COMPUTING THE SPECTRUM AND REPRESENTING THE RESOLVENT

Olavi Nevanlinna



TEKNILLINEN KORKEAKOULU TEKNISKA HÖGSKOLAN HELSINKI UNIVERSITY OF TECHNOLOGY TECHNISCHE UNIVERSITÄT HELSINKI UNIVERSITE DE TECHNOLOGIE D'HELSINKI

COMPUTING THE SPECTRUM AND REPRESENTING THE RESOLVENT

Olavi Nevanlinna

Helsinki University of Technology Faculty of Information and Natural Sciences Department of Mathematics and Systems Analysis **Olavi Nevanlinna**: *Computing the spectrum and representing the resolvent*; Helsinki University of Technology Institute of Mathematics Research Reports A568 (2009).

Abstract: We discuss computing the spectrum of a bounded operator and representing its resolvent operator. The results include a general convergence theorem for the polynomial convex hull of the spectrum and explicit representations for the resolvent outside. The results are formulated and proved in general Banach algebras.

AMS subject classifications: 47A10, 47A12, 47A66

Keywords: resolvent operator, spectrum, polynomial numerical hull, numerical range, semidecidable, quasialgebraic operator, residue calculus, perturbation

Correspondence

Olavi Nevanlinna Helsinki University of Technology Department of Mathematics and Systems Analysis P.O. Box 1100 FI-02015 TKK Finland

olavi.nevanlinna@tkk.fi

I dedicate this to Gennadi Vainikko

ISBN 978-951-22-9850-1 (print) ISBN 978-951-22-9851-8 (PDF) ISSN 0784-3143 (print) ISSN 1797-5867 (PDF)

Helsinki University of Technology Faculty of Information and Natural Sciences Department of Mathematics and Systems Analysis P.O. Box 1100, FI-02015 TKK, Finland email: math@tkk.fi http://math.tkk.fi/

1 Introduction

If $A \in M_d(\mathbb{C})$ is a complex $d \times d$ matrix, and $\sigma(A)$ denotes its eigenvalues, then the *resolvent* of A is the analytic $M_d(\mathbb{C})$ -valued function

$$\lambda \mapsto (\lambda I - A)^{-1} \tag{1.1}$$

defined in $\mathbb{C} \setminus \sigma(A)$. The resolvent is usually given in one of the following three ways.

The series

$$(\lambda I - A)^{-1} = \sum_{j=0}^{\infty} A^j \lambda^{-1-j}$$
(1.2)

is simple to state but it converges only for

$$|\lambda| > \rho(A) = \max_{\lambda \in \sigma(A)} |\lambda|.$$

Secondly, assuming that the *characteristic polynomial* π_A is available one can form a polynomial q in two variables, see (1.14) below, so that

$$(\lambda I - A)^{-1} = \frac{q(\lambda, A)}{\pi_A(\lambda)},\tag{1.3}$$

which is valid everywhere. However, this is practical only if the dimension d is moderate. The third obvious possibility is to work out at some discrete points λ_i the matrix (or a numerical approximation to it)

$$R(\lambda_j) = (\lambda_j I - A)^{-1}$$

and then write for λ near λ_j

$$(\lambda I - A)^{-1} = \sum_{k=0}^{\infty} R(\lambda_j)^{k+1} (\lambda_j - \lambda)^k.$$
(1.4)

However, then we need to work with several different local representations and, for example, Cauchy integrals do not reduce to residue calculus.

In this paper we give a construction which produces smaller and smaller inclusion sets for the spectrum in such a way that outside of each such inclusion set a single explicit representation for the resolvent converges. The construction is formulated in general Banach algebras.

Let \mathcal{A} be a complex Banach algebra with unit. We shall write all formulas without explicitly expressing the unit in the algebra, e.g. we write $(\lambda - a)^{-1}$ to denote the resolvent.

We assume that for given $a \in \mathcal{A}$ we can compute polynomials of it, and norms of the polynomials generated. Thus, all our calculations stay in the subalgebra generated by the element and thus the spectrum we can compute is at best w.r.t to this subalgebra. In particular, the resolvent exists in the subalgebra only outside the polynomially convex hull of the spectrum $\sigma(a)$, that is for $\lambda \notin \hat{\sigma}(a)$. Recall, that the polynomially convex hull of a compact set $K \subset \mathbb{C}$ is

$$\hat{K} = \{ z \in K : |p(z)| \le \max_{w \in K} |p(w)| \text{ holds for all polynomials} \}.$$

We denote by \mathbb{P} the set of polynomials with complex coefficients. Given $p \in \mathbb{P}$ and $a \in \mathcal{A}$ we put

$$V_p(a) = \{ \lambda \in \mathbb{C} : |p(\lambda)| \le ||p(a)|| \}.$$
 (1.5)

In [11], we defined the polynomial numerical hull of a as

$$V(a) = \bigcap_{p \in \mathbb{P}} V_p(a) \tag{1.6}$$

and showed that it always equals the polynomial convex hull of the spectrum:

$$V(a) = \hat{\sigma}(a). \tag{1.7}$$

Further, if we denote by \mathbb{P}_k polynomials of degree at most k, then setting

$$V^{k}(a) = \bigcap_{p \in \mathbb{P}_{k}} V_{p}(a)$$
(1.8)

we obtain the polynomial numerical hull of degree k. These form a nested sequence $V^{k+1}(a) \subset V^k(a)$ of compact sets such that

$$V(a) = \bigcap_{k} V^{k}(a).$$

In [11] this was discussed for bounded operators in Banach spaces. If A is a bounded operator in a Hilbert space, then $V^1(A)$ agrees with the closure of the numerical range W(A) of A, [11]

$$V^1(A) = cl \ W(A).$$

On polynomial numerical hulls see additionally [1], [2], [3], [4], [5], [6], [7], [12].

One immediate consequence of (1.7) is that we can formulate a simple (but slow) procedure to find out whether $\lambda \notin \hat{\sigma}(a)$. In fact, enumerate all monic polynomials with coefficients having rational real and imaginary parts and check whether

$$|p(\lambda)| > ||p(a)||.$$
(1.9)

If $\lambda \notin \hat{\sigma}(a)$, then eventually there comes a polynomial p_j such that (1.9) holds, and you get an answer. On the other hand, if $\lambda \in \hat{\sigma}(a)$, then (1.9) never holds and you never get an answer. Situations like this are sometimes called *semidecidable*:

Proposition 1.1. There exists a sequence of monic polynomials $\{p_k\}$ with the following property. If \mathcal{A} is a Banach algebra with unit and $a \in \mathcal{A}$ and $\lambda \in \mathbb{C}$ are given, then there exists j such that

$$|p_j(\lambda)| > ||p_j(a)||$$
 (1.10)

if and only if $\lambda \notin \hat{\sigma}(a)$.

