based on the resolvent directly.

Rolf Nevanlinna [20], [21] introduced in 1925 a characteristic function T to measure meromorphic functions. In particular it has the following property: if f is meromorphic and $f(0) = 1$, then (by the first main theorem)

$$T(r, f) = T(r, \frac{1}{f}). \tag{9}$$

This can be generalized to operator valued functions. The obvious way is to replace the absolute value by the norm of the (operator valued) function. This leads to a generalization denoted here by $T_\infty(r, F)$, see below. Another useful generalization $T_1(r, F)$ is introduced in [17], and is discussed shortly in Section 6. We only mention here that $T_1(r, F)$ satisfies the identity (9), and that we always have $T_\infty \leq T_1$. The function T_∞ is defined as follows.

Definition 2 *Let $F(z)$ be a meromorphic operator valued function as in Definition 1. If $-h < 0$, then F has a pole at z_0 of order h, otherwise F is analytic at z_0. Denote $h(z_0) := \max\{h, 0\}$ and define*

$$n_\infty(r, F) := \sum_{|b| \leq r} h(b).$$

Thus n_∞ counts the poles in $\{z \mid |z| < r\}$ together with their orders. Furthermore define

$$N_\infty(r, F) := \int_0^r \frac{n_\infty(t, F) - n_\infty(0, F)}{t} dt + n_\infty(0, F) \log r$$

and

$$m_\infty(r, F) := \frac{1}{2\pi} \int_{-\pi}^{\pi} \log^+ \|F(re^{i\varphi})\| d\varphi,$$

where

$$\log^+ t = \begin{cases} \log t, & t \geq 1 \\ 0, & t \leq 1. \end{cases}$$

Finally define $T_\infty(r, F) = m_\infty(r, F) + N_\infty(r, F)$.

Here we only deal with functions such that $n_\infty(0, F) = 0$, in which case the above expression for $N_\infty(r, F)$ is simplifies to $N_\infty(r, F) = \int_0^r \frac{n_\infty(t,F)}{t} \mathrm{d}t$.

By Theorem 2.1 in [19] the following holds.

Theorem 4 *Let F and G be meromorphic for $|z| < R \leq \infty$. Then $T_\infty(r, F)$ is a nonnegative and nondecreasing function in r for $0 \leq r < R$ which is convex in the variable $\log r$. The following inequality holds:*

$$T_\infty(r, FG) \leq T_\infty(r, F) + T_\infty(r, G).$$

For an analytic function $T_\infty(r,F)$ and $M_\infty(r,F)$ are related in the following way:

Theorem 5 *If F is an operator valued analytic function for $|z| < R$ and $0 < r < \theta r < R$, then*

$$T_\infty(r,F) \le \log^+ M_\infty(r,F) \le \frac{\theta+1}{\theta-1} T_\infty(\theta r, F).$$

This is Theorem 2.2 in [19]. It allows one to move back and forth between the maximum norm analysis and logarithmic averages when working with Φ_A.

Furthermore we have the following result.

Theorem 6 *Let χ_A and Φ_A be as in Theorem 1. Then for $r < R$*

$$\begin{aligned} T_\infty(r,\Phi_A(z)) &\le T(r,\chi_A(z)) + T_\infty(r,(1-zA)^{-1}), & (10)\\ T_\infty(r,(1-zA)^{-1}) &\le T(r,\chi_A(z)) + T_\infty(r,\Phi_A(z)), & (11)\\ T(r,\chi_A(z)) &\le T_\infty(r,\Phi_A(z)) + \log(1+r\|A\|). & (12)\end{aligned}$$

Proof. The inequality (10) follows directly by applying Theorem 4 to $\Phi_A(z) = \chi_A(z)(1-zA)^{-1}$. To get (11), apply Theorem 4 to $(1-zA)^{-1} = \frac{1}{\chi_A(z)}\Phi_A(z)$ and notice that by (9) we have $T(r,\frac{1}{\chi_A}) = T(r,\chi_A)$. To get (12), again apply Theorem 4, this time to $\chi_A(z) = (1-zA)\Phi_A(z)$. As $N_\infty(r,1-zA) = 0$, we have

$$T_\infty(r,1-zA) = m_\infty(r,1-zA) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log^+ \|1-re^{i\varphi}A\| d\varphi \le \log(1+r\|A\|),$$

which completes the proof. □

Remark 2 Trivially $T(r,\chi_A) \ll T_\infty(r,(1-zA)^{-1})$ is possible. To see this, consider the case where $(1-zA)^{-1}$ is entire in z. However, Theorem 6 does not give us a bound for $T(r,\chi_A)$ in terms of $T_\infty(r,(1-zA)^{-1})$, quite simply because such a bound does not exist. To see this consider Example 4 in Section 7. It is however possible to choose for each $\eta < R$ the function $\chi_A(z) = \chi_{A,\eta}(z)$, valid for $r \le \eta$. This has been done in Section 7. Then it is possible to show that

$$T(r,\chi_A) \le N_\infty(\eta,(1-zA)^{-1}) \le T_\infty(\eta,(1-zA)^{-1}).$$

4 Almost algebraic operators

In this section we assume that $R = \infty$. Recall that an operator A is called *quasinilpotent* if

$$\|A^n\|^{1/n} \to 0 \quad \text{as } n \to 0,$$

which is equivalent with $\sigma(A) = \{0\}$. This has been generalized to *quasialgebraic* operators [4]. Quasialgebraicity again is a smallness property of the spectrum alone: the logarithmic capacity of $\sigma(A)$ vanishes.

Quasinilpotent operators have also a function theoretic property: their resolvents are entire. There is a natural generalization of quasinilpotent (and of algebraic) operators which form a subclass of quasialgebraic operators.

Definition 3 *A bounded operator A is called almost algebraic if there exists a sequence $\{a_j\} \subset \mathbb{C}$ such that*

$$||q_j(A)||^{1/j} \to 0 \quad \textit{as } j \to 0,$$

where

$$q_j(z) = z^j + a_1 z^{j-1} + \ldots + a_j.$$

It was shown in [15, Section 5.7] that an operator A is almost algebraic if and only if the resolvent operator $(1 - zA)^{-1}$ is meromorphic in $|z| < \infty$. This also follows from Theorem 2 in Section 2 of this paper, by choosing $R = \infty$.

Compact, quasinilpotent and algebraic operators are all examples of almost algebraic operators. The whole class is closed under perturbing with finite rank operators, which follows from a general perturbation result in [19], see also Section 6 below.

Remark 3 It also follows that an operator is almost algebraic if and only if $T_\infty(r, (1 - zA)^{-1})$ is finite for all finite r. In the next section we demonstrate that if

$$T_\infty(r, (1 - zA)^{-1}) = \mathcal{O}(\log r)$$

then A is algebraic. The growth of $T_\infty(r, (1 - zA)^{-1})$ thus divides operators into subclasses.

