Delft University of Technology
Faculty of Electrical Engineering, Mathematics and Computer Sciences
Delft Institute of Applied Mathematics

A conjecture on the complete boundedness of Schur multipliers.
(Nederlandse titel: Een vermoeden over de complete begrensdheid van Schur multipliers.)

A thesis submitted to the
Delft Institute of Applied Mathematics
in partial fulfillment of the requirements

of the degree of

BACHELOR OF SCIENCE
in
APPLIED MATHEMATICS

by

A.G. Wildschut

Delft, The Netherlands
June 2018

Copyright © 2018 by Guillermo Wildschut. All rights reserved.

BSc Thesis APPLIED MATHEMATICS

A conjecture on the complete boundedness of Schur multipliers.
(Nederlandse titel: Een vermoeden over de complete begrensdheid van Schur multipliers.)

A.G. Wildschut

Delft University of Technology

Supervisor

Dr. M. Caspers

Other committee members

Dr. D. Gijswijt

Dr. B. van den Dries

June, 2018

Delft

Abstract

Schur multipliers are a concept from functional analysis that have various uses in mathematics. In this thesis we provide an introduction of the aforementioned Schur multipliers and the associated Schatten p-classes. We prove a number of results and introduce some concepts of functional analysis in order to get to the central topic: a conjecture by Pisier regarding Schur multipliers. For $p \in \{1, 2, \infty\}$ all Schur bounded multipliers are completely bounded, and the completely bounded norm of a Schur multiplier is in fact equal to its operator norm. On the other hand, for $p \notin \{1, 2, \infty\}$ Pisier conjectures that there exist bounded, but not completely bounded Schur multipliers.
Whereas the first part of the thesis is spent on studying the theoretical nature of the problem, in the second part we perform a number of numerical computations yielding insight into the problem. For a number of random finite-dimensional Schur multipliers and various p we approximated the operator norm using the BFGS minimization algorithm. This resulted in us posing a new conjecture that the completely bounded norm is equal to the norm for any Schur multiplier *for any* $1 \leq p \leq \infty$, i.e. we suggest Pisier's conjecture is false. Finally, we suggest further studies that can be done.

Contents

1

Introduction

The classical spaces ℓ^p and L^p have widely been used in Fourier analysis and harmonic analysis. Given a Hilbert space, using the linear operators on H we can construct a non-commutative analogue to the ℓ^p spaces. We call these the Schatten p-classes, denoted by $S_p(H)$, which will play an important role in this thesis. In the preliminaries chapter we introduce the necessary concepts from functional analysis, before we define and study Schatten classes in Chapter 3.

Following this, we define the Schur multiplier, which is a linear map on a Schatten p-class that "multiplies" an element from $S_p(H)$, i.e. a linear map from H to H, by a matrix. These Schur multipliers have applications in Fourier analysis for instance, and can be linked to Fourier multipliers, see [2], [4], [7], and [10]. Schur multipliers can be used in Quantum Information Theory as well, see [12]. In Chapter 3 we look at finite-dimensional Schur multipliers, in addition to infinite-dimensional Schur multipliers, and show that we can approximate infinite-dimensional Schur multipliers using finite-dimensional Schur multipliers.

Next, we introduce the concept of complete boundedness for Schur multipliers, a property that is relevant in the previously mentioned applications. Strongly related to the complete boundedness of Schur multipliers is the following conjecture posed by Pisier in [10]:

Conjecture. *Let $1 \leq p \leq \infty$ and H be a Hilbert space. Pisier conjectures the following in [10]:*

For $1 < p < \infty$ and $p \neq 2$ there exists a Schur multiplier on $S_p(H)$ that is bounded but not completely bounded.

In Chapter 4 we give a number of equalities and inequalities pertaining to the completely bounded norm and Pisier's conjecture, in the effort to simplify the problem. Moreover, we look at how unitary operators are of use here.

