Math. Nachr. **000** (0000), 0 – 0

Estimates for norms of resolvents and an application to the perturbation of spectra

By Oscar F. Bandtlow of London

(Received July 11, 2002; accepted May 4, 2003)

Abstract. Let A belong to the Schatten-von Neumann ideal S_p for $0 < p < \infty$. We give an upper bound for the operator norm of the resolvent $(zI - A)^{-1}$ of A in terms of the departure from normality of A and the distance of z to the spectrum of A. As an application we provide an upper bound for the Hausdorff distance of the spectra of two operators belonging to S_p.

1. Introduction

Resolvent estimates for linear operators on Hilbert spaces have a comparatively long history dating back to the early days of functional analysis. The first such estimate appeared in a celebrated paper of Carleman from 1921 [Car]. There he showed that for any Hilbert-Schmidt operator A on a Hilbert space H the resolvent $(zI - A)^{-1}$ satisfies the following inequality

$$\left\|\det_2(I - z^{-1}A)(zI - A)^{-1}\right\| \leq \frac{1}{|z|} \exp\left(\frac{1}{2}\frac{\|A\|_2^2}{|z|^2} + \frac{1}{2}\right). \tag{1.1}$$

Here, $\|.\|$ denotes the operator norm on H, $\|.\|_2$ the Hilbert-Schmidt norm, and $\det_2(I - z^{-1}A)$ the regularised determinant of order 2 given by

$$\det_2(I - z^{-1}A) = \prod_{k=1}^{\infty}\left(1 - \frac{\lambda_k}{z}\right)\exp\left(\frac{\lambda_k}{z}\right),$$

where λ_k denotes the k-th eigenvalue of A (repeated according to multiplicities). Later, this result has been generalized to include operators belonging to the Schatten-von Neumann ideals S_p, $0 < p < \infty$ (see, for example, [DS2, Sim]) and to the Banach space setting (see, for example, [Bur]).

1991 *Mathematics Subject Classification.* Primary: 47A10; Secondary: 47 B 10
Keywords and phrases. Resolvent growth, Schatten ideals

Estimates of this type have a number of important applications in spectral theory, ranging from the problem of establishing the completeness of eigenvectors and root vectors of operators to problems in perturbation theory (see, for example, [DS2, GK, Kat]).

A particular feature of resolvent estimates of the form (1.1) is that complete knowledge of the spectrum of A is required. In other words, in order to find an upper bound for $\|(zI-A)^{-1}\|$ all the eigenvalues λ_k of A need to be known. In a number of applications, however, it is desirable to have an upper bound which only involves knowledge of the distance $d(z,\sigma(A))$ of z to $\sigma(A)$, the spectrum of A. Estimates of this type have been obtained by Gil' in a series of papers begun in 1979 (see [Gil] and references therein) and by Dechevski and Persson [DP1] in the mid-nineties.

The results of Gil' yield an upper bound for $\|(zI-A)^{-1}\|$ with A belonging to the Schatten-von Neumann ideals S_p for p an even integer in terms of a power series involving $d(z,\sigma(A))$ and a quantity measuring the departure from normality of A. An interesting feature of his bound is that, for A normal, it reduces to the usual estimate

$$\|(zI-A)^{-1}\| \leq \frac{1}{d(z,\sigma(A))}.$$

His proof relies on an adaptation of finite-dimensional arguments which can be traced to work of Henrici from the early sixties [Hen].

The estimate of Dechevski and Persson, obtained by different methods, does not rely on a power series and is valid for $A \in S_p$ for any $0<p<\infty$:

$$\|(zI-A)^{-1}\| \leq \frac{1}{d(z,\sigma(A))} \exp\left(c_p \frac{\|A\|_p^p}{d(z,\sigma(A))^p} + b_p \right), \tag{1.2}$$

where c_p and b_p are positive constants depending on p only. Unfortunately, as noted in the Mathematical Reviews, their proof of (1.2) applies only to operators for which the generalized eigenspaces are mutually orthogonal (see [MR, 95f:47008]).