In this paper we aim to show that we can in an effective way construct a nested sequence of compact sets K_k such that

$$\hat{\sigma}(a) = \bigcap_{k} K_k \tag{1.11}$$

and for every k an explicit representation for the resolvent in $\mathbb{C} \setminus K_k$.

Given an arbitrary monic polynomial $p \in \mathbb{P}_d$

$$p(\lambda) = \lambda^d + a_1 \lambda^{d-1} + \dots + a_d \tag{1.12}$$

we put for j = 0, 1, ..., d - 1

$$Q_j(\lambda) = \lambda^j + a_1 \lambda^{j-1} + \dots + a_j \tag{1.13}$$

and then, given $a \in \mathcal{A}$

$$q(\lambda, a) = \sum_{j=0}^{d-1} Q_{d-1-j}(\lambda) a^j.$$
 (1.14)

One checks easily that then

$$(\lambda - a)q(\lambda, a) = p(\lambda) - p(a).$$
(1.15)

Suppose now that λ is such that (1.9) holds. Then $p(\lambda) - p(a)$ has an inverse in the form of a convergent power series and we obtain from (1.15) a representation for the resolvent in an explicit form:

$$(\lambda - a)^{-1} = \frac{q(\lambda, a)}{p(\lambda)} \sum_{j=0}^{\infty} \frac{p(a)^j}{p(\lambda)^j} \text{ for all } \lambda \notin V_p(a).$$
(1.16)

We start with the following result.

Theorem 1.2. Given an element a in a Banach algebra \mathcal{A} and an open set U containing $\hat{\sigma}(a)$ there exists a polynomial p such that $\hat{\sigma}(a) \subset V_p(a) \subset U$. In particular, the representation (1.16) holds in $\mathbb{C} \setminus U$.

This result is weak in two ways. It assumes the knowledge of the spectrum and it does not indicate how to find the representing polynomial. Our main result is summarized in the following result. **Theorem 1.3.** There exists a procedure such that, given an element a in a Banach algebra \mathcal{A} it produces a sequence of compact sets K_k and polynomials p_k satisfying the following: $K_{k+1} \subset K_k$, $V_{p_k}(a) \subset K_k$ and

$$\hat{\sigma}(a) = \bigcap_{k} K_k. \tag{1.17}$$

In particular, the representation (1.16) holds in $\mathbb{C} \setminus K_k$ with $p = p_k$.

The computation of K_k in Theorem 1.3 assumes a finite number of minimization problems to be carried out. Theorem 1.2 is proved in Section 2 and the construction of sets and polynomials in Theorem 1.3 in Sections 3 to 5. Section 6 discusses "stagnation": the polynomial convex hull of the spectrum has been found but one does not know it. To the end we make two remarks on the representation: one on low rank perturbation theory and another on holomorphic functional calculus. Preliminary versions of these ideas were reported in [14].

2 Proof of Theorem 1.2 and convergence of the representations

The first part of Theorem 1.2 is the following.

Lemma 2.1. If for a given $a \in \mathcal{A}$ one knows that an open $U \subset \mathbb{C}$ is such that $\hat{\sigma}(a) \subset U$, then there exists a polynomial p such that

$$\hat{\sigma}(a) \subset V_p(a) \subset U. \tag{2.1}$$

Proof. If $a \in \mathcal{A}$ is algebraic, then there exists p such that p(a) = 0 and $V_p(a) = \sigma(a)$. So, suppose a is not algebraic so that $p(a) \neq 0$ for all nonzero polynomials. By Hilbert Lemniscate Theorem, see Theorem 5.5.8 in [17], there exists a polynomial q such that

$$|q(\lambda)| > \sup_{z \in \sigma(a)} |q(z)| \text{ for } \lambda \in \mathbb{C} \setminus U.$$

But then there exists $\delta > 0$ such that

$$|q(\lambda)| > (1+\delta) \sup_{z \in \sigma(a)} |q(z)| \text{ for } \lambda \in \mathbb{C} \setminus U.$$
(2.2)

By the spectral radius formula

$$\sup_{z\in\sigma(a)}|q(z)| = \lim_{k\to\infty} \|q(a)^k\|^{1/k}$$

and thus for large enough m we have

$$\sup_{z \in \sigma(a)} |q(z)| > \frac{1 + \delta/2}{1 + \delta} ||q(a)^m||^{1/m}.$$
(2.3)

Combining (2.2) and (2.3) we obtain

$$|q(\lambda)^m| > (1 + \delta/2)^m ||q(a)^m||$$
 for $\lambda \in \mathbb{C} \setminus U$.

Thus we have (2.1) with $p = q^m$.

Theorem 1.2 follows from Lemma 2.1 as the series $\sum \frac{\|p(a)^k\|}{|p(\lambda)^k|}$ converges in $\mathbb{C} \setminus U$. In fact, in the notation of the proof we have

$$\sum \frac{\|p(a)^k\|}{|p(\lambda)^k|} < \frac{1}{1 - (1 + \delta/2)^{-m}}.$$

It is of interest to study more careful the convergence of the series. Let us put

$$R_m(\lambda, a, p) = \frac{q(\lambda, a)}{p(\lambda)} \sum_{k=0}^m \frac{p(a)^k}{p(\lambda)^k}$$
(2.4)

as a rational approximation to the resolvent and ask about the error

$$\|(\lambda - a)^{-1} - R_m(\lambda, a, p)\|$$
(2.5)

as $m \to \infty$. We shall see that the error can be controlled by the Green's function for the set $\mathbb{C} \setminus V_p(a)$.