Finally, in Chapter 5 we numerically compute the operator norm of various random Schur multipliers, aided by theoretical results from the previous chapters. We utilize the BFGS minimization algorithm to approximate the operator norm of (amplified) Schur multipliers, with the hope that these results give us more insight. These insights will in turn hopefully allow us to strengthen (or weaken) the aforementioned conjecture, or possibly pose a new conjecture.

2

Preliminaries

In order to study Schur multipliers and Schatten classes, we need to create the necessary foundation in functional analysis, which is the purpose of this chapter. To facilitate this, the reader is assumed to be familiar with the concepts of Hilbert spaces and dual spaces, orthonormal bases, and the underlying notions of these. Throughout this thesis we will assume that any vector space is over $\mathbb{C}$, unless explicitly stated otherwise.

Notation

Below we list some notation commonly used within this thesis.

1. $\mathrm{R}(f)$ denotes the range of a function f.
2. $M_{n\times m}(\mathbb{C})$ denotes the $n \times m$ matrices with complex entries. If $n = m$ we write $M_n(\mathbb{C})$.
3. I_n is the $n \times n$ identity matrix in $M_n(\mathbb{C})$.
4. For matrices we use capital letters, i.e. $A \in M_{n\times m}(\mathbb{C})$. For the entry at position (i, j) we use lower case letters, i.e. $a_{i,j}$.
5. If $A \in M_{nk\times mk}(\mathbb{C})$, we can partition A into $k \times k$ blocks. We indicate a block at position (i, j) (where $1 \leq i \leq n, 1 \leq j \leq m$) with $A_{i,j}$.
6. We denote the adjoint of a bounded linear operator or matrix with *, i.e. T^* and A^*.

Definition 2.1. *Let X, Y be normed vector spaces, and $T : X \to Y$ be a linear transformation. We say that T is bounded if there exists some $M \geq 0$ such that for all $x \in X : \|T(x)\| \leq K\|x\|$. We denote the space of all such functions by $B(X, Y)$. If $X = Y$ then we use $B(X)$ instead.*

Intuitively, this means that growth of $\|T(x)\|$ is bounded by the growth of $\|x\|$. Furthermore, we shall introduce the notation $T(x) = Tx$.

Example. *Let X be any normed vector space, and $I : X \to X$ be the identity. Then for any $\alpha \in \mathbb{C}$ the linear transformation $\alpha I : X \to X$ is bounded, as for all $x \in X : \|(\alpha I)x\| = |\alpha|\|Ix\| = |\alpha|\|x\|$.*

Definition 2.2. *Let X, Y be normed vector spaces, and $T \in B(X, Y)$. We define the operator norm of T as follows:*

$$\|T\|_{op} = \inf\{M \geq 0 : \text{ for all } x \in X : \|T(x)\| \leq M\|x\|\}.$$

Should it be clear that we are taking the operator norm, we may choose to simply write $\|T\|$ to indicate the operator norm of T. Furthermore, as unitary operators are surjective isometries, one finds that the operator norms is invariant under unitary operators, i.e. $\|UT\| = \|TU\| = \|T\|$. Lastly, if Y is a Banach space and we equip $B(X, Y)$ with the operator norm, then $B(X, Y)$ is a Banach space as well.

Remark 2.3. *If $X \neq 0$, then $\|T\| = \sup_{x\neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|\leq 1}\|Tx\| = \sup_{\|x\|=1}\|Tx\|$.*

Generally we will use this last equality if applicable. In the previous example, we have that $\|\alpha I\| = |\alpha|$ which is found easily using $\|T\| = \sup_{\|x\|=1}\|Tx\|$. However, in the finite dimensional case, this supremum is a maximum as well, which will be a result of following lemma.

Lemma 2.4. *Let X be a finite-dimensional normed vector space. Then the unit sphere is compact.*

Proof. We prove it for a normed vector space over $\mathbb{R}$ first. Let $n = \dim X$. By identifying bases of X and $\mathbb{R}^n$ respectively, we find X and $\mathbb{R}^n$ to be isomorphic, where $\mathbb{R}^n$ is equipped with some norm. Consider the unit sphere S in $\mathbb{R}^n$ under said norm. By definition S is bounded, and by the continuity of the norm it follows that it is closed as well. However, as $\mathbb{R}^n$ is finite-dimensional, all norms are equivalent, and thus the open and closed sets are equal for all norms. In particular, S is therefore closed and bounded in $\mathbb{R}^n$ equipped with the standard Euclidian metric, and by the Heine-Borel Theorem we find that S is compact. Consequently, we find that the unit sphere in X is compact as well.