In this article we shall focus on deriving resolvent estimates which combine features of that of Gil' and that of Dechevski and Persson. Our method of proof follows Henrici and Gil', while our upper bounds are expressible in closed form similar to those of Dechevski and Persson, which turn out to be a special case of ours. As an application of our estimate we show that it can be used to derive an upper bound for the Hausdorff distance of the spectra of two operators belonging to S_p $(0<p<\infty)$.

2. The Schatten-von Neumann ideals

Let H be a separable Hilbert space with scalar product $(.,.)$. We use $L(H)$ to denote the Banach algebra of bounded linear operators on H equipped with the uniform operator norm $\|.\|$. The spectrum and the resolvent set of $A \in L(H)$ will be denoted by $\sigma(A)$ and $\varrho(A)$, respectively. An operator $A \in L(H)$ is said to be *quasi-nilpotent* if $\sigma(A)=\{0\}$.

We use S_∞ for the closed two-sided ideal of compact operators in $L(H)$. If $A \in S_\infty$, then $\{\lambda_k(A)\}$ denotes the sequence of eigenvalues of A, each eigenvalue being repeated according to its multiplicity. We also assume that $\{\lambda_k(A)\}$ is ordered by magnitude, so that $|\lambda_1(A)| \geq |\lambda_2(A)| \geq \cdots$.

For $0 < p < \infty$, the quasi-normed operator ideal S_p, known as the *Schatten-von Neumann ideal*, is defined in the usual way by

$$S_p := \left\{ A \in S_\infty \;\middle|\; \|A\|_p := \left(\sum_{k=1}^{\infty} s_k(A)^p \right)^{1/p} < \infty \right\},$$

where $s_k(A)$ denotes the k-th singular number of A. Note that by convention $\|.\|_\infty := \|.\|$. More information about singular numbers and the Schatten-von Neumann ideals can be found in [Pie, GK, DS2, Rin].

The following theorem, a specialization of Carleman-type resolvent estimates to quasi-nilpotent operators, is crucial to our analysis.

Theorem 2.1. *Let $0 < p < \infty$. Then there are positive constants a_p and b_p such that for every quasi-nilpotent $A \in S_p$*

$$\left\|(I - A)^{-1}\right\| \leq \exp(a_p \|A\|_p^p + b_p).$$

Proof. See [Rin, Theorem 3.4.6]. See also [DS2, Corollary XI.9.25], where the result is erroneously stated with $b_p = 0$ (see Remark (ii) below), and [Sim, Corollary 7.7], where an elegant proof for $p \in \mathbb{N}$ is given. The method in [Sim] also works for general p and yields better constants a_p and b_p than the procedure in [Rin] (see Remark (iii) below). □

Remark 2.2.

(i) For certain values of p the constants a_p and b_p are known

  (a) $a_1 = 1$, $b_1 = 0$; (see [Sim]);

  (b) $a_2 = \frac{1}{2}$, $b_2 = \frac{1}{2}$; this is Carleman's original inequality (see [Car] or [DS2, Corollary XI.6.28]);

(ii) $b_p \neq 0$ for $p > 1$. To see this let $A \in L(\mathbb{C}^2)$ be defined by

$$Ax = \epsilon(x, e_1)e_2,$$

where $\epsilon \in \mathbb{R}$ and $\{e_1, e_2\}$ is an orthonormal basis for $\mathbb{C}^2$. A short calculation shows that $\|A\|_p = |\epsilon|$, while

$$\begin{aligned} \left\|(I - A)^{-1}\right\|^2 &= \frac{1}{2}\left(2 + |\epsilon|^2 + |\epsilon| \sqrt{|\epsilon|^2 + 4}\right) \\ &\geq 1 + |\epsilon| > \exp(2K |\epsilon|^p) = \exp(2K \|A\|_p^p) \end{aligned}$$

for any $p > 1$, any $K > 0$, and $|\epsilon| > 0$ small.