Definition 2.2. Given a nonempty compact $K \subset \mathbb{C}$ such that $G = \mathbb{C} \setminus K$ is connected we call g_G a Green's function for G if

$$g_G$$
 is harmonic in G ,
 $g_G(\lambda) = \log |\lambda| + O(1)$ as $\lambda \to \infty$

and

$$g_G(\lambda) \to 0 \quad as \ \lambda \to \zeta \quad for \ n.e. \ \zeta \in \partial G.$$

For $G = \mathbb{C} \setminus K$ such a Green's function exists if and only if K has positive capacity, cap(K) > 0. The function is unique, see [17] and the capacity satisfies

$$\log cap(K) = \log |\lambda| - g_G(\lambda) + o(1)$$
 as $\lambda \to \infty$

Lemma 2.3. Let p be a monic polynomial of degree d, and suppose $p(a) \neq 0$. Then the Green's function for $\mathbb{C} \setminus V_p(a)$, denoted here by $g(\lambda, a, p)$ is

$$g(\lambda, a, p) = \frac{1}{d} \log \frac{|p(\lambda)|}{\|p(a)\|}.$$
(2.6)

In particular, the capacity of $V_p(a)$ is

$$cap(V_p(a)) = ||p(a)||^{1/d}.$$
 (2.7)

Proof. Since $|p(\lambda)| > ||p(a)||$ outside of $V_p(a)$, $g(\lambda, p, a)$ is there harmonic. It vanishes along $\partial V_p(a)$ and

$$g(\lambda, p, a) = \log|\lambda| + \frac{1}{d}\log\frac{1}{\|p(a)\|} + o(1) \quad \text{for } \lambda \to \infty.$$
(2.8)

Consider now estimating the error in (2.5). The truncation of the sum satisfies

$$\left\|\sum_{k=m+1}^{\infty} \frac{p(a)^{k}}{p(\lambda)^{k}}\right\| \le \frac{e^{-(m+1)dg(\lambda,a,p)}}{1 - e^{-dg(\lambda,a,p)}}$$

We need also the following.

Lemma 2.4. Assume $p(a) \neq 0$. The function

$$R_0(\lambda, a, p) = \frac{q(\lambda, a)}{p(\lambda)}$$
(2.9)

is rational of degree d and satisfies for $\lambda \in \mathbb{C} \setminus V_p(a)$ the following estimates:

$$||R_0(\lambda, a, p)|| \le 2||(\lambda - a)^{-1}||$$
 (2.10)

and

$$||R_0(\lambda, a, p)|| \le C \tag{2.11}$$

where

$$C = \max_{\lambda \in V_p(a)} \frac{\|q(\lambda, a)\|}{\|p(a)\|}.$$
(2.12)

Proof. The first estimate (2.10) follows immediately from (1.15) and from the fact that for $\lambda \notin V_p(a)$

$$\|1 - \frac{p(a)}{p(\lambda)}\| \le 2.$$

To conclude (2.11) notice that R_0 is analytic and bounded in $\mathbb{C} \setminus V_p(a)$ and by maximum principle its maximum value is obtained on the boundary. However, on the boundary $|p(\lambda)| = ||p(a)||$ so all we really need, is the maximum of the polynomial q in $V_p(a)$.

We formulate still another lemma for later reference.

Lemma 2.5. Assume $p(a) \neq 0$. For $\theta > 1$ put

$$V_p(a,\theta) = \{\lambda \in \mathbb{C} : |p(\lambda)| \le \theta ||p(a)||\}.$$
(2.13)

Then for $\lambda \notin V_p(a, \theta)$

$$\|(\lambda - a)^{-1}\| \le C \ \frac{\theta}{\theta - 1},\tag{2.14}$$

where C is given in (2.12).

We can now summarize the convergence discussion as follows.

Theorem 2.6. Let an element a be given in a Banach algebra \mathcal{A} , p be monic of degree d such that $p(a) \neq 0$, q be given by (1.14) and R_m by (2.4). Then we have

$$\|(\lambda - a)^{-1} - R_m(\lambda, a, p)\| \le M(\lambda) \frac{\eta(\lambda)^{m+1}}{1 - \eta(\lambda)} \quad for \ \lambda \in \mathbb{C} \setminus V_p(a), \quad (2.15)$$

where

$$M(\lambda) = \min\{\max_{\lambda \in V_p(a)} \frac{\|q(\lambda, a)\|}{\|p(a)\|}, 2\|(\lambda - a)^{-1}\|\}$$
(2.16)

and

$$\eta(\lambda) = \frac{\|p(a)\|}{|p(\lambda)|} = e^{-dg(\lambda, p, a)}.$$
(2.17)

Remark 2.7. Notice that both $\eta(\lambda)$ and $M(\lambda)$, or more precisely C in (2.12), can be computed based on knowing p and a. Thus no knowledge of the location of the spectrum enters as an assumption.

Remark 2.8. Notice that if we normalize the convergence speed wrt evaluations of the element, which is natural in particular when dealing with operators operating on vectors, $\eta(\lambda)$ is not the true convergence factor but rather $\eta(\lambda)^{1/d} = e^{-g(\lambda, a, p)}$ as evaluation of p requires d operations with an operator on a vector. In any case the speed is *linear* and generally *not superlinear* even if the element a would be quasialgebraic. It would be superlinear only in the unlikely case that $\sigma(a)$ is finite and p would happen to vanish on it exactly.

Remark 2.9. Recall that *a* is quasialgebraic if $cap(\sigma(a)) = 0$, [8], [16]. For example, all compact operators are quasialgebraic. For such elements Theorem 1.3 provides polynomials for which the capacity of $V_p(a)$ becomes arbitrarily small. Likewise, the convergence factor $e^{-g(\lambda,a,p)}$ can be brought arbitrarily close to 0. On the other hand, if $cap(\sigma(a)) > 0$, then the Green's function for $\mathbb{C} \setminus \hat{\sigma}(a)$ provides the lower bound for the convergence factor with any polynomial *p*, but on the other hand by again Theorem 1.3 provides polynomials for which the speed becomes arbitrarily close to the lower bound.

