Queen Mary, University of London

School of Mathematical Sciences

Quantitative Perturbation Theory for Compact Operators on a Hilbert Space

Thesis submitted in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

Ayşe Güven

Supervisors:

Dr Oscar F. Bandtlow

Prof Cho-Ho Chu

August, 2016

Statement of Originality

I, Ayşe Güven, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Signature:

Date: 4th March 2016

Details of collaboration and publications:

Chapter 4 of this thesis has been published in the following paper, written in collaboration with my supervisor, Dr Oscar F. Bandtlow.

- O. F. Bandtlow and A. Güven, Explicit upper bounds for the spectral distance of two trace class operators, Linear Algebra Appl. **466** (2015), 329–342.

Acknowledgements

I would first like to express my sincere gratitude to my supervisor, Dr Oscar F. Bandtlow, who first introduced me to many of the ideas studied in this thesis, for his wonderful supervision and inspiration over the past four years. I would also like to thank the Turkish Government and the Turkish Ministry of National Education for funding my both master's and doctoral studies in London. Many thanks to Prof Cho-Ho Chu for his help, particularly during my first year, and to Dr Wolfram Just for our enlightening discussions during the second year of my study. I am also very grateful to Tayyibe Erten and Julia Slipantschuk for their kind personal support over the years. Additionally, I would like to acknowledge the help of all my friends for keeping me sane throughout my doctoral program.

My special thanks go out to my family—my dad, mum, and sister—who have continually supported and encouraged me. Finally, words cannot express the gratitude I owe to my husband, Ali Sarıhan, whose love has provided endless encouragement. He has always supported me, even when he has needed help for his own doctoral process, and I have been fortunate enough to share this experience with him. It is to him; and to the memory of my brother, Ahmet Güven, and my grandparents, Esma & Ahmet Güven, Günay Gür, that I dedicate this thesis.

Abstract

This thesis makes novel contributions to a problem of practical and theoretical importance, namely how to determine explicitly computable upper bounds for the Hausdorff distance of the spectra of two compact operators on a Hilbert space in terms of the distance of the two operators in operator norm.

It turns out that the answer depends crucially on the speed of decay of the sequence of singular values of the two operators. To this end, 'compactness classes', that is, collections of operators the singular values of which decay at a certain speed, are introduced and their functional analytic properties studied in some detail.

The main result of the thesis is an explicit formula for the Hausdorff distance of the spectra of two operators belonging to the same compactness class. Along the way, upper bounds for the resolvents of operators belonging to a particular compactness class are established, as well as novel bounds for determinants of trace class operators.

Contents

CHAPTER 1

Introduction

Perturbation theory is the study of the behaviour of characteristic data of a mathematical object when replacing it by a similar nearby object. More narrowly, spectral perturbation theory is concerned with the change of spectral data of linear operators (such as their spectrum, their eigenvalues and corresponding eigenvectors) when the operators are subjected to a small perturbation.

There are two sides to spectral perturbation theory, a qualitative one and a quantitative one. Qualitative perturbation theory focusses on questions such as the continuity, differentiability and analyticity of eigenvalues and eigenvectors, while quantitative perturbation theory attempts to provide computationally accessible bounds for the smallness of the change in the spectral data in terms of the smallness of the perturbation.

The book by Kato [Kat76] is the main reference for spectral perturbation theory, focussing mostly on the qualitative part of the theory. Qualitative and quantitative aspects are discussed in the article and book by Chatelin [Cha81, Cha83] and the book by Hinrichsen and Pritchard [HP11].

The present thesis, located at the interface of functional analysis and linear algebra, addresses the following problem of fundamental importance in both qualitative and quantitative perturbation theory.

If A and B are two compact operators acting on a separable Hilbert space which are close, then how close are their spectra $\sigma(A)$ and $\sigma(B)$?

In order to make this question more precise we need to specify metrics to measure distances of operators and spectra. Distances of operators will typically be given by the underlying operator norm $\|\cdot\|$, while distances of spectra will be determined by the Hausdorff metric (see below).

A standard result in qualitative perturbation theory tells us that if A and B are compact operators and $\|A-B\|$ becomes vanishingly small, then so does the Hausdorff distance of their spectra (see, for example, [New51, Theorem 3]). However, this result does not give any quantitative information on how large the Hausdorff distance of $\sigma(A)$ and $\sigma(B)$ is when $\|A-B\|$ is small but non-zero.

Quantitative information of this type is interesting in situations where one wants to determine the spectrum of an arbitrary compact operator A on a separable Hilbert space by numerical means. The standard approach to solving this infinite-dimensional problem is to reduce it to a finite-dimensional one. This can, for example, be achieved as follows. Fix an orthonormal basis $(e_n)_{n\in\mathbb{N}}$ for the Hilbert space and define orthogonal projections onto the space spanned by the first k basis vectors by setting

$$P_k x = \sum_{n=1}^{k} (x, e_n) e_n \,,$$

where $(\cdot,\cdot)$ denotes the inner product of H. Now

$$A_k = P_k A P_k$$

is a finite rank operator, the spectrum of which is in principle computable, at least to arbitrary precision, since it boils down to the computation of the eigenvalues of a matrix. Moreover, it is possible to show that this sequence of finite rank operators $(A_k)_{k\in\mathbb{N}}$ converges to A in operator norm (see, for example, [ALL01, Theorem 4.1]). Thus, if quantitative bounds for the Hausdorff distance of the spectra of two compact operators are available, then the spectrum of A can be computed to arbitrary precision.

The main concern of the present thesis is to provide explicit upper bounds for the Hausdorff distance of the spectra of two arbitrary compact operators A and B on a separable Hilbert space in terms of the distance of the two operators A and B in operator norm.

In the finite-dimensional setting this problem has a long history going back more than 50 years. Before surveying it we briefly digress to fix notation.

For $z \in \mathbb{C}$ and a compact subset $\sigma \subset \mathbb{C}$ let

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

denote the distance of z to σ.

The *Hausdorff distance* $\operatorname{Hdist}(\cdot, \cdot)$, also known as the *Pompeiu-Hausdorff distance* (see [BP13] for some historical background), is the following metric defined on the set of compact subsets of $\mathbb{C}$

$$\operatorname{Hdist}(\sigma_1, \sigma_2) = \max\{\hat{d}(\sigma_1, \sigma_2), \hat{d}(\sigma_2, \sigma_1)\}$$

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 $\mathbb{C}$.

It is easy to see that the Hausdorff distance is a metric on the set of compact subsets of $\mathbb{C}$.

Now recall the following notions from matrix perturbation theory: for two bounded operators A and B, the *spectral variation of A with respect to B* is defined to be

$$\hat{d}(\sigma(A), \sigma(B)) \,,$$

while the *spectral distance of A and B* is

$$\mathrm{Hdist}(\sigma(A), \sigma(B))$$

(see, for example, [Gil03, Chapter 8, Definition 8.4.1]).

For arbitrary matrices, the first bound for the spectral distance is due to Ostrowski [Ost57] who approached this problem by bounding the Hausdorff distance of the zeros of the characteristic polynomials of the corresponding matrices. To be precise, he showed that if $A = (a_{ij})$ and $B = (b_{ij})$ are any two $n \times n$ matrices, then

$$\mathrm{Hdist}(\sigma(A), \sigma(B)) \leq (n+2) M_{Ost}^{1-\frac{1}{n}} \left\| A - B \right\|_{Ost}^{\frac{1}{n}} \,, \tag{1.1.1}$$

where

$$M_{Ost} = \max_{1 \leq i,j \leq n} \{ |a_{ij}| \,, |b_{ij}| \} \,, \quad \|A\|_{Ost} = \sum_{1 \leq i,j \leq n} |a_{ij}| \;.$$

Ostrowski later improved this result (see [Ost73, Appendix K]) by showing that the inequality (1.1.1) holds with the norm $\|\cdot\|_{Ost}$ replaced by

$$\|A\|_{Ost_{73}} = \frac{1}{n} \sum_{1 \leq i,j \leq n} |a_{ij}| \;.$$

Other bounds of this type can be found in [BM79, BF81] using different norms.

A sharp bound was found by Elsner in 1985 (see [Els85]) using a different approach. As his proof is very simple and has provided the inspiration for some of our results in Chapter 4 of this thesis we shall give it here.

Theorem 1.1.1. *Let A and B be $n \times n$ matrices. Then we have*

$$\mathrm{Hdist}(\sigma(A), \sigma(B)) \leq (\|A\| + \|B\|)^{1-\frac{1}{n}} \|A - B\|^{\frac{1}{n}} , \tag{1.1.2}$$

where $\|\cdot\|$ denotes the operator norm corresponding to the Euclidean norm on $\mathbb{C}^n$.

Proof. Since the right-hand side of (1.1.2) is symmetric in A and B it suffices to show that

$$\hat{d}(\sigma(A), \sigma(B)) \leq (\|A\| + \|B\|)^{1-\frac{1}{n}} \|A - B\|^{\frac{1}{n}} .$$

Let $\lambda \in \sigma(A)$ be chosen such that $\hat{d}(\sigma(A), \sigma(B)) = \inf_{\mu \in \sigma(B)} |\lambda - \mu|$. Now fix an orthonormal basis $e_1, \ldots, e_n$ with $Ae_1 = \lambda e_1$. Then

$$\left(\hat{d}(\sigma(A), \sigma(B))\right)^n = \left(\inf_{\mu \in \sigma(B)} |\lambda - \mu|\right)^n \leq \prod_{i=1}^{n} |\lambda - \mu_i| = |\det(\lambda I - B)| ,$$

where $\mu_1, \ldots, \mu_n$ denote the eigenvalues of B each repeated according to its algebraic multiplicity. Using the above and Hadamard's inequality (see, for example, [Pie86, A.4.5]) we have

$$\begin{aligned}
\left(\hat{d}(\sigma(A), \sigma(B))\right)^n &\leq \|(\lambda I - B)e_1\| \cdot \|(\lambda I - B)e_2\| \cdots \|(\lambda I - B)e_n\| \\
&= \|(A - B)e_1\| \cdot \|(\lambda I - B)e_2\| \cdots \|(\lambda I - B)e_n\| \\
&\leq \|(A - B)\| \, (\|A\| + \|B\|)^{n-1} ,
\end{aligned}$$

as $\|(\lambda I - B)e_k\| \leq |\lambda| + \|Be_k\| \leq \|A\| + \|B\|$, for $k = 2, 3, \ldots, n$. The theorem follows. □

Remark 1.1.2.

(i) Elsner's formula (1.1.2) is sharp in the following sense. Let a and b be positive real numbers and let A and B be $n \times n$ matrices, given by

$$A = aI\,, \quad Be_k = \begin{cases} -be_1 & \text{if } k = 1, \\ 0 & \text{if } 1 < k \leq n, \end{cases}$$

where $(e_k)_{k=1}^n$ is an orthonormal basis for $\mathbb{C}^n$. Since

$$\det(zI - A) = (z-a)^n\,, \quad \det(zI - B) = (z+b)\cdot z^{n-1}\,,$$

we have

$$\lambda_k(A) = a\,, \quad 1 \leq k \leq n\,,$$

$$\lambda_1(B) = -b\,, \quad \lambda_k(B) = 0\,, \quad 1 < k \leq n\,,$$

so

$$\mathrm{Hdist}(\sigma(A), \sigma(B)) = a + b\,.$$

Noting that $\|A - B\| = a + b$, $\|A\| = a$, $\|B\| = b$ we see that for these matrices there is equality in Elsner's formula.

(ii) The dimension-dependent exponent $1/n$ in Elsner's formula (1.1.2) cannot be improved. In order to see this let ϵ be a positive real number and let A and B_ϵ be $n \times n$ matrices given by

$$Ae_k = \begin{cases} e_{k+1} & \text{if } 1 \leq k \leq n-1, \\ 0 & \text{if } k = n, \end{cases} \qquad B_\epsilon e_k = \begin{cases} e_{k+1} & \text{if } 1 \leq k \leq n-1, \\ \epsilon e_1 & \text{if } k = n, \end{cases}$$

where $(e_k)_{k=1}^n$ is an orthonormal basis for $\mathbb{C}^n$. It is not difficult to see that

$$\det(zI - A) = z^n\,, \quad \det(zI - B_\epsilon) = z^n - \epsilon\,,$$

so

$$\lambda_k(A) = 0\,, \quad |\lambda_k(B_\epsilon)| = \epsilon^{1/n}\,, \quad 1 \leq k \leq n\,.$$

Thus

$$\operatorname{Hdist}(\sigma(A), \sigma(B_\epsilon)) = \epsilon^{1/n}\,.$$

Noting that $\|A - B_\epsilon\| = \epsilon$ we see that the exponent $1/n$ in Elsner's formula cannot be improved, as claimed.

In [Els85], Elsner conjectured that the inequality (1.1.2) holds for all operator norms taken with respect to any other (non-Euclidean) norm on $\mathbb{C}^n$. However, this conjecture was later disproved by Langlois and Ransford (see [LR02]).

A simple generalisation of the above bound to infinite dimensions valid for all compact operators is not possible due to the presence of the exponent $1/n$. Note, however, that versions of (1.1.2) hold for very special operators such as for A and B algebraic elements of a Banach algebra of degree at most n (see, for example, [CNR00]).

Using a more careful approach, to be presented in Chapter 4 of the present thesis, Elsner's proof can be modified so as to give a bound for the spectral variation of two matrices which generalises to trace class operators, but unfortunately not to arbitrary compact operators.

A different approach to estimating the spectral distance of matrices was discovered by Henrici in 1962 (see [Hen62]). As his approach can be generalised to the infinite-dimensional setting, and will play a fundamental role in

Chapter 3 of this thesis, we shall outline the main ideas.

Henrici's approach relies on two steps. The first step is to find an upper bound for the norm of the resolvent $(zI - A)^{-1}$ of a matrix A which depends only on the distance of z to the spectrum of A. Once this has been achieved, explicit upper bounds for the spectral distance of two matrices can be obtained in the second step, by using an argument going back to Bauer and Fike [BF60].

We shall now look at each of these steps in more detail. The principal idea to achieve the first step is to write a matrix A as a sum of a normal matrix D having the same spectrum as A and a nilpotent matrix N.

Lemma 1.1.3. *If A is an $n \times n$ matrix, then A can be written*

$$A = D + N\,,$$

where D and N are $n \times n$ matrices with the following properties:

(i) D is normal and $\sigma(D) = \sigma(A)$;

(ii) N is nilpotent of order at most n, that is, $N^n = 0$;

(iii) for any $z \notin \sigma(D) = \sigma(A)$ the matrix $(zI - D)^{-1}N$ is nilpotent of order at most n.

Proof. By a theorem of Schur (see, for example, [Mir55, Theorem 10.4.1]) there is a unitary matrix U such that $\tilde{A} = U^*AU$ is an upper triangular matrix. We now let $\tilde{D}$ denote the diagonal matrix, the main diagonal of which coincides with that of $\tilde{A}$, and write $\tilde{N} = \tilde{A} - \tilde{D}$. Thus

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

where $\tilde{D}$ is a normal matrix having the same eigenvalues as $\tilde{A}$, while $\tilde{N}$ is strictly upper triangular, which implies $\tilde{N}^n = 0$. Moreover, for any $z \notin \sigma(\tilde{D})$ the matrix $(zI - \tilde{D})^{-1}\tilde{N}$ is also strictly upper triangular, so must also be nilpotent of order at most n. Letting $D = U\tilde{D}U^*$ and $N = U\tilde{N}U^*$ we obtain matrices with the desired properties. □

The lemma motivates the following definition.

Definition 1.1.4. Let A be an $n \times n$ matrix. A decomposition of the form

$$A = D + N\,,$$

with D and N having the properties (i–iii) listed in the lemma above, is called a *Schur decomposition of* A. The matrix D will be referred to as the *normal part* and N as the *nilpotent part of the Schur decomposition of* A.

Note that a matrix may have more than one Schur decomposition. For an example see Remark 3.2.6.

Using the idea of a Schur decomposition allows us to quantify how far a matrix is from being normal.

Definition 1.1.5. Let A be a square matrix. We call

$$\nu(A) = \inf\{\|N\| : N \,\text{is the nilpotent part of a Schur decomposition of}\, A\}$$

the *departure from normality of* A.

The terminology introduced above is justified since A is normal if and only if $\nu(A) = 0$.

The departure from normality is not easy to calculate in practice. However, there is a simple upper bound already mentioned in Henrici's original paper. Using the fact that for an upper triangular matrix $\tilde{A}$ with normal part $\tilde{D}$ and nilpotent part $\tilde{N}$ we have

$$\left\|\tilde{A}\right\|_{HS}^2 = \left\|\tilde{D}\right\|_{HS}^2 + \left\|\tilde{N}\right\|_{HS}^2 ,$$

where $\|\cdot\|_{HS}$ denotes the Hilbert-Schmidt norm, it follows by unitary invariance of the Hilbert-Schmidt norm that for any square matrix A we have

$$\nu(A) \leq \sqrt{\|A\|_{HS}^2 - \sum_{i=1}^{n} |\lambda_i|^2} ,$$

where $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of A, with each eigenvalue repeated according to its algebraic multiplicity.

Before stating and proving Henrici's bound for the resolvent of an arbitrary matrix we recall the following elementary resolvent bounds for normal and nilpotent matrices. We start with the bound for normal matrices.

Lemma 1.1.6. *Let D be a normal matrix. Then*

$$\left\|(zI - D)^{-1}\right\| = \frac{1}{d(z, \sigma(D))} \quad (\forall z \notin \sigma(D)) . \tag{1.1.3}$$

Proof. Recall that if D is normal, then $\|D\| = r(D)$, where $r(D)$ denotes the spectral radius of D (see, for example, [Kat76, p. 55]). Now observe that if $z \notin \sigma(D)$, then $(zI - D)^{-1}$ is also normal. Thus

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

□

Remark 1.1.7. The lemma remains valid, with the same proof, for bounded normal operators on a separable Hilbert space.

Next we turn to the bound for nilpotent matrices.

Lemma 1.1.8. *Let N be nilpotent of order n. Then*

$$\left\|(I-N)^{-1}\right\| \le \sum_{k=0}^{n-1} \|N\|^k . \tag{1.1.4}$$

Proof. Since N is nilpotent of order n, we have $N^n = 0$, and it follows that

$$(I-N)^{-1} = \sum_{k=0}^{n-1} N^k ,$$

from which the bound follows. □

Henrici's idea for the first step was to combine these two bounds to produce an upper bound for the resolvent of an arbitrary square matrix.

Theorem 1.1.9. *Let A be an $n \times n$ matrix. Then*

$$\left\|(zI-A)^{-1}\right\| \le \frac{1}{d(z,\sigma(A))} f\left(\frac{\nu(A)}{d(z,\sigma(A))}\right) \quad (\forall z \notin \sigma(A)) . \tag{1.1.5}$$

Here $f : [0,\infty) \to [0,\infty)$ is the function defined by

$$f(r) = \sum_{k=0}^{n-1} r^k .$$

Proof. Let $A = D + N$ be a Schur decomposition of A and fix $z \notin \sigma(A)$. Then $(zI-D)^{-1}N$ is nilpotent of order at most n. Using (1.1.3), it follows that

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

thus, using (1.1.4), we obtain

$$\left\|(I-(zI-D)^{-1}N)^{-1}\right\| \le f\left(\frac{\|N\|}{d(z,\sigma(A))}\right) .$$

Since

$$(zI - A) = (zI - D)(I - (zI - D)^{-1}N)\,,$$

we conclude that

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

Taking the infimum over all Schur decompositions of A the theorem follows. □

Remark 1.1.10. Note that the estimate (1.1.5) is sharp in the sense that if A is normal, then (1.1.5) reduces to the standard estimate (1.1.3).

We now turn to discussing the second step in Henrici's proof, which is based on a result by Bauer and Fike [BF60]. Their simple but powerful argument shows how to obtain spectral variation bounds from bounds for resolvents. The formulation below is based on [Ban08, Proposition 4.1].