(iii) For $0 < p \leq 1$ it is possible to choose (see [DS2, Theorem XI.9.26])

$$a_p = \sup_{z \in \mathbb{C}} |z|^{-p} \log |(1+z)| \text{ and } b_p = 0.$$

Closer inspection of Simon's method reveals that for $p > 1$ it is possible to choose a_p to be any real number strictly larger than Γ_p, where

$$\Gamma_p := \sup_{z \in \mathbb{C}} |z|^{-p} \log \Big| (1+z) \exp\Big(\sum_{k=1}^{\lceil p \rceil - 1} \frac{(-z)^k}{k} \Big) \Big|,$$

and $\lceil p \rceil := \inf \left\{ n \in \mathbb{N} \,\middle|\, n \geq p \right\}$. The value of b_p in turn depends on a_p.

A simple consequence of the theorem above is the following estimate for the growth of the resolvent of a quasi-nilpotent $A \in S_p$.

Corollary 2.3. *Let $A \in S_p$ be quasi-nilpotent. Then*

$$\left\| (zI - A)^{-1} \right\| \leq \frac{1}{|z|} \exp\Big(a_p \frac{\|A\|_p^p}{|z|^p} + b_p\Big).$$

Proof. Follows from $(zI - A)^{-1} = z^{-1}(I - z^{-1}A)^{-1}$. □

3. The Schur decomposition

One possibility of obtaining growth estimates of the resolvent of an operator A is to consider A as a perturbation of a normal operator D having the same spectrum as A by a quasi-nilpotent N. In the finite-dimensional case such a perturbation is easily seen to exist by an argument going back to Henrici [Hen]: if A is any matrix then, by a classical result due to Schur, A is unitarily equivalent to an upper-triangular matrix $\tilde{A}$,

$$A = U\tilde{A}U^* \quad (U \text{ unitary}).$$

Writing

$$\tilde{A} = \tilde{D} + \tilde{N},$$

where D denotes the diagonal matrix whose main diagonal coincides with that of $\tilde{A}$, it is not difficult to see that

$$A = D + N,$$

where $D := U^*\tilde{D}U$, $N := U^*\tilde{N}U$, is the desired perturbation with D normal, N quasi-nilpotent, and $\sigma(D) = \sigma(A)$.

It turns out that with a bit of care this argument also works in the infinite-dimensional setting. Before we proceed we require the following version of Schur's result.

Theorem 3.1. *Let $A : H \to H$ belong to S_∞. Let E_A denote the closed span (in H) of all eigenvectors and generalized eigenvectors of A corresponding to non-zero eigenvalues of A. Then*

(i) *(*'Schur's Lemma'*) E_A has an orthonormal basis $\{b_1, b_2, \ldots\}$ such that*

$$Ab_i = \sum_{j=1}^{k} a_{ij} b_j, \quad a_{kk} = \lambda_k(A) \quad (\forall k \in \mathbb{N});$$

(ii) *with respect to the orthogonal sum $H = E_A \oplus E_A^\perp$ the operator A can be written as a 2×2 matrix with S_∞ entries*

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} : E_A \oplus E_A^\perp \to E_A \oplus E_A^\perp,$$

such that

(a) *A_{11} and A have the same non-zero eigenvalues;*

(b) *A_{22} is quasi-nilpotent.*

Proof. See [GGK, Lemmata II.3.3 and II.3.4]. □

We are now ready to extend Henrici's idea to the infinite-dimensional setting.

Theorem 3.2. *Let $A \in S_\infty$. Then A can be written as a sum*

$$A = D + N,$$

such that

(i) *$D \in S_\infty$, $N \in S_\infty$;*

(ii) *D is normal, $\lambda_k(D) = \lambda_k(A)$ ($k \in \mathbb{N}$), and $\sigma(D) = \sigma(A)$;*

(iii) *N is quasi-nilpotent;*

(iv) *if f is a function analytic on $\sigma(D)$ ($\sigma(D^*)$) and $f(D)$ ($f(D^*)$) is given by Dunford's analytic functional calculus [DS1, VII.3], then $Nf(D)$ and $f(D)N$ ($Nf(D^*)$ and $f(D^*)N$) are quasi-nilpotent;*