Lastly, if X is a normed vector space over $\mathbb{C}$, we can identify X with $\mathbb{C}^n \cong \mathbb{R}^{2n}$, and conclude that the unit sphere is compact similarly. $\square$

The following theorem gives a useful relation between the continuity of linear operators and their boundedness:

Theorem 2.5. *Let X, Y be normed vector spaces, and $T : X \to Y$ be a linear transformation. Then T is (uniformly) continuous if and only if T is bounded.*

Proof. If $X = 0$, the result is trivial, so we can assume this not to be the case. First, assume that T is continuous. In particular, this means T is continuous in 0. Then, by definition, for every $\varepsilon > 0$ there exists a $\delta > 0$ such that if $\|x\| < \delta$ then $\|Tx\| < \varepsilon$. Choose $\varepsilon = 1$, and set $M = \frac{2}{\delta}$. Now let $x \in X$ arbitrarily. In the case that $x = 0$, we have that $Tx = 0$ by linearity, and thus $\|Tx\| = 0 \leq M\|x\|$. Now if $x \neq 0$, we can consider the element $y = \frac{x}{M\|x\|}$. Since $\|y\| < \delta$, by continuity we have that $\|Ty\| < 1$. But using the linearity of T we find:

$$\|Ty\| < 1 \;\Leftrightarrow\; \|T\frac{x}{M\|x\|}\| < 1 \;\Leftrightarrow\; \|T\| < M\|x\|.$$

Thus, we see that T is bounded.
Conversely, assume that T is bounded. If $T = 0$ then the result is trivial, so assume $T \neq 0$. Let $x \in X$ and $\varepsilon > 0$, and set $\delta = \frac{\varepsilon}{\|T\|}$, where $\|T\| > 0$. Now if for $y \in X : \|x - y\| < \delta = \frac{\varepsilon}{\|T\|}$, since T is bounded and linear we obtain:

$$\|Tx - Ty\| = \|T(x - y)\| \leq \|T\|\|x - y\| < \|T\|\frac{\varepsilon}{\|T\|} = \varepsilon.$$

That is, T is (uniformly) continuous. $\square$

Theorem 2.6. *Let X, Y are normed vector spaces, where X is finite-dimensional. Now suppose $T : X \to Y$ is a linear transformation. Then T is continuous.*

Proof. By Theorem 2.5 is suffices to show that T is bounded. Since on a finite-dimensional space all norms are equivalent, we only need to show that it holds for a single norm. Let $\{b_1, \ldots, b_n\}$ be a basis for X. If we now write $x = \lambda_1 b_1 + \cdots + \lambda_n b_n$, consider the norm given by

$$\|x\|_1 = \sum_{i=1}^{n} |\lambda_i|$$

It is easily checked that this is well-defined, and indeed a norm. Now applying T, using the linearity of T and the triangle inequality, we find:

$$\|Tx\|_1 = \left\| T\left(\sum_{i=1}^{n} \lambda_i b_i\right)\right\|_1 = \left\|\sum_{i=1}^{n} \lambda_i T b_i\right\|_1 \leq \sum_{i=1}^{n} |\lambda_i| \|T b_i\|_1.$$

If we now set $M = \max_{1 \leq i \leq n}\|T(b_i)\|_1$, we find that $\|Tx\|_1 \leq \sum_{i=1}^{n} |\lambda_i| \cdot M = M\|x\|_1$. $\square$

Definition 2.7. *Let X, Y be normed vector spaces, and T a linear operator. T is called a compact operator if for any bounded subset $V \subset X$ the closure of the image, $\overline{T[V]}$, is compact.*

During this thesis we will often look at compact operators because they have numerous useful properties that will become apparent later. We shall denote the space of compact operators from X to Y by $K(X, Y)$ and similarly as for the bounded linear operators we write $K(X) := K(X, X)$. One may note that a compact operator is necessarily bounded (since the image of the unit sphere is bounded), and thus $K(X, Y) \subseteq B(X, Y)$. Moreover, Rynne [11] proves that $K(X, Y)$ is in fact a closed subspace $B(X, Y)$, and that an operator is compact if and only if its adjoint is. Lastly, compositions of compact operators are compact as well.