Theorem 1.1.11. *Let A be an $n \times n$ matrix. Suppose that there is a strictly monotonically increasing surjective function $g : [0,\infty) \to [0,\infty)$ and a positive constant C such that*

$$\left\|(zI - A)^{-1}\right\| \leq \frac{1}{C} g\left(\frac{C}{d(z,\sigma(A))}\right) \quad (\forall z \notin \sigma(A))\,.$$

Then, for any $n \times n$ matrix B, we have

$$\hat{d}(\sigma(B),\sigma(A)) \leq Ch\left(\frac{\|A - B\|}{C}\right)\,.$$

Here, the function $h : [0,\infty) \to [0,\infty)$ is given by

$$h(r) = (\tilde{g}(r^{-1}))^{-1}\,,$$

where $\tilde{g} : [0,\infty) \to [0,\infty)$ *is the inverse of the function* g.

Proof. Assume $B - A \neq 0$, since otherwise there is nothing to prove. We start by establishing the following statement:

$$\text{if } z \in \sigma(B), \text{ but } z \notin \sigma(A), \text{ then } \|B - A\|^{-1} \leq \left\|(zI - A)^{-1}\right\|. \tag{1.1.6}$$

This is done by contradiction. Let $z \in \sigma(B)$ and $z \notin \sigma(A)$. Assume to the contrary that

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

Then $(I - (zI - A)^{-1}(B - A))$ is invertible. It follows that

$$(zI - B) = (zI - A)(I - (zI - A)^{-1}(B - A))$$

is invertible. Therefore $z \notin \sigma(B)$ which contradicts $z \in \sigma(B)$. Hence statement (1.1.6) holds.

In order to prove the theorem it suffices to show that if $z \in \sigma(B)$, then

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

Let $z \in \sigma(B)$. If $z \in \sigma(A)$, then the left-hand side of the above inequality is zero, hence there is nothing to prove. Now assume $z \notin \sigma(A)$. By (1.1.6) and the hypothesis we have

$$\frac{1}{\|B - A\|} \leq \left\|(zI - A)^{-1}\right\| \leq \frac{1}{C} g\left(\frac{C}{d(z, \sigma(A))}\right).$$

Since g is strictly monotonically increasing, so is $\tilde{g}$. Therefore

$$\tilde{g}\left(\frac{C}{\|B - A\|}\right) \leq \frac{C}{d(z, \sigma(A))}\,,$$

and so

$$d(z, \sigma(A)) \leq \frac{C}{\tilde{g}\left(\frac{C}{\|B-A\|}\right)} = Ch\left(\frac{\|B-A\|}{C}\right),$$

as desired. □

Remark 1.1.12. Using the same proof, the above theorem is also valid for A and B bounded operators on a Banach space.

By combining Theorems 1.1.9 and 1.1.11 we are finally able to state Henrici's spectral variation and spectral distance formulae.

Theorem 1.1.13. *Let A and B be $n \times n$ matrices.*

(i) If A is not normal, then

$$\hat{d}(\sigma(B), \sigma(A)) \leq \nu(A) h\left(\frac{\|A-B\|}{\nu(A)}\right).$$

(ii) If neither A nor B are normal, then

$$\mathrm{Hdist}(\sigma(A), \sigma(B)) \leq mh\left(\frac{\|A-B\|}{m}\right),$$

where $m := \max\{\nu(A), \nu(B)\}$.

Here, the function $h : [0, \infty) \to [0, \infty)$ is given by

$$h(r) = (\tilde{g}(r^{-1}))^{-1},$$

where $\tilde{g} : [0, \infty) \to [0, \infty)$ is the inverse of the function

$$g(r) = \sum_{k=1}^{n} r^k.$$

Proof.

(i) By Theorem 1.1.9 we have

$$\left\|(zI-A)^{-1}\right\| \le \frac{1}{\nu(A)} g\left(\frac{\nu(A)}{d(z,\sigma(A))}\right) \quad (\forall z \notin \sigma(A)),$$

so the assertion follows from Theorem 1.1.11.

(ii) Similarly, by Theorem 1.1.9 we have

$$\left\|(zI-A)^{-1}\right\| \le \frac{1}{m} g\left(\frac{m}{d(z,\sigma(A))}\right) \quad (\forall z \notin \sigma(A)),$$

and

$$\left\|(zI-B)^{-1}\right\| \le \frac{1}{m} g\left(\frac{m}{d(z,\sigma(B))}\right) \quad (\forall z \notin \sigma(B)),$$

so again the assertion follows from Theorem 1.1.11.

□

Remark 1.1.14. Note that if A is a normal $n \times n$ matrix, then combining Lemma 1.1.6 and Theorem 1.1.11 yields, for any $n \times n$ matrix B, the spectral variation bound

$$\hat{d}(\sigma(B), \sigma(A)) \le \|A - B\| .$$

The same argument shows that if A and B are normal matrices, then we have the sharp bound

$$\text{Hdist}(\sigma(A), \sigma(B)) \le \|A - B\| . \tag{1.1.7}$$

Note that these two bounds can be thought of as limiting cases of the theorem above, since, as is easily seen, we have for any $r \ge 0$

$$\lim_{C \downarrow 0} Ch\left(\frac{r}{C}\right) = r .$$

We also note that the spectral distance bound (1.1.7) does not depend on the dimension of the underlying space and turns out to be valid, more generally, for any two bounded normal operators on a separable Hilbert space. This can for example be seen by combining Remarks 1.1.7 and 1.1.12.

The main aim of this thesis is to extend the spectral variation and spectral distance bounds for matrices, which we have just surveyed, to arbitrary compact operators on a separable Hilbert space. So far, infinite-dimensional analogues of these bounds have been obtained only for certain subclasses of compact operators. To the best of our knowledge, the first results in this direction are due to Gil′, who, in a series of papers begun in 1979, has obtained spectral variation and distance bounds mostly for operators in the Schatten classes (see [Gil95, Gil03] and references therein) and more recently for operators with inverses in the Schatten classes (see [Gil12, Gil14]). Pokrzywa [Pok85] has found similar bounds for operators in symmetrically normed ideals, while Bandtlow obtained bounds, simpler and sharper than those of Gil′ and Pokrzywa, for Schatten class operators [Ban04] and for operators in certain subclasses of trace class operators, termed exponential classes [Ban08]. All three authors essentially use Henrici's approach to obtain their bounds, by first deriving resolvent bounds for quasi-nilpotent operators and then using the perturbation argument outlined in the proof of Theorem 1.1.9.

The present thesis is organised as follows. Chapter 2 deals with the technical background which we will require throughout the thesis. Basic properties of compact operators, including their eigenvalues and singular values will be covered, as well as the connection between the two, known as Weyl inequalities. We also discuss trace class operators and the notions of a trace and a determinant, which can be defined for these operators, and finish with a short discussion of Schatten class operators and their properties.

Chapters 3 and 4 contain the main results of the thesis. In Chapter 3 we introduce new classes of operators, termed compactness classes, which generalise Bandtlow's exponential classes from [Ban08]. The basic idea here is to group together all operators the singular values of which decay at a certain speed. We shall study the functional analytic properties of these classes in some detail. In particular we shall find sufficient conditions guaranteeing that these classes of operators form quasi-Banach operator ideals in the sense of Pietsch. We shall then use a theorem of Dostanić to produce bounds for the resolvents of quasi-nilpotent operators in a given compactness class. Using the technique of Henrici discussed earlier on (see Theorem 1.1.9) we then obtain an upper bound for $\|(zI - A)^{-1}\|$ for an arbitrary operator A in a given compactness class, which depends only on the asymptotics of the singular values of A and the distance of z to the spectrum of A (see Theorem 3.2.12).

Using the Bauer-Fike argument discussed earlier (see Theorem 1.1.11), these resolvent bounds then yield the main result of the thesis: an explicit upper bound for the spectral distance of two operators in a given compactness class involving only the distance of the two operators in operator norm and the asymptotics of the singular values of the two operators (see Theorem 3.4.1). To the best of our knowledge, no bound for the spectral distance applicable to arbitrary compact operators has appeared in the literature yet.

We finish the chapter with an application of the resolvent bounds producing explicitly computable circular inclusion regions for pseudospectra of a given compact operator.

In Chapter 4, the main part of which has been published in [BG15], we will revisit Elsner's determinant based proof of Theorem 1.1.1 and show, how, with a bit of care, it can be made to work in the infinite-dimensional setting, producing spectral variation and distance bounds without recourse to Henrici's approach. We shall establish upper and lower bounds for determinants of trace class operators, which are of independent interest and do not seem to have appeared in the literature yet. Using these determinant bounds we obtain a version of Elsner's formula (see Theorem 4.2.3) which holds for trace class operators and which produces new bounds even in the finite-dimensional setting. Finally we will compare our bound to Elsner's original one and show that it reproduces or improves existing bounds in [Ban04, Ban08].

CHAPTER 2

Basic Spectral Properties of Compact Operators

In this chapter we provide the technical background for the research to be presented in the following chapters. Most of the material is assembled from the books [GK69], [GGK90] and [GGK03]. For convenience of the reader we shall provide details of proofs of some of the results, whenever these are instructive. Singular values play a key role in this study of compact operators since they characterize the approximability of a given operator by operators of finite rank. We shall focus on the basic properties of eigenvalues and singular values of compact operators, the connection between eigenvalues and singular values of compact operators, known as Weyl inequalities, and introduce trace class operators and more general classes known as Schatten classes.

2.1 Eigenvalues of compact operators

Notation 2.1.1. The symbols $\mathbb{N}$, $\mathbb{R}^+$ and $\mathbb{C}$ will denote the set of all positive integers, positive real numbers and complex numbers, respectively. We will use the abbreviation $\mathbb{R}_0^+ := \mathbb{R}^+ \cup \{0\}$.

Throughout the thesis H or H_i will denote separable Hilbert spaces over $\mathbb{C}$. The corresponding scalar products will be denoted by $(\cdot,\cdot)$ and the induced norm by $\|\cdot\| = \sqrt{(\cdot,\cdot)}$.

We write $L(H_1, H_2)$ to denote the Banach space of bounded linear operators from H_1 to H_2 equipped with the operator norm $\|\cdot\|$ and $S_\infty(H_1, H_2) \subset L(H_1, H_2)$ to denote the closed subspace of compact operators from H_1 to H_2. If $H = H_1 = H_2$ we use the short-hands $L(H)$ and $S_\infty(H)$ for $L(H_1, H_2)$ and $S_\infty(H_1, H_2)$, respectively.

For $A \in L(H)$ the spectrum, the set of eigenvalues and the resolvent set will be denoted by $\sigma(A)$, $\sigma_p(A)$ and $\rho(A)$, respectively. The resolvent of an operator A is defined by $R(A; z) := (zI - A)^{-1}$ for every $z \notin \sigma(A)$.

Let $A \in S_\infty(H)$. It is known that the spectrum $\sigma(A)$ is either a finite set or a sequence converging to zero. The point zero is the only possible accumulation point of $\sigma(A)$ and each non-zero element of the spectrum is an eigenvalue of A of finite algebraic multiplicity (see, for example, [Kre89, Chapter 8, Theorems 8.3-1 and 8.4-4]).

We now let $\lambda(A) = (\lambda_k(A))_{k\in\mathbb{N}}$ denote the eigenvalue sequence of A with the following convention:

1. Each non-zero eigenvalue of A is repeated in the sequence $\lambda(A)$ as many times as the value of its algebraic multiplicity.

2. The eigenvalues are ordered by decreasing modulus so that

$$|\lambda_1(A)| \geq |\lambda_2(A)| \geq \cdots \geq 0\,.$$

3. The number of non-zero eigenvalues of A is finite if and only if A is a finite rank operator. In that case we extend the sequence $\lambda(A)$ by adding zero elements so that it is an infinite sequence in all cases. In particular,

if $A \in S_\infty(H)$ is quasi-nilpotent, that is, $\sigma(A) = \{0\}$, then $\lambda_k(A) = 0$ for every $k \in \mathbb{N}$.

Occasionally, we shall use $|\lambda(A)|$ for the sequence $(|\lambda_k(A)|)_{k\in\mathbb{N}}$.

2.2 Singular values of compact operators

Another sequence of numbers associated to a compact operator is its sequence of singular values which we now turn to.

Let $A \in S_\infty(H_1, H_2)$. Then A^*A belongs to $S_\infty(H_1)$ and is positive since

$$(A^*Ax, x) = (Ax, Ax) = \|Ax\|^2 \geq 0 \quad (\forall x \in H_1)\,. \tag{2.2.1}$$

Thus, the sequence of non-zero eigenvalues of A^*A ordered according to our convention satisfies

$$\lambda_1(A^*A) \geq \lambda_2(A^*A) \geq \cdots \geq 0\,.$$

Definition 2.2.1. Let $A \in S_\infty(H_1, H_2)$. For $k \in \mathbb{N}$, the *k-th singular value of A* is defined by

$$s_k(A) = \sqrt{\lambda_k(A^*A)}\,.$$

Remark 2.2.2. The definition immediately implies the following properties:

(i) $s_k(A) \geq s_{k+1}(A)$ for every $k \in \mathbb{N}$.

(ii) $s_k(A) \to 0$ as $k \to \infty$.

(iii) Let $\alpha \in \mathbb{C}$. Then, for every $k \in \mathbb{N}$,

$$s_k(\alpha A) = |\alpha|\, s_k(A)\,. \tag{2.2.2}$$

In what follows a number of alternative characterisations of eigenvalues and singular values will prove to be useful.

Theorem 2.2.3. *Suppose that* $A \in S_\infty(H)$ *is positive. Then we have*

$$\lambda_k(A) = \min_{\substack{M \subset H \\ \dim M = k-1}} \max_{\substack{\|x\|=1 \\ x \perp M}} (Ax, x) \quad (\forall k \in \mathbb{N}),$$

where the min *is taken over all closed subspaces* M *of* H *of dimension* $k-1$.

Proof. See [GGK03, Chapter IV, Theorem 9.1]. □

Theorem 2.2.4. *Let* $A \in S_\infty(H_1, H_2)$. *Then we have*

$$s_k(A) = \min_{\substack{M \subset H_1 \\ \dim M = k-1}} \max_{\substack{\|x\|=1 \\ x \perp M}} \|Ax\| \quad (\forall k \in \mathbb{N}),$$

where the min *is taken over all closed subspaces* M *of* H_1 *of dimension* $k-1$.

Proof. Follows from equality (2.2.1) and Theorem 2.2.3 applied to A^*A. □

Singular values play an important role in the following representation of compact operators.

Theorem 2.2.5. *Let* $A \in S_\infty(H_1, H_2)$ *and let* $N \in \mathbb{N} \cup \{\infty\}$ *denote the number of non-zero singular values of* A. *Then there exist orthonormal systems* $(e_k)_{k=1}^N \subseteq H_1$ *and* $(f_k)_{k=1}^N \subseteq H_2$ *such that*

$$Ax = \sum_{k=1}^{N} s_k(A)(x, e_k) f_k \quad (\forall x \in H_1). \tag{2.2.3}$$

Proof. See [GGK03, Chapter X, Theorem 4.1]. □

Remark 2.2.6. Formula (2.2.3) is called a *Schmidt representation of* A.

Before continuing we note that if $A \in S_\infty(H_1, H_2)$ has a Schmidt representation of the form (2.2.3) then its adjoint $A^* \in S_\infty(H_2, H_1)$ has a Schmidt representation of the form

$$A^*y = \sum_{k=1}^{N} s_k(A)(y, f_k)e_k \quad (\forall y \in H_2)\,, \tag{2.2.4}$$

since for any $x \in H_1$ and $y \in H_2$, we have

$$(Ax, y) = \sum_{k=1}^{N} s_k(A)(x, e_k)(f_k, y) = (x, \sum_{k=1}^{N} s_k(A)(y, f_k)e_k)\,.$$

As a result we obtain the following corollary.

Corollary 2.2.7. *Let $A \in S_\infty(H_1, H_2)$. Then the operator A and its adjoint A^* have the same singular values.*

Proof. Let A have a Schmidt representation of the form (2.2.3). A short calculation using (2.2.4) shows that

$$AA^*y = \sum_{k=1}^{N} s_k^2(A)(y, f_k)f_k \quad (\forall y \in H_2)\,,$$

implying

$$s_k(A^*) = \sqrt{\lambda_k(AA^*)} = s_k(A) \quad (\forall k \in \mathbb{N})\,,$$

as required. □

Remark 2.2.8. The corollary also follows from the fact that if A and B are two compact operators on a Hilbert space, then AB and BA have the same non-zero eigenvalues, counting algebraic multiplicities (see, for example, [Pie86, Lemma 3.3.1]).

The following result will be used several times in this thesis, for example, to obtain part of the ideal structure of the space of trace class operators in Proposition 2.4.1.

Proposition 2.2.9. *Let $A \in S_\infty(H_3, H_2)$, $B \in L(H_1, H_2)$ and $C \in L(H_4, H_3)$. Then we have*

$$s_k(BAC) \leq \|B\| \, s_k(A) \, \|C\| \qquad (\forall k \in \mathbb{N}) \, .$$

Proof. It suffices to show that, for every $k \in \mathbb{N}$,

$$s_k(BA) \leq \|B\| \, s_k(A) \quad \text{and} \quad s_k(AC) \leq s_k(A) \, \|C\| \; .$$

Using the fact that $\|BAx\| \leq \|B\| \, \|Ax\|$ and the min-max characterisation of the singular values in Theorem 2.2.4 we obtain

$$s_k(BA) = \min_{\substack{M \subset H_3 \\ \dim M = k-1}} \max_{\substack{\|x\|=1 \\ x \perp M}} \|BAx\| \leq \min_{\substack{M \subset H_3 \\ \dim M = k-1}} \max_{\substack{\|x\|=1 \\ x \perp M}} \|B\| \, \|Ax\| \; ,$$

which shows that, for every $k \in \mathbb{N}$,

$$s_k(BA) \leq \|B\| \, s_k(A) \, .$$

Using the result above and Corollary 2.2.7 we get

$$s_k(AC) = s_k((AC)^*) = s_k(C^* A^*) \leq \|C^*\| \, s_k(A^*) = \|C\| \, s_k(A)$$

and the theorem is proved. □

Corollary 2.2.10. *Let $A \in S_\infty(H)$. Then for any self-adjoint projection $P \in L(H)$ we have*

$$s_k(PA|_{P(H)}) \leq s_k(A) \quad (\forall k \in \mathbb{N}) \, .$$

Proof. This follows from the observation that $PA|_{P(H)} : P(H) \to P(H)$ and $PAP : P(H) \to P(H)$ have the same singular values together with Proposition 2.2.9 and the fact that $\|P\| \leq 1$, since

$$s_k(PAP) \leq \|P\|\, s_k(A)\, \|P\| \leq s_k(A) \quad (\forall k \in \mathbb{N})\,,$$

as desired. □

The following theorem provides an alternative characterisation of the singular values of an operator A, linking them to the degree of approximability of A by finite rank operators.

Theorem 2.2.11. *Let $A \in S_\infty(H_1, H_2)$. Then we have*

$$s_k(A) = \inf\{\, \|A - F\| \,:\, F \in L(H_1, H_2),\, \mathrm{rank}(F) \leq k - 1\,\} \quad (\forall k \in \mathbb{N})\,.$$

Proof. Fix $k \in \mathbb{N}$ and suppose that $\mathrm{rank}(F) = m \leq k-1$. Now $\dim(\mathrm{Ker}(F))^\perp = \dim \mathrm{Ran}(F) = m$. Therefore by Theorem 2.2.4 with $M = (\mathrm{Ker}(F))^\perp$,

$$s_k(A) \leq s_{m+1}(A) \leq \max_{\substack{\|x\|=1 \\ x \in \mathrm{Ker}(F)}} \|Ax\| = \max_{\substack{\|x\|=1 \\ x \in \mathrm{Ker}(F)}} \|(A - F)x\| \leq \|A - F\|\,.$$

Thus $s_k(A) \leq \|A - F\|$ for any F with $\mathrm{rank}(F) \leq k - 1$. It remains to prove that the infimum is attained and is equal to $s_k(A)$. Let N denote the number of non-zero singular values of A and let $A = \sum_{j=1}^{N} s_j(A)(., e_j) f_j$ be a Schmidt representation of A. For any $k < N + 1$, define the operator

$$F_k x = \begin{cases} 0 & \text{if } k = 1, \\ \sum_{j=1}^{k-1} s_j(A)(x, e_j) f_j & \text{if } k > 1. \end{cases}$$

Since $\operatorname{rank}(F_k) = k - 1$ and

$$\|(A - F_k)x\|^2 = \left\| \sum_{k}^{N} s_j(A)(x, e_j) f_j \right\|^2 \leq s_k(A)^2 \|x\|^2 ,$$

it follows that $\|A - F_k\| \leq s_k(A)$. Thus the infimum is attained at $F = F_k$ and is equal to $s_k(A)$. If $k > N$, then $\operatorname{rank}(A) \leq k - 1$ and $s_k(A) = 0$. Hence in this case the infimum is attained at $F = A$. □

Remark 2.2.12. If in the proof of the theorem above we use the spectral theorem for normal operators (see, for example, [EE87, Chapter II, Theorem 5.2]) in place of the Schmidt representation, it follows that for any compact normal operator A we have

$$s_k(A) = |\lambda_k(A)| \quad (\forall k \in \mathbb{N}) .$$

As a first application of the above alternative characterisation of singular values in Theorem 2.2.11, we now give the following inequality originally due to Fan [Fan51].