Proof. (i–iii) Using the notation of the previous theorem define $D_{11} : E_A \to E_A$ by

$$D_{11} x = \sum_{k=1}^{\infty} (Ab_k, b_k)(x, b_k) b_k.$$

It is easily verified that D_{11} is compact, normal, and has the same non-zero eigenvalues (counting multiplicities) as A. Next define $N_{11} : E_A \to E_A$ by $N_{11} := A_{11} - D_{11}$. Then N_{11} is compact. Now define

$$D := \begin{bmatrix} D_{11} & 0 \\ 0 & 0 \end{bmatrix}$$

$$N := \begin{bmatrix} N_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}.$$

It is easily verified that D and N satisfy (i) and (ii). Property (iii) follows from (iv) by choosing $f \equiv 1$.

(iv) We shall only prove that $f(D^*)N$ is quasi-nilpotent. The other cases are similar. In order to do so it suffices to show that $(zI - D^*)^{-1}N$ is quasi-nilpotent for $z \in \varrho(D^*)$. The general case then follows from the observation that the Dunford integral converges in the uniform operator topology and the fact that the limit of a sequence of compact quasi-nilpotent operators converging in the uniform operator topology is quasi-nilpotent (see [GGK, Corollary II.2.4]). Fix $z \in \varrho(D^*)$. Before proving that $(zI - D^*)^{-1}N$ is quasi-nilpotent we shall show that $T := (zI_{E_A} - D_{11}^*)^{-1}N_{11}$ is quasi-nilpotent. To see this let P_n be the orthogonal projection onto $\bigvee_{i=1}^n \{b_i\}$. Then $P_n \to I_{E_A}$ in the strong operator topology. Since T is compact, it follows that

$$P_nTP_n \to T \quad \text{in the uniform operator topology} \tag{3.1}$$

(see, for example, [ALL, Theorem 4.1]). Now observe that each $P_nTP_n = P_n(zI_{E_A} - D_{11}^*)^{-1}N_{11}P_n$ has a strictly upper triangular representation w.r.t. the basis $\{b_i\}$ and is thus nilpotent. This together with (3.1) implies that $(zI_{E_A} - D_{11}^*)^{-1}N_{11}$ is quasi-nilpotent. In order to complete the proof we only need to show that $(z - D^*)^{-1}N$ is quasi-nilpotent. This, however, follows from

$$(z - D^*)^{-1}N = \begin{bmatrix} (zI_{E_A} - D_{11}^*)^{-1}N_{11} & (zI_{E_A} - D_{11}^*)^{-1}A_{12} \\ 0 & z^{-1}A_{22} \end{bmatrix},$$

and a short calculation using the fact that both $(zI_{E_A} - D_{11}^*)^{-1}N_{11}$ and $z^{-1}A_{22}$ are quasi-nilpotent.

□

Remark 3.3.

(i) A similar decomposition holds in a more general context, where A is merely assumed to be a Riesz operator, that is, an operator whose essential spectral radius vanishes (see [Dow, Chapter 3]). In this case (i)–(iii) of the theorem above remain valid with the modification that N need no longer belong to S_∞ (see [Dow, Theorem 3.33]).

(ii) Assertions (i–iii) are also proved in [Gil, Theorem 2.2.1].

This result motivates the following definition.

Definition 3.4. Let $0 < p \leq \infty$ and $A \in S_p$. A decomposition

$$A = D + N$$

with D and N enjoying properties (i–iv) of the previous theorem is called a *Schur decomposition of A*. The operators D and N will be referred to as the *normal* and the *quasi-nilpotent part of the Schur decomposition of A*, respectively.

Remark 3.5.

(i) The decomposition is not unique. Witness the following example:

$$A := \begin{bmatrix} 2 & 2 & 2 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} = \underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{=:D_1} + \underbrace{\begin{bmatrix} 0 & 2 & 2 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}}_{=:N_1}$$

$$= \underbrace{\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{=:D_2} + \underbrace{\begin{bmatrix} 1 & 1 & 2 \\ -1 & -1 & 2 \\ 0 & 0 & 0 \end{bmatrix}}_{=:N_2}.$$

It is easily verified that D_1 and D_2 are normal and that N_1 and N_2 are nilpotent. Furthermore $\sigma(A) = \sigma(D_1) = \sigma(D_2) = \{2, 0\}$.