Definition 4 *Let F be meromorphic for $|z| < \infty$. We say that as a meromorphic function $F(z)$ is of order ω_{mer} and (if $0 < \omega_{mer} < \infty$) of type τ_{mer} where*

$$\omega_{mer} := \limsup_{r \to \infty} \frac{\log T_\infty(r, F)}{\log r}$$

$$\tau_{mer} := \limsup_{r \to \infty} r^{-\omega_{mer}} T_\infty(r, F).$$

Similarly, if F is entire, it is of order ω and type τ, where

$$\omega := \limsup_{r \to \infty} \frac{\log \log M_\infty(r, F)}{\log r}$$

$$\tau := \limsup_{r \to \infty} r^{-\omega} \log M_\infty(r, F).$$

With a slight abuse of language we say that in the above case $T_\infty(r,F)$ grows with order ω_{mer} and type τ_{mer}, and $M_\infty(r,F)$ grows with order ω and type τ. Furthermore, we sometimes wish to distinguish between the growth of $N_\infty(r,F)$ and $m_\infty(r,F)$. They we say that $N_\infty(r,F)$ grows with order ω_N and type τ_N, where

$$\omega_N := \limsup_{r\to\infty} \frac{\log N_\infty(r,F)}{\log r}$$
$$\tau_N := \limsup_{r\to\infty} r^{-\omega_N} N_\infty(r,F),$$

and $m_\infty(r,F)$ grows with order ω_m and type τ_m, where ω_m and τ_m are obtained by replacing N by m in the expressions above.

For an entire function the growth of $T_\infty(r,F)$ and $M_\infty(r,F)$ are related in the following way.

Theorem 7 *If $F(z)$ is an operator valued entire function, then its orders as an entire and a meromorphic function are equal: $\omega_{mer} = \omega$. Furthermore the types satisfy the following inequality:*

$$\tau_{mer} \le \tau \le (2\omega+1)\mathrm{e}\tau_{mer}.$$

Proof. Following the proof of the scalar valued case in [22, pp. 181-182], choose $\theta = (1+\frac{1}{\omega})$ in Theorem 5 to get

$$\begin{aligned}\log^+ M_\infty(r,F) &\le \frac{(2+\frac{1}{\omega})r}{\frac{1}{\omega}r} T_\infty((1+\frac{1}{\omega})r,F)\\ &= (2\omega+1)\frac{T_\infty((1+\frac{1}{\omega})r,F)}{((1+\frac{1}{\omega})r)^\omega}(1+\frac{1}{\omega})^\omega r^\omega.\end{aligned}$$

Divide by r^ω and let $r\to\infty$ to get

$$\tau < (2\omega+1)\tau_{mer}(1+\frac{1}{\omega})^\omega < (2\omega+1)\mathrm{e}\tau_{mer}.$$

□

This means that instead of $M_\infty(r,\Phi_A)$ we can look at $T_\infty(r,\Phi_A)$, as these both grow with the same order and their types are related as indicated by Theorem 7.

In particular we see that the orders ω and ω_{mer} are both either zero, finite or infinite. Following the tradition in scalar valued theory we distinguish the following type-classes:

Definition 5 *We say that $F(z)$ has*

1. *minimal type if $\tau_{mer} = 0$,*

2. *mean type if $0 < \tau_{mer} < \infty$,*

3. maximal type if $\tau_{mer} = \infty$.

Definition 6 *We say that an almost algebraic operator* F *belongs to the class* $\mathcal{AA}(\omega_{mer}, m)$ *if* F *is of order* ω_{mer} *and type-class* m*, where* m *stands for minimal, mean or maximal. Moreover we say that* F *belongs to the class* $\mathcal{AA}(exp)$ *if it is of order* $\omega_{mer} \leq 1$ *and when* $\omega_{mer} = 1$ *then* $\tau_{mer} = 0$.

Example 1 Let A be a compact operator such that $\|A\|_1 := \sum_j \sigma_j(A) < \infty$ where $\sigma_j(A)$ indicate the singular values of A. Then by Theorem 5.1 in [19] $(1 - zA)^{-1} \in \mathcal{AA}(\text{exp})$.

Theorem 8 *The growth of* $T_\infty(r, F)$ *divides the operators* F *into subclasses as in Definition 6 such that each class is closed under finite rank perturbations.*

Proof. This follows from Theorem 12. □

5 Algebraic operators

An operator A is said to be *algebraic* if there exists a polynomial q, which we require to be monic and of smallest degree, such that $q(A) = 0$. The polynomial is called the minimal polynomial of A and we say that A is algebraic of degree $\deg A = d$ where d is the degree of q. The Arnoldi process terminates (in a finite number of steps) at every starting vector if and only if A is algebraic. This is a simple consequence of Kaplansky's theorem, see Sections 2.8 and 5.4 in [15]. Furthermore, an operator has a rational resolvent if and only if it is algebraic (Theorem 2.8.9. in [15]). In this section we discuss the growth of resolvents for algebraic operators.

Assume that A is algebraic. If the minimal polynomial is

$$q(z) = z^d + a_1 z^{d-1} + \ldots + a_d,$$

then we put

$$\chi_A(z) = 1 + a_1 z + \ldots + a_n z^n,$$

where $a_n \neq 0$, $n \leq d$ and $\chi_A(1/z)z^d = q(z)$. Now the resolvent takes the form

$$(1 - zA)^{-1} = \frac{1}{\chi_A(z)} \sum_{j=0}^{d-1} q_j(A) z^j.$$

(This follows from Theorem 2 since $q_k(A) = A^{k-d} q(A) = 0$ for $k \geq d$.) This representation implies immediately that

$$T_\infty(r, (1 - zA)^{-1}) \leq T(r, \chi_A) + T_\infty(r, \Phi_A) \leq n \log^+ r + (d - 1) \log^+ r + \mathcal{O}(1),$$

as χ_A and Φ_A are polynomials of degree n and $d - 1$ respectively. Actually, the growth is slower, as we show below that for all $r > 0$

$$T_\infty(r, (1 - zA)^{-1}) = \max\{n, d - 1\} \log^+ r + \mathcal{O}(1).$$

Definition 7 *Given an algebraic A with* $\deg A = d$*, call the number*

$$\max\{n, d-1\} =: \hat{d}$$

the degree of the rational function $(1 - zA)^{-1}$.

Observe that

$$\hat{d} = \begin{cases} d & \text{if } A \text{ is nonsingular,} \\ d-1 & \text{if } A \text{ is singular.} \end{cases}$$

In fact, if A is nonsingular, $n = d$, otherwise $n < d$.