Corollary 2.2.13. *Let $A, B \in S_\infty(H_1, H_2)$. Then we have*

$$s_{m+n-1}(A + B) \leq s_m(A) + s_n(B) \quad (\forall m, n \in \mathbb{N}) .$$

Proof. Fix $m, n \in \mathbb{N}$. Now choose $F, G \in L(H_1, H_2)$ with $\operatorname{rank}(F) \leq m - 1$ and $\operatorname{rank}(G) \leq n - 1$. Since $\operatorname{rank}(F + G) \leq m + n - 2$, we have

$$s_{m+n-1}(A + B) \leq \|(A + B) - (F + G)\| \leq \|A - F\| + \|B - G\| .$$

Taking the infimum over all operators F with $\operatorname{rank}(F) \leq m - 1$ and all operators G with $\operatorname{rank}(G) \leq n - 1$ we obtain the required inequality. □

An immediate consequence of the above corollary is the following result, which shows that the singular values are continuous functions of compact operators in $L(H_1, H_2)$.

Corollary 2.2.14. *Let* $A, B \in S_\infty(H_1, H_2)$. *Then we have*

$$|s_k(A) - s_k(B)| \le \|A - B\| \quad (\forall k \in \mathbb{N}).$$

Proof. Fix $k \in \mathbb{N}$. By Corollary 2.2.13, we have

$$s_k(A) = s_k(B + A - B) \le s_k(B) + s_1(A - B) = s_k(B) + \|A - B\| ,$$

so

$$s_k(A) - s_k(B) \le \|A - B\| , \tag{2.2.5}$$

and by symmetry,

$$-(s_k(A) - s_k(B)) \le \|A - B\| . \tag{2.2.6}$$

Combining (2.2.5) and (2.2.6), the assertion follows. □

Corollary 2.2.15. *Let* $A, A_n \in S_\infty(H_1, H_2)$. *If* $\lim_{n\to\infty} A_n = A$ *in operator norm, then*

$$\lim_{n\to\infty} s_k(A_n) = s_k(A) \quad (\forall k \in \mathbb{N}).$$

Proof. Follows from Corollary 2.2.14 since $|s_k(A_n) - s_k(A)| \le \|A_n - A\|$. □

2.3 Weyl inequalities

In this section we briefly discuss a number of relationships between eigenvalues and singular values of compact operators originally due to Weyl (see [Wey49]). We start with the most important one, known as the *multiplicative Weyl inequality*.

Theorem 2.3.1. *Let* $A \in S_\infty(H)$. *Then we have*

$$\prod_{k=1}^{n} |\lambda_k(A)| \le \prod_{k=1}^{n} s_k(A) \quad (\forall n \in \mathbb{N}).$$

Proof. See [GGK90, Chapter VI, Theorem 2.1]. □

Remark 2.3.2. While the proof of multiplicative Weyl inequality is a bit involved, there is a very special case, which is easily understood. If A is an $n \times n$ matrix, then

$$\prod_{k=1}^{n} |\lambda_k(A)| = \prod_{k=1}^{n} s_k(A).$$

In order to see this, note that

$$\prod_{k=1}^{n} |\lambda_k(A)|^2 = |\det(A)|^2 = \det(A^*)\det(A) = \det(A^*A) = \prod_{k=1}^{n} s_k(A)^2.$$

The following inequality in Corollary 2.3.3, known as *additive Weyl inequality*, which can be derived from the multiplicative one, plays an important role in extending the notion of the trace of a matrix to certain compact operators, known as 'trace class operators', to be discussed in the next section.

Corollary 2.3.3. *Let* $A \in S_\infty(H)$. *Then we have*

$$\sum_{k=1}^{n} |\lambda_k(A)| \le \sum_{k=1}^{n} s_k(A) \quad (\forall n \in \mathbb{N}).$$

Proof. See [GGK90, Chapter VI, Corollary 2.4]. □

We finish with an inequality, which can also be derived from the multiplicative Weyl inequality and which plays an important role in extending the notion of a determinant of a matrix to trace class operators.

Corollary 2.3.4. *Let $A \in S_\infty(H)$. Then we have*

$$\prod_{k=1}^{n}(1+r|\lambda_k(A)|) \leq \prod_{k=1}^{n}(1+rs_k(A)) \quad (\forall r \geq 0, \forall n \in \mathbb{N})\,.$$

Proof. See [GGK90, Chapter VI, Corollary 2.5]. □

2.4 Trace class operators

The notion of a trace class operator grew out of attempts to extend the notions of 'trace' and 'determinant' familiar from linear algebra, to operators acting on infinite-dimensional spaces. To motivate what is to follow, we briefly recall the finite-dimensional case.

For $A \in L(\mathbb{C}^n)$ we can define the trace and determinant of the operator A in the usual way

$$\operatorname{tr}(A) = \sum_{k=1}^{n}(Ae_k, e_k)\,, \quad \det(I+A) = \det(\delta_{lk} + (Ae_k, e_l))_{l,k=1}^{n}\,,$$

where $e_1, \ldots, e_n$ is an arbitrary orthonormal basis in $\mathbb{C}^n$. We add that it is possible to show that the trace and determinant defined above are well-defined in the sense that they do not depend on the chosen basis of $\mathbb{C}^n$.

If we choose $e_1, \ldots, e_n$ so that the corresponding matrix representation of A is upper triangular we see that the eigenvalues of A, taking algebraic multiplicities into account, are precisely the diagonal elements of the resulting matrix. Thus, in particular, we have

$$\operatorname{tr}(A) = \sum_{k=1}^{n}\lambda_k(A), \quad \det(I+A) = \prod_{k=1}^{n}(1+\lambda_k(A))\,. \tag{2.4.1}$$

In order to extend these concepts to the infinite-dimensional case, we need to ensure that the expressions in (2.4.1) converge. In order to do this we define the following collection of operators

$$S_1(H_1, H_2) = \{\, A \in S_\infty(H_1, H_2) \,:\, \|A\|_1 = \sum_{k=1}^{\infty} s_k(A) < \infty \,\}\,.$$

The elements of $S_1(H_1, H_2)$ are called *trace class operators*, while the norm $\|\cdot\|_1$ is referred to as *trace norm*. As usual, we will use the shorthand $S_1(H)$ if $H = H_1 = H_2$.

It turns out that for any operator $A \in S_1(H)$ one can define a trace $\mathrm{tr}(A)$ and a determinant $\det(I + A)$ using (2.4.1) as a definition, observing that the series and infinite product converge by Corollaries 2.3.3 and 2.3.4.

It is possible to show that $S_1(H)$ is a Banach space and that finite rank operators are dense in $S_1(H)$ with respect to the trace norm $\|\cdot\|_1$ (see [GGK90, Chapter VI, Theorem 4.1]). We also note that $S_1(H)$, on top of being a Banach space, is a two-sided ideal in the algebra $L(H)$, as the following proposition shows.

Proposition 2.4.1. *Let $A \in S_1(H)$ and $B, C \in L(H)$. Then $BAC \in S_1(H)$ and*

$$\|BAC\|_1 \leq \|B\| \, \|A\|_1 \, \|C\| \,.$$

Proof. Follows from Proposition 2.2.9. □

These two properties of $S_1(H)$ are captured in the following definition, originally due to Pietsch (see, for example, [Pie80, 1.1.1]).

Definition 2.4.2. A *two-sided operator ideal* $S(H)$ is a subclass of $L(H)$ with the following properties:

(i) $S(H)$ is a linear subspace of $L(H)$.

(ii) If $A \in S(H)$ and $B, C \in L(H)$, then $BAC \in S(H)$.

If, in addition, $S(H)$ has a norm turning it into a Banach space, then $S(H)$ is called a *Banach operator ideal.*

It turns out that the techniques developed in Chapter 3 of this thesis yield slightly more general structures than Banach operator ideals, known as 'quasi-Banach operator ideals' also introduced by Pietsch (see, for example, [Pie86, D.1]), which we now briefly describe.

Definition 2.4.3. A *quasi-norm* $\|\cdot\|$ defined on a linear space X is a real-valued function with the following properties:

(i) For any $x \in X$, we have $\|x\| = 0$ if and only if $x = 0$.

(ii) For any $x \in X$ and any $\lambda \in \mathbb{C}$, we have $\|\lambda x\| \leq |\lambda| \, \|x\|$.

(iii) There is a constant c_X such that

$$\|x\| + \|y\| \leq c_X(\|x\| + \|y\|) \quad (\forall x, y \in X)\,. \tag{2.4.2}$$

Remark 2.4.4. Inequality (2.4.2) is known as *quasi-triangle inequality.* If $c_X = 1$, then $\|\cdot\|$ is a *norm* and (2.4.2) becomes the well-known *triangle inequality.*

Definition 2.4.5. A two-sided operator ideal $S(H)$ of $L(H)$ is called a *quasi-Banach operator ideal* if it has a quasi-norm, turning it into a complete metric space.

For examples of Banach and quasi-Banach operator ideals, let $p \in (0, \infty)$ and define

$$S_p(H_1, H_2) = \left\{ A \in S_\infty(H_1, H_2) \,:\, \|A\|_p = \left(\sum_{k=1}^{\infty} s_k(A)^p \right)^{\frac{1}{p}} < \infty \right\}.$$

The elements of $S_p(H_1, H_2)$ are called *Schatten-von Neumann operators*. As usual, we write $S_p(H)$ for $S_p(H, H)$. It turns out that, if $p \geq 1$, then $\|\cdot\|_p$ is a norm turning $S_p(H)$ into a Banach operator ideal, while, if $p < 1$, then $\|\cdot\|_p$ is only a quasi-norm turning $S_p(H)$ into a quasi-Banach operator ideal (see, for example, [Pie86, 2.11.20]).

CHAPTER 3

Compactness Classes

The main purpose of this chapter is to show that the approach of Henrici [Hen62] for obtaining spectral distance bounds for matrices outlined in the introduction can be made to work for arbitrary compact operators on a separable Hilbert space. In order to do this we shall first introduce new classes of operators, termed 'compactness classes', determined by the speed of decay of the singular values of the operators in the class. This will be done in Section 3.1, where we give the precise definition of these classes and where we study some of their functional analytic properties.

In Section 3.2 we shall derive an explicit upper bound for the norm of the resolvent $R(A;z)$ of an operator A in a given compactness class in terms of the distance of z to the spectrum of A and a number measuring the departure from normality of A (see Theorem 3.2.12).

The following Section 3.3 is devoted to studying the behaviour of the bound for the norm of resolvents derived in the previous section for two particular families of compactness classes already in the literature.

In Section 3.4, the general resolvent bounds obtained in Section 3.2 together with the Bauer-Fike argument discussed in the introduction (see Theorem 1.1.11) will yield the main result of this thesis, an explicit upper bound for the spectral distance of two operators in a given compactness class, depending only on the distance in operator norm of the operators and their respective departures from normality (see Theorem 3.4.1). A particular feature of this result is that it turns out to be sharp for normal operators (see Remark 3.4.2 (iii)).

In the final section we will briefly discuss an application of the main result giving circular inclusion regions for pseudospectra of an operator in a given compactness class (see Theorem 3.5.2).

3.1 Compactness classes

The basic idea to define these classes is to group together all compact operators on a separable Hilbert space the singular values of which decay at a certain speed, quantified by a given 'weight sequence' (see Definition 3.1.2).

The aim of this section is to examine the behaviour of compactness classes under addition and multiplication, to show that these classes are quasi-Banach operator ideals under suitable conditions on the weight sequence and to determine the decay rate of the eigenvalue sequence of an operator in a given compactness class.

We start by defining the notion of a weight sequence.

Definition 3.1.1. Let

$$\mathcal{W} = \left\{ w : \mathbb{N} \to \mathbb{R}_0^+ \,:\, w_k \geq w_{k+1},\ \forall k \in \mathbb{N} \text{ and } \lim_{k\to\infty} w_k = 0 \right\}.$$

Elements of $\mathcal{W}$ will be referred to as *weight sequences*, or simply *weights*.

Every $w \in \mathcal{W}$ now gives rise to a compactness class as follows.

Definition 3.1.2. Let $w \in \mathcal{W}$. An operator $A \in S_\infty(H_1, H_2)$ is said to be *w-compact* if there is a constant $M \geq 0$ such that

$$s_k(A) \leq M w_k \quad (\forall k \in \mathbb{N})\,. \tag{3.1.1}$$

The infimum over all M such that (3.1.1) holds will be referred to as the *w-gauge of A* and will be denoted by $|A|_w$.

The collection of all w-compact operators $A \in S_\infty(H_1, H_2)$ will be denoted by $E_w(H_1, H_2)$ or simply by $E_w(H)$ in case $H = H_1 = H_2$.

For later use we also define the following sequence space analogues of compactness classes.

Definition 3.1.3. Given $w \in \mathcal{W}$, let $\mathcal{E}_w$ denote the set of all complex-valued sequences $(x_n)_{n\in\mathbb{N}}$ for which there is a constant $M \geq 0$ such that

$$|x_k| \leq M w_k \quad (\forall k \in \mathbb{N})\,. \tag{3.1.2}$$

The infimum over all M such that (3.1.2) holds will be referred to as the *w-gauge of x* and will be denoted by $|x|_w$.

Remark 3.1.4. It is not difficult to see that $\mathcal{E}_w$ is a Banach space when equipped with the w-gauge $|\cdot|_w$. The situation is different for E_w, which need not even be a linear space in general (see Proposition 3.1.12).

Compactness classes generalise classes that have already appeared in the literature, such as the Schatten-Lorentz ideals $S_{p,\infty}$ (see, for example, [Pel85, p. 481]), which correspond to the weights $w_k = k^{-1/p}$ with $p \in (0, \infty)$ or the 'exponential classes' studied by Bandtlow (see [Ban08]), which correspond to weights of the form $w_k = \exp(-ak^\alpha)$ with $a \in (0, \infty)$ and $\alpha \in (0, \infty)$.

We shall now explore some of the properties of $E_w(H_1, H_2)$ for a general weight w. We start with the following elementary observation.

Proposition 3.1.5. *Let $v, w \in \mathcal{W}$. If there exists $M \geq 0$ such that $v_k \leq M w_k$ for every $k \in \mathbb{N}$ and $A \in E_v(H_1, H_2)$, then $A \in E_w(H_1, H_2)$ and $|A|_w \leq M|A|_v$.*

Proof. Suppose $A \in E_v(H_1, H_2)$ and there exists $M \geq 0$ such that $v_k \leq M w_k$ for every $k \in \mathbb{N}$. Then we have, for every $k \in \mathbb{N}$,

$$s_k(A) \leq |A|_v v_k \leq |A|_v M w_k \, .$$

Hence we obtain $A \in E_w(H_1, H_2)$ and $|A|_w \leq M|A|_v$. □

The observation above motivates defining a partial order on $\mathcal{W}$ as follows

$$v \preceq w :\Longleftrightarrow \exists M \geq 0 \quad \text{such that} \quad v_k \leq M w_k \quad (\forall k \in \mathbb{N}) \, .$$

We shall also define an equivalence relation on $\mathcal{W}$ by setting

$$v \asymp w :\Longleftrightarrow v \preceq w \quad \text{and} \quad w \preceq v \, .$$

Using the above partial order we obtain the following inclusion.

Proposition 3.1.6. *Let $\dim H_1 = \dim H_2 = \infty$ and let $v, w \in \mathcal{W}$. Then*

$$v \preceq w \iff E_v(H_1, H_2) \subseteq E_w(H_1, H_2) \, .$$

Proof. For the forward implication we need to show that if $v \preceq w$ then $E_v(H_1, H_2) \subseteq E_w(H_1, H_2)$. This, however, follows directly from Proposition 3.1.5.

For the converse, suppose that $E_v(H_1, H_2) \subseteq E_w(H_1, H_2)$. We need to show that $v \preceq w$. Fix orthonormal bases $(e_k)_{k\in\mathbb{N}}$ for H_1 and $(f_k)_{k\in\mathbb{N}}$ for H_2 .

Define an operator $A \in L(H_1, H_2)$ by setting $Ae_k = v_k f_k$ for every $k \in \mathbb{N}$. We clearly have $s_k(A) = v_k$ for every $k \in \mathbb{N}$, so $A \in E_v(H_1, H_2)$. But since $E_v(H_1, H_2) \subseteq E_w(H_1, H_2)$, we have $A \in E_w(H_1, H_2)$. Thus there exists $M \geq 0$ such that $v_k = s_k(A) \leq M w_k$ for every $k \in \mathbb{N}$, so $v \preceq w$ and the backwards implication is proved as well. □

Corollary 3.1.7. *Let* $\dim H_1 = \dim H_2 = \infty$ *and let* $v, w \in \mathcal{W}$. *Then*

$$v \asymp w \iff E_v(H_1, H_2) = E_w(H_1, H_2)\,.$$

Although $E_w(H_1, H_2)$ is not a linear space in general, it is closed under multiplication by scalars and operators, as we shall see presently.

Lemma 3.1.8. *If* $A \in E_w(H_1, H_2)$ *and* $\alpha \in \mathbb{C}$, *then*

$$\alpha A \in E_w(H_1, H_2) \quad \textit{and} \quad |\alpha A|_w = |\alpha|\, |A|_w\,.$$

Proof. Let $A \in E_w(H_1, H_2)$ and $\alpha \in \mathbb{C}$. By (2.2.2) we have

$$s_k(\alpha A) = |\alpha| s_k(A) \leq |\alpha| |A|_w w_k \quad (\forall k \in \mathbb{N})\,.$$

Thus, for every $\alpha \in \mathbb{C}$,

$$\alpha A \in E_w(H_1, H_2)\,,$$

and

$$|\alpha A|_w \leq |\alpha| |A|_w\,. \tag{3.1.3}$$

It remains to prove that $|\alpha A|_w \geq |\alpha| |A|_w$ for every $\alpha \in \mathbb{C}$. If $\alpha = 0$, then there is nothing to prove. If $\alpha \neq 0$, then using (3.1.3) we have

$$|A|_w = |\alpha^{-1} \alpha A|_w \leq |\alpha^{-1}| |\alpha A|_w\,.$$

Therefore we obtain, for every $\alpha \in \mathbb{C}$,

$$|\alpha||A|_w \leq |\alpha A|_w .$$

□

Proposition 3.1.9. *If* $B \in L(H_2, H_1)$, $A \in E_w(H_3, H_2)$ *and* $C \in L(H_4, H_3)$, *then* $|BAC|_w \leq \|B\| \, |A|_w \, \|C\|$. *And*

$$L(H_2, H_1)E_w(H_3, H_2)L(H_4, H_3) \subseteq E_w(H_4, H_1) .$$

Proof. Let $A \in E_w(H_3, H_2)$. By Proposition 2.2.9, we obtain

$$s_k(BAC) \leq \|B\| \, s_k(A) \, \|C\| \leq \|B\| \, |A|_w \, \|C\| \, w_k \quad (\forall k \in \mathbb{N}) .$$

Thus we have $BAC \in E_w(H_4, H_1)$ and $|BAC|_w \leq \|B\| \, |A|_w \, \|C\|$. □

Remark 3.1.10. Note that Proposition 3.1.9 implies that

$$L(H)E_w(H)L(H) \subseteq E_w(H) .$$

Hence $E_w(H)$ satisfies the second condition of the definition of an operator ideal (see Definition 2.4.2), though not necessarily the first one, concerned with linearity. Thus $E_w(H)$ is what is sometimes referred to as a pre-ideal (see, for example, [Nel82]).