Example 2.8. *Let $T \in B(X, Y)$ be such that T is of finite rank, and $V \subset X$ be a bounded subset. Then since $T \in B(X, Y)$, $T[V]$ is bounded. Hence, $\overline{T[V]}$ is bounded and closed. Thus, by Bolzano-Weierstrass, we find that T is a compact operator.*

Now for matrices we have one last definition:

Definition 2.9. *Let $A \in M_n(\mathbb{C})$ and consider A^*A. As the latter is self-adjoint, by the Spectral Theorem it can be written as $A^*A = UDU^*$, for U a unitary matrix and D a diagonal matrix with real elements. We now define:*

$$|A| := \sqrt{A^*A} := UD^{1/2}U^*, \quad \textit{where } (d^{1/2})_{i,i} = \sqrt{d_{i,i}}.$$

We note that $|A|^2 = UDU^* = A^*A$, which justifies the notation $\sqrt{A^*A}$. However, we should check if we can indeed take the (non-complex) square root of the of the diagonal elements of D, i.e. the eigenvalues of A^*A. Let λ be an eigenvalue of A^*A and v an associated eigenvector. The following computation show that λ is indeed non-negative:

$$\lambda\|v\|_H^2 = \langle \lambda v, v\rangle_H = \langle A^*Av, v\rangle_H = \langle Av, Av\rangle_H = \|Av\|_H^2, \tag{2.1}$$

where it remains to divide by $\|v\|_H^2$.

Remark 2.10. *$|A|$ is self-adjoint, which follows directly from the definition $|A| := UD^{1/2}U^*$.*

Theorem 2.11. *(Spectral Theorem for compact operators) Let H be a separable Hilbert space and T be a self-adjoint compact operator on H. Then the eigenvectors $(e_n)_{n=1}^{\infty}$ form an orthonormal basis, and we can write:*

$$T = \sum_{n=1}^{\infty} \lambda_n \langle \cdot, e_n\rangle e_n,$$

where λ_n are the (non-zero) eigenvalues and the eigenvalues are real. Moreover, $\lambda_n \to 0$ as $n \to \infty$.

Proof. See Conway [3]. ☐

Remark 2.12. *For any compact operator T on a Hilbert space H we can define $|T| := \sqrt{T^*T}$ analogously to how we defined $|A|$. First, we consider T^*T. As $(T^*T)^* = T^*T$ we find that it is self-adjoint, and as a composition of compact operators it is a compact operator too. We can now diagonalize T^*T by Theorem 2.11, i.e. $T^*T = \sum_{n=1}^{\infty} \lambda_n \langle e_n, \cdot\rangle e_n$. As equation (2.1) holds for T as well, we can define $\sqrt{T^*T} := \sum_{n=1}^{\infty} \sqrt{\lambda_n}\langle e_n, \cdot\rangle e_n$, and from this diagonalization we can see that $|T|$ is self-adjoint. Lastly, we claim that it is a compact operator as well. Indeed, for any bounded sequence $(x_n)_{n\in\mathbb{N}}$ in H, using that $\sqrt{\lambda_n} \to 0$ when $n \to \infty$ one checks that $(|T|x_n)_{n\in\mathbb{N}}$ has a convergent sequence, which in turns implies the definition of a compact operator.*

We have one more important proposition to state, namely the polar decomposition of bounded linear operators.