Theorem 9 *Suppose* $(1 - zA)^{-1}$ *is meromorphic for all* $|z| < \infty$ *and that*

$$T_\infty(r, (1 - zA)^{-1}) = \mathcal{O}(\log r) \quad \text{as } r \to \infty. \tag{13}$$

Then A is algebraic and

$$T_\infty(r, (1 - zA)^{-1}) = \hat{d} \log r + \mathcal{O}(1) \quad \text{as } r \to \infty.$$

Proof. Since (13) holds it means that there is a constant K such that say for $r \geq 2$

$$N_\infty(r, (1 - zA)^{-1}) \leq K \log r.$$

Furthermore

$$\begin{aligned} N_\infty(r, (1 - zA)^{-1}) &\geq \int_{\sqrt{r}}^{r} \frac{n_\infty(t, (1 - zA)^{-1})}{t} \mathrm{d}t \geq n_\infty(\sqrt{r}, (1 - zA)^{-1}) \log \frac{r}{\sqrt{r}} \\ &= \frac{1}{2} n_\infty(\sqrt{r}, (1 - zA)^{-1}) \log r, \end{aligned}$$

so $n_\infty(\sqrt{r}, (1 - zA)^{-1}) \leq 2K$ for all r. Thus, the number of poles of the resolvent is bounded. That is, there exists a (monic) polynomial $p(z)$ such that $\Phi(z) := p(z)(1 - zA)^{-1}$ has no poles and is therefore entire.

Next conclude that

$$T_\infty(r, \Phi(z)) \leq T(r, p(z)) + T_\infty(r, (1 - zA)^{-1}) = \mathcal{O}(\log r).$$

Entire (vector valued or not) functions growing this slow are polynomials. In fact Theorem 5 implies

$$M_\infty(r, \Phi(z)) \leq \mathcal{O}(r^{\text{const}}) \quad \text{as } r \to \infty,$$

and such functions are polynomials by the maximum principle (see Lemma 5.4.2. in [15]). But then

$$(1 - zA)^{-1} = \frac{1}{p(z)} \Phi(z)$$

is rational and A is algebraic (Theorem 2.8.9. in [15]). Since A is algebraic, we can choose $\chi_A(z) = p(z)$ and $\Phi_A(z) = \Phi(z)$.

Since the number of poles of the resolvent of A is bounded, $n_\infty(r,(1-zA)^{-1}) = n$ for large enough r. Therefore

$$N_\infty(r,(1-zA)^{-1}) = n\log^+ r + \mathcal{O}(1). \tag{14}$$

On the other hand

$$\|\Phi_A(z)\| = \|\sum_{j=0}^{d-1} q_j(A)z^j\| = \|q_{d-1}(A)\|r^{d-1}(1+o(1))$$

and

$$|\chi_A(z)| = |a_n|r^n(1+o(1))$$

which imply

$$\begin{aligned} m_\infty(r,(1-zA)^{-1}) &= \log^+\left[\frac{\|q_{d-1}(A)\|}{|a_n|}r^{d-1-n}(1+o(1))\right] \\ &= \begin{cases} 0 \quad \text{for } r \text{ large when } n=d \\ \mathcal{O}(1) \quad \text{when } n=d-1 \\ (d-1-n)\log^+ r + \mathcal{O}(1) \quad \text{when } n<d-1 \end{cases} \end{aligned} \tag{15}$$

Summing up (14) and (15) gives the result. □

Remark 4 Notice that if A is nonsingular then $(1-zA)^{-1}$ is analytic at $z=\infty$ and either A is algebraic or $(1-zA)^{-1}$ has a singularity at a finite point which is not a pole. If A is both nonsingular and algebraic (for example, a nonsingular matrix in a finite dimensional space), then

$$T_\infty(r,(1-zA)^{-1}) = N_\infty(r,(1-zA)^{-1}) = m(r,\chi_A).$$

Theorem 9 suggests to look for a constant $\hat{c}$ such that

$$T_\infty(r,(1-zA)^{-1}) = \hat{d}\log^+ r + \hat{c} + o(1).$$

Observe that both $\hat{d}$ and $\hat{c}$ are given (in the nonsingular case) strictly by the spectrum and $o(1)$ may be very large.

6 Operators of small rank

The sum of two algebraic operators need not be algebraic, or even quasialgebraic. For example, the bilateral shift S can be represented as a sum of two nilpotent operators, $S = S_e + S_o$, where S_e shifts the evenly indexed coordinates and S_o the odd ones.

In contrast, perturbations of finite rank within classes of e.g. algebraic operators, almost algebraic operators, operators with meromorphic resolvents for $|z| < R$ etc. stay within the class.

As previously seen, these classes correspond to

$$\begin{cases} T_\infty(r,(1-zA)^{-1}) = \mathcal{O}(\log^+ r), & r\to\infty, \\ T_\infty(r,(1-zA)^{-1}) < \infty, & r<\infty, \\ T_\infty(r,(1-zA)^{-1}) < \infty, & r<R. \end{cases}$$

We formulate a perturbation result and a "characterization" of finite rank operators. These results are based on another generalization of $T(r,f)$ for operator valued functions, which we only shortly mention.

The function $T_\infty(r,F)$ is obtained by replacing $\log^+|f|$ in the scalar theory by $\log^+\|F\|$. An alternative generalization is given in [19] for finitely S_1-meromorphic functions F. Here it suffices to know that if $F(0)=0$ then (unlike in the case of T_∞) the first main theorem holds

$$T_1(r,1-F) = T_1(r,(1-F)^{-1}),$$

and that, therefore, when applied to the resolvent operator $(1-zA)^{-1}$ its value can be computed directly from $(1-zA)$:

$$\begin{aligned} T_1(r,(1-zA)^{-1}) &= T_1(r,1-zA) \\ &= m_1(r,1-zA) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_j \log^+(\sigma_j(1-re^{i\varphi}A))d\varphi, \end{aligned}$$

where σ_j indicates the singular values. Since $\sigma_j(1-re^{i\varphi}A)\le 1+r\sigma_j(A)$ so that $\log^+(\sigma_j(1-re^{i\varphi}A))\le r\sigma_j(A)$, we always have

$$T_1(r,1-zA)\le r\|A\|_1,$$

where $\|A\|_1$ is the trace norm of A:

$$\|A\|_1 = \sum \sigma_j(A).$$

Moreover, we have always $T_\infty(r,1-F)\le T_1(r,1-F)$. This allows us to formulate the following simple result:

Theorem 10 *If $\|A\|_1<\infty$, then for all $r>0$*

$$T_\infty(r,(1-zA)^{-1})\le T_1(r,1-zA)\le r\|A\|_1.$$

Theorem 11 *If $T_1(r,1-zA)=\mathcal{O}(\log^+ r)$, then A is of finite rank and as $r\to\infty$*

$$T_1(r,1-zA) = k\log^+ r + \sum_{j=1}^{k}\log\sigma_j(A) + o(1), \tag{16}$$

where $k = rankA$.