We shall now investigate the behaviour of compactness classes under addition (see Proposition 3.1.12). Before doing so we require the following definition.

Definition 3.1.11. Let $w \in \mathcal{W}$. Then $\dot{w}$ is the sequence obtained from w by doubling each entry, that is, $\dot{w} = (w_1, w_1, w_2, w_2, w_3, w_3, \ldots)$. More precisely,

$\dot{w}$ is the sequence given by

$$\dot{w}_k = \begin{cases} w_{\frac{k}{2}} & \text{if } k \quad \text{is even} \\ w_{\frac{k+1}{2}} & \text{if } k \quad \text{is odd.} \end{cases}$$

We are now ready to investigate how compactness classes behave under addition.

Proposition 3.1.12. *Let $w \in \mathcal{W}$. Then the following assertions hold.*

(i) If $A, B \in E_w(H_1, H_2)$, then $A + B \in E_{\dot{w}}(H_1, H_2)$ with $|A + B|_{\dot{w}} \leq |A|_w + |B|_w$.

(ii) If $\dim H_1 = \dim H_2 = \infty$, then assertion (i) is sharp in the sense that if there is $v \in \mathcal{W}$ such that $A+B \in E_v(H_1, H_2)$ for all $A, B \in E_w(H_1, H_2)$, then $\dot{w} \preceq v$.

Proof.

(i) Suppose $A, B \in E_w(H_1, H_2)$. Using Corollary 2.2.13 we have

$$s_{2k-1}(A + B) \leq s_k(A) + s_k(B) \leq (|A|_w + |B|_w)w_k = (|A|_w + |B|_w)\dot{w}_{2k-1}$$

since $\dot{w}_{2k-1} = w_k$ for every $k \in \mathbb{N}$. As the singular values are monotonically decreasing and $\dot{w}_{2k} = w_k$ for every $k \in \mathbb{N}$, we obtain

$$s_{2k}(A + B) \leq s_{2k-1}(A + B) \leq (|A|_w + |B|_w)w_k = (|A|_w + |B|_w)\dot{w}_{2k} \,.$$

Hence we have

$$s_k(A + B) \leq (|A|_w + |B|_w)\dot{w}_k \quad (\forall k \in \mathbb{N}) \,.$$

Therefore

$$A + B \in E_{\dot{w}} \quad \text{and} \quad |A + B|_{\dot{w}} \leq |A|_w + |B|_w \,.$$

(ii) Since both H_1 and H_2 are infinite-dimensional we can choose orthonormal bases $(e_k)_{k\in\mathbb{N}}$ for H_1 and $(f_k)_{k\in\mathbb{N}}$ for H_2. Define an operator $A \in L(H_1, H_2)$ by setting

$$Ae_k = \begin{cases} 0 & \text{if } k \text{ is even,} \\ w_{\frac{k+1}{2}} f_k & \text{if } k \text{ is odd,} \end{cases}$$

for every $k \in \mathbb{N}$, so that the matrix representation of A with respect to the chosen bases is

$$\begin{pmatrix} w_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & w_2 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & 0 & \dots & \end{pmatrix}.$$

Furthermore, define an operator $B \in L(H_1, H_2)$ by setting

$$Be_k = \begin{cases} w_{\frac{k}{2}} f_k & \text{if } k \text{ is even,} \\ 0 & \text{if } k \text{ is odd,} \end{cases}$$

for every $k \in \mathbb{N}$, so that the matrix representation of B with respect to the chosen bases is

$$\begin{pmatrix} 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & w_1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & 0 & \dots & \end{pmatrix}.$$

Clearly, we have

$$s_k(A) = s_k(B) = w_k \quad (\forall k \in \mathbb{N}),$$

so $A, B \in S_\infty(H_1, H_2)$. Moreover, the matrix representation of $A + B$ with respect to the chosen bases is

$$\begin{pmatrix} w_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & w_1 & 0 & 0 & \dots & 0 \\ 0 & 0 & w_2 & 0 & \dots & 0 \\ 0 & 0 & 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & 0 & \dots & \end{pmatrix},$$

and

$$s_k(A + B) = \dot{w}_k \quad (k \in \mathbb{N}).$$

Therefore if $A + B \in E_v(H_1, H_2)$, then, using the observation above, there exists $M \geq 0$ such that, for every $k \in \mathbb{N}$,

$$\dot{w}_k = s_k(A + B) \leq M v_k,$$

which means $\dot{w} \preceq v$.

□

The proposition above implies that $E_w(H_1, H_2)$ is not a linear space in general. However, it points towards a simple sufficient condition guaranteeing linearity.

Corollary 3.1.13. *If $\dot{w} \asymp w$, then $E_w(H_1, H_2)$ is a linear space and $|\cdot|_w$ is a quasi-norm.*

Proof. By Lemma 3.1.8, we have $\alpha A \in E_w(H_1, H_2)$ for every $\alpha \in \mathbb{C}$ and $A \in E_w(H_1, H_2)$. Moreover, using Proposition 3.1.12 and the assumption $\dot{w} \asymp w$ we have

$$A + B \in E_{\dot{w}}(H_1, H_2) = E_w(H_1, H_2)\,.$$

Thus $E_w(H_1, H_2)$ is a linear space. It remains to show that $|\cdot|_w$ is a quasi-norm.

(i) $|A|_w \geq 0$ is clear from the definition.

(ii) We need to show that $|A|_w = 0$ if and only if $A = 0$. In order to see this note that if $|A|_w = 0$, then $s_1(A) = \|A\| = 0$, so $A = 0$. The converse is trivial.

(iii) We need to show that $|\alpha A|_w = |\alpha||A|_w$ for any $A \in E_w(H_1, H_2)$ and $\alpha \in \mathbb{C}$. This, however, follows from Lemma 3.1.8.

(iv) We need to show that there is $M \geq 1$ such that

$$|A + B|_w \leq M(|A|_w + |B|_w) \quad (\forall A, B \in E_w(H_1, H_2))\,.$$

In order to see this note that, since $\dot{w} \asymp w$ there exists $M \geq 1$ such that

$$\frac{1}{M}|A|_w \leq |A|_{\dot{w}} \leq M|A|_w$$

for every $A \in E_w(H_1, H_2) = E_{\dot{w}}(H_1, H_2)$. Since $A, B \in E_w(H_1, H_2)$ then, by Proposition 3.1.12, we have $|A + B|_{\dot{w}} \leq |A|_w + |B|_w$. It follows that if $A, B \in E_w(H_1, H_2)$, then $A + B \in E_{\dot{w}}(H_1, H_2) = E_w(H_1, H_2)$ and

$$\frac{1}{M}|A + B|_w \leq |A + B|_{\dot{w}} \leq |A|_w + |B|_w .$$

□

Proposition 3.1.14. *If $\dot{w} \asymp w$, then $E_w(H_1, H_2)$ is complete with respect to the quasi-norm $|\cdot|_w$.*

Proof. Let $(A_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $E_w(H_1, H_2)$. First we note that $(A_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $S_\infty(H_1, H_2)$ with respect to the operator norm $\|\cdot\|$, since $\|A_n - A_m\| \leq |A_n - A_m|_w w_1$. As $S_\infty(H_1, H_2)$ is complete there is an $A \in S_\infty(H_1, H_2)$ such that $A_n \to A$ as $n \to \infty$ in the operator norm $\|\cdot\|$. We need to prove that $A \in E_w(H_1, H_2)$ and $|A_n - A|_w \to 0$ as $n \to \infty$. Fix $\epsilon \geq 0$. Since $(A_n)_{n\in\mathbb{N}}$ is Cauchy in $|\cdot|_w$, there exists $N_\epsilon \in \mathbb{N}$ such that

$$s_k(A_n - A_m) \leq |A_n - A_m|_w w_k \leq \epsilon w_k \quad (\forall n, m \geq N_\epsilon, \forall k \in \mathbb{N}) .$$

Letting $m \to \infty$ in the above and using Corollary 2.2.15 we obtain

$$s_k(A_n - A) \leq \epsilon w_k \quad (\forall n \geq N_\epsilon, \forall k \in \mathbb{N}) ,$$

and so

$$|A_n - A|_w \leq \epsilon \quad (\forall n \geq N_\epsilon) . \tag{3.1.4}$$

The above implies that $|A_n - A|_w \to 0$ as $n \to \infty$. It remains to show that $A \in E_w(H_1, H_2)$. In order to see this, fix $n \geq N_\epsilon$. Inequality (3.1.4) now implies that $A_n - A$ is an element of $E_w(H_1, H_2)$. Since A_n is also an element of $E_w(H_1, H_2)$ and $E_w(H_1, H_2)$ is linear by Corollary 3.1.13, we then obtain $A \in E_w(H_1, H_2)$. □

Proposition 3.1.15. *If $\dot{w} \asymp w$, then E_w is a quasi-Banach operator ideal.*

Proof. Follows from Proposition 3.1.9, Corollary 3.1.13 and Proposition 3.1.14. □

We now turn to studying the rate of decay of the eigenvalue sequence of an operator in a given compactness class. In order to do this we require the following notation.

Definition 3.1.16. Let $w \in \mathcal{W}$. Then we define $\bar{w}$ as the sequence of successive geometric means of w, that is,

$$\bar{w}_k = (w_1 \cdots w_k)^{\frac{1}{k}} \quad (\forall k \in \mathbb{N}) .$$

Proposition 3.1.17. *Let $A \in E_w(H_1, H_2)$. Then*

$$\lambda(A) \in \mathcal{E}_{\bar{w}} \quad \textit{with} \quad |\lambda(A)|_{\bar{w}} \leq |A|_w .$$

Proof. Let $A \in E_w(H_1, H_2)$. By the multiplicative Weyl inequality (see Theorem 2.3.1) we have, for every $k \in \mathbb{N}$,

$$|\lambda_k(A)|^k \leq \prod_{l=1}^{k} |\lambda_l(A)| \leq \prod_{l=1}^{k} s_l(A) \leq |A|_w w_1 \cdots |A|_w w_k \leq |A|_w^k \, w_1 \cdots w_k .$$

Thus

$$|\lambda_k(A)| \leq |A|_w (w_1 \cdots w_k)^{\frac{1}{k}} = |A|_w \bar{w}_k \quad (\forall k \in \mathbb{N}) ,$$

and we obtain

$$\lambda(A) \in \mathcal{E}_{\bar{w}} \quad \text{and} \quad |\lambda(A)|_{\bar{w}} \leq |A|_w ,$$

as desired.

□

3.2 General resolvent bounds

The first bound for the norm of the resolvent of a linear operator on an infinite-dimensional Hilbert space was derived by Carleman (see [Car21]), who obtained a bound for Hilbert-Schmidt operators. His result was later generalised to Schatten-von Neumann operators (see, for example, [DS63, Sim77]). For more information about generalised Carleman type estimates see also [DP94, DP96].

In this section we shall derive an upper bound for the norm of the resolvent $R(A;z)$ of $A \in E_w(H_1, H_2)$ in terms of the distance of z to the spectrum of A and the w-departure from normality of A, a number measuring the non-normality of A. As already mentioned, we shall generalise the approach of Henrici in [Hen62] outlined in the introduction to the infinite-dimensional setting. The basic idea will be to write A as a sum of a normal operator D with $\sigma(D) = \sigma(A)$ and a quasi-nilpotent operator N, that is, an operator the spectrum of which consists of the point 0 only, and to consider A as a perturbation of D by N.

We start with a bound for powers of quasi-nilpotent operators, due to Dostanić.

Theorem 3.2.1. *There is a constant $C \geq \pi/2$ such that for any quasi-*

nilpotent $A \in S_\infty(H)$ and for every $k \in \mathbb{N}$ we have

$$\left\|A^{2k}\right\| \leq C^{2k}(s_1(A)\cdots s_k(A))^2 .$$

Proof. See [Dos01, Theorem 1]. □

Given $w \in \mathcal{W}$, we define a function $F_w : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ by setting

$$F_w(r) = (1 + rw_1)\left(1 + \sum_{k=1}^{\infty}(w_1\cdots w_k)^2(Cr)^{2k}\right) . \tag{3.2.1}$$

It is not difficult to see that F_w is well-defined (for example by using the ratio test), real-analytic and strictly monotonically increasing. We are now ready to deduce resolvent bounds for quasi-nilpotent operators.

Proposition 3.2.2. *Let $w \in \mathcal{W}$ and let $A \in E_w(H)$ be a quasi-nilpotent operator. Then*

$$\left\|(I - A)^{-1}\right\| \leq F_w(|A|_w) .$$

Proof. Suppose $A \in E_w(H)$ is a quasi-nilpotent operator. Using a Neumann series and Theorem 3.2.1, we have

$$\left\|(I - A)^{-1}\right\| \leq \sum_{k=0}^{\infty}\left\|A^k\right\| = \sum_{k=0}^{\infty}(\left\|A^{2k}\right\| + \left\|A^{2k+1}\right\|) ,$$

$$\leq (1 + \|A\|)\left(1 + \sum_{k=1}^{\infty}\left\|A^{2k}\right\|\right) ,$$

$$\leq (1 + s_1(A))\left(1 + \sum_{k=1}^{\infty}C^{2k}(s_1(A)\cdots s_k(A))^2\right) .$$

Therefore we obtain

$$\left\|(I - A)^{-1}\right\| \leq (1 + |A|_w w_1)\left(1 + \sum_{k=1}^{\infty}(w_1\cdots w_k)^2(C|A|_w)^{2k}\right) ,$$

as required. □

An immediate consequence of the previous proposition is the following estimate for the growth of the resolvent of a quasi-nilpotent operator $A \in E_w(H)$.

Corollary 3.2.3. *Let $w \in \mathcal{W}$ and let $A \in E_w(H)$ be quasi-nilpotent. Then for any $z \neq 0$*

$$\|R(A;z)\| \leq |z|^{-1} F_w(|z|^{-1} |A|_w) .$$

The following is the infinite-dimensional analogue of the Schur decomposition discussed in the introduction.

Theorem 3.2.4. *Let $A \in S_\infty(H)$. Then A can be written as a sum*

$$A = D + N ,$$

such that

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

(ii) D is normal and $\lambda(D) = \lambda(A)$;

(iii) N and $(zI - D)^{-1}N$ are quasi-nilpotent for every $z \in \rho(D) = \rho(A)$.

Proof. See [Ban04, Theorem 3.2]. □

The theorem above motivates the following definition.

Definition 3.2.5. Let $A \in S_\infty(H)$. A decomposition

$$A = D + N$$

with D and N satisfying the properties (i–iii) of the previous theorem is called a *Schur decomposition of A*. We call the operators D and N the *normal* and the *quasi-nilpotent part of the Schur decomposition of A*, respectively.

Remark 3.2.6. The decomposition is not unique, as can be seen from the following example taken from [Ban04, Remark 3.5 (i)]. Consider

$$A := \begin{pmatrix} 2 & 2 & 2 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} = \underbrace{\begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{=:D_1} + \underbrace{\begin{pmatrix} 0 & 2 & 2 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}}_{=:N_1}$$
$$= \underbrace{\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{=:D_2} + \underbrace{\begin{pmatrix} 1 & 1 & 2 \\ -1 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix}}_{=:N_2}.$$

It is easy to see that D_1 and D_2 are normal and that N_1 and N_2 are nilpotent. Moreover $\sigma(A) = \sigma(D_1) = \sigma(D_2) = \{2, 0\}$. Furthermore, both $(zI - D_1)^{-1}N_1$ and $(zI - D_2)^{-1}N_2$ are nilpotent for any $z \in \rho(A)$. Thus A has two different Schur decompositions.

Note that the normal parts are obviously unitarily equivalent. However, the nilpotent parts are not. In order to see this observe that

$$\|N_1\|_4^4 = 112 \neq 80 = \|N_2\|_4^4 \,,$$

where $\|\cdot\|_4$ is the norm of the Schatten class $S_4(\mathbb{C}^3)$.

In the following proposition we determine an upper bound for the singular values of the normal part and the quasi-nilpotent part of a Schur decomposition of an operator in a given compactness class.

Proposition 3.2.7. *Let $A \in E_w(H)$. If $A = D + N$ is a Schur decomposition of A with normal part D and quasi-nilpotent part N, then*

(i) $D \in E_{\bar{w}}(H)$ with $|D|_{\bar{w}} \le |A|_w$, where $\bar{w}_k = (w_1 \cdots w_k)^{\frac{1}{k}}$.

(ii) $N \in E_{\dot{w}}(H)$ *with* $|N|_{\dot{w}} \leq 2|A|_w$, *where* $\dot{w} = (w_1, w_1, (w_1w_2)^{\frac{1}{2}}, (w_1w_2)^{\frac{1}{2}}, \ldots)$.

Proof. Let $A \in E_w(H)$. Since D is normal, it follows from Remark 2.2.12 that its singular values coincide with the moduli of its eigenvalues, which also coincide with the moduli of the eigenvalues of A. Using Proposition 3.1.17 and the fact that D is normal we obtain

$$s_k(D) \leq |A|_w \bar{w}_k \, ,$$

so $D \in E_{\bar{w}}(H)$ and $|D|_{\bar{w}} \leq |A|_w$, as required.

For the second part, observe that since $w \preceq \bar{w}$, we have $|A|_{\bar{w}} \leq |A|_w$ via Proposition 3.1.5. Then we also have $A \in E_{\bar{w}}(H)$ by Proposition 3.1.6. Thus, using Proposition 3.1.12, we have

$$N = A - D \in E_{\dot{w}} \, ,$$

$$|A - D|_{\dot{w}} \leq |A|_{\bar{w}} + |D|_{\bar{w}}$$

and so, using assertion (i), we obtain

$$|N|_{\dot{w}} \leq |A|_w + |A|_w = 2|A|_w \, ,$$

as desired. □

We now define the analogue of Henrici's departure from normality for operators in a given compactness class.

Definition 3.2.8. Let $w \in \mathcal{W}$ and $A \in E_w(H)$. Then

$$\nu_w(A) = \inf\{\, |N|_{\dot{w}} \, : \, N \text{ is the quasi-nilpotent part of a Schur decomposition of } A \,\}$$

is called the *w-departure from normality of A.*

Remark 3.2.9. Note that by the previous proposition, the w-departure from normality of an operator in $E_w(H)$ is always finite.

The term 'departure from normality' is justified in view of the following proposition.

Proposition 3.2.10. *Let $A \in E_w(H)$. Then*

$$A \text{ is normal} \iff \nu_w(A) = 0\,.$$

Proof. The forward implication is trivial. For the backwards implication, let $\nu_w(A) = 0$. Then there exists a sequence of Schur decompositions with quasi-nilpotent parts N_n such that $|N_n|_{\dot{w}} \to 0$ as $n \to \infty$. But

$$\|A - D_n\| = \|N_n\| = s_1(N_n) \leq \dot{w}_1 |N_n|_{\dot{w}}\,,$$

where D_n are the corresponding normal parts, so $\lim_{n\to\infty} \|A - D_n\| = 0$. Hence A is a limit of normal operators which converge in operator norm. Thus A is normal. □

Since the departure from normality is difficult to calculate for a given $A \in E_w(H)$, we now give a simple upper bound.

Proposition 3.2.11. *Let $A \in E_w(H)$. Then*

$$\nu_w(A) \leq 2\,|A|_w\,.$$

Proof. Follows from Proposition 3.2.7 (ii). □

We are now able to obtain growth estimates for the resolvents of operators in a given compactness class. Before doing so we recall the bound for the resolvent of a normal operator already mentioned in the introduction (see Remark 1.1.7). If D is a normal operator on a separable Hilbert space, then

$$\|R(D;z)\| = \frac{1}{d(z,\sigma(D))} \quad (\forall z \in \rho(D))\,. \tag{3.2.2}$$

The following is the main result of this section.