Proposition 2.13. *(Polar decomposition) Let H be a separable Hilbert space and $T \in B(H)$. Then we can decompose T as $T = US$, with U a linear isometry, and S a self-adjoint operator. Moreover, if H is finite-dimensional, then H is unitary.*

Proof. First of all, set $S := |T|$. We now construct U such that $T = U|T|$. If T is invertible, so will $|T|$ be, and in that case we can simply take $U = T|T|^{-1}$, which proves sufficient. However, in general we cannot assume that T is invertible, and instead we start with creating a map $\widetilde{U_1} : \mathrm{R}(|T|) \to \mathrm{R}(T)$.

Let $v \in \mathrm{R}(|T|)$. Then we can write $v = |T|w$ for some $w \in H$. Now define the well-defined linear map:

$$\widetilde{U_1}v := Aw.$$

We validate that this map is isometric. Let $v, v' \in \mathrm{R}(|T|)$ where $v = |T|w, v' = |T|w'$. It suffices to show that $\widetilde{U_1}$ preserves the inner product. Using the definition of $\widetilde{U_1}$ and that of the adjoint operator, and that $|T|$ is self-adjoint, we have:

$$\begin{aligned} \langle \widetilde{U_1}v, \widetilde{U_1}v' \rangle &= \langle Aw, Aw' \rangle = \langle T^* Aw, w' \rangle \\ &= \langle |A|^2 w, w' \rangle = \langle |T|w, |T|w' \rangle = \langle v, v' \rangle. \end{aligned} \tag{2.2}$$

Thus, we see that $\widetilde{U_1}$ preserves the inner product and is thus an isometry. Furthermore, it follows that $\dim \mathrm{R}(|T|) = \dim \mathrm{R}(T)$.

The question now becomes, how can we extend $\widetilde{U_1}$ to all of H? First of all, by continuity we can extend $\widetilde{U_1}$ to a map $\widetilde{U_1'} : \overline{\mathrm{R}(|T|)} \to \overline{\mathrm{R}(T)}$. As $\overline{\mathrm{R}(T)}$ is a closed linear subspace of H, we can now write $H = \overline{\mathrm{R}(T)} \oplus \overline{\mathrm{R}(T)}^\perp$, where $\overline{\mathrm{R}(T)}^\perp$ is the orthogonal complement of $\overline{\mathrm{R}(T)}$. Similarly, $H = \overline{\mathrm{R}(|T|)} \oplus \overline{\mathrm{R}(|T|)}^\perp$. We would now like to construct an linear isometry $\widetilde{U_2'} : \overline{\mathrm{R}(|T|)}^\perp \to \overline{\mathrm{R}(T)}^\perp$. Since $\dim \overline{\mathrm{R}(|T|)} = \dim \overline{\mathrm{R}(T)}$ the same holds for the orthogonal complements. Let $k = \dim \mathrm{R}(|T|)^\perp = \dim \mathrm{R}(T)^\perp$ (which is possibly infinite). Now let $\{b_i\}_{i=1}^k$ and $\{b_i'\}_{i=1}^k$ be **orthonormal** bases for $\overline{\mathrm{R}(|T|)}^\perp$ and $\overline{\mathrm{R}(T)}^\perp$ respectively. If we now identify the bases by $\widetilde{U_2'}$, i.e. $\widetilde{U_2} b_i = b_i'$, by the orthonormality of the bases we have that $\widetilde{U_2'}$ preserves the inner product. Subsequently, we define $U : H \to H$. Any $v \in H$ we can uniquely write as $v = w_1 \oplus w_2$, with $w_1 \in \overline{\mathrm{R}(|T|)}$ and $w_2 \in \overline{\mathrm{R}(|T|)}^\perp$. We now define $Uv := \widetilde{U_1'} w_1 \oplus \widetilde{U_2'} w_2$, which is again again preserves the inner product. Although U need not be surjective for infinite-dimensional H, should H be finite-dimensional it follows from the rank–nullity theorem that U is in fact surjective, and hence unitary. Lastly, note that by definition of $\widetilde{U_1'}$ we have that $T = US = U|T|$. □