Proof. Suppose $T_1(r, 1-zA) \leq K \log r$ for $r \geq 2$ and $\sigma_l(A) > 0$ for some $l > K$. Then (as the singular values are assumed to be ordered nonincreasingly)

$$\sigma_j(1 - zA) \geq r\sigma_j(A) - 1,$$

so that for $j \leq l$ and $r \geq 2/\sigma_l(A)$

$$\begin{aligned} \log^+ \sigma_j(1 - zA) &\geq \log(r\sigma_j(A)) + \log(1 - \frac{1}{r\sigma_j(A)}) \\ &\geq \log r + \log \sigma_j(A) + \log(1 - \frac{1}{r\sigma_l(A)}). \end{aligned}$$

Thus, as $r \to \infty$,

$$\begin{aligned} m_1(r, 1 - zA) &\geq \sum_{j=1}^{l} \frac{1}{2\pi} \int \log^+ \sigma_j(1 - re^{i\varphi}A) d\varphi \\ &\geq l \log r + \sum_{j=1}^{l} \log \sigma_j(A) + l \log(1 - \frac{1}{r\sigma_l(A)}) \\ &= l \log r + \sum_{j=1}^{l} \log \sigma_j(A) + o(1). \end{aligned}$$

This contradicts the assumption and $\operatorname{rank} A \leq K$.

Let $k = \operatorname{rank} A$. Then

$$\sigma_j(1 - zA) \leq r\sigma_j(A) + 1$$

implies for all $r > 0$

$$m_1(r, 1 - zA) \leq \sum_{j=1}^{k} \log^+(r\sigma_j(A) + 1) = k \log r + \sum \log \sigma_j(A) + o(1)$$

as $r \to \infty$. □

Observe that for nonsingular finite dimensional operators $\prod_{j=1}^{k} \sigma_j(A) = |\det A|$ and thus $\sum \log \sigma_j(A) = \log |\det A|$. Thus

$$T_1(r, 1 - zA) = k \log r + \log |\det A| + o(1)$$

where the $o(1)$-term contains also "nonspectral" information. The function T_1 was designed for perturbation analysis [19]. We give a simple result concerning resolvents.

Theorem 12 *Let A be a bounded operator such that $(1-zA)^{-1}$ is meromorphic for $|z| < R$ and let B be a finite rank operator of rank k. Then*

$$T_\infty(r, (1-z(A+B))^{-1}) \leq (k+1)T_\infty(r, (1-zA)^{-1}) + k(\log^+ r + \log^+ \|B\| + \log 2). \tag{17}$$

Proof. Factor $1 - z(A+B) = (1-zA)(1-(1-zA)^{-1}zB)$ so that

$$T_\infty(r,(1-z(A+B))^{-1}) \le T_\infty(r,(1-zA)^{-1}) + T_\infty(r,(1-(1-zA)^{-1}zB)^{-1}).$$

By definition, $T_\infty(r,F) \le T_1(r,F)$ for any F, so we have

$$T_\infty(r,(1-(1-zA)^{-1}zB)^{-1}) \le T_1(r,(1-(1-zA)^{-1}zB)^{-1}).$$

But

$$T_1(r,(1-(1-zA)^{-1}zB)^{-1}) = T_1(r,(1-(1-zA)^{-1}zB)).$$

Now, we have

$$\sigma_j(1-(1-zA)^{-1}zB) \le \|1-(1-zA)^{-1}zB\|$$

for $j = 1, 2, \ldots, k$ and thus

$$T_1(r, 1-(1-zA)^{-1}zB) \le kT_\infty(r, 1-(1-zA)^{-1}zB).$$

But

$$\log^+\|1-(1-zA)^{-1}zB\| \le \log^+\|(1-zA)^{-1}\| + \log^+ r + \log^+\|B\| + \log 2$$

and therefore

$$T_\infty(r, 1-(1-zA)^{-1}zB) \le T_\infty(r,(1-zA)^{-1}) + \log^+ r + \log^+\|B\| + \log 2.$$

□

Example 2 If $A = I$ and $B = \text{diag}(\beta_1, \ldots, \beta_k, 0, \ldots)$ with

$$\beta_1 > \beta_2 > \ldots > \beta_k > 0,$$

then $(1-z(A+B))^{-1}$ has poles at 1 and at $\frac{1}{1+\beta_j}$ so that for $r \ge 1$ we have

$$\begin{aligned} T_\infty(r,(1-z(A+B))^{-1}) &= \log^+ r + \sum_{j=1}^{k} \log^+((1+\beta_j)r) \\ &= (k+1)\log^+ r + \sum_{j=1}^{k} \log^+(1+\beta_j). \end{aligned}$$

For $r \ge 2$ we have $T_\infty(r,(1-zA)^{-1}) = \log^+ r$ and therefore for $r \ge 2$

$$T_\infty(r,(1-z(A+B))^{-1}) = (k+1)T_\infty(r,(1-zA)^{-1}) + \sum_{j=1}^{k} \log^+(1+\beta_j).$$

Example 3 Define

$$S = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & 0 \end{pmatrix} \quad \text{and} \quad L = \begin{pmatrix} & & \\ & & \\ 1 & & \end{pmatrix}$$

where S and L are of equal size and let

$$A = \begin{pmatrix} S & & & \\ & S & & \\ & & \ddots & \\ & & & S \end{pmatrix} \quad B = \begin{pmatrix} 0 & L & & \\ & \ddots & \ddots & \\ & & & L \\ & & & 0 \end{pmatrix}$$

so that

$$A + B = \begin{pmatrix} S & L & & \\ & S & \ddots & \\ & & \ddots & L \\ & & & S \end{pmatrix}$$

where A is block diagonal with $(k+1)$ identical blocks S of dimension d on the diagonal, so B is of rank k. By Theorem 9

$$T_\infty(r, (1 - zA)^{-1}) = (d-1)\log^+ r + \mathcal{O}(1)$$

while

$$T_\infty(r, (1 - z(A+B))^{-1}) = ((k+1)d - 1)\log^+ r + \mathcal{O}(1).$$

Compare this to the fact that by Theorem 12 we have

$$\begin{aligned} T_\infty(r, (1 - z(A+B))^{-1}) &\leq (k+1)T_\infty(r, (1-zA)^{-1}) + k\log^+ r + \mathcal{O}(1) \\ &= (k+1)(d-1)\log^+ r + k\log^+ r + \mathcal{O}(1) \\ &= ((k+1)d - 1)\log^+ r + \mathcal{O}(1), \end{aligned}$$

so in fact in this example equality holds in (17).