3 Convergence of an idealized Arnoldi-type algorithm

We shall start now discussing procedures with aim to produce polynomials p_j such that

$$\hat{\sigma}(a) = \bigcap_{j} V_{p_j}(a).$$

The major difficulty is of course that in the beginning of the computation $\hat{\sigma}(a)$ is not known. We first consider a sequence which imitates the Arnoldi

process. These polynomials can be thought of as solutions of minimization problems of the form

$$\min_{q\in\mathbb{P}_{j-1}}\|a^j-q(a)\|.$$

It is worth of reminding how the standard Arnoldi method works, here formulated in a Hilbert space H. Given a bounded operator $A \in B(H)$ and a vector $b \in H$ one creates polynomials p_j of degree j such that $\{p_j(A)b\}$ becomes an orthogonal basis for the Krylov space

$$\mathcal{K}(A,b) = cl \ span_{j\geq 0}\{A^jb\}.$$
(3.1)

Denote the restriction of A to the invariant subspace $\mathcal{K}(A, b)$ by $A_{[b]}$. We have always

$$\hat{\sigma}(A_{[b]}) \subset \hat{\sigma}(A)$$

and the set of vectors b for which the inclusion is equality is of second category, [10], [11]. In standard Arnoldi process one approximates the spectrum of $A_{[b]}$ by considering the accumulation of the zeros of the polynomials p_j .

There are two kinds of problems with this procedure.

(i) The zeros can accumulate outside of $\hat{\sigma}(A_{[b]})$,

and

(ii) parts of the boundary of $\hat{\sigma}(A_{[b]})$ can stay away form the zeros.

Example 3.1. Suppose K is a compact set of the complex plane and X is the space of bounded complex functions on K with supremum-norm. If A denotes the multiplication operator

$$(Af)(z) = zf(z)$$

then clearly $\sigma(A) = K$. Consider $K = [-2, -1] \cup [1, 2]$ and suppose b(z) = 1. Then the polynomials minimizing $||p_j(A)b||$ over all monic polynomials of degree j have a zero at origin for every odd j. Such an example can naturally also be formulated in a Hilbert space.

Example 3.2. Consider the unitary shift S in l_2 with initial vector $b = e_0$. Then $S_{[e_0]}$ is the forward shift with unit disc as spectrum, the polynomials are simply λ^j with zeros staying at the origin while the spectrum of S is the unit circle.

Example 3.3. In the previous example the initial vector e_0 is in the kernel of $(S_{[e_0]})^*$ while $||S_{[e_0]}x|| = ||x||$. This can happen only when the operator is not quasitriangular. Recall that A is called quasitriangular if there exists a sequence $\{P_n\}$ of finite rank projections converging strongly to identity, $P_nX \subset P_{n+1}X$ and such that

$$||AP_n - P_n AP_n|| \to 0.$$

Iteration operators in the waveform relaxation or Picard-Lindelöf iteration are examples of operators which are not quasitriangular, if considered as operating on infinite intervals. More on this, see [9]. **Remark 3.4.** If $A_{[b]}$ is quasialgebraic, then (i) cannot happen, if one selects a suitable subsequence of polynomials. In fact, if $h_{j\ k}$ denotes the elements of $A_{[b]}$ when represented in Hessenberg form using the orthogonal basis created, then

$$(\prod_{1}^{n} h_{j \ j-1})^{1/n} \to 0$$

and in particular, there exists a subsequence such that

$$h_{j_k \ j_k-1} \to 0,$$

see Theorem 4.3, [12]. If P_n denotes the orthogonal projection onto

$$span\{b, Ab, \ldots, A^{n-1}b\}$$

then

$$h_{n \ n-1} = \|P_n A_{[b]} P_n - A_{[b]} P_n\|$$

and we see that a quasialgebraic $A_{[b]}$ is quasitriangular. If now $\mu_n \sigma(P_n A_{[b]} P_n)$ and if a subsequence $\mu_{n_m} \to \mu_0$ while $h_{n_m n_m-1} \to 0$ then $\mu_0 \in \sigma(A_{[b]})$. In fact, if $x_n = P_n x_n$ is of unit length and such that

$$P_n A_{[b]} P_n x_n = \mu_n x_n$$

then

$$||Ax_n - \mu_0 x_n|| \le h_{n \ n-1} + |\mu_n - \mu_0|$$

implies that μ_0 is in the approximate point spectrum of $A_{[b]}$ and hence in $\sigma(A)$ as well. More on this, see [12], [15].

For theoretical reasons there has been some interest also on an *idealized* Arnoldi method. At each j compute the monic polynomial of degree j minimizing the operator norm ||p(A)||. It is clear from the examples above that this method suffers from the same weaknesses as the true Arnoldi, that is both (i) and (ii) can happen. Our first result concerns a modification of this idealized Arnoldi so that the true outer boundary of the spectrum becomes uniformly close to the outside of the inclusion set constructed. The discussion is formulated for general Banach algebras and it applies to the idealized Arnoldi as such, taking the algebra to be B(H).

Theorem 3.5. Let \mathcal{A} be a Banach algebra and $a \in \mathcal{A}$ be given. Let p_j be a monic polynomial such that for all monic polynomials p of degree j one has

$$||p_j(a)|| \le ||p(a)||. \tag{3.2}$$

Denoting

$$Z = \bigcap_{j>0} V_{p_j}(a)$$
$$Z_n = \bigcap^n V_{p_j}(a),$$

and

i=1

we have

$$\hat{\sigma}(a) \subset Z \tag{3.3}$$

and

$$\sup_{\lambda \in \partial \hat{\sigma}(a)} dist(\lambda, \ \mathbb{C} \setminus Z_n) \to 0 \ as \ n \to \infty.$$
(3.4)

In order to prove the theorem we need a simple lemma.

Lemma 3.6. Let $E, F \subset \mathbb{C}$ be compact sets such that $E = \hat{E}, F = \hat{F}$ and $E \cap F = \emptyset$. If cap(F) > 0 then

$$cap(E) < cap(E \cup F). \tag{3.5}$$

Proof. If E is polar and cap(E) = 0, then (3.5) is trivial. Otherwise, let g be the Green's function for $\mathbb{C} \setminus E$ and h that for $\mathbb{C} \setminus (E \cup F)$, with logarithmic singularities at ∞ . Then u = g - h is harmonic in $\mathbb{C} \setminus (E \cup F)$ and bounded near ∞ , hence harmonic and bounded in $\mathbb{C}_{\infty} \setminus (E \cup F)$. Since $E \cap F = \emptyset$ we conclude that $u \ge 0$ on $\mathbb{C} \setminus (E \cup F)$ and in particular

$$\lim_{\lambda \to \zeta} u(\lambda) > 0 \text{ for n.e. } \zeta \in \partial F.$$
(3.6)

If $cap(E) = cap(E \cup F)$, then $u(\lambda) \to 0$ as $\lambda \to \infty$. As u is harmonic this would imply that $u(\lambda) = 0$ in all of $\mathbb{C} \setminus (E \cup F)$ contradicting (3.6). \Box