This example also illustrates the fact that, while the normal parts corresponding to different Schur decompositions are necessarily unitarily equivalent, the quasi-nilpotent parts need not be unitarily equivalent in general. To see this observe that, for example, $\|N_1\|_4^4 = 112 \neq 80 = \|N_2\|_4^4$.

(ii) It is a simple consequence of Weyl's inequality (see, for example, [DS2, Corollary XI.9.7]) that if $A \in S_p$ for some $0 < p \leq \infty$ and $A = D + N$ is a Schur decomposition with normal part D and quasi-nilpotent part N, then

$$\|D\|_p \leq \|A\|_p .$$

In particular, $D \in S_p$ and $N \in S_p$.

For later use we define the following quantities originally introduced by Henrici [Hen].

Definition 3.6. Let $0 < p \leq \infty$ and $A \in S_p$. Then

$$\nu_p(A) := \inf \left\{ \|N\|_p \,\middle|\, N \text{ is a quasi-nilpotent part of } A \right\}$$

is called the (p-)*departure from normality of A*.

The term 'departure from normality' is justified since $\nu_p(A) = 0$ if and only if A is normal.

For general p, the p-departure from normality is difficult to calculate in practice. For $p = 2$, however, we have the following formula originally due to Henrici (see [Hen]) in the finite-dimensional setting and to Gil' in the infinite-dimensional case.

Proposition 3.7. *Let $A \in S_2$. Then*

$$\nu_2(A) = \sqrt{\|A\|_2^2 - \sum_k |\lambda_k(A)|^2}.$$

Proof. See [Gil, Lemma 2.4.6]. □

Simple but somewhat crude upper bounds for $\nu_p(A)$ can be obtained by invoking the (p-)triangle inequality.

Proposition 3.8.

$$\nu_p(A) \leq \begin{cases} 2\,\|A\|_p & \text{for } 1 \leq p \leq \infty; \\ 2^{1/p}\,\|A\|_p & \text{for } 0 < p < 1. \end{cases}$$

Proof. Let $A = D + N$ be a Schur decomposition of A. Recall that

$$\|D\|_p \leq \|A\|_p$$

by Remark 3.5 (ii).

If $0 < p < 1$, then

$$\|N\|_p^p = \|A - D\|_p^p \leq \|A\|_p^p + \|D\|_p^p \leq 2\,\|A\|_p^p\,.$$

Thus $\nu_p(A) \leq 2^{1/p}\,\|A\|_p$.

If $1 \leq p \leq \infty$, then

$$\|N\|_p = \|A - D\|_p \leq \|A\|_p + \|D\|_p \leq 2\,\|A\|_p\,.$$

Thus $\nu_p(A) \leq 2\,\|A\|_p$. □

Remark 3.9. For $p = 2m$, $m \in \mathbb{N}$, the estimates above have been obtained by Gil' [Gil, Lemma 2.6.6].

4. Resolvent estimates

Using the results of the previous section we are now able to derive upper bounds for the resolvent of $A \in S_p$. The proof relies on the following well-known fact. If $D \in L(H)$ is normal, then

$$\left\|(zI - D)^{-1}\right\| = \frac{1}{d(z, \sigma(D))} \quad (z \in \varrho(D)), \tag{4.1}$$

where for $z \in \mathbb{C}$ and $\sigma \subset \mathbb{C}$ closed,

$$d(z, \sigma) := \inf_{\lambda \in \sigma} |z - \lambda|$$

denotes the distance of z to σ.