7 Decay of monic polynomials

One of the topics in this paper is to estimate a basic convergence speed and to show that it is insensitive to low rank perturbations. In the previous chapters this has been done by showing that given a growth estimate for $T_\infty(r, (1 - zA)^{-1})$ there exists a sequence of monic polynomials decaying with at least the given speed. The Arnoldi method (if implemented formally in a Hilbert space) would produce not these polynomials but minimal ones. Here we demonstrate that also the basic speed of convergence of the Arnoldi method is essentially unchanged by low rank perturbations. Consider updating A by a rank one operator uv^*:

$$A_+ = A + uv^*.$$

If we denote by $K_n(A, b)$ the Krylov subspaces

$$K_n(A, b) = \operatorname{span}\{A^j b\}_{j=0}^{n-1},$$

then the following subspaces are identical:

$$K_n(A_+, u) = K_n(A, u).$$

Thus the monic polynomials of given degree minimizing $\|q_n(A)u\|$ and $\|q_n(A_+)u\|$ would in general be different but they would represent the same vector and the norms would be equal. The remaining question concerns problems related to the relative size of the terms $\inf \|q_n(A)u\|$, $\sup_{\|b\|=1} \inf \|q_n(A)b\|$ and $\inf \|q_n(A)\|$. Here the infimum is taken over all monic polynomials of degree n. First, for a fixed n and any constant $\varepsilon > 0$ there are matrices A such that

$$\sup_{\|b\|=1} \inf \|q_n(A)b\| \le \varepsilon \inf \|q_n(A)\|.$$

See [23], [3]. Yet, when we look at the speed with which these decay when n increases, the picture changes. An important result was that of V. Müller [14] which says that if

$$\limsup(\inf \|q_n(A)\|)^{1/n} > 0$$

then there is a dense set of initial vectors b so that

$$\limsup(\inf \|q_n(A)b\|)^{1/n} > 0.$$

See also [6], [15].

Now let q_j be as in Theorem 3. Obtaining an upper bound for these polynomials gives a bound for the monic polynomials generated by the Arnoldi method. By Theorem 3 we have

$$\|q_j(A)\| \le M_\infty(r, \Phi_A) r^{-j},$$

where $\Phi_A(z) = \chi_A(z)(1 - zA)^{-1}$ and $\chi_A(z)$ are as in Theorem 1.

Let $\eta < R$ and let $\{z_j\}_{j=1}^n$ be the poles of $(1 - zA)^{-1}$ (with multiplicities) up to $|z| \le \eta$. Denote by $B_\eta(z)$ the Blaschke product

$$B_\eta(z) = \prod_{j=1}^n \eta \frac{|z_k|}{z_k} \frac{z_k - z}{\eta^2 - \bar{z}_k z}.$$

$B_\eta(z)(1 - zA)^{-1}$ is analytic in $|z| \le \eta$ and $M(\eta, B_\eta) = 1$. Moreover,

$$B_\eta(0) = \prod_{k=1}^n \frac{|z_k|}{\eta}$$

so

$$\log \frac{1}{B_\eta(0)} = \sum \log^+ \frac{\eta}{|z_k|} = N_\infty(\eta, (1-zA)^{-1}).$$

Therefore $\chi_A(z) := \frac{1}{B_\eta(0)} B_\eta(z)$ is analytic for $|z| \le \eta$ and so is $\Phi_A(z)$. Furthermore for $r \le \eta$

$$T(r, \chi_A) \le \log \frac{1}{B_\eta(0)} = N_\infty(\eta, (1-zA)^{-1}).$$

To summarize, we have the following.

Theorem 13 *Let A be a bounded linear operator such that $(1-zA)^{-1}$ is meromorphic for $|z| < R$. Then there exist, for every $\eta < R$, functions $\chi_A = \chi_{A,\eta}$ and $\Phi_A = \Phi_{A,\eta}$ satisfying the requirements of Theorem 1 such that for $r \le \eta$*

$$T(r, \chi_A) \le N_\infty(\eta, (1-zA)^{-1}) \tag{18}$$

and

$$T_\infty(r, \Phi_A) \le N_\infty(\eta, (1-zA)^{-1}) + T_\infty(r, (1-zA)^{-1}). \tag{19}$$

Theorem 14 *Let $\eta < R$ be fixed. Then there exists a sequence of monic polynomials $\{q_j\}$, each of degree j, such that for all $\theta > 1$*

$$\|q_j(A)\| \le \exp\Big(2\frac{\theta+1}{\theta-1} T_\infty(\eta, (1-zA)^{-1})\Big)(\frac{\theta}{\eta})^j.$$

Proof. This follows from Theorem 3 together with Theorem 5 by applying (19) for $\theta r = \eta$ and remembering that $T_\infty(\eta, (1-zA)^{-1}) \ge N_\infty(\eta, (1-zA)^{-1})$. $\square$

Consider now the case $R = \infty$. We have the following:

Theorem 15 *Let $R = \infty$, and assume that for all $r \ge 1$*

$$T_\infty(r, (1-zA)^{-1}) \le \tau r^\omega.$$

Then for $j \ge \hat{\tau}\omega$ there are monic polynomials such that

$$\|q_j(A)\| \le (\frac{\hat{\tau} e \omega}{j})^{j/\omega}, \tag{20}$$

where

$$\hat{\tau} = c(\omega)\tau, \quad c(\omega) = \frac{2(1+\omega+\sqrt{1+\omega^2})(1+\sqrt{1+\omega^2})^\omega}{(1-\omega+\sqrt{1+\omega^2})\omega^\omega}.$$

Proof. By Theorem 14 for every η

$$\|q_j(A)\| \le \exp\Big(2\frac{\theta+1}{\theta-1} T_\infty(\eta, (1-zA)^{-1})\Big)(\frac{\theta}{\eta})^j$$

holds for all $\theta > 1$. Write $r = \frac{\eta}{\theta}$ to get

$$\|q_j(A)\| \leq \exp\Big(2\frac{\theta+1}{\theta-1}T_\infty(\theta r, (1-zA)^{-1})\Big)r^{-j}.$$