Proof. With the help of the lemma we can now prove Theorem 3.5. Note first that (3.3) follows immediately from the fact that each $V_p(a)$ is an inclusion set for $\hat{\sigma}(a)$. Suppose now that (3.2) holds. Since the capacity is a monotone set function we have

$$cap(\hat{\sigma}(a)) \le cap(Z) \le cap(V_{p_i}(a))$$

However, a result of Halmos, [8] implies that

$$cap(\hat{\sigma}(a)) = \liminf \|p_j(a)\|^{1/j}$$

and therefore

$$cap(\hat{\sigma}(a)) = cap(Z) \tag{3.7}$$

If (3.4) would not hold, by compactness of boundary of $\hat{\sigma}(a)$ there would exist a $\lambda_0 \in \partial \hat{\sigma}(a)$ such that

$$dist(\lambda_0, \mathbb{C} \setminus Z) > 0,$$

implying the existence of a small closed line segment L inside $Z \setminus \hat{\sigma}(a)$. Taking L to be F and $\hat{\sigma}(a)$ to be E in the previous lemma, we get

$$cap(\hat{\sigma}(a)) < cap(\hat{\sigma}(a) \cup L) \le cap(Z)$$

contradicting (3.7).

4 A convergent algorithm

The Arnoldi type algorithm discussed in the last section seems to be efficient in practise, extreme points of the outer boundary of the spectrum and the boundaries of $V_{p_j}(a)$ are close already in early stages of the iteration. However, Theorem 3.5 leaves open the possibility that $Z \setminus \hat{\sigma}(a)$ could be nonempty. Isolated points in $Z \setminus \hat{\sigma}(a)$ can easily be detected using the representations for the resolvent but we prefer here to present a different type of approach for detecting and eliminating $Z \setminus \hat{\sigma}(a)$.

Let p_j be as before a monic polynomial of degree j which minimizes the norm ||p(a)|| over all such polynomials, and

$$V_p(a) = \{\lambda \in \mathbb{C} : |p(\lambda)| \le ||p(a)||\}.$$

We create a sequence of compact sets W_n such that the sequence is decreasing, $W_{n+1} \subset W_n$, and the following holds

$$\hat{\sigma}(a) \subset W_n \subset Z_n = \bigcap_{j=1}^n V_{p_j}(a) \tag{4.1}$$

and

$$\hat{\sigma}(a) = \bigcap_{n>0} W_n. \tag{4.2}$$

Description of the algorithm

Initially, put $W_1 = V_{p_1}(a)$.

Cover W_j with open balls $B_{j,k}$ of radius ||a||/j. By compactness of W_j this can be done with a finite number of balls. In each such ball B compute a polynomial $p = p_{j,k}$ which maximizes, over all monic polynomials of at most degree j the following expression

$$\frac{|p|_B}{\|p(a)\|}\tag{4.3}$$

where

$$|p|_B = \min_{\lambda \in \partial B} |p(\lambda)|.$$

Let J(j) contain all k's for which

$$|p_{j,k}|_{B_{j,k}} > ||p_{j,k}(a)||.$$

Then set

$$W_{j+1} = W_j \cap V_{p_j}(a) \cap \bigcap_{k \in J(j)} V_{p_{j,k}}(a).$$
(4.4)

Notice that by construction, $B_{j,k} \cap V_{p_{j,k}}(a) = \emptyset$ whenever $k \in J(j)$.

Theorem 4.1. The nested sequence of compact sets $\{W_j\}$ satisfies

$$\hat{\sigma}(a) = \bigcap_{j>0} W_j. \tag{4.5}$$

Proof. By construction $\hat{\sigma}(a) \subset W_{j+1} \subset W_j$. If (4.5) would not hold, there would exist $\mu \in \bigcap_{j>0} W_j$ such that $\mu \notin \hat{\sigma}(a)$. But since $\hat{\sigma}(a) = V(a)$, there would exist p such that for $\lambda = \mu$

$$|p(\lambda)| > ||p(a)||.$$
(4.6)

By continuity of p we there exists an open neighborhood U such that (4.6) holds for all $\lambda \in U$. Since $\mu \in W_j$ for all j, there exists for every j an index k(j) such that

 $\mu \in B_{j,k(j)}.$

As the radii of the balls become arbitrarily small, and U is a fixed open neighborhood of μ , there exists an index j_0 such that for $j \ge j_0$ the closures of the balls $B_{j,k(j)}$ are inside of U. As soon as $j \ge \max\{j_0, \deg(p)\}$ where pis the polynomial which is large on U we get for $B = B_{j,k(j)}$

$$\max_{q \in \mathbb{P}_j} \frac{|q|_B}{\|q(a)\|} \ge \frac{|p|_B}{\|p(a)\|} > 1$$

which means that $k(j) \in J(j)$ and by (4.4) W_{j+1} no longer contains μ which is a contradiction.

Remark 4.2. It is clear from the proof that the sizes of the balls, and the degrees of polynomials p_j and $p_{j,k}$ can be modified. One can monitor the size of isolated components in Z_n and use the safeguard procedure adaptively and not at every iteration round. In the other extreme, one can bypass the computation of p_j 's completely, by starting with $W_1 = \{\lambda \in \mathbb{C} : |\lambda| \leq ||a||\}$ and computing only $p_{j,k}$'s.

5 Construction of a representing polynomial sequence

We shall complete our proof of Theorem 1.3 by showing how one can, based on a nested sequence $\{W_j\}$ satisfying (4.5) compute monic polynomials p_j , generally of degree d_j , and another nested sequence K_j such that

$$V_{p_i}(a) \subset K_j$$

and such that (1.17) holds. Assume W_j has been computed as in the last section, using e.g. the procedure described in the last section. Put

$$K_j = \{\lambda \in \mathbb{C} : dist(\lambda, W_j) \le ||a||/j\}.$$
(5.1)

We shall consider *Fekete polynomials* for the set W_j . Recall that f_n is a Fekete polynomial for a compact set K of degree n if

$$f_n(z) = \prod_{j=1}^n (z - z_j)$$
(5.2)

where z_j are such that the expression

$$\delta_n = (\prod_{j < k} |w_j - w_k|)^{2/n(n-1)}$$

with $z_j = w_j$ is maximal over all choices satisfying $w_j \in K$. This construction forces the zeros of the polynomials to lie in the desired set.