For a short proof of (4.1) recall that if D is normal, then $\|D\| = r(D)$, where $r(D)$ denotes the spectral radius of D. Let $z \in \varrho(D)$. Then $(zI - D)^{-1}$ is normal. Thus

$$\left\|(zI - D)^{-1}\right\| = r\big((zI - D)^{-1}\big) = \sup_{\lambda \in \sigma(D)} \frac{1}{|z - \lambda|} = \frac{1}{d(z, \sigma(D))}.$$

The proof of our main result relies on the following idea. Suppose that $A \in S_p$ $(0 < p < \infty)$. Using a Schur decomposition we can write A as a sum of a normal operator D with $\sigma(D)=\sigma(A)$ and a quasi-nilpotent operator N. We can think of A as a perturbation of D by N. Using standard arguments from perturbation theory together with the fact that the influence of the quasi-nilpotent perturbation N can be controlled by Theorem 2.1 it is possible to find upper bounds for the resolvent of A in terms of the standard bounds (4.1) for the resolvent of the normal part D. More precisely, we have the following.

Theorem 4.1. *Let $0<p<\infty$. Then for every $A \in S_p$*

$$\left\|(zI-A)^{-1}\right\| \leq \frac{1}{d(z,\sigma(A))}\exp\left(a_p\frac{\nu_p(A)^p}{d(z,\sigma(A))^p}+b_p\right),$$

where a_p and b_p are the positive constants (depending on p only) given by Theorem 2.1.

Proof. If $z\in\sigma(A)$, the RHS of the inequality is infinite, so there is nothing to prove. We may thus assume $z\in\varrho(A)$. Let $A=D+N$ be a Schur decomposition of A with normal part D and quasi-nilpotent part N. Noting that $\sigma(A)=\sigma(D)$ we see that $(zI-D)^{-1}$ exists. Furthermore, $(zI-D)^{-1}N$ is quasi-nilpotent and belongs to S_p with

$$\left\|(zI-D)^{-1}N\right\|_p \leq \left\|(zI-D)^{-1}\right\| \|N\|_p = \frac{\|N\|_p}{d(z,\sigma(D))}.$$

Thus, $(I-(zI-D)^{-1}N)$ is invertible in $L(H)$ and

$$\left\|(I-(zI-D)^{-1}N)^{-1}\right\| \leq \exp\left(a_p\frac{\|N\|_p^p}{d(z,\sigma(D))^p}+b_p\right),$$

by Theorem 2.1. Now, since $(zI-A)=(zI-D)(I-(zI-D)^{-1}N)$, we conclude that $(zI-A)$ is invertible in $L(H)$ and

$$\begin{aligned}\left\|(z-A)^{-1}\right\| &\leq \left\|(I-(zI-D)^{-1}N)^{-1}\right\|\left\|(zI-D)^{-1}\right\| \\ &\leq \frac{1}{d(z,\sigma(D))}\exp\left(a_p\frac{\|N\|_p^p}{d(z,\sigma(D))^p}+b_p\right).\end{aligned}$$

Taking the infimum over all Schur decompositions while using $\sigma(A)=\sigma(D)$, the result follows. □

Remark 4.2.

(i) The estimate remains valid if $\nu_p(A)$ is replaced by something larger, for example by the upper bound given in Proposition 3.8. In this case we recover the estimate of Dechevski and Persson [DP1].

(ii) For normal A our estimate reduces to the standard estimate (4.1) — apart from the factor $\exp(b_p)$, which equals 1 for $0<p\leq 1$ by Remark 2.2 (iii).

(iii) For $p=2m$, $m\in\mathbb{N}$ the estimates above can be sharpened (see [Gil, Lemma 3.3.1]).

(iv) An interesting discussion of various notions of sharpness for the resolvent estimates given above can be found in [DP2].

5. An application to the perturbation of spectra

As an application of our resolvent estimates we now consider the problem of finding upper bounds for the Hausdorff distance of the spectra of two operators belonging to S_p. Recall that the *Hausdorff distance* $h(.,.)$ is the following metric defined on the space of compact subsets of $\mathbb{C}$

$$h(\sigma_1,\sigma_2) := \max\{\hat{d}(\sigma_1,\sigma_2),\hat{d}(\sigma_2,\sigma_1)\},$$

where

$$\hat{d}(\sigma_1,\sigma_2) := \sup_{\lambda\in\sigma_1} d(\lambda,\sigma_2)$$

and σ_1 and σ_2 are two compact subsets of $\mathbb{C}$.