Define $\alpha = 2\frac{\theta+1}{\theta-1}\tau\theta^\omega$ and consider the function

$$f(r) = \mathrm{e}^{\alpha r^\omega} r^{-j}. \tag{21}$$

Now

$$f'(r) = \mathrm{e}^{\alpha r^\omega} r^{-j-1}(\alpha\omega r^\omega - j) = 0$$

for $r = (\frac{j}{\alpha\omega})^{1/\omega}$, which is a minimum of $f(r)$. As $r \geq 1$ we must have $j \geq \alpha\omega$. Insert this r into (21) to get

$$f((\frac{j}{\alpha\omega})^{1/\omega}) = (\frac{2\omega\tau\mathrm{e}(\theta+1)\theta^\omega}{j(\theta-1)})^{j/\omega} = (\frac{2\omega\tau\mathrm{e}}{j})^{j/\omega}g(\theta)^{j/\omega}, \tag{22}$$

where $g(\theta) = \frac{(\theta+1)\theta^\omega}{(\theta-1)}$. Now $g(\theta)$ has a minimum at $\theta = \frac{1+\sqrt{1+\omega^2}}{\omega}$. Insert this into (22) to get

$$f = (\frac{c(\omega)\tau\mathrm{e}\omega}{j})^{j/\omega}, \quad c(\omega) = \frac{2(1+\omega+\sqrt{1+\omega^2})(1+\sqrt{1+\omega^2})^\omega}{(1-\omega+\sqrt{1+\omega^2})\omega^\omega},$$

and (20) holds for $j \geq \tau\omega c(\omega)$. □

This sequence $\{q_j\}$ is obtained by using a different $\chi_A(r)$, valid for $r \leq \eta < R$, for each η. It is in fact possible to find a $\chi_A(r)$, valid for all $r < R$. Using such a $\chi_A(r)$ would lead to a sequence $\{q_j\}$ which gives the Taylor coefficients of Φ_A, but the speed bound obtained this way is coarser than the one obtained above.

Theorem 16 *Let $(1-zA)^{-1}$ be meromorphic in $|z| < \infty$ with the poles $\{z_k\}$ and let the infimum of α for which*

$$\sum_{k=1}^{\infty} |z_k|^{-\alpha} \tag{23}$$

converges be ν, and let the smallest integer for which (23) converges be γ. Moreover, define for any positive integer q the Weierstrass primary factor

$$E(z,q) = (1-z)\mathrm{e}^{z+z^2/2+\ldots+z^q/q}$$

and

$$E(z,0) = 1-z.$$

Then the product

$$\prod_{k=1}^{\infty} E(z/z_k, \gamma-1)$$

converges absolutely and uniformly in any bounded part of the plane to an entire function χ_A satisfying the requirements of Theorem 1 and having the same order as $N_\infty(r,(1-zA)^{-1})$, $\omega=\nu$, and the same type-class if the order is not an integer. Further $\chi_A(z)$ satisfies the inequality

$$\log|\Pi(z)| \le \gamma C(\gamma)\Big(|z|^{\gamma-1}\int_0^{|z|} t^{-\gamma}n(t)\mathrm{d}t + |z|^{\gamma}\int_{|z|}^{\infty} t^{-\gamma-1}n(t)\mathrm{d}t\Big), \qquad (24)$$

where $C(\gamma)=1$ if $\gamma=1$, $C(\gamma)=2(2+\log\gamma)$, for $\gamma>1$.

This follows directly from Theorem 1.11 in [5]. See also Lemma 1.4 in [5] and Theorem 2.6.5 in [1].

Suppose that $n_\infty(t,\chi_A) < ct^\nu$ for $t > t_0$, where $\gamma-1<\nu<\gamma$. Then (24) gives for $|z|=r>t_0$

$$\log^+|\chi_A(z)| \le \gamma C(\gamma)c\Big(\frac{1}{\nu-\gamma+1}+\frac{1}{\gamma-\nu}\Big)r^\nu + \mathcal{O}(r^{\gamma-1})$$

from which we see that the type of $\chi_A(z)$ as a meromorphic function satisfies

$$\tau \le \gamma C(\gamma)c\Big(\frac{1}{\nu-\gamma+1}+\frac{1}{\gamma-\nu}\Big).$$

Moreover the type of $N_\infty(r,(1-zA)^{-1})$ satisfies

$$\tau \le c\nu.$$

Note that if the order of $\chi_A(z)$ is an integer then it need not have the same type-class as $N_\infty(r,(1-zA)^{-1})$. Consider the following example.

Example 4 Let A be a bounded, self-adjoint operator the eigenvalues of which are $\lambda_n=\frac{1}{n}$. By Theorem 5.5 in [19] we know that since A is bounded and self-adjoint

$$m_\infty(r,(1-zA)^{-1}) \le \log 2,$$

so

$$T_\infty(r,(1-zA)^{-1}) = N_\infty(r,(1-zA)^{-1}) + \mathcal{O}(1) = r + \mathcal{O}(1).$$

$T_\infty(r,(1-zA)^{-1})$ grows with order $\omega=1$ and type $\tau_{mer}=1$. The function

$$\chi_A(z) = \prod_{n=1}^{\infty}\Big(1+\frac{z}{n}\Big)\mathrm{e}^{-z/n} \qquad (25)$$

satisfies the requirements of Theorem 1 and from [5, p. 29] we know that this is of order $\omega=1$ and of maximal type. Here the weight functions $\mathrm{e}^{-z/n}$ are necessary for the product 25 to converge. Naturally the weights can be chosen differently. Observe however, that to get an entire function convergent in the complex plane with the zeros $z_n=n$ some weights are needed, and these will cause the function χ_A to grow faster than the resolvent.

We see that $N_\infty(r,(1-zA)^{-1})$ and $T(r,\chi_A)$ will always grow with the same order, so the order of $T(r,\chi_A)$ never exceeds that of $T_\infty(r,(1-zA)^{-1})$. However, it is indeed possible that the orders of $T(r,\chi_A)$ and $T_\infty(r,(1-zA)^{-1})$ are equal (and integer) and the type-class of $T(r,\chi_A)$ exceeds that of $T_\infty(r,(1-zA)^{-1})$.

8 Decay of normalized polynomials

We have so far dealt with the problem of finding a bound for $\inf \|q_j(A)\|$, where q_j is a monic polynomial of degree j. However, applying a Krylov method to the fixed point problem $x = Ax + b$ leads to a sequence of polynomials $\{p_j\}$ such that $\deg p_j = j$, $p_j(1) = 1$ and the error of the iteration on the j^{th} iteration step is $e_j = p_j(A)e_0$, where e_0 is the initial error. So in fact we wish to find bound for the decay of $\inf \|p_j(A)\|$, as $j \to \infty$, where the infimum is taken over all polynomials such that $\deg p_j = j$, $p_j(1) = 1$. Trivially this is sensitive to low rank perturbations, as such perturbations may cause the problem to become singular. Here we show in Theorem 17 that this sensitivity is all due to the spectrum.