Theorem 3.2.12. *Let $A \in E_w(H)$. Then*

$$\|R(A;z)\| \leq \frac{1}{d(z,\sigma(A))} F_{\dot{w}}\left(\frac{\nu_w(A)}{d(z,\sigma(A))}\right) \quad (\forall z \in \rho(A))\,. \tag{3.2.3}$$

Proof. Fix $z \in \rho(A)$. By Proposition 3.2.7, the operator A has a Schur decomposition with normal part D and quasi-nilpotent part N. Thus, we know that $\sigma(A) = \sigma(D)$, that $(zI-D)^{-1}$ exists and that $(zI-D)^{-1}N$ is quasi-nilpotent. Furthermore

$$\begin{aligned} s_k((zI-D)^{-1}N) &\leq \left\|(zI-D)^{-1}\right\| s_k(N) \\ &= \frac{s_k(N)}{d(z,\sigma(D))} \leq \frac{|N|_{\dot{w}}\dot{w}_k}{d(z,\sigma(D))} = \frac{|N|_{\dot{w}}\dot{w}_k}{d(z,\sigma(A))}\,, \end{aligned}$$

using (3.2.2) as well as Propositions 2.2.9 and 3.2.7. Now $(I-(zI-D)^{-1}N)$ is invertible in $L(H)$ and, using Proposition 3.2.2, it follows that

$$\left\|(I-(zI-D)^{-1}N)^{-1}\right\| \leq F_{\dot{w}}\left(\frac{|N|_{\dot{w}}}{d(z,\sigma(A))}\right)\,.$$

Since

$$(zI-A) = (zI-D)(I-(zI-D)^{-1}N)\,,$$

we can conclude that $(zI - A)$ is invertible in $L(H)$ and

$$\begin{aligned}\|R(A;z)\| &\leq \|R(D;z)\| \left\|(I-(zI-D)^{-1}N)^{-1}\right\| \\ &\leq \frac{1}{d(z,\sigma(A))} F_{\dot{\bar{w}}}\left(\frac{|N|_{\dot{\bar{w}}}}{d(z,\sigma(A))}\right).\end{aligned}$$

Taking the infimum over all Schur decompositions the theorem follows. □

Remark 3.2.13.

(i) Another look at the above proof shows that the bound (3.2.3) also holds if we replace $\nu_w(A)$ by a larger quantity, say by the upper bound given in Proposition 3.2.11.

(ii) The bound (3.2.3) is optimal for normal A, as it reduces to the sharp bound (3.2.2).

3.3 Resolvent bounds for particular classes

As we saw in the last section, the growth of the resolvent of an operator belonging to a given compactness class E_w in the vicinity of a spectral point is, by Theorem 3.2.12, controlled by the behaviour of the function $F_{\dot{\bar{w}}}$ at infinity. In this section we shall study the asymptotics of this function for particular compactness classes, namely the Schatten-Lorentz ideals, given by $w_k = k^{-1/p}$ with $p \in (0, \infty)$ and the exponential classes, given by $w_k = \exp(-ak^\alpha)$ with $a \in (0, \infty)$ and $\alpha \in (0, \infty)$.

Before starting with the Schatten-Lorentz ideals we briefly recall Stirling's approximation for the factorial in the form

$$\sqrt{2\pi k}\left(\frac{k}{\mathrm{e}}\right)^k \leq k! \leq \sqrt{\mathrm{e}^2 k}\left(\frac{k}{\mathrm{e}}\right)^k \quad (\forall k \in \mathbb{N}).$$

Lemma 3.3.1. *Let $p \in (0,\infty)$, and let $w_k = k^{-1/p}$ for $k \in \mathbb{N}$. Then the following inequalities hold:*

$$\exp\left(-\frac{1}{p\sqrt{k}}\right)\frac{\mathrm{e}^{1/p}}{k^{1/p}} \le \bar{w}_k \le \frac{\mathrm{e}^{1/p}}{k^{1/p}} \quad (\forall k \in \mathbb{N}), \tag{3.3.1}$$

$$\exp\left(-\frac{3}{p\sqrt{k}}\right)\frac{(2\mathrm{e})^{1/p}}{k^{1/p}} \le \dot{\bar{w}}_k \le \frac{(2\mathrm{e})^{1/p}}{k^{1/p}} \quad (\forall k \in \mathbb{N}), \tag{3.3.2}$$

$$\exp\left(-\frac{6}{p}\sqrt{k}\right)\frac{(2\mathrm{e})^{1/p}}{(k!)^{1/p}} \le \prod_{n=1}^{k}\dot{\bar{w}}_n \le \frac{(2\mathrm{e})^{1/p}}{(k!)^{1/p}} \quad (\forall k \in \mathbb{N}). \tag{3.3.3}$$

Proof. We start with the case $p = 1$, that is, we set $w_k = k^{-1}$ and show that

$$\exp\left(-\frac{1}{\sqrt{k}}\right)\frac{\mathrm{e}}{k} \le \bar{w}_k \le \frac{\mathrm{e}}{k} \quad (\forall k \in \mathbb{N}), \tag{3.3.4}$$

$$\exp\left(-\frac{3}{\sqrt{k}}\right)\frac{2\mathrm{e}}{k} \le \dot{\bar{w}}_k \le \frac{2\mathrm{e}}{k} \quad (\forall k \in \mathbb{N}), \tag{3.3.5}$$

$$\exp\left(-6\sqrt{k}\right)\frac{2\mathrm{e}}{k!} \le \prod_{n=1}^{k}\dot{\bar{w}}_n \le \frac{2\mathrm{e}}{k!} \quad (\forall k \in \mathbb{N}). \tag{3.3.6}$$

Now, the upper bound in (3.3.4) follows from Stirling's approximation by observing that for all $k \in \mathbb{N}$ we have

$$\bar{w}_k^k = \frac{1}{k!} \le \left(\frac{\mathrm{e}}{k}\right)^k .$$

For the lower bound in (3.3.4) we again use Stirling's approximation to obtain

$$\bar{w}_k^k = \frac{1}{k!} \ge \frac{1}{\sqrt{\mathrm{e}^2 k}}\left(\frac{\mathrm{e}}{k}\right)^k ,$$

and we see that we are done if we can show that

$$\frac{1}{\sqrt{\mathrm{e}^2 k}} \ge \exp(-\sqrt{k}) \quad (\forall k \in \mathbb{N}). \tag{3.3.7}$$

The above, however, is true since, using the inequality $1 + x \le \exp(x)$ which holds for all real x, we see that for all $k \in \mathbb{N}$ we have

$$\sqrt{k} \le \exp(\sqrt{k} - 1)$$

from which

$$\sqrt{\mathrm{e}^2 k} \le \exp(\sqrt{k})\,,$$

which implies (3.3.7).

We now turn to (3.3.5). For the upper bound we note that, for $k \in \mathbb{N}$ even, (3.3.4) implies

$$\dot{\bar{w}}_k = \bar{w}_{\frac{k}{2}} \le \frac{2\mathrm{e}}{k}\,,$$

while for $k \in \mathbb{N}$ odd, (3.3.4) implies

$$\dot{\bar{w}}_k = \bar{w}_{\frac{k+1}{2}} \le \frac{2\mathrm{e}}{k+1} \le \frac{2\mathrm{e}}{k}\,.$$

For the lower bound we note that, for $k \in \mathbb{N}$ even, (3.3.4) implies

$$\dot{\bar{w}}_k = \bar{w}_{\frac{k}{2}} \ge \exp\left(-\frac{\sqrt{2}}{\sqrt{k}}\right)\frac{2\mathrm{e}}{k} \ge \exp\left(-\frac{3}{\sqrt{k}}\right)\frac{2\mathrm{e}}{k}\,,$$

while for $k \in \mathbb{N}$ odd, (3.3.4) implies

$$\dot{\bar{w}}_k = \bar{w}_{\frac{k+1}{2}} \ge \exp\left(-\frac{\sqrt{2}}{\sqrt{k+1}}\right)\frac{2\mathrm{e}}{k+1}\,,$$

and we are done if we can show that for all $k \in \mathbb{N}$ we have

$$\exp\left(-\frac{\sqrt{2}}{\sqrt{k+1}}\right)\frac{1}{k+1} \ge \exp\left(-\frac{3}{\sqrt{k}}\right)\frac{1}{k}\,,$$

which, in turn, is equivalent to

$$\left(1+\frac{1}{k}\right)\exp\left(-\frac{3}{\sqrt{k}}+\frac{\sqrt{2}}{\sqrt{k+1}}\right)\leq 1 \quad (\forall k\in\mathbb{N})\,. \tag{3.3.8}$$

The above, however, follows by observing that we have for all $k \in \mathbb{N}$

$$\begin{aligned}\left(1+\frac{1}{k}\right)\exp\left(-\frac{3}{\sqrt{k}}+\frac{\sqrt{2}}{\sqrt{k+1}}\right) &\leq \left(1+\frac{1}{k}\right)\exp\left(-\frac{1}{\sqrt{k}}\right)\\ &\leq \exp\left(\frac{1}{k}-\frac{1}{\sqrt{k}}\right)\leq \exp\left(-\frac{\sqrt{k}-1}{k}\right)\leq 1\,.\end{aligned}$$

This finishes the proof of (3.3.5).

Finally, the upper bound in (3.3.6) is obvious, while the lower one follows from

$$\prod_{n=1}^{k}\dot{\varpi}_n \geq \exp\left(-3\sum_{n=1}^{k}\frac{1}{\sqrt{n}}\right)\frac{(2\mathrm{e})^k}{k!}\geq \exp\left(-6\sqrt{k}\right)\frac{(2\mathrm{e})^k}{k!}\,,$$

where we have used that $\sum_{n=1}^{k} n^{-1/2} \leq \int_0^k t^{-1/2} = 2k^{1/2}$ for every $k \in \mathbb{N}$.

This finishes the proof of the lemma for $p = 1$. The general case follows by taking p-th roots in (3.3.4), (3.3.5) and (3.3.6). □

In order to be able to study the behaviour of $F_{\dot{\varpi}}$ we require another auxiliary result. Before stating it we introduce some more notation. If f and g are two real-valued functions defined on a neighbourhood of ∞, we write

$$f(r) \sim g(r) \text{ as } r \to \infty$$

if

$$\lim_{r\to\infty}\frac{f(r)}{g(r)} = 1\,.$$

For later use, we note the following relation between the asymptotics of a function and that of its inverse.

Lemma 3.3.2. *Let $a, b \in (0, \infty)$ and let I and J be neighbourhoods of ∞. Suppose that $f : I \to J$ is a bijection with inverse $f^{-1} : J \to I$. Then the following assertions hold.*

(i) If

$$f(r) \sim ar^b \text{ as } r \to \infty$$

then

$$f^{-1}(r) \sim \left(\frac{r}{a}\right)^{1/b} \text{ as } r \to \infty\,.$$

(ii) If

$$\log f(r) \sim ar^b \text{ as } r \to \infty$$

then

$$f^{-1}(r) \sim \left(\frac{\log r}{a}\right)^{1/b} \text{ as } r \to \infty\,.$$

(iii) If

$$\log f(r) \sim a(\log r)^b \text{ as } r \to \infty$$

then

$$\log f^{-1}(r) \sim \left(\frac{\log r}{a}\right)^{1/b} \text{ as } r \to \infty\,.$$

Proof.

(i) This follows from

$$\lim_{r\to\infty} \frac{(r/a)^{1/b}}{f^{-1}(r)} = \lim_{r\to\infty} \frac{(f(r)/a)^{1/b}}{f^{-1}(f(r))} = \left(\lim_{r\to\infty} \frac{f(r)}{ar^b}\right)^{1/b} = 1\,.$$

(ii) If

$$(\log \circ f)(r) \sim ar^b \text{ as } r \to \infty\,,$$

then by (i) we have

$$(\log \circ f)^{-1}(r) \sim \left(\frac{r}{a}\right)^{1/b} \quad \text{as } r \to \infty\,,$$

so

$$(f^{-1} \circ \exp)(r) \sim \left(\frac{r}{a}\right)^{1/b} \quad \text{as } r \to \infty\,,$$

hence

$$f^{-1}(r) \sim \left(\frac{\log r}{a}\right)^{1/b} \quad \text{as } r \to \infty\,.$$

(iii) If

$$(\log \circ f)(r) \sim a(\log r)^b \text{ as } r \to \infty\,,$$

then

$$(\log \circ f \circ \exp)(r) \sim ar^b \text{ as } r \to \infty\,,$$

so by (i) we have

$$(\log \circ f \circ \exp)^{-1}(r) \sim \left(\frac{r}{a}\right)^{1/b} \quad \text{as } r \to \infty\,,$$

hence

$$(\log \circ f^{-1} \circ \exp)(r) \sim \left(\frac{r}{a}\right)^{1/b} \quad \text{as } r \to \infty\,,$$

whence

$$(\log \circ f^{-1})(r) \sim \left(\frac{\log r}{a}\right)^{1/b} \quad \text{as } r \to \infty\,.$$

□

We are now able to state the following result.

Lemma 3.3.3. *Suppose that $p, b \in (0, \infty)$. Let $\Phi_p^{L,u}$ and $\Phi_{p,b}^{L,l}$ be two functions* *given by the power series*

$$\Phi_p^{L,u}(r) = \sum_{k=0}^{\infty} \frac{1}{(k!)^{1/p}} r^k$$

$$\Phi_{p,b}^{L,l}(r) = \sum_{k=0}^{\infty} \frac{\exp(-b\sqrt{k})}{(k!)^{1/p}} r^k \, .$$

Then $\Phi_p^{L,u}$ and $\Phi_{p,b}^{L,l}$ extend to entire functions with the following asymptotics

$$\log \Phi_p^{L,u}(r) \sim \log \Phi_{p,b}^{L,l}(r) \sim \frac{1}{p} r^p \text{ as } r \to \infty \, .$$

Proof. Using Stirling's approximation we see that both $\Phi_p^{L,u}$ and $\Phi_{p,b}^{L,l}$ extend to entire functions. Since $\Phi_{p,b}^{L,l}(r) \leq \Phi_p^{L,u}(r)$ for all $r \in (0,\infty)$ the remaining assertions will hold if we can show that

$$\limsup_{r\to\infty} pr^{-p} \log \Phi_p^{L,u}(r) \leq 1 \tag{3.3.9}$$

and

$$\liminf_{r\to\infty} pr^{-p} \log \Phi_{p,b}^{L,l}(r) \geq 1 \, . \tag{3.3.10}$$

We start with (3.3.9). For $p \leq 1$ we have, using the ℓ_p-ℓ_1 inequality

$$\sum_{k=0}^{\infty} x_k \leq \left(\sum_{k=0}^{\infty} x_k^p \right)^{1/p} ,$$

which holds for all positive sequences $(x_k)_{k=0}^{\infty}$, the bound

$$\Phi_p^{L,u}(r) \leq \left(\sum_{k=0}^{\infty} \frac{r^{pk}}{k!} \right)^{1/p} = \exp\left(\frac{1}{p} r^p \right) ,$$

and (3.3.9) holds in this case. For $p > 1$ we split the sum as follows

$$\Phi_p^{L,u}(r) = \sum_{k<2\mathrm{e}r^p} \frac{r^k}{(k!)^{1/p}} + \sum_{k\geq 2\mathrm{e}r^p} \frac{r^k}{(k!)^{1/p}} .$$

In order to bound the first term we use Hölder's inequality to obtain

$$\begin{aligned}\sum_{k<2\mathrm{e}r^p} \frac{r^k}{(k!)^{1/p}} &\leq \left(\sum_{k<2\mathrm{e}r^p} \frac{r^{pk}}{k!}\right)^{1/p} \left(\sum_{k<2\mathrm{e}r^p} 1\right)^{(p-1)/p} \\ &\leq (1+2\mathrm{e}r^p)^{(p-1)/p} \exp\left(\frac{1}{p} r^p\right) .\end{aligned}$$

For the second term, we use Stirling's approximation and obtain

$$\sum_{k\geq 2\mathrm{e}r^p} \frac{r^k}{(k!)^{1/p}} \leq \sum_{k\geq 2\mathrm{e}r^p} \left(\frac{\mathrm{e}r^p}{k}\right)^{k/p} \leq \sum_{k\geq 2\mathrm{e}r^p} 2^{-k/p} \leq 2\cdot 2^{-(2\mathrm{e}r^p)/p} .$$

Combining these two estimates, the bound (3.3.9) follows for $p > 1$ as well.

We now turn to the proof of (3.3.10). For a given $r \geq 1$ choose $k \in \mathbb{N}$ such that

$$r^p - 1 < k \leq r^p .$$

Since all terms in the sum defining $\Phi_{p,b}^{L,l}$ are positive it follows that

$$\Phi_{p,b}^{L,l}(r) \geq \frac{\exp(-b\sqrt{k})}{(k!)^{1/p}} r^k .$$

Now

$$r^k \geq r^{r^p-1}$$

and, using Stirling's approximation,

$$(k!)^{1/p} \leq (\mathrm{e}^2 k)^{1/(2p)} \left(\frac{k}{\mathrm{e}}\right)^{k/p} \leq (\mathrm{e}^2 r^p)^{1/(2p)} \frac{r^{r^p}}{\mathrm{e}^{\frac{1}{p}r^p}} .$$

Furthermore, we have

$$\exp(-b\sqrt{k}) \geq \exp(-br^{p/2})\,.$$

Thus, combining all previous estimates and simplifying we have

$$\Phi_p^{L,u}(r) \geq \frac{\exp(-br^{p/2})}{\mathrm{e}^{1/p} r^{3/2}} \left(\frac{1}{p} r^p\right)\,,$$

and the bound (3.3.10) follows. □

We are now ready to give upper and lower bounds for $F_{\dot{w}}$ as well as its asymptotics for w generating the Schatten-Lorentz ideal.

Proposition 3.3.4. *Let $p \in (0,\infty)$ and let $w_k = k^{-1/p}$ for $k \in \mathbb{N}$. Then for all $r > 0$ we have*

$$(1+r)\Phi_{p/2,12/p}^{L,l}\left((2\mathrm{e})^{2/p}(Cr)^2\right) \leq F_{\dot{w}}(r) \leq (1+r)\Phi_{p/2}^{L,u}\left((2\mathrm{e})^{2/p}(Cr)^2\right) \quad (3.3.11)$$

Moreover

$$\log F_{\dot{w}}(r) \sim \frac{4\mathrm{e}C^p}{p} r^p \;\text{ as } r \to \infty.$$

Proof. By Lemma 3.3.1 we have for all $k \in \mathbb{N}$

$$\exp\left(-\frac{12}{p}\sqrt{k}\right) \frac{(2\mathrm{e})^{2k/p}}{(k!)^{2/p}} \leq \prod_{n=1}^{k} \dot{w}_n^2 \leq \frac{(2\mathrm{e})^{2k/p}}{(k!)^{2/p}}\,.$$

Using the definition of $F_{\dot{w}}$ in (3.2.1) the inequalities in (3.3.11) follow, which, using Lemma 3.3.3, imply the remaining assertion. □

We now turn our attention to the exponential cases, which are compactness classes E_w with weights of the form $w_k = \exp(-ak^\alpha)$ with $a, \alpha \in (0,\infty)$. We start with two technical lemmas.