For $A, B \in S_p$, our aim is to find upper bounds for $h(\sigma(A),\sigma(B))$ expressible in terms of $\|A-B\|$. Results of this type are useful in providing computationally accessible *a posteriori* error estimates for spectral approximation procedures (see [BJ]). The finite-dimensional prototype of our result is the Ostrowski-Elsner formula, which states that

$$h(\sigma(A),\sigma(B)) \leq (2M)^{1-1/n}\|A-B\|^{1/n} \text{ for } A, B \in L(\mathbb{C}^n),$$

where $M := \max\{\|A\|,\|B\|\}$ (see [Els]).

In order to extend this formula to operators in S_p we shall use the resolvent bounds obtained in the previous section together with a simple but powerful argument usually credited to Bauer and Fike [BF] who first employed it in a finite-dimensional context.

Lemma 5.1. *Let A, $B \in L(H)$. Then*

$$z \in \sigma(A)\cap\varrho(B) \Longrightarrow \|A-B\|^{-1} \leq \left\|(zI-B)^{-1}\right\|.$$

Proof. By contradiction. Let $z \in \sigma(A)\cap\varrho(B)$ and assume to the contrary that

$$\left\|(zI-B)^{-1}\right\|\|A-B\| < 1.$$

Then $\big(I-(zI-B)^{-1}(A-B)\big)$ is invertible in $L(H)$, which implies that $(zI-A) = (zI-B)\big(I-(zI-B)^{-1}(A-B)\big)$ is invertible in $L(H)$. Thus $z \in \varrho(A)$ which contradicts $z \in \sigma(A)$. □

We are now ready to prove the main result of this section.

Theorem 5.2. *Let $0 < p < \infty$ and let A, $B \in S_p$. Then*

(i)

$$\hat{d}(\sigma(A),\sigma(B)) \leq \frac{\nu_p(B)}{f_p\left(\frac{\nu_p(B)}{\|A-B\|}\right)};$$

(ii)

$$h(\sigma(A),\sigma(B)) \leq \frac{\max\{\nu_p(A),\nu_p(B)\}}{f_p\left(\frac{\max\{\nu_p(A),\nu_p(B)\}}{\|A-B\|}\right)}.$$

Here, $f_p:\mathbb{R}_0^+\to\mathbb{R}_0^+$ is the inverse of the function

$$g_p:\mathbb{R}_0^+\to\mathbb{R}_0^+,$$

$$g_p(x) := x\exp(a_p x^p + b_p),$$

and a_p and b_p are the constants given by Theorem 2.1.

Proof. (i) It suffices to show that

$$z\in\sigma(A) \Longrightarrow d(z,\sigma(B)) \leq \frac{\nu_p(B)}{f_p\left(\frac{\nu_p(B)}{\|A-B\|}\right)}. \tag{5.1}$$

In what follows we shall use the following abbreviations:

$$d := d(z,\sigma(B)), \quad E := \|A-B\|.$$

In order to prove (5.1) let $z\in\sigma(A)$. If $z\in\sigma(B)$ there is nothing to prove. We may thus assume that $z\in\varrho(B)$. Hence

$$\frac{1}{\|A-B\|} \leq \left\|(zI-B)^{-1}\right\| \leq \frac{1}{d}\exp\left(a_p\frac{\nu_p(B)^p}{d^p}+b_p\right) \tag{5.2}$$

$$= \frac{1}{\nu_p(B)}g_p\left(\frac{\nu_p(B)}{d}\right), \tag{5.3}$$

where the first inequality follows from the previous lemma and the second from Theorem 4.1. Since g_p is strictly monotonically increasing, so is f_p. Thus

$$f_p\left(\frac{\nu_p(B)}{E}\right) \leq \frac{\nu_p(B)}{d},$$

whence

$$d(z,\sigma(B)) \leq \frac{\nu_p(B)}{f_p\left(\frac{\nu_p(B)}{\|A-B\|}\right)}.$$

(ii) It is easily verified that the implication (5.1) remains true, if $\nu_p(B)$ is replaced by $\max\{\nu_p(A),\nu_p(B)\}$. To see this, note that (5.2) and hence (5.3) hold with $\max\{\nu_p(A),\nu_p(B)\}$ in place of $\nu_p(B)$. Thus $\hat{d}(\sigma(A),\sigma(B))$ is bounded by the RHS of (ii), and so is $\hat{d}(\sigma(B),\sigma(A))$, by symmetry. The desired inequality follows. □

Remark 5.3.