Assume $1 - A$ has a bounded inverse and let $q(\lambda)$ be an arbitrary polynomial and set $p(\lambda) := 1 - (1 - \lambda)q(\lambda)$. Then by Proposition 1.6.1 in [15] we have

$$\frac{1}{\|1 - A\|}\|p(A)\| \leq \|(1 - A)^{-1} - q(A)\| \leq \|(1 - A)^{-1}\|\|p(A)\|. \tag{26}$$

Now by Theorem 3 we have

$$(1 - zA)^{-1} = \frac{1}{\chi_A(z)} \sum_{j=0}^{\infty} q_j(A)z^j,$$

so

$$(1 - A)^{-1} = \frac{1}{\chi_A(1)} \sum_{j=0}^{\infty} q_j(A)$$

and we get approximations $\tilde{p}_k(A)$ for $(1 - A)^{-1}$ from this, where

$$\tilde{p}_k(A) = \frac{1}{\chi_A(1)} \sum_{j=0}^{k} q_j(A). \tag{27}$$

The error of this approximation is

$$(1 - A)^{-1} - \tilde{p}_k(A) = \frac{1}{\chi_A(1)} \sum_{j=k+1}^{\infty} q_j(A). \tag{28}$$

Both (27) and (28) divide into two parts in a natural way: the part $\frac{1}{\chi_A(1)}$ is sensitive to low rank perturbations, while the summation part $\sum q_j(A)$ is not. As in Section 7, we can choose for each η a function $\chi_A(z) = \chi_{A,\eta}(z)$ valid for $r \leq \eta$. In fact, fix $1 < \gamma < \eta$ and define

$$c_\gamma(z) = \prod_{k=1}^{n_\gamma}(1 - \frac{z}{z_k}), \tag{29}$$

where $n_\gamma = n_\infty(\gamma, (1 - zA)^{-1})$. Then Theorem 17 gives an upper bound for the decay of the normalized polynomials generated by GMRES, such that the bound is a product of two terms, of which $\frac{\|1-A\|}{|c_\gamma(1)|}$ is sensitive to low rank perturbations, while the other term is essentially determined by the decay of the monic polynomials q_j, which, as discussed in Section 7, is insensitive to low rank perturbations.

Theorem 17 *Let A be a bounded linear operator such that $(1-zA)^{-1}$ is meromorphic for $|z|<\infty$ and assume that for all $r \geq 1$*

$$T_\infty(r,(1-zA)^{-1}) \leq \tau r^\omega.$$

Fix $\alpha > 0$, $\beta > 1$ and $\gamma > 1$. Then there exists a sequence of polynomials p_j such that $p_j(1) = 1$ and for $j \geq \tilde{\tau}\omega$

$$\|p_j(A)\| \leq \frac{\|1-A\|}{|c_\gamma(1)|}\Big(\frac{\tilde{\tau}\mathrm{e}\omega}{j}\Big)^{j/\omega},$$

where

$$\begin{aligned}\tilde{\tau} &= \hat{\tau}(\frac{\gamma(1+\alpha)}{\gamma-1})^{a\omega}, \\ a &= \frac{\tau}{\log\beta}(\frac{\beta}{\alpha})^\omega\end{aligned}$$

and $\hat{\tau} = c(\omega)\tau$ is as in Theorem 15.

Proof. Define $\tilde{p}_j(\lambda) = \frac{1}{\chi_A(1)}\sum_{k=0}^j q_k(\lambda)$ where $q_k(\lambda)$ are as in Theorem 3, and set $p_j(\lambda) = 1-(1-\lambda)\tilde{p}_{j-1}(\lambda)$. From (26) and (28) we have

$$\frac{1}{\|1-A\|}\|p_j(A)\| \leq \|(1-A)^{-1}-\tilde{p}_j(A)\| \leq \frac{1}{|\chi_A(1)|}\|q_j(A)\|. \tag{30}$$

From Theorem 15 we have for $j > \tau\omega c(\omega)$

$$\|q_j(A)\| \leq (\frac{\hat{\tau}\mathrm{e}\omega}{j})^{j/\omega}. \tag{31}$$

We need a lower bound for $|\chi_A(1)|$. Now as in Section 7 we define for $|z| \leq \eta$

$$\chi_{A,\eta}(z) = \frac{B_\eta(z)}{B_\eta(0)} = \prod_{k=1}^n \frac{\eta^2}{z_j}\frac{z_k - z}{\eta^2 - \bar{z}_k z},$$

where $n = n_\infty(\eta,(1-zA)^{-1})$. So

$$\chi_{A,\eta}(1) = \prod_{k=1}^{n_\gamma}\frac{\eta^2}{z_k}\frac{z_k-1}{\eta^2-\bar{z}_k}\prod_{k=n_\gamma+1}^{n}\frac{\eta^2}{z_k}\frac{z_k-1}{\eta^2-\bar{z}_k},$$

where $1<\gamma<\eta$ and $n_\gamma = n_\infty(\gamma,(1-zA)^{-1})$. For $\gamma \leq |z_k| \leq \eta$ we have

$$|1-\frac{1}{z_k}| \geq |1-\frac{1}{\gamma}| = 1-\frac{1}{\gamma}$$

and for $|z_k| \le \eta$

$$|1 - \frac{\bar{z}_k}{\eta^2}| \le 1 + \frac{|\bar{z}_k|}{\eta^2} \le 1 + \frac{1}{\eta},$$

so

$$\prod_{k=n_\gamma+1}^{n} |\frac{\eta^2}{z_k} \frac{z_k - 1}{\eta^2 - \bar{z}_k}| \ge (\frac{1 - \frac{1}{\gamma}}{1 + \frac{1}{\eta}})^{n-n_\gamma},$$

and

$$\prod_{k=1}^{n_\gamma} |\frac{\eta^2}{z_k} \frac{z_k - 1}{\eta^2 - \bar{z}_k}| \ge \prod_{k=1}^{n_\gamma} |1 - \frac{1}{z_k}| (\frac{1}{1 + \frac{1}{\eta}})^{n_\gamma}.$$

Finally we have

$$|\chi_{A,\eta}(1)| \ge (\frac{1 - \frac{1}{\gamma}}{1 + \frac{1}{\eta}})^n \prod_{k=1}^{n_\gamma} |1 - \frac{1}{z_k}|,$$

where $n = n_\infty(\eta, (1 - zA)^{-1})$ and $n_\gamma = n_\infty(\gamma, (1 - zA)^{-1})$. Let $\beta > 1$. Then

$$\begin{aligned} N_\infty(\beta r, (1 - zA)^{-1}) &= \int_0^r \frac{n_\infty(t, (1 - zA)^{-1})}{t} \mathrm{d}t + \int_r^{\beta r} \frac{n_\infty(t, (1 - zA)^{-1})}{t} \mathrm{d}t \\ &\ge n_\infty(r, (1 - zA)^{-1})(\log \beta r - \log r) \\ &= n_\infty(r, (1 - zA)^{-1}) \log \beta \end{aligned}$$

so

$$n_\infty(r, (1 - zA)^{-1}) \le \frac{N_\infty(\beta r, (1 - zA)^{-1})}{\log \beta} \le \frac{T_\infty(\beta r, (1 - zA)^{-1})}{\log \beta} \le \frac{\tau(\beta r)^\omega}{\log \beta}.$$

By choosing $\eta = \frac{j^{1/\omega}}{\alpha}$, $\alpha > 0$, we have

$$n = n_\infty(\eta, (1 - zA)^{-1}) \le \frac{\tau(\beta\eta)^\omega}{\log \beta} \frac{\tau\beta^\omega}{\alpha^\omega \log \beta} j.$$

Moreover, with this choice of η we have

$$1 + \frac{1}{\eta} = 1 + \frac{\alpha}{j^{1/\omega}} \le 1 + \alpha.$$

We have

$$|\chi_{A,\eta}(1)| \ge (\frac{1 - \frac{1}{\gamma}}{1 + \alpha})^{aj} \prod_{k=1}^{n_\gamma} |1 - \frac{1}{z_k}|, \tag{32}$$

where $a = \frac{\tau\beta^\omega}{\alpha^\omega \log \beta}$. Insert (31) and (32) into (30) to complete the proof. □

References

[1] R. P. Boas Jr., *Entire functions*, Academic Press, 1954.