Lemma 3.3.5. *Let $a, \alpha \in (0, \infty)$, and let $w_k = \exp(-ak^\alpha)$ for $k \in \mathbb{N}$. Then there are strictly positive real constants $\bar{c}_{a,\alpha}$, $\dot{\bar{c}}_{a,\alpha}$ and $c_{a,\alpha}$ such that the following inequalities hold for every $k \in \mathbb{N}$*

$$\exp\left(-\frac{a}{\alpha+1}k^\alpha - \bar{c}_{a,\alpha}k^{\alpha-1/2}\right) \leq \bar{w}_k \leq \exp\left(-\frac{a}{\alpha+1}k^\alpha\right), \tag{3.3.12}$$

$$\exp\left(-\frac{2^{-\alpha}a}{\alpha+1}k^\alpha - \dot{\bar{c}}_{a,\alpha}k^{\alpha-1/2}\right) \leq \dot{\bar{w}}_k \leq \exp\left(-\frac{2^{-\alpha}a}{\alpha+1}k^\alpha\right), \tag{3.3.13}$$

$$\exp\left(-\frac{2^{-\alpha}a}{(\alpha+1)^2}k^{\alpha+1} - c_{a,\alpha}k^{\alpha+1/2}\right) \leq \prod_{n=1}^{k} \dot{\bar{w}}_n \leq \exp\left(-\frac{2^{-\alpha}a}{(\alpha+1)^2}k^{\alpha+1}\right). \tag{3.3.14}$$

Proof. We start with (3.3.12). First we note that

$$\bar{w}_k^k = \exp\left(-a\sum_{n=1}^{k} n^\alpha\right) \quad (\forall k \in \mathbb{N}).$$

Since $\int_0^k t^\alpha\,dt \leq \sum_{n=1}^k n^\alpha \leq \int_0^{k+1} t^\alpha\,dt$, we have

$$\frac{1}{\alpha+1}k^{\alpha+1} \leq \sum_{n=1}^{k} n^\alpha \leq \frac{1}{\alpha+1}(k+1)^{\alpha+1} \quad (\forall k \in \mathbb{N}), \tag{3.3.15}$$

from which the upper bound of (3.3.12) readily follows, while the lower bound can be obtained by observing that there is a constant $K_1 > 0$ such that

$$\frac{(k+1)^{\alpha+1}}{k} \leq k^\alpha + K_1 k^{\alpha-1/2} \quad (\forall k \in \mathbb{N}).$$

For the next pair of inequalities (3.3.13) we note that, using (3.3.12), we have for $k \in \mathbb{N}$ even

$$\dot{\bar{w}}_k = \bar{w}_{\frac{k}{2}} \leq \exp\left(-\frac{2^{-\alpha}a}{\alpha+1}k^\alpha\right),$$

while for $k \in \mathbb{N}$ odd

$$\dot{\bar{w}}_k = \bar{w}_{\frac{k+1}{2}} \leq \exp\left(-\frac{2^{-\alpha}a}{\alpha+1}(k+1)^\alpha\right) \leq \exp\left(-\frac{2^{-\alpha}a}{\alpha+1}k^\alpha\right),$$

and the upper bound follows. For the lower bound we note that by (3.3.12), we have for $k \in \mathbb{N}$ even

$$\dot{\bar{w}}_k = \bar{w}_{\frac{k}{2}} \geq \exp\left(-\frac{2^{-\alpha}a}{\alpha+1}k^\alpha - 2^{-\alpha+1/2}\bar{c}_{a,\alpha}k^{\alpha-1/2}\right),$$

while for $k \in \mathbb{N}$ odd we have

$$\dot{\bar{w}}_k = \bar{w}_{\frac{k+1}{2}} \geq \exp\left(-\frac{2^{-\alpha}a}{\alpha+1}(k+1)^\alpha - 2^{-\alpha+1/2}\bar{c}_{a,\alpha}(k+1)^{\alpha-1/2}\right),$$

from which the lower bound follows for all $k \in \mathbb{N}$ by observing that for any $\beta > 0$ and any $K_2 > 0$ there is a constant $K_3 > 0$ such that

$$(k+1)^\beta + K_2(k+1)^{\beta-1/2} \leq k^\beta + K_3 k^{\beta-1/2} \quad (\forall k \in \mathbb{N}). \tag{3.3.16}$$

Finally, using (3.3.13) and (3.3.15), the upper bound in (3.3.14) follows, since we have for all $k \in \mathbb{N}$

$$\prod_{n=1}^{k} \dot{\bar{w}}_n \leq \exp\left(-\frac{2^{-\alpha}a}{\alpha+1}\sum_{n=1}^{k} n^\alpha\right) \leq \exp\left(-\frac{2^{-\alpha}a}{(\alpha+1)^2}k^{\alpha+1}\right).$$

The lower bound in turn follows from

$$\begin{aligned}\prod_{n=1}^{k} \dot{\bar{w}}_n &\geq \exp\left(-\frac{2^{-\alpha}a}{\alpha+1}\sum_{n=1}^{k} n^\alpha - \dot{\bar{c}}_{a,\alpha}\sum_{n=1}^{k} n^{\alpha-1/2}\right)\\ &\geq \exp\left(-\frac{2^{-\alpha}a}{(\alpha+1)^2}(k+1)^{\alpha+1} - \frac{2\dot{\bar{c}}_{a,\alpha}}{2\alpha+1}(k+1)^{\alpha+1/2}\right)\end{aligned}$$

and (3.3.16). □

Lemma 3.3.6. *Suppose that* $a, \alpha, b \in (0, \infty)$. *Let* $\Phi^{E,u}_{a,\alpha}$ *and* $\Phi^{E,l}_{a,\alpha,b}$ *be two functions given by the power series*

$$\Phi^{E,u}_{a,\alpha}(r) = \sum_{k=0}^{\infty} \exp(-ak^{\alpha+1}) r^k \,, \tag{3.3.17}$$

$$\Phi^{E,l}_{a,\alpha,b}(r) = \sum_{k=0}^{\infty} \exp(-ak^{\alpha+1} - bk^{\alpha+1/2}) r^k \,. \tag{3.3.18}$$

Then $\Phi^{E,u}_{a,\alpha}$ *and* $\Phi^{E,l}_{a,\alpha,b}$ *extend to entire functions with the following asymptotics*

$$\log \Phi^{E,u}_{a,\alpha}(r) \sim \log \Phi^{E,l}_{a,\alpha,b}(r) \sim a^{-1/\alpha} \frac{\alpha}{(\alpha+1)^{1+1/\alpha}} (\log r)^{1+1/\alpha} \quad \textit{as } r \to \infty \,.$$

Proof. It is not difficult to see that both $\Phi^{E,u}_{a,\alpha}$ and $\Phi^{E,l}_{a,\alpha,b}$ extend to entire functions. As in the proof of the analogous result for the Schatten-Lorentz ideal, we note that since $\Phi^{E,l}_{a,\alpha,b}(r) \le \Phi^{E,u}_{a,\alpha}(r)$ for all $r \in (0, \infty)$, the remaining assertions will hold if we can show that

$$\limsup_{r\to\infty} a^{1/\alpha} \frac{(\alpha+1)^{1+1/\alpha}}{\alpha} (\log r)^{-1-1/\alpha} \log \Phi^{E,u}_{a,\alpha}(r) \le 1 \tag{3.3.19}$$

and

$$\liminf_{r\to\infty} a^{1/\alpha} \frac{(\alpha+1)^{1+1/\alpha}}{\alpha} (\log r)^{-1-1/\alpha} \log \Phi^{E,l}_{a,\alpha,b}(r) \ge 1 \,. \tag{3.3.20}$$

We start with (3.3.19). Fix $r \ge 1$. Let $\mu(r)$ denote the maximal term of the series (3.3.17), that is,

$$\mu(r) = \max_{k\in\mathbb{N}} \left\{ \exp(-ak^{\alpha+1}) r^k \right\} \,,$$

and note that

$$\log \mu(r) \le a^{-1/\alpha} \frac{\alpha}{(\alpha+1)^{1+1/\alpha}} (\log r)^{1+1/\alpha} , \tag{3.3.21}$$

which follows from a short calculation. Next, let

$$k(r) = \left(\frac{\log(2r)}{a}\right)^{1/\alpha} ,$$

and observe that

$$\exp(-ak^{\alpha+1}) \le (2r)^{-k} \quad (\forall k \ge k(r)) .$$

Thus, for every $r \ge 1$ we have

$$\begin{aligned}
\Phi^{E,u}_{a,\alpha}(r) &= \sum_{k<k(r)} \exp(-ak^{\alpha+1}) r^k + \sum_{k \ge k(r)} \exp(-ak^{\alpha+1}) r^k \\
&\le \sum_{k<k(r)} \mu(r) + \sum_{k \ge k(r)} \frac{1}{2^k} \\
&\le (k(r)+1)\mu(r) + \frac{1}{2^{k(r)}} ,
\end{aligned}$$

from which (3.3.19) follows.

We now turn to the proof of (3.3.20). For a given $r \ge 1$ choose $k \in \mathbb{N}$ such that

$$k \le \left(\frac{\log r}{a(\alpha+1)}\right)^{1/\alpha} < k+1 .$$

Since all terms in the sum defining $\Phi^{E,l}_{a,\alpha,b}$ are positive we have

$$\begin{aligned}
\Phi^{E,l}_{a,\alpha,b}(r) &\ge \exp(-ak^{\alpha+1} - bk^{\alpha+1/2}) r^k \\
&\ge \frac{1}{r} \exp\left(-b \left(\frac{\log r}{a(\alpha+1)}\right)^{1+1/(2\alpha)}\right) \exp\left(\frac{\alpha (\log r)^{1+1/\alpha}}{a^{1/\alpha}(\alpha+1)^{1+1/\alpha}}\right) ,
\end{aligned}$$

from which the bound (3.3.20) follows. □

We are now able to give upper and lower bounds as well as the precise asymptotics of $F_{\dot{w}}$ for weights generating exponential classes.

Proposition 3.3.7. *Let $a, \alpha \in (0, \infty)$ and let $w_k = \exp(-ak^\alpha)$ for $k \in \mathbb{N}$. Then for all $r \geq 1$ we have*

$$(1+r)\Phi^{E,l}_{a',\alpha,2c_{a,\alpha}}\left((Cr)^2\right) \leq F_{\dot{w}}(r) \leq (1+r)\Phi^{E,u}_{a',\alpha}\left((Cr)^2\right), \qquad (3.3.22)$$

where $c_{a,\alpha}$ is the constant occurring in Lemma 3.3.5 and

$$a' = \frac{2^{1-\alpha}a}{(\alpha+1)^2}.$$

Moreover

$$\log F_{\dot{w}}(r) \sim 4\left(\frac{\alpha+1}{a}\right)^{1/\alpha} \frac{\alpha}{\alpha+1} (\log r)^{1+1/\alpha} \quad \text{as } r \to \infty.$$

Proof. The inequalities in (3.3.22) follow from Lemma 3.3.5 and the definition of $F_{\dot{w}}$ in (3.2.1). The remaining assertion follows from (3.3.22) and Lemma 3.3.6. □

3.4 Bounds for the spectral distance

As already mentioned in the introduction, the resolvent bounds deduced in Section 3.2 together with the Bauer-Fike argument allow us to derive the main result of this thesis: upper bounds for the spectral distance of two operators belonging to $E_w(H)$ expressible in terms of the distance of the two operators in operator norm and their w-departures from normality.

Theorem 3.4.1. *Let $w \in \mathcal{W}$.*

(i) If $A \in E_w(H)$ is not normal, then

$$\hat{d}(\sigma(B), \sigma(A)) \leq \nu_w(A) H_w \left(\frac{\|A - B\|}{\nu_w(A)} \right) \quad (\forall B \in L(H)) \,. \tag{3.4.1}$$

(ii) If $A, B \in E_w(H)$ and neither A nor B are normal, then

$$\mathrm{Hdist}(\sigma(A), \sigma(B)) \leq m H_w \left(\frac{\|A - B\|}{m} \right) , \tag{3.4.2}$$

where $m := \max\{\nu_w(A), \nu_w(B)\}$.

Here, the function $H_w : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ is defined by

$$H_w(r) = \frac{1}{\tilde{F}_{\dot{w}}^{-1}(\frac{1}{r})} \,,$$

where $\tilde{F}_{\dot{w}}^{-1}$ is the inverse of $\tilde{F}_{\dot{w}} : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ defined by

$$\tilde{F}_{\dot{w}}(r) = r F_{\dot{w}}(r) \,,$$

and $F_{\dot{w}}$ is the function defined in (3.2.1).

Proof.

(i) By Theorem 3.2.12,

$$\|R(A; z)\| \leq \frac{1}{\nu_w(A)} \tilde{F}_{\dot{w}} \left(\frac{\nu_w(A)}{d(z, \sigma(A))} \right) \quad (\forall z \in \rho(A)) \,,$$

so the assertion follows from the Bauer-Fike argument (see Remark 1.1.12).

(ii) Similarly, by Theorem 3.2.12 and Remark 3.2.13 we have

$$\|R(A;z)\| \leq \frac{1}{m}\tilde{F}_{\dot{w}}\left(\frac{m}{d(z,\sigma(A))}\right) \quad (\forall z \in \rho(A)),$$

and

$$\|R(B;z)\| \leq \frac{1}{m}\tilde{F}_{\dot{w}}\left(\frac{m}{d(z,\sigma(B))}\right) \quad (\forall z \in \rho(B)),$$

so the assertion again follows by invoking the Bauer-Fike argument.

□

Remark 3.4.2.

(i) Note that $\lim_{r\downarrow 0} H_w(r) = 0$, thus the bounds for the spectral variation and spectral distance become small when $\|A - B\|$ is small.

(ii) The bounds (3.4.1) and (3.4.2) remain valid if we replace $\nu_w(A)$ and $\nu_w(B)$ by a larger quantity, say by the upper bounds given in Proposition 3.2.11.

(iii) Combining Remarks 1.1.7 and 1.1.12 it follows that if A is a bounded normal operator on H, then

$$\hat{d}(\sigma(B),\sigma(A)) \leq \|A - B\| \quad (\forall B \in L(H)).$$

Moreover, by symmetry it follows from the above that if both A and B are bounded normal operators then

$$\text{Hdist}(\sigma(A),\sigma(B)) \leq \|A - B\| .$$

Note that these two bounds can be thought of as limiting cases of the previous theorem, since, as is easily seen, we have for any $r \geq 0$

$$\lim_{C \downarrow 0} C H_w\left(\frac{r}{C}\right) = r\,.$$

It is in this respect that the bounds (3.4.1) and (3.4.2) are sharp.

Remark 3.4.3. In the case of the Schatten-Lorentz ideal and the exponential classes it is possible to give rather precise estimates for the behaviour of the general bound in Theorem 3.4.1 for two operators which are close in operator norm.

We start with the Schatten-Lorentz ideal. Let $p \in (0,\infty)$ and let $w_k = k^{-1/p}$ for $k \in \mathbb{N}$. Then Proposition 3.3.4 yields

$$\log \tilde{F}_{\dot{w}}(r) \sim \frac{4C^p \mathrm{e}}{p} r^p \text{ as } r \to \infty$$

which, using Lemma 3.3.2, implies that

$$\tilde{F}_{\dot{w}}^{-1}(r) \sim \frac{1}{C}\left(\frac{p}{4\mathrm{e}}\right)^{1/p} (\log r)^{1/p} \text{ as } r \to \infty,$$

which, in turn, gives

$$H_w(r) \sim C\left(\frac{4\mathrm{e}}{p}\right)^{1/p} |\log r|^{-1/p} \text{ as } r \downarrow 0.$$

We now turn to the exponential classes. Let $a, \alpha \in (0,\infty)$ and let $w_k = \exp(-ak^\alpha)$ for $k \in \mathbb{N}$. Now, Proposition 3.3.7 yields

$$\log \tilde{F}_{\dot{w}}(r) \sim 4\left(\frac{\alpha+1}{a}\right)^{1/\alpha} \frac{\alpha}{\alpha+1} (\log r)^{1+1/\alpha} \text{ as } r \to \infty$$

which, using Lemma 3.3.2, implies that

$$\log \tilde{F}_{\dot{w}}^{-1}(r) \sim 4^{-\alpha/(\alpha+1)} \left(\frac{a}{\alpha+1}\right)^{1/(\alpha+1)} \left(\frac{\alpha+1}{\alpha}\right)^{\alpha/(\alpha+1)} (\log r)^{\alpha/(\alpha+1)} \quad \text{as } r \to \infty\,,$$

which, in turn, gives

$$\log H_w(r) \sim -4^{-\alpha/(\alpha+1)} \left(\frac{a}{\alpha+1}\right)^{1/(\alpha+1)} \left(\frac{\alpha+1}{\alpha}\right)^{\alpha/(\alpha+1)} |\log r|^{\alpha/(\alpha+1)} \quad \text{as } r \downarrow 0.$$

3.5 An application to inclusion regions for pseudospectra

Pseudospectra play an important role in numerical linear algebra and perturbation theory (see, for example, [Tre97, Dav07]). They are defined as follows.

Definition 3.5.1. Let $A \in L(H)$ and $\epsilon > 0$. The *ϵ-pseudospectrum of A* is defined by

$$\sigma_\epsilon(A) = \sigma(A) \cup \{\, z \in \rho(A) \,:\, \|R(A;z)\| > 1/\epsilon \,\}\,. \tag{3.5.1}$$

The motivation behind this definition is the observation that for any $A \in L(H)$ and any $\epsilon > 0$ we have

$$\sigma_\epsilon(A) = \bigcup_{\substack{B \in L(H) \\ \|A-B\| < \epsilon}} \sigma(B) \tag{3.5.2}$$

as is easily seen using standard perturbation theory. In other words, the ϵ-pseudospectrum of a bounded linear operator is equal to the union of the spectra of all perturbed operators with perturbations that have norms strictly less than ϵ.

It turns out that if in the definition of the pseudospectrum (3.5.1) the strict inequality is replaced by a non-strict one, then the alternative characterisation (3.5.2) holds with the strict inequality replaced by a non-strict one. Curiously enough, this is no longer necessarily true for operators on Banach spaces (see [Sha09]).

While there exist efficient methods to compute pseudospectra of matrices (see, for example, [Tre97, Section 4], for a brief overview), the same is not true for operators on infinite-dimensional spaces, where the exact computation of pseudospectra can be a very challenging a task. As an application of our resolvent bounds obtained in Section 3.2, we shall now provide circular inclusion regions for the pseudospectra of operators in a given compactness class.

Theorem 3.5.2. *Let* $\epsilon > 0$.

(i) If $A \in L(H)$*, then*

$$\{ z \in \mathbb{C} : d(z, \sigma(A)) < \epsilon \} \subseteq \sigma_\epsilon(A) .$$

(ii) If $A \in E_w(H)$ *is not normal, then*

$$\sigma_\epsilon(A) \subseteq \left\{ z \in \mathbb{C} : d(z, \sigma(A)) < \nu_w(A) H_w \left(\frac{\epsilon}{\nu_w(A)} \right) \right\} ,$$

where H_w *is the function defined in Theorem 3.4.1.*

Proof.

(i) The inclusion relation follows immediately from the following lower bound for the resolvent of an operator

$$\|R(A; z)\| \geq \frac{1}{d(z, \sigma(A))} ,$$

which in turn follows from

$$\frac{1}{d(z,\sigma(A))} = \sup_{\lambda\in\sigma(A)} |z-\lambda|^{-1} = r(R(A;z)) \leq \|R(A;z)\| \,.$$

(ii) By Theorem 3.2.12 we have the resolvent bound

$$\|R(A;z)\| \leq \frac{1}{\nu_w(A)}\tilde{F}_{\dot{w}}\left(\frac{\nu_w(A)}{d(z,\sigma(A))}\right) \qquad (\forall z\in\rho(A))\,,$$

where $\tilde{F}_{\dot{w}}(r) = rF_{\dot{w}}(r)\,.$

If $z\in\sigma_\epsilon(A)$, then

$$\frac{1}{\epsilon} < \|R(A;z)\| \leq \frac{1}{\nu_w(A)}\tilde{F}_{\dot{w}}\left(\frac{\nu_w(A)}{d(z,\sigma(A))}\right)\,,$$

and a short calculation shows that

$$d(z,\sigma(A)) < \nu_w(A)H_w\left(\frac{\epsilon}{\nu_w(A)}\right)\,,$$

as desired.

□

Remark 3.5.3.

(i) Note that the inclusion (ii) above also follows from the characterisation (3.5.2) and Theorem 3.4.1 (i).

(ii) Note that the inclusion (ii) is sharp in the limiting case of normal A, since it reduces to

$$\sigma_\epsilon(A) = \{\, z\in\mathbb{C} \,:\, d(z,\sigma(A)) < \epsilon \,\}\,.$$

CHAPTER 4

Trace Class Operators

In this chapter, which is based on the paper [BG15], we shall return to Elsner's bound for the spectral distance of two matrices discussed in the introduction (see Theorem 1.1.1), and show that his determinant-based proof can be modified so as to produce spectral distance bounds for trace class operators.