(i) It is not difficult to see, for example by arguing as in the proof of part (ii) of the theorem, that the inequalities (i) and (ii) above remain valid if $\nu_p(A)$ or $\nu_p(B)$ is replaced by something larger — for example, by the upper bounds given in Proposition 3.8.

(ii) If both operators are normal, then (ii) reduces to

$$h(\sigma(A), \sigma(B)) \leq \|A - B\| \exp(b_p).$$

Note that the factor $\exp(b_p)$ equals 1 for $0 < p \leq 1$ by Remark 2.2 (iii).

(iii) For $p = 2m$, $m \in \mathbb{N}$ estimates of the Hausdorff distance of the spectra of two operators belonging to S_p similar to the above can be found in [Gil, Section 4.1].

Acknowledgements

I would like to thank Philip Spain, Michael Dritschel, Oliver Jenkinson, Jeff Webb and Simon Wassermann for stimulating discussions and feedback during the preparation of this article.

References

[ALL] M Ahues, A Largillier, BV Limaye (2001) *Spectral Computations for Bounded Operators*; Roca Baton, Chapman & Hall/CRC

[BJ] OF Bandtlow and O Jenkinson (2003) Finite sections of holomorphic transfer operators; *preprint*

[BF] FL Bauer, CT Fike (1960) Norms and exclusion theorems; *Num. Math.* **2**, 42–53

[Bur] J Burgoyne (1995) Denseness of the generalized eigenvectors of a discrete operator in a Banach space; *J. Operator Theory* **33**, 279–297

[Car] T Carleman (1921) Zur Theorie der linearen Intergralgleichungen; *Math. Z.* **9**, 196–217

[DP1] LT Dechevski, LE Persson (1994) Sharp generalized Carleman inequalities with minimal information about the spectrum; *Math. Nachr.* **168**, 61–77

[DP2] LT Dechevski, LE Persson (1996) On sharpness, applications and generalisations of some Carleman type inequalities; *Tôhoku Math. J.* **48**, 1–22

[Dow] HR Dowson (1978) *Spectral Theory of Linear Operators*; London, Academic Press

[DS1] N Dunford, JT Schwartz (1958) *Linear Operators, Vol. 1*; New York, Interscience

[DS2] N Dunford, JT Schwartz (1963) *Linear Operators, Vol. 2*; New York, Interscience

[Els] L Elsner (1985) An optimal bound for the spectral variation of two matrices; *Linear Algebra and Appl.* **71**, 77–80

[Gil] MI Gil' (1995) *Norm Estimations for Operator-Valued Functions and Applications*; New York, Marcel Dekker

[GGK] I Gohberg, S Goldberg, MA Kaashoek (1990) *Classes of Linear Opeators Vol. 1*; Basel, Birkhäuser

[GK] I Gohberg, MG Krein (1969) *Introduction to the Theory of Linear Non-Selfadjoint Operators*; Providence, AMS

[Hen] P Henrici (1962) Bounds for iterates, inverses, spectral variation and fields of values of non-normal matrices; *Num. Math.* **4**, 24–40

[Kat] T Kato (1966) *Perturbation Theory for Linear Operators*; Berlin, Springer

[MR] *Mathematical Reviews*; Providence, AMS

[Pie] A Pietsch (1988) *Eigenvalues and s-Numbers*; Cambridge, CUP

[Rin] JR Ringrose (1971) *Compact Non-Self-Adjoint Operators*; London, van Nostrand

[Sim] B Simon (1977) Notes on infinite determinants of Hilbert space operators; *Adv. in Math.*, **24**, 244–273

School of Mathematical Sciences
Queen Mary, University of London
Mile End Road
London E1 4NS
United Kingdom

E – mail:
o.bandtlow@qmul.ac.uk