Let now $f = f_{j,n(j)}$ be a Fekete polynomial of smallest degree n = n(j) such that

$$\{\lambda \in \mathbb{C} : |f(\lambda)| \le (1+\frac{1}{n}) \max_{z \in W_j} |f(z)|\} \subset K_j.$$

Such a polynomial always exists, see Theorem 5.5.8 in [17]. By the spectral mapping theorem and $\hat{\sigma}(a) \subset W_j$ we have

$$\|f_{j,n(j)}(a)^m\|^{1/m} \to \max_{\lambda \in \hat{\sigma}(a)} |f_{j,n(j)}(\lambda)| \le \max_{\lambda \in W_j} |f_{j,n(j)}(\lambda)|.$$

Choose m large enough so that

$$||f_{j,n(j)}(a)^{m}|| \le (1 + \frac{1}{n(j)})^{m} \max_{\lambda \in W_{j}} |f_{j,n(j)}(\lambda)^{m}|,$$

then with $p_j = f_{j,n(j)}^m$ we have

$$V_{p_j}(a) \subset \{\lambda : |p_j(\lambda)| \le (1 + \frac{1}{n(j)})^m \max_{\lambda \in W_j} |f_{j,n(j)}(\lambda)^m|\} \subset K_j.$$

This completes the proof of Theorem 1.3 setting.

6 Finite termination and stagnation

Recall that $a \in \mathcal{A}$ is called *algebraic* if there exists a monic polynomial p such that p(a) = 0; if such a p is of lowest degree it is called the *minimal* polynomial and a is said to be algebraic of degree deg(p).

Consider the sequence of monic polynomials $\{p_j\}$ as in Theorem 3.4 where p_j satisfies

$$||p_j(a)|| \le ||p(a)|| \tag{6.1}$$

for monic polynomials p of degree j. As we can monitor $||p_j(a)||$ we can *terminate* the algorithm as soon as $p_j(a) = 0$. Thus we can formulate

Proposition 6.1. The sequence $\{p_j\}$ contains the minimal polynomial if and only if a is algebraic. If a is of degree d then $p_d(a) = 0$, and

$$\sigma(a) = V_{p_d}(a)$$

and

$$(\lambda - a)^{-1} = R_0(\lambda, a, p_d).$$

Suppose now that $a \in \mathcal{A}$ is not algebraic, that is ||p(a)|| > 0 for all nontrivial polynomials and the construction does not terminate. However, the process can still *stagnate*. Assume there exists a monic polynomial p such that

$$\hat{\sigma}(a) = V_p(a). \tag{6.2}$$

In that case there again exists such a monic polynomial \tilde{p} of smallest degree, let it be of degree d.

Lemma 6.2. The polynomial \tilde{p} is unique and satisfies $\tilde{p} = p_d$.

Proof. We have

$$\hat{\sigma}(a) = V_{\tilde{p}}(a) \subset V_{p_d}(a)$$

and if the inclusion is proper then $cap(V_{\tilde{p}}(a)) < cap(V_{p_d}(a))$. However, by Lemma 2.3 this means that

$$\|\tilde{p}(a)\| < \|p_d(a)\|$$

contradicting the definition of p_d . Then $\tilde{p} = p_d$ as the polynomials are monic, of the same degree and the Green's function for $V_p(a)$ is unique.

This allows us to se up a definition.

Definition 6.3. Suppose $a \in \mathcal{A}$ is such that there exists a monic p for which (6.2) holds. Then we say that a is polynomially round and the monic polynomial of smallest degree for which (6.2) holds is spectrally minimal for a. Finally, a is polynomially round of degree d if the spectrally minimal polynomial is of degree d.

Proposition 6.4. If $\{p_j\}$ is defined as in (6.1), then $a \in \mathcal{A}$ is polynomially round of degree d if and only if for k = 1, 2, ...

$$||p_{kd}(a)|| = ||p_d(a)||^k.$$
(6.3)

Proof. If p_d is spectrally minimal, then

$$||p_d(a)||^{1/d} \le ||p_{kd}(a)||^{1/kd} \le ||p_d(a)^k||^{1/kd}$$

where the first inequality follows from the monotonicity of capacity and the second from the definition of p_{kd} .

Suppose therefore that there exists a j such that we have for all k

$$||p_{kj}(a)|| = ||p_j(a)||^k$$

Put $b = p_j(a)$. We claim that

$$\hat{\sigma}(b) = \{\mu : |\mu| \le \|b\|\}.$$
 (6.4)

In fact, if there is a point μ_0 of absolute value ||b|| which is not part of spectrum of b then there exists a monic polynomial q such that the capacity of $V_q(b)$ is less than ||b||. If q is of degree n then we have

$$||q(p_j(a))|| < ||p_j(a)||^n = ||p_{nj}(a)|$$

contradicting the minimality of p_{nj} . Thus (6.4) holds and it follows from the spectral theory that

$$\hat{\sigma}(p_j(a)) = V_{p_j}(a).$$

Here is another characterization of a being polynomially round.

Proposition 6.5. Given $a \in A$ and a polynomial p we have

$$\hat{\sigma}(a) = V_p(a) \tag{6.5}$$

if and only if

$$\|(p(\lambda) - p(a))^{-1}\| = (|p(\lambda)| - \|p(a)\|)^{-1}$$
(6.6)

holds for all $\lambda \notin V_p(a)$.

Proof. It is clear that (6.6) implies (6.5) as $||(p(\lambda) - p(a))^{-1}||$ blows up when λ approaches the boundary of $V_p(\lambda)$.

Reversely, assuming (6.5) we obtain immediately that for $\lambda \notin V_p(a)$

$$\|(p(\lambda) - p(a))^{-1}\| \le \sum_{k=0}^{\infty} \frac{\|p(a)\|^k}{|p(\lambda)|^{k+1}} = (|p(\lambda)| - \|p(a)\|)^{-1}$$

On the other hand, if μ_0 satisfies $|\mu_0| = ||p(a)||$, then $\mu_0 \in \sigma(p(a))$. Put $\mu = t\mu_0$ with t > 1. Then

$$\|(\mu - p(a))^{-1}\| \ge \frac{1}{dist(\mu, \sigma(p(a)))} = \frac{1}{(t-1)\|p(a)\|}$$

For $p(\lambda) = \mu$ this is

$$||(p(\lambda) - p(a))^{-1}|| \ge (|p(\lambda)| - ||p(a)||)^{-1}.$$

Propositions 6.4 and 6.5 give some necessary conditions for stagnation. Here is still one such.