In Section 4.1 we first derive a lower bound for $\det(I - z^{-1}A)$ of a trace class operator A involving solely the distance of z to the spectrum of A and the singular values of A. Using the assumption that z is an eigenvalue of a bounded operator B we then derive an upper bound for $\det(I - z^{-1}A)$ in terms of the distance of z to the spectrum of A and the distance of A and B in operator norm. These lower and upper bounds are of independent interest and do not seem to have appeared in the literature yet. In Section 4.2 we combine the upper and lower bounds to provide an upper bound for the spectral distance of two trace class operators A and B expressible only in terms of $\|A - B\|$ and the singular values of A and B, which produces new bounds even in the finite-dimensional setting. In Section 4.3 we compare our bound to Elsner's and show that it reproduces or improves the bounds in [Ban04, Ban08]. We also compare it to the bound obtained in the previous chapter for operators in the Schatten-Lorentz ideals and in the exponential classes.

4.1 Upper and lower bounds for determinants

The upper and lower bounds for the determinant of a trace class operator, which are the main ingredients for our spectral distance bounds to be presented later in this chapter, will be expressible in terms of a certain function of the singular values of a given trace class operator, which we now define.

Definition 4.1.1. Given $A \in S_1(H)$, let $F_A : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ be defined by

$$F_A(r) = \prod_{k=1}^{\infty} (1 + r s_k(A)) \, .$$

Note that since the singular values of a trace class operator A are by definition summable, the function F_A is well-defined. Moreover, it is not difficult to see that it is also real-analytic and strictly monotonically increasing.

We start with the lower bound. In order to motivate it, we observe that the non-zero eigenvalues of a trace class operator A are zeros of the function $z \mapsto \det(I - z^{-1}A)$ by the very definition of the determinant of A. We thus expect that $z \mapsto \det(I - z^{-1}A)$ should have a lower bound which can be expressed in terms of the distance of z to the spectrum of A. This is indeed the case, as the following proposition shows.

Proposition 4.1.2. *Suppose that $A \in S_1(H)$. Then, for any $z \in \rho(A)$ with $z \neq 0$,*

$$\left|\det(I - z^{-1}A)\right|^{-1} \leq F_A\left(\frac{1}{d(z, \sigma(A))}\right) \, .$$

Proof. We shall use the abbreviations $\lambda_k = \lambda_k(A)$ and $s_k = s_k(A)$. For $z \in$

$\rho(A)$ with $z \neq 0$ we get

$$\left|\det(I - z^{-1}A)\right|^{-1} = \prod_{k=1}^{\infty} \left|\frac{1}{1 - z^{-1}\lambda_k}\right| = \prod_{k=1}^{\infty} \left|1 + \frac{\lambda_k}{z - \lambda_k}\right|$$
$$\leq \prod_{k=1}^{\infty} \left(1 + \frac{|\lambda_k|}{|z - \lambda_k|}\right) \leq \prod_{k=1}^{\infty} \left(1 + \frac{|\lambda_k|}{d(z, \sigma(A))}\right) \leq \prod_{k=1}^{\infty} \left(1 + \frac{s_k}{d(z, \sigma(A))}\right),$$

where the final inequality follows from Corollary 2.3.4. □

We now turn to the upper bound. As we are ultimately interested in spectral distance bounds expressible in terms of the distance in operator norm of two trace class operators, we shall now derive an upper bound for the determinant $\det(I - z^{-1}A)$ in case z is an eigenvalue of a bounded operator B, which involves the distance of z to $\sigma(A)$ and the distance of A and B in operator norm.

Proposition 4.1.3. *Let* $\dim H = \infty$ *and suppose that* $A, B \in L(H)$*. If* A *is a finite rank operator with* $\operatorname{rank}(A) = n$*, then, for any* $z \in \rho(A) \cap \sigma_p(B)$*, we have*

$$\left|\det(I - z^{-1}A)\right| \leq \frac{\|A - B\|}{d(z, \sigma(A))} \prod_{k=1}^{n} \left(1 + \frac{s_k(A)}{d(z, \sigma(A))}\right).$$

Proof. Fix $z \in \rho(A) \cap \sigma_p(B)$. Note that since $\dim H = \infty$, we have $0 \in \sigma(A)$, thus $z \neq 0$. Assume now that e is an eigenvector of B corresponding to z and let E be the smallest closed linear span containing the range of A and e. It is easy to see that E is an invariant subspace of A with $\nu := \dim E \leq n + 1$. Let now A_E, B_E and I_E denote the restrictions of $A, B, I \in L(H)$ to E, respectively. We then have

$$\lambda_k(A) = \begin{cases} \lambda_k(A_E) & \text{if } k \leq \nu, \\ 0 & \text{if } k > \nu. \end{cases}$$

Therefore by Remark 2.3.2 we now have

$$\left|\det(I - z^{-1}A)\right| = \prod_{k=1}^{\nu} \left|\lambda_k(I_E - z^{-1}A_E)\right| = \prod_{k=1}^{\nu} s_k(I_E - z^{-1}A_E)\,. \tag{4.1.1}$$

For $k < \nu$ we use the alternative characterisation of the singular values in Theorem 2.2.11 giving

$$s_k(I_E - z^{-1}A_E) \leq 1 + |z|^{-1}\, s_k(A_E) \leq 1 + |z|^{-1}\, s_k(A)\,. \tag{4.1.2}$$

For the ν-th factor we define a finite rank operator

$$K : E \to E \quad \text{by} \quad K = I_E - z^{-1}PB_E\,,$$

where P denotes the orthogonal projection onto E. Since $Ke = 0$ we have $\operatorname{rank}(K) < \nu$. Thus using Theorem 2.2.11 and Corollary 2.2.10 we get

$$\begin{aligned} s_\nu(I_E - z^{-1}A_E) &\leq \left\|(I_E - z^{-1}PA_E) - K\right\| \\ &= |z|^{-1}\left\|PB_E - PA_E\right\| \leq |z|^{-1}\left\|A - B\right\|\,. \end{aligned} \tag{4.1.3}$$

In order to complete the proof we combine (4.1.1), (4.1.2), (4.1.3) and note that $|z| \geq d(z, \sigma(A))$ since $0 \in \sigma(A)$. This yields

$$\left|\det(I - z^{-1}A)\right| \leq \frac{\|A - B\|}{d(z,\sigma(A))} \prod_{k=1}^{\nu-1}\left(1 + \frac{s_k(A)}{d(z,\sigma(A))}\right),$$

as desired. □

Remark 4.1.4. In the above proof we have only used the hypothesis that H is infinite-dimensional to conclude that $|z| \geq d(z, \sigma(A))$. Another look at the above proof shows that it also works for operators acting on finite-dimensional spaces. The finite-dimensional version of the above result is the following.

Let $A, B \in L(\mathbb{C}^n)$. Using the fact that both the eigenvector of B and the range of A trivially belong to the same n-dimensional space, we have, for any $z \in \rho(A) \cap \sigma(B)$ with $z \neq 0$,

$$\left|\det(I - z^{-1}A)\right| \leq \frac{\|A - B\|}{d(z, \sigma(A) \cup \{0\})} \prod_{k=1}^{n-1} \left(1 + \frac{s_k(A)}{d(z, \sigma(A) \cup \{0\})}\right).$$

We now extend the result obtained in the previous proposition to any trace class operator using an approximation argument.

Proposition 4.1.5. *Suppose that* $\dim H = \infty$. *If* $A \in S_1(H)$ *and* $B \in L(H)$, *then, for any* $z \in \rho(A) \cap \sigma_p(B)$, *we have*

$$\left|\det(I - z^{-1}A)\right| \leq \frac{\|A - B\|}{d(z, \sigma(A))} F_A \left(\frac{1}{d(z, \sigma(A))}\right).$$

Proof. Let $A \in S_1(H)$. If A is a finite rank operator, then the assertion follows from Proposition 4.1.3, so we now assume that A does not have finite rank. By Theorem 2.2.5 it has a Schmidt representation of the form

$$Ax = \sum_{k=1}^{\infty} s_k(A)(x, e_k) f_k \quad (\forall x \in H),$$

where $(e_k)_{k\in\mathbb{N}}$ and $(f_k)_{k\in\mathbb{N}}$ are orthonormal systems in H.

For any $n \in \mathbb{N}$, let $A_n : H \to H$ be a finite rank operator defined by

$$A_n x = \sum_{k=1}^{n} s_k(A)(x, e_k) f_k \quad (\forall x \in H).$$

After a short calculation we find that

$$s_k(A_n) = \begin{cases} s_k(A) & \text{for } k \le n, \\ 0 & \text{for } k > n, \end{cases}$$

which shows that

$$\lim_{n\to\infty} \|A_n - A\|_1 = 0 \tag{4.1.4}$$

and

$$F_{A_n}(r) \le F_A(r) \quad (\forall r \in \mathbb{R}_0^+)\,. \tag{4.1.5}$$

Fix $z \in \rho(A) \cap \sigma_p(B)$. Since $z \in \rho(A)$ there is $N \in \mathbb{N}$ such that $z \in \rho(A_n)$ for all $n \ge N$ (see, for example, [GGK90, Chapter II, Theorem 4.1]). Therefore using Proposition 4.1.3 and the inequality (4.1.5) we have

$$\left|\det(I - z^{-1}A_n)\right| \le \frac{\|A_n - B\|}{d(z,\sigma(A_n))} F_A\left(\frac{1}{d(z,\sigma(A_n))}\right) \quad (\forall n \ge N)\,. \tag{4.1.6}$$

We now combine a number of arguments in order to obtain the desired inequality.

First we note that the determinant is Lipschitz-continuous on the unit ball of $S_1(H)$. More precisely, by [Sim77, Theorem 6.5], we have

$$|\det(I-z^{-1}A_n)-\det(I-z^{-1}A)| \le |z^{-1}|\,\|A_n - A\|_1 \exp\left(|z^{-1}|\,(\|A_n\|_1 + \|A\|_1) + 1\right)\,.$$

Using the above and (4.1.4), it follows that

$$\lim_{n\to\infty} \det(I - z^{-1}A_n) = \det(I - z^{-1}A)\,.$$

Besides, we obviously have

$$\lim_{n\to\infty} \|A_n - B\| = \|A - B\| \ .$$

Moreover since the spectrum of A is discrete, the spectrum of A_n converges to the spectrum of A in the Hausdorff metric (see [New51, Theorem 3]). Thus we have

$$\lim_{n\to\infty} d(z, \sigma(A_n)) = d(z, \sigma(A))\,.$$

Taking the limit on both sides of (4.1.6) the desired inequality follows. □

4.2 Bounds for the spectral distance

In this section we combine the upper and lower bounds for determinants deduced in the previous section in order to obtain an explicitly computable upper bound for the spectral distance of two trace class operators. Before doing so, we shall introduce some notation.

Let $\mathcal{F}$ denote the set of all continuous strictly monotonically increasing functions

$$F : \mathbb{R}_0^+ \to \mathbb{R}_0^+ \quad \text{with} \quad \lim_{r\to\infty} F(r) = \infty\,,$$

and let $\mathcal{F}_0$ denote the subset of those $F \in \mathcal{F}$ with $F(0) = 0$. Both $\mathcal{F}$ and $\mathcal{F}_0$ are partially ordered by defining

$$F_1 \leq F_2 :\Leftrightarrow F_1(r) \leq F_2(r) \quad (\forall r \geq 0)\,.$$

We now define an operation H on $\mathcal{F}$, which will play an important role in the formulation of our spectral distance bounds:

$$H : \mathcal{F} \to \mathcal{F}_0$$

$$F \mapsto H_F \,,$$

where H_F is given by

$$H_F(r) = \begin{cases} \frac{1}{\tilde{F}^{-1}\left(\frac{1}{r}\right)} & \text{if } r > 0, \\ 0 & \text{if } r = 0, \end{cases}$$

and $\tilde{F}^{-1}$ is the inverse of $\tilde{F} \in \mathcal{F}_0$ defined by $\tilde{F}(r) = rF(r)^2$.

The following lemma collects some useful properties of H to be used later when deducing the main result of this chapter.

Lemma 4.2.1. *Suppose that $F_1, F_2 \in \mathcal{F}$ and let $M_m \in \mathcal{F}$ be defined by $M_m(r) = mr$ for $m > 0$. Then we have*

(i) If $F_1 \leq F_2$, then $H_{F_1} \leq H_{F_2}$.

(ii) If $F_1 = F_2 \circ M_m$, then $H_{F_1} = M_m \circ H_{F_2} \circ M_m^{-1}$.

Proof. The proof of the first assertion is straightforward and will be skipped. For the second assertion let $F_1(r) = F_2 \circ M_m(r) = F_2(mr)$ then by definition we obtain $\tilde{F}_2(mr) = m\tilde{F}_1(r)$ which means

$$\tilde{F}_2 \circ M_m(r) = M_m \circ \tilde{F}_1(r) \,. \tag{4.2.1}$$

Taking inverses in (4.2.1) we obtain

$$M_m^{-1} \circ \tilde{F}_2^{\,-1} \circ M_m(r) = \tilde{F}_1^{\,-1}(r) \,. \tag{4.2.2}$$

By the construction of H_{F_1} and equality (4.2.2) we have

$$H_{F_1}(r) = \frac{1}{\tilde{F_1}^{-1}\left(\frac{1}{r}\right)} = \frac{1}{\frac{1}{m}\tilde{F_2}^{-1}\left(\frac{m}{r}\right)}$$

$$= mH_{F_2}(\frac{r}{m}) = M_m \circ H_{F_2} \circ M_m^{-1}(r)\,,$$

and the assertion follows. □

We are now ready to state and prove an upper bound for the spectral variation of a bounded operator with respect to a trace class operator.

Theorem 4.2.2. *Let* $\dim H = \infty$ *and suppose that* $A \in S_1(H)$ *and* $B \in L(H)$. *Then we have*

$$\hat{d}(\sigma_p(B), \sigma(A)) \leq H_{F_A}(\|A - B\|)\,.$$

Proof. For a given $B \in L(H)$ it suffices to show that

$$d(z, \sigma(A)) \leq H_{F_A}(\|A - B\|) \quad (\forall z \in \sigma_p(B))\,.$$

Since the bound above is trivial for $z \in \sigma(A)$ we shall assume $z \in \rho(A) \cap \sigma_p(B)$. Noting that z is not zero (since $\dim H = \infty$), we use Propositions 4.1.2 and 4.1.5 to obtain

$$\frac{1}{F_A\left(\frac{1}{d(z,\sigma(A))}\right)} \leq \left|\det(I - z^{-1}A)\right| \leq \frac{\|A - B\|}{d(z, \sigma(A))} F_A\left(\frac{1}{d(z, \sigma(A))}\right)\,.$$

This gives

$$\frac{1}{\|A - B\|} \leq \tilde{F}_A\left(\frac{1}{d(z, \sigma(A))}\right)\,,$$

where $\tilde{F}_A(r) = rF_A(r)^2$. Hence we get

$$\tilde{F}_A^{-1}\left(\frac{1}{\|A - B\|}\right) \leq \frac{1}{d(z, \sigma(A))}\,,$$

which means

$$d(z,\sigma(A)) \le \frac{1}{\tilde{F}_A^{-1}\left(\frac{1}{\|A-B\|}\right)} = H_{F_A}(\|A-B\|)\,,$$

as required. ☐

The preceding theorem immediately implies the following theorem, the main result of this chapter: an explicitly computable bound for the spectral distance of two trace class operators.

Theorem 4.2.3. *Let* $\dim H = \infty$ *and suppose that* $A, B \in S_1(H)$. *Then we have*

$$\mathrm{Hdist}(\sigma(A),\sigma(B)) \le H_{F_{A,B}}(\|A-B\|)\,,$$

where $F_{A,B} \in \mathcal{F}$ *is given by*

$$F_{A,B}(r) = \max\{F_A(r), F_B(r)\}\,.$$

Proof. Using the fact that $B \in S_1(H)$ as well as Theorem 4.2.2 together with Lemma 4.2.1 we obtain

$$\hat{d}(\sigma(B),\sigma(A)) = \hat{d}(\sigma_p(B),\sigma(A)) \le H_{F_A}(\|A-B\|) \le H_{F_{A,B}}(\|A-B\|)\,.$$

By symmetry, we also obtain

$$\hat{d}(\sigma(A),\sigma(B)) \le H_{F_{A,B}}(\|A-B\|)\,,$$

from which the theorem follows. ☐

Remark 4.2.4. In a spirit similar to that of Remark 4.1.4 we note that there are finite-dimensional versions of the previous two theorems, which can be proved along the same lines.

Let $A, B \in L(\mathbb{C}^n)$. Then we have

$$\hat{d}(\sigma(B), \sigma(A) \cup \{0\}) \leq H_{F_A}(\|A - B\|)\,,$$

and

$$\text{Hdist}(\sigma(A) \cup \{0\}, \sigma(B) \cup \{0\}) \leq H_{F_{A,B}}(\|A - B\|)\,.$$

4.3 Comparison with other bounds

For applications of the bound deduced in Theorem 4.2.3 we note that precise information about the singular values of the operators considered is not always necessary. In fact, using Lemma 4.2.1, it follows that upper bounds for the singular values suffice in the sense that any upper bound for the singular values of two operators translates into a bound for the spectral distance of the two operators.

By specialising to particular types of upper bounds we are able to compare our bound in Theorem 4.2.3 to the bounds derived in the previous chapter as well as to existing bounds in the literature, in particular to Elsner's original bound in [Els85]. We shall also show that it reproduces or improves the bounds in [Ban04, Ban08].

We start with a bound expressible in terms of one free parameter, namely the trace norm. Given $A \in S_1(H)$, we have

$$F_A(r) = \prod_{k=1}^{\infty}(1 + r s_k(A)) \leq \exp(r\,\|A\|_1)\,.$$

Theorem 4.2.3 and Lemma 4.2.1 now imply the following result which reproduces the known bound from [Ban04, Theorem 5.2] derived by a technique similar to the one we used in Chapter 3 of this thesis.

Corollary 4.3.1. *Suppose that* $\dim H = \infty$ *and that* $A, B \in S_1(H)$, *with not both of them the zero operator. Then we have*

$$\mathrm{Hdist}(\sigma(A), \sigma(B)) \leq mH_{G^S}\left(\frac{\|A-B\|}{m}\right),$$

where $G^S(r) = \exp(r)$ *and* $m = \max\{\|A\|_1, \|B\|_1\}$.

Proof. It is not difficult to see that

$$F_{A,B}(r) \leq G^S(mr) = G^S \circ M_m(r),$$

where $m = \max\{\|A\|_1, \|B\|_1\}$. Using Lemma 4.2.1 we obtain

$$H_{F_{A,B}}(\|A-B\|) \leq H_{G^S \circ M_m}(\|A-B\|) = M_m \circ H_{G^S} \circ M_m^{-1}(\|A-B\|). \quad (4.3.1)$$

The desired inequality now follows by applying Theorem 4.2.3 to (4.3.1). □

We shall now consider how the bound obtained in this chapter compares to the bound in the previous chapter for operators in the Schatten-Lorentz ideal, the definition of which we briefly recall.

Definition 4.3.2. Let $p \in (0, \infty)$. Then

$$S_{p,\infty}(H) = \{\, A \in S_\infty(H) \,:\, |A|_p := \sup_{k \in \mathbb{N}} s_k(A) k^{1/p} < \infty \,\},$$

is called the *Schatten-Lorentz ideal of type* p.

Choosing $w_k = k^{-1/p}$ we see that $E_w(H) = S_{p,\infty}(H)$ and that $|A|_w = |A|_p$ for every $A \in E_w(H)$. Moreover, by Proposition 3.1.12, the gauge $|\cdot|_p$ turns out to be a quasi-norm on the linear space $S_{p,\infty}(H)$.

It is not difficult to see that for $p < 1$ every operator in $S_{p,\infty}(H)$ is trace class, while for $p \geq 1$ this is no longer the case. For this reason we have to restrict to the case $p < 1$ in the following.

Now, Theorem 4.2.3 and Lemma 4.2.1 yield the following result.

Corollary 4.3.3. *Let $p \in (0,1)$. If $\dim H = \infty$ and $A, B \in S_{p,\infty}(H)$, not both of them the zero operator, then*

$$\mathrm{Hdist}(\sigma(A), \sigma(B)) \leq m H_{G_p^L}\left(\frac{\|A - B\|}{m}\right),$$

where

$$G_p^L(r) = \prod_{k=1}^{\infty}(1 + rk^{-1/p}) \tag{4.3.2}$$

and $m = \max\{|A|_p, |B|_p\}$.