[2] A. Greenbaum and L. N. Trefethen, *GMRES/CR and Arnoldi/Lanczos as matrix approximation problems,* SIAM J. Sci. Comput. 15 (1994), pp. 359–368.

[3] A. Greenbaum, *Iterative methods for solving linear systems,* Frontiers in Applied Mathematics, SIAM, Philadelphia, 1997.

[4] P. Halmos, *Capacity in Banach algebras,* Ind. Univ. Math. J. Vol 20, No. 9, 1971.

[5] W.K. Hayman, *Meromorphic functions*, Oxford mathematical monographs, Oxford University Press, 1964.

[6] M. Huhtanen, *Local operators in Krylov subspaces,* Num. Funct. An. Opt. 17 (1996), pp. 123–142.

[7] M. Huhtanen and O. Nevanlinna, *Minimal decompositions and iterative methods,* Centro Internacional de Matemática, Preprint No 3, 1998.

[8] S. Hyvönen, *Numerical computation of the Nevanlinna characteristic function,* Institut Mittag-Leffler, Report 17, 1997/98.

[9] S. Hyvönen, *Case studies on growth properties of meromorphic resolvents,* Institut Mittag-Leffler, Report 18, 1997/98.

[10] S. Hyvönen, *Polynomial acceleration of the Picard-Lindelöf iteration*, IMA Journal of Numerical Analysis 18(4) (1998), pp. 519–543.

[11] S. Hyvönen, *Growth of resolvents of certain infinite matrices,* Helsinki University of Technology, Institute of Mathematics, Research Reports A402, 1998.

[12] J. Malinen, *On properties for iteration of a compact operator with unstructured perturbation,* Helsinki University of Technology, Institute of Mathematics, Research Report A360, 1996.

[13] A. I. Markushevich, *Theory of functions of a complex variable*, Volume 3, Chelsea Publishing Company, New York, 1977.

[14] V. Müller, *Local behaviour of the polynomial calculus of operators,* J. reine angew. Math. 430 (1992), pp. 61–68.

[15] O. Nevanlinna, *Convergence of iterations for linear equations*, Birkhäuser Verlag, Basel, 1993.

[16] O. Nevanlinna, *Convergence of Krylov methods for sums of two operators,* BIT 36(4) (1996), pp. 775–785.

[17] O. Nevanlinna, *A characteristic function for matrix valued meromorphic functions,* XVIth Rolf Nevanlinna Colloquium, Eds.: Laine/Martio, Walter Gruyter and Co., Berlin (1996), pp. 171–179.

[18] O. Nevanlinna, *On the growth of the resolvent operators for power bounded operators,* Banach Center Publications, Vol. 38 (1997), pp. 247–264.

[19] O. Nevanlinna, *Growth of operator valued meromorphic functions*, Helsinki University of Technology, Institute of Mathematics, Research Reports A389, 1997.

[20] R. Nevanlinna, *Zur Theorie der meromorphen Funktionen*, Acta Math. 46 (1925), pp. 1–99.

[21] R. Nevanlinna, *Analytic functions,* Springer-Verlag, 1970.

[22] S. Segal, *Nine introductions in complex analysis*, North-Holland Publishing Company, 1981.

[23] C. H. Toh, *GMRES vs. ideal GMRES,* SIAM J. Matrix Anal. Appl., 18 (1997), pp. 30–36.

(continued from the back cover)

A395 Matti Lassas and Erkki Somersalo
Analysis of the PML Equations in General Convex Geometry, 1998

A393 Jukka Tuomela and Teijo Arponen
On the numerical solution of involutive ordinary differential equation systems, 1998

A392 Hermann Brunner, Arvet Pedas, Gennadi Vainikko
The Piecewise Polynomial Collocation Method for Nonlinear Weakly Singular Volterra Equations, 1997

A391 Kari Eloranta
The bounded Eight-Vertex Model, 1997

A390 Kari Eloranta
Diamond Ice, 1997

A389 Olavi Nevanlinna
Growth of Operator Valued Meromorphic Functions, 1997

A388 Jukka Tuomela
On the Resolution of Singularities of Ordinary Differential Systems, 1997

A387 Gennadi Vainikko
Fast Solvers of the Lippman-Schwinger Equation, 1997

A386 Ville Turunen and Gennadi Vainikko;
On Symbol Analysis of Periodic Pseudodifferential Operators, 1997

A385 Jarmo Malinen
Minimax Control of Distributed Discrete Time Systems through Spectral Factorization; 1997

A384 Jarmo Malinen
Well-Posed Discrete Time Linear Systems and Their Feedbacks, 1997

A383 Kalle Mikkola
On the Stable H^2 and H∞ Infinite-Dimensional Regular Problems and Their Algebraic Riccati Equations, 1997

A382 Nikolai Volkov
Algebraizations of the First Order Logic, 1997

HELSINKI UNIVERSITY OF TECHNOLOGY INSTITUTE OF MATHEMATICS
RESEARCH REPORTS

The list of reports is continued inside. Electronical versions of the reports are available at *http://www.math.hut.fi/reports/* .

A403 Saara Hyvönen and Olavi Nevanlinna
Robust bounds for Krylov method, Nov 1998

A402 Saara Hyvönen
Growth of resolvents of certain infinite matrice, Nov 1998

A399 Otso Ovaskainen
Asymptotic and Adaptive Approaches to thin Body Problems in Elasticity

A398 Jukka Liukkonen
Uniqueness of Electromagnetic Inversion by Local Surface Measurements, Aug 1998

A397 Jukka Tuomela
On the Numerical Solution of Involutive Ordinary Differential Systems, 1998

A396 Clement Ph., Gripenberg G. and Londen S-O
Hölder Regularity for a Linear Fractional Evolution Equation, 1998

ISBN 951-22-4319-9
ISSN 0784-3143
Edita, Espoo, 1998