Proposition 6.6. Assume that $a \in \mathcal{A}$ is polynomially round and p_d is the spectrally minimal polynomial. If p_d has at least one simple zero, then for 0 < j < d

$$\|p_{d+j}(a)\|^{\frac{1}{d+j}} > \|p_d(a)\|^{1/d}.$$
(6.7)

Proof. Let λ_0 be a simple zero of p_d and write $p_d(\lambda) = (\lambda - \lambda_0)q_0(\lambda)$ where $q_0(\lambda_0) \neq 0$. Further, write likewise $p_{d+j} = (\lambda - \lambda_0)^{m_j}q_j(\lambda)$, with $q_j(\lambda_0) \neq 0$. Let us assume that (6.7) does not hold for some 0 < j < d. Since p_d is spectrally minimal we must have equality:

$$||p_{d+j}(a)||^{\frac{1}{d+j}} = ||p_d(a)||^{1/d}.$$

As in the proof of Lemma 6.2. we conclude that p_d and p_{d+j} determine the same Green's function. This means that

$$p_d^{d+j} = p_{d+j}^d, (6.8)$$

which near λ_0 implies

$$(\lambda - \lambda_0)^{d+j} = (\lambda - \lambda_0)^{m_j d}.$$

In particular, $m_j = (1 + j/d)$. But this is a contradiction as m_j must be an integer.

7 Spectrum in low rank perturbation

We point out that the explicit representation of the resolvent allows applications to perturbation theory. We demonstrate it in simplest form by considering rank-1 perturbations of bounded operators in a Hilbert space. We developed a general perturbation theory for low rank perturbations in [13], but without explicit representations of the resolvent.

Consider rank-1 perturbations of $A \in B(H)$

$$A_{\mu} = A + \mu u v^* \tag{7.1}$$

where $\mu \in \mathbb{C}$ and $u, v \in H$ are unit vectors. Assume we are interested in knowing for what μ the spectrum of A_{μ} lies in an open set Ω provided we know that this is the case for $\mu = 0$:

$$\sigma(A) \subset \Omega. \tag{7.2}$$

It follows from Theorem 1.3 that we may further assume that we know a polynomial p such that

$$V_p(A) \subset \Omega. \tag{7.3}$$

By continuity of p and compactness of $V_p(A)$ there exists $\theta > 1$ such that

$$V_p(A,\theta) = \{\lambda \in \mathbb{C} : |p(\lambda)| \le \theta ||p(A)||\} \subset \Omega.$$
(7.4)

Lemma 7.1. If A_{μ} is as in (7.1), then

$$\sigma(A_{\mu}) \subset \sigma(A) \cup Z_{\mu} \tag{7.5}$$

where Z_{μ} denotes the set of solutions of

$$\varphi(\lambda) = \frac{1}{\mu} \tag{7.6}$$

and

$$\varphi(\lambda) = v^*(\lambda - A)^{-1}u, \text{ for } \lambda \notin \sigma(A).$$
 (7.7)

Proof. We have for $\lambda \notin \sigma(A)$

$$\lambda - A_{\mu} = (\lambda - A)(1 - \mu(\lambda - A)^{-1}uv^{*})$$
(7.8)

so that

$$(\lambda - A_{\mu})^{-1} = (1 + \frac{\mu}{1 - \mu\varphi(\lambda)}uv^*)(\lambda - A)^{-1},$$
(7.9)

from which the claim follows.

We are interested in the part of the spectrum of A_{μ} which is not inside $V_p(A, \theta)$, that is $Z_{\mu} \cap (\mathbb{C} \setminus V_p(A, \theta))$. To start with, observe that we can calculate a t_{θ} such that for $|\mu| < t_{\theta}$ the intersection is empty. In fact, for $\lambda \notin V_p(A, \theta)$ we have by Lemma 2.5

$$|\varphi(\lambda)| \le C \frac{\theta}{\theta - 1},\tag{7.10}$$

Then, (7.6) cannot have solutions outside $V_p(A, \theta)$ as long as

$$C\frac{\theta}{\theta-1} < \frac{1}{|\mu|}$$

Denoting $t_{\theta} = \frac{\theta - 1}{C\theta}$ we thus have the following.

Proposition 7.2. If $|\mu| \leq t_{\theta}$, then $\sigma(A_{\mu}) \subset V_p(A, \theta)$.

From the value distribution theory of meromorphic functions we know that the solutions of (7.6), that is, the poles of

$$\frac{1}{\varphi(\lambda) - \frac{1}{\mu}}$$

are tied with the speed of growth of $|\varphi|$ as $|p(\lambda)|$ decays, [13]. As $|\mu|$ grows over t_{θ} solutions branches $\lambda_j(\mu)$ can appear outside K_{θ} , but their number is always finite and following each branch is in principally easy as φ is analytic and explicitly given. Notice further that when one moves away from $V_p(A, \theta)$, $|p(\lambda)|$ decays and the accurate computation of φ becomes easier making it possible to study when

$$\lambda_j(\mu) \in \partial\Omega.$$

8 Holomorphic functional calculus using residues

Suppose $\Omega \subset \mathbb{C}$ is open and f is a holomorphic scalar valued function in Ω . Suppose $a \in \mathcal{A}$ and p is a polynomial such that

$$V_p(a) \subset \Omega. \tag{8.1}$$

If γ is a contour in Ω surrounding $V_p(a)$, then

$$f(a) = \frac{1}{2\pi i} \int_{\gamma} f(\lambda) (\lambda - a)^{-1} d\lambda.$$
(8.2)

Let R_m be as in (2.4). Then we can set

$$f_m(a) = \frac{1}{2\pi i} \int_{\gamma} f(\lambda) R_m(\lambda, a, p) d\lambda.$$
(8.3)

Let $\theta > 1$ be such that $V_p(a, \theta) \subset \Omega$. With help of Theorem 2.6 we know that there exists $\eta \leq 1/\theta$ and a constant C such that for all m = 0, 1, ...

$$||f_m(a) - f(a)|| \le C\eta^{m+1}.$$
(8.4)

The interesting thing here is that R_m is a rational function with known singularities. Thus $f_m(a)$ can be computed using residues at zeros of p, without knowing the singularities of the resolvent.