Proof. This follows by arguments similar to those used in the previous corollary. ☐

In order to compare the bound above with the bound obtained in the previous chapter we require the following lemma.

Lemma 4.3.4. *Let $p \in (0,1)$ and let G_p^L be the function defined in (4.3.2), that is,*

$$G_p^L(r) = \prod_{k=1}^{\infty}(1 + rk^{-1/p}).$$

Then

$$\log G_p^L(r) \sim \frac{C_p}{p} r^p \text{ as } r \to \infty, \tag{4.3.3}$$

where

$$C_p = \int_0^\infty \frac{t^p}{(1+t)^2}\, dt,$$

with

$$\frac{1}{4}\frac{1}{1-p} \le C_p \le \frac{3-p}{2}\frac{1}{1-p}\,.$$

Moreover, the function $p \mapsto C_p^{1/p}$ *is monotonically increasing on* $(0,1)$.

Proof. We start by observing that for $r \ge 0$ we have

$$\log G_p^L(r) = \sum_{k=1}^{\infty} \log(1 + rk^{-1/p})\,,$$

so

$$\int_1^\infty \log(1+rt^{-1/p})\,dt \le \log G_p^L(r) \le \int_0^\infty \log(1+rt^{-1/p})\,dt\,. \tag{4.3.4}$$

Now, changing variables and integrating by parts twice in the integral on the right-hand side of the inequality above we have

$$\begin{aligned}\int_0^\infty \log(1+rt^{-1/p})\,dt &= r^p \int_0^\infty pt^{p-1}\log(1+t^{-1})\,dt\\ &= r^p\left(\left[t^p\log(1+t^{-1})\right]_0^\infty + \int_0^\infty t^{p-1}\frac{1}{1+t}\,dt\right)\\ &= r^p\left(\left[\frac{t^p}{p}\frac{1}{1+t}\right]_0^\infty + \int_0^\infty \frac{t^p}{p}\frac{1}{(1+t)^2}\,dt\right)\\ &= \frac{C_p}{p}r^p\,.\end{aligned}$$

Next we observe that there is $K_p > 0$ such that $\log(1+x) \le K_p x^{p/2}$ for all $x \ge 0$. Thus

$$\int_0^1 \log(1+rt^{-1/p})\,dt \le K_p \int_0^1 \frac{r^{p/2}}{\sqrt{t}}\,dt = 2K_p r^{p/2}\,. \tag{4.3.5}$$

Combining (4.3.4) and (4.3.5) we have for all $r \ge 0$

$$\frac{C_p}{p}r^p - 2K_p r^{p/2} \le \log G_p^L(r) \le \frac{C_p}{p}r^p\,,$$

from which the desired asymptotics (4.3.3) follows.

For the bounds on the constant C_p we note that

$$C_p = \int_0^1 \frac{t^p}{(1+t)^2}\,dt + \int_1^\infty \frac{t^p}{(1+t)^2}\,dt \le \int_0^1 \frac{1}{(1+t)^2}\,dt + \int_1^\infty \frac{t^{p-2}}{(1+t^{-1})^2}\,dt$$
$$\le \int_0^1 \frac{1}{(1+t)^2}\,dt + \int_1^\infty t^{p-2}\,dt = \frac{3-p}{2}\frac{1}{1-p}$$

and that

$$C_p \ge \int_1^\infty \frac{t^p}{(1+t)^2}\,dt = \int_1^\infty \frac{t^{p-2}}{(1+t^{-1})^2}\,dt \ge \int_1^\infty \frac{t^{p-2}}{4}\,dt = \frac{1}{4}\frac{1}{1-p}.$$

For the final assertion notice that for $p, q \in (0,1)$ with $p < q$ we have using Hölder's inequality

$$\int_0^\infty t^p \cdot 1 \frac{dt}{(1+t)^2} \le \left(\int_0^\infty (t^p)^{q/p} \frac{dt}{(1+t)^2} \right)^{p/q} \left(\int_0^\infty 1^{q/(q-p)} \frac{dt}{(1+t)^2} \right)^{(q-p)/q}$$

which implies

$$\left(\int_0^\infty t^p \frac{dt}{(1+t)^2} \right)^{1/p} \le \left(\int_0^\infty t^q \frac{dt}{(1+t)^2} \right)^{1/q}$$

from which the assertion follows. □

Using Lemma 3.3.2 the asymptotics given in the lemma above now implies that the asymptotics of the bound occurring in Corollary 4.3.3 is given by

$$H_{G_p^L}(r) \sim \left(\frac{2C_p}{p} \right)^{1/p} |\log r|^{-1/p} \ \text{ as } r \downarrow 0. \tag{4.3.6}$$

Let now $w_k = k^{-1/p}$ for $k \in \mathbb{N}$ and recall that the bound for the spectral distance of two operators in $S_{p,\infty}$ obtained in the previous chapter is governed by the function H_w, which, by Remark 3.4.3, satisfies

$$H_w(r) \sim C \left(\frac{4\mathrm{e}}{p}\right)^{1/p} |\log r|^{-1/p} \text{ as } r \downarrow 0, \tag{4.3.7}$$

where C is the constant occurring in Theorem 3.2.1. Comparing (4.3.6) and (4.3.7) the limit of the quotient of $H_w(r)$ and $H_{G_p^L}(r)$ as r tends to zero from above exists and we write

$$C(p) := \lim_{r\downarrow 0} \frac{H_w(r)}{H_{G_p^l}(r)} = C \left(\frac{2\mathrm{e}}{C_p}\right)^{1/p}$$

for $p \in (0,1)$.

Using the bounds for C_p given in Lemma 4.3.4 we see that

$$\lim_{p\uparrow 1} C(p) = 0 \text{ and } \lim_{p\downarrow 0} C(p) = \infty\,,$$

so for all nearby operators the seemingly cruder bound for the spectral distance obtained in the previous chapter fares better if p is close to 1, that is, close to the applicability of the method in this chapter, while the bound in this chapter is better if p is close to zero. The exact cross-over of the two regimes cannot be determined at this stage since the exact value of the constant C in Theorem 3.2.1 is not known. However, as $C \geq \pi/2$ we have

$$C(1/2) = C \left(\frac{4\mathrm{e}}{\pi}\right)^2 \geq \frac{8\mathrm{e}^2}{\pi} > 1\,.$$

Moreover, by the last assertion of Lemma 4.3.4, the function $p \mapsto C(p)$ is monotonically decreasing on $(0,1)$ so we certainly know that for nearby operators the bound in this chapter is better than the one in the previous chapter for all $p \leq 1/2$.

We now turn to the exponential classes introduced in [Ban08], which, as already mentioned, fit into the framework considered in this thesis.

Definition 4.3.5. Let $a > 0$ and $\alpha > 0$. Then

$$E(a, \alpha; H) = \{\, A \in S_1(H) \,:\, |A|_{a,\alpha} := \sup_{k\in\mathbb{N}} s_k(A)\exp(ak^\alpha) < \infty \,\},$$

is called the *exponential class of type* (a, α). The number $|A|_{a,\alpha}$ is called the (a, α)-*gauge* or simply *gauge of* A.

Clearly, choosing $w_k = \exp(-ak^\alpha)$ we see that $E_w(H) = E(a, \alpha; H)$ and that $|A|_w = |A|_{a,\alpha}$ for every $A \in E_w(H)$.

Theorem 4.2.3 and Lemma 4.2.1 now yield the following corollary.

Corollary 4.3.6. *Suppose that* $\dim H = \infty$ *and that* $A, B \in E(a, \alpha; H)$*, with not both of them the zero operator. Then we have*

$$\mathrm{Hdist}(\sigma(A), \sigma(B)) \le m H_{G^E_{a,\alpha}}\left(\frac{\|A - B\|}{m}\right),$$

where

$$G^E_{a,\alpha}(r) = \prod_{k=1}^{\infty}(1 + r\exp(-ak^\alpha))$$

and $m = \max\{|A|_{a,\alpha}, |B|_{a,\alpha}\}$.

Proof. This follows by arguments similar to those used in the previous corollary. □

By [Ban08, Proposition 3.1] we have

$$\log G^E_{a,\alpha}(r) \sim a^{-1/\alpha}\frac{\alpha}{\alpha+1}(\log r)^{1+1/\alpha} \quad \text{as } r \to \infty,$$

and a short calculation using Lemma 3.3.2 shows that

$$\log H_{G^E_{a,\alpha}}(r) \sim -2^{-\alpha/(\alpha+1)} a^{1/(\alpha+1)} \left(\frac{\alpha+1}{\alpha}\right)^{\alpha/(\alpha+1)} |\log r|^{\alpha/(\alpha+1)} \quad \text{as } r \downarrow 0\,.$$

Thus our Corollary 4.3.6 improves the bound obtained in the previous chapter (see Remark 3.4.3) as well as the bound in [Ban08, Theorem 4.2], both derived using the Henrici method.

Using the same method as in the previous two corollaries, we finally provide a bound involving only the first singular value, that is, the operator norm. In order to compare it to Elsner's bound we recall Remark 4.2.4, which gives the following.

Corollary 4.3.7. *Suppose that $A, B \in L(\mathbb{C}^n)$, not both of them the zero matrix. Then we have*

$$\mathrm{Hdist}(\sigma(A) \cup \{0\}, \sigma(B) \cup \{0\}) \leq m H_{G^F}\left(\frac{\|A-B\|}{m}\right),$$

where

$$G^F(r) = (1+r)^n$$

and $m = \max\{\|A\|, \|B\|\}$.

Proof. Follows by arguments similar to those used in the proof of Corollary 4.3.1. □

Corollary 4.3.7 gives a bound very similar to Elsner's bound, which for any $A, B \in L(\mathbb{C}^n)$ can be written in the form

$$\mathrm{Hdist}(\sigma(A), \sigma(B)) \leq m H\left(\frac{\|A-B\|}{m}\right),$$

where $m = \|A\| + \|B\|$ and $H(r) = r^{1/n}$.

Our bound in Corollary 4.3.7 is weaker than Elsner's bound (apart from trivial cases) since

$$H_{G^F}(r) \sim r^{1/(2n+1)} \quad \text{as } r \downarrow 0\,.$$

On the other hand, in the derivation of the above bound only the first singular value is used. We can improve our bound by using more information, for example about the first and the second singular values of the operators involved.

Corollary 4.3.8. *Let $A, B \in L(\mathbb{C}^n)$. Then we have*

$$\operatorname{Hdist}(\sigma(A) \cup \{0\}, \sigma(B) \cup \{0\}) \leq H_{G^F_{s_1,s_2}}(\|A - B\|)\,,$$

where

$$G^F_{s_1,s_2}(r) = (1 + rs_1)(1 + rs_2)^{n-1}\,,$$

with $s_1 = \max\{s_1(A), s_1(B)\}$ and $s_2 = \max\{s_2(A), s_2(B)\}$.

Proof. Follows by arguments similar to those used in the proof of Corollary 4.3.1. □

A short calculation shows that

$$H_{G^F_{s_1,s_2}}(r) \sim s_1^{2/(2n+1)} s_2^{(2n-2)/(2n+1)} r^{1/(2n+1)} \quad \text{as } r \downarrow 0.$$

Unlike the previous corollary our bound in Corollary 4.3.8 is able to compete with Elsner's bound. In order to see this we define weighted shift matrices

$A_\epsilon, B_\epsilon \in L(\mathbb{C}^n)$ by setting

$$A_\epsilon e_k = \begin{cases} e_2 & \text{if } k = 1, \\ \epsilon e_{k+1} & \text{if } 1 < k < n, \\ 0 & \text{if } k = n, \end{cases} \quad B_\epsilon e_k = \begin{cases} e_2 & \text{if } k = 1, \\ \epsilon e_{k+1} & \text{if } 1 < k < n, \\ \epsilon e_1 & \text{if } k = n, \end{cases}$$

where $(e_k)_{k=1}^n$ is an orthonormal basis of $\mathbb{C}^n$. Then

$$\|A_\epsilon - B_\epsilon\| = \epsilon\,,$$

$$s_1(A_\epsilon) = s_1(B_\epsilon) = 1\,,$$

and

$$s_2(A_\epsilon) \le s_2(B_\epsilon) = \epsilon$$

provided that $n > 1$. While Elsner's bound gives

$$H(\|A_\epsilon - B_\epsilon\|) \sim 2^{1-1/n}\epsilon^{1/n} \quad \text{as } \epsilon \downarrow 0\,,$$

our bound in Corollary 4.3.8 yields

$$H_{G^F_{1,\epsilon}}(\|A_\epsilon - B_\epsilon\|) \sim \epsilon^{(2n-1)/(2n+1)} \quad \text{as } \epsilon \downarrow 0\,.$$

Clearly, our bound is better than Elsner's in this case. This is since we have used information about the first and the second singular values of the matrices involved.

Using the preceding discussion it can be seen that upper bounds for the singular values of A and B turn to upper bounds for the function $H_{F_{A,B}}$, which bounds the Hausdorff distance of the spectra of A and B. Furthermore the faster the bounds for the singular values decay to zero, the slower $F_{A,B}$ will grow at infinity therefore the faster $H_{F_{A,B}}$ will tend to zero at zero.

Bibliography

[ALL01] M. Ahues, A. Largillier, B. V. Limaye, *Spectral Computations for Bounded Operators*, Roca Baton, Chapman & Hall/CRC, 2001.

[Ban04] O. F. Bandtlow, *Estimates for norms of resolvents and an application to the perturbation of spectra*, Math. Nachr. **267(1)** (2004), 3–11.

[Ban08] O. F. Bandtlow, *Resolvent estimates for operators belonging to exponential classes*, Integral Equations Operator Theory, **61** (2008), 21–43.

[BG15] O. F. Bandtlow, A. Güven, *Explicit upper bounds for the spectral distance of two trace class operators*, Linear Algebra Appl. **466** (2015), 329–342.

[BF60] F. L. Bauer, C. T. Fike, *Norms and exclusion theorems*, Num. Math. **2** (1960), 42–53.

[BP13] V. Berinde, M. Păcurar, *The role of the Pompeiu-Hausdorff metric in fixed point theory*, Creative Mathematics and Informatics, **22(2)** (2013), 143–150.

[Bha97] R. Bhatia, *Matrix Analysis*, New York, Springer, 1997.

[BF81] R. Bhatia, S. Friedland, *Variation of Grassmann powers and spectra*, Linear Algebra Appl. **40** (1981), 1–18.

[BM79] R. Bhatia, K. K. Mukherjea, *On the rate of change of spectra of operators*, Linear Algebra Appl. **27** (1979), 147–157.

[Car21] T. Carleman, *Zur Theorie der linearen Integralgleichungen*, Math. Z. **9** (1921), 196–217.

[Cha81] F. Chatelin, *The spectral approximation of linear operators with applications to the computation of eigenelements of differential and integral operators*, SIAM Review, **23(4)** (1981), 495–522.

[Cha83] F. Chatelin, *Spectral Approximation of Linear Operators*, New York, Academic Press, 1983.

[CNR00] Y. Chen, A. Nokrane, T. Ransford, *Estimates for the spectrum near algebraic elements*, Linear Algebra Appl. **308** (2000), 153–161.

[Dav07] E. B. Davies, *Linear Operators and their Spectra Vol. 106*, Cambridge, CUP, 2007.

[DP94] L. T. Dechevski, L. E. Persson, *Sharp generalized Carleman inequalities with minimal information about the spectrum*, Math. Nachr. **168** (1994), 61–77.

[DP96] L. T. Dechevski, L. E. Persson, *On sharpness, applications and generalisations of some Carleman type inequalities*, Tôhoku Math. J. **48** (1996), 1–22.

[Dos01] M. R. Dostanić, *Estimates for the norm of powers of Volterra's operator through its singular values*, Math. Z. **236(3)** (2001), 453–459.

[DS63] N. Dunford, J. T. Schwartz, *Linear Operators Vol. 2*, New York, Interscience, 1963.

[EE87] D. E. Edmunds, W. D. Evans, *Spectral Theory and Differential Operators*, Oxford, Clarendon Press, 1987.

[Els85] L. Elsner, *An optimal bound for the spectral variation of two matrices*, Linear Algebra Appl. **71** (1985), 77–80.

[Fan51] K. Fan, *Maximum properties and inequalities for the eigenvalues of completely continuous operators*, Proc. Nat. Acad. Sci. USA **37(11)** (1951), 760–766.

[Gil95] M. I. Gil′, *Norm Estimations for Operator-Valued Functions and Applications*, New York, Marcel Dekker, 1995.

[Gil03] M. I. Gil′, *Operator Functions and Localization of Spectra*, Berlin, Springer, 2003.

[Gil12] M. I. Gil′, *Norm estimates for resolvents of non-selfadjoint operators having Hilbert-Schmidt inverse ones*, Math. Commun. **17** (2012), 599–611.

[Gil14] M. I. Gil′, *Resolvents of operators inverse to Schatten-von Neumann ones*, Ann. Univ. Ferrara Sez. VII Sci. Mat. **60(2)** (2014), 363–376.

[GGK90] I. Gohberg, S. Goldberg, M. A. Kaashoek, *Classes of Linear Operators Vol. 1*, Basel, Birkhäuser Verlag, 1990.

[GGK03] I. Gohberg, S. Goldberg, M. A. Kaashoek, *Basic Classes of Linear Operators*, Basel, Birkhäuser, 2003.

[GK69] I. Gohberg, M. G. Krein, *Introduction to the Theory of Linear Non-Selfadjoint Operators*, Providence, AMS, 1969.

[Hen62] P. Henrici, *Bounds for iterates, inverses, spectral variation and fields of values of non-normal matrices*, Num. Math. **4** (1962), 24–40.

[HP11] D. Hinrichsen, A. J. Pritchard, *Mathematical Systems Theory I: Modelling, State Space Analysis, Stability and Robustness Vol. 48*, Heidelberg, Springer, 2011.

[Kat76] T. Kato, *Perturbation Theory for Linear Operators*, Berlin, Springer, 1976.

[Kre89] E. Kreyszig, *Introductory Functional Analysis with Applications*, New York, Wiley, 1989.

[LR02] Y. Langlois, T. Ransford, *Counterexample to a conjecture of Elsner on the spectral variation of matrices*, Linear Algebra Appl. **349** (2002), 193–195.

[Mir55] L. Mirsky, *An Introduction to Linear Algebra*, Oxford, OUP, 1955.

[Nel82] E. Nelimarkka, *On $\lambda(P,N)$-nuclearity and operator ideals*, Math. Nachr. **99** (1982), 231–237.

[New51] J. D. Newburgh, *The variation of spectra*, Duke Math. J. **18** (1951), 165–176.

[Ost57] A. Ostrowski, *Über die Stetigkeit von charakteristischen Wurzeln in Abhängigkeit von den Matrizenelementen*, Jber. Deutsch. Math. Verein **60** (1957), 40–42.

[Ost73] A. M. Ostrowski, *Solution of Equations in Euclidean and Banach Spaces*, New York-London, Academic Press, 1973.

[Pel85] V. V. Peller, *A description of Hankel operators of the class $\mathfrak{S}_p$ for $p > 0$, an investigation of the rate of rational approximation and other applications*, Math. USSR Sbornik **50(2)** (1985), 465–494.

[Pie80] A. Pietsch, *Operator Ideals*, Amsterdam-New York, North-Holland, 1980.

[Pie86] A. Pietsch, *Eigenvalues and s-Numbers*, Cambridge, CUP, 1986.

[Pok85] A. Pokrzywa, *On continuity of spectra in norm ideals*, Linear Algebra Appl. **69** (1985), 121–130.

[Sha09] E. Shargorodsky, *On the definition of pseudospectra*, Bull. London Math. Soc. **41** (2009), 524–534.

[Sim77] B. Simon, *Notes on infinite determinants of Hilbert space operators*, Adv. Math. **24** (1977), 244–273.

[Tre97] L. N. Trefethen, *Pseudospectra of linear operators*, SIAM Review, **39(3)** (1997), 383–406.

[Wey49] H. Weyl, *Inequalities between the two kinds of eigenvalues of a linear transformation*, Proc. Nat. Acad. Sci. USA **35** (1949), 408–411.