It is natural also to look at f(a) as expanded as a power series in p(a). If we put

$$c_j(a) = \frac{1}{2\pi i} \int_{\gamma} f(\lambda) \frac{q(\lambda, a)}{p(\lambda)^{j+1}} d\lambda, \qquad (8.5)$$

then $c_j(a)$ is a polynomial in a of degree one less than p. Notice that computing $c_j(a)$ can be performed by computing the residues at zeros of p and hence, the computation of f(a) can be carried out by performing at most a *countable* number of residues, independent of the cardinality of the spectrum of a.

We close this by a convergence estimate for the power series.

Proposition 8.1. Let $\theta > 1$ and p be such that

$$V_p(a,\theta) = \{\lambda : |p(\lambda)| \le \theta \|p(a)\|\} \subset \Omega$$

and suppose γ surrounds $V_p(a, \theta)$ inside Ω . Put

$$M = \frac{1}{2\pi} \int_{\gamma} |f|$$

and let C be as in Lemma 2.5. Then

$$f(a) = \sum_{j=0}^{\infty} c_j(a) p(a)^j,$$
(8.6)

where the coefficients $c_j(a)$ satisfy

$$\|c_j(a)\| \le MC \frac{1}{(\theta \| p(a)\|)^j}$$

References

[1] J. Burke, A. Greenbaum: Characterizations of the polynomial numerical hull of degree k, Lin. Alg. Appl. 419 (2006), pp. 37-47

[2] E.B. Davies: Spectral bounds using higher order numerical ranges, LMS J. Comput. Math. 8, 17-45 (2005)

[3] E.B. Davies: Linear Operators and their Spectra, Cambridge studies in advanced mathematics 106, Cambridge University Press (2007)

[4] Ch. Davis, A.Salemi: On polynomial numerical hull of normal matrices, Lin. Alg. Appl. 383 (2004) pp. 151-161

[5] V. Faber, A. Greenbaum, and D. E. Marshall, The polynomial numerical hulls of Jordan blocks and related matrices, Linear Algebra Appl., 374 (2003), pp. 231-246

[6] A. Greenbaum: Generalizations of the Field of Values Useful in the Study of Polynomial Functions of a Matrix, Lin. Alg. Appl. 347 (2002), pp. 233-249

[7] A. Greenbaum: Card Shuffling and the Polynomial Numerical Hull of Degree k SIAM J. Sci. Comput. 25 (2004), pp. 408-416

[8] P.R.Halmos: Capacity in Banach Algebras, Indiana Univ.Math. 20, pp.855-863 (1971)

[9] S.Hyvönen: Convergence of the Arnoldi process when applied to the Pickard-Lindelöf operator, J.Comp.Appl. Math. 87 (1997) 303-320

[10] K.B. Laursen, M.M. Neumann: An introduction to local spectral theory, London Math. Soc. Monograph, 20, Oxford, Clarendon Press (2000)

[11] O.Nevanlinna: Convergence of Iterations for Linear Equations, Birkhäuser, Basel (1993)

[12] O.Nevanlinna: Hessenberg matrices in Krylov subspaces and the computation of the spectrum, Numer. Funct. Anal. and Optimiz., 16 (3,4), 443-473 (1995)

[13] O.Nevanlinna: Meromorphic Functions and Linear Algebra, AMS Fields Institute Monograph 18 (2003)

[14] O. Nevanlinna: Resolvent and polynomial numerical hull. Math. Reports A546 (2008-05-13), TKK (http://math.tkk.fi/reports/)

[15] O.Nevanlinna, G.Vainikko: Limit Spectrum of Discretely Converging Operators, Numer. Funct. Anal. and Optimiz.17, (7, 8) (1996) 797-808

[16] O-P. Piirilä: Questions and notions related to quasialgebraicity in Banach algebras, Ann. Acad. Sci. Fenn. Math. Diss. 88 (1993)

[17] Th. Ransford: Potential Theory in the Complex Plane, London Math.Soc Student Texts 28, Cambridge Univ Press, 1995

(continued from the back cover)

- A562 Antti H. Niemi A bilinear shell element based on a refined shallow shell model December 2008
- A561 Antti Hannukainen, Sergey Korotov, Michal Krizek On nodal superconvergence in 3D by averaging piecewise linear, bilinear, and trilinear FE approximations December 2008
- A560 Sampsa Pursiainen Computational methods in electromagnetic biomedical inverse problems November 2008
- A559 Sergey Korotov, Michal Krizek, Jakub Solc On a discrete maximum principle for linear FE solutions of elliptic problems with a nondiagonal coefficient matrix November 2008
- A558 José Igor Morlanes, Antti Rasila, Tommi Sottinen Empirical evidence on arbitrage by changing the stock exchange December 2008
- A557 Mika Juntunen, Rolf Stenberg Analysis of finite element methods for the Brinkman problem April 2009
- A556 Lourenço Beirão da Veiga, Jarkko Niiranen, Rolf Stenberg A posteriori error analysis for the Morley plate element with general boundary conditions December 2008
- A555 Juho Könnö, Rolf Stenberg Finite element analysis of composite plates with an application to the paper cockling problem December 2008
- A554 Lasse Leskelä Stochastic relations of random variables and processes October 2008

HELSINKI UNIVERSITY OF TECHNOLOGY INSTITUTE OF MATHEMATICS RESEARCH REPORTS

The reports are available at *http://math.tkk.fi/reports/*. The list of reports is continued inside the back cover.

- A567 Antti Hannukainen, Sergey Korotov, Michal Krizek On a bisection algorithm that produces conforming locally refined simplicial meshes April 2009
- A566 Mika Juntunen, Rolf Stenberg A residual based a posteriori estimator for the reaction-diffusion problem February 2009
- A565 Ehsan Azmoodeh, Yulia Mishura, Esko Valkeila On hedging European options in geometric fractional Brownian motion market model February 2009
- A564 Antti H. Niemi Best bilinear shell element: flat, twisted or curved? February 2009
- A563 Dmitri Kuzmin, Sergey Korotov Goal-oriented a posteriori error estimates for transport problems February 2009

ISBN 978-951-22-9850-1 (print) ISBN 978-951-22-9851-8 (PDF) ISSN 0784-3143 (print) ISSN 1797-5867 (PDF)