3

Schatten classes and Schur multipliers

3.1. The Schatten p-norm

In this section we define the so-called Schatten p-classes, a subspace of the compact linear operators on a Hilbert space H, equipped with Schatten p-norm. However, before we can give this norm explicitly we must define the singular values:

Definition 3.1. *Let H, H' be Hilbert spaces, and $T : H \to H'$ a compact bounded linear operator. We define the singular values of T for $n \in \mathbb{N}$ as follows:*

$$s_n(T) := \inf\{\|T - S\| \mid S \in K(H, H') \text{ and } \operatorname{rank}(S) < n\}.$$

In the case that $n = 1$ then $s_1(T) = \|T\|$, and if $n > \operatorname{rank}(T)$ then $s_n(T) = 0$. Furthermore, since $\{\|T-S\| \mid \operatorname{rank}(S) < n\} \subset \{\|T-S\| \mid \operatorname{rank}(S) < n+1\}$, we have that $s_n(T) \geq s_{n+1}(T)$, i.e. $(s_n(T))_{n=1}^{\infty}$ is a (non-negative) decreasing sequence. Before we define our Schatten-p-norm, we take a look at the following lemma, which will proven useful later:

Lemma 3.2. *Let H be a non-trivial finite-dimensional Hilbert space and let $T : H \to H$ be a (compact) bounded linear operator. Then if we represent T by the matrix $A \in M_n(\mathbb{C})$ we have that the singular values of T coincide with the eigenvalues of $|A|$.*

Proof. Let $n = \dim H$. After choosing a basis we can represent T with a matrix A, which we from now on shall identify with T. By the polar decomposition we can decompose it as follows: $A = U|A|$, for U some unitary matrix. With U a unitary matrix, it is easily validated that it does not change the singular values, and so the singular values of A and $|A|$ are equal. However, since $|A| = VDV^*$ for V a unitary matrix the singular values of $|A|$ are in turn equal to the (non-negative) diagonal elements of D (the square roots of the eigenvalues of A^*A). Lastly, the singular values of D are simply the diagonal elements. □

We are now ready to define the **Schatten classes**.

Definition 3.3. *Let $p \in [1, \infty]$ and H be a separable Hilbert space. For $p < \infty$ we define the Schatten class $S_p(H)$ as follows:*

$$S_p(H) := \{T \in K(H) \mid \sum_{k=1}^{\infty} s_k(T)^p < \infty\}.$$

Lastly, $S_\infty(H) := K(H)$.

Remark 3.4. *For any finite-dimensional H we have that $S_p(H) = K(H) = B(H)$, where the first equality follow from the sum in the definition of $S_p(H)$ being finite, and the second equality follows from Example 2.8.*

Remark 3.5. *Whenever $p < \infty$ from the definition of $S_p(H)$ it follows that for $T \in S_p(H) : s_n(T) \to 0$ as $n \to \infty$. In fact, as Rynne [11] proves, this is true in general for any $T \in K(H)$, and thus for $S_\infty(H)$ as well.*

Definition 3.6. *Let $p \in [1, \infty]$ and H be a separable Hilbert space. For $T \in S_p(H)$ and $p < \infty$ we define the **Schatten p-norm** as*

$$\|T\|_p := \left(\sum_{k=1}^{\infty} s_k(T)^p \right)^{1/p} .$$

Analogously to the ℓ^p norm, if $p = \infty$ then we define $\|A\|_\infty = \sup_k s_k(T) = s_1(T) = \|T\|$, whereas the second equality holds as $(s_k(T))_{k=1}^{\infty}$ is a decreasing sequence.

Remark 3.7. *If we see it as the ℓ^p norm of $(s_n(T))_{n=1}^{\infty} \in \ell^p$ we see that for any $p' \geq p : T \in S_{p'}(H)$ and that $\|T\|_{p'} \leq \|T\|_p$.*

Furthermore, any linear operator on $S_p(H)$ we call a Schatten class operator. We are interested in showing that the above defined $\|\cdot\|_p$ is indeed a norm on $S_p(H)$ (simultaneously showing that $S_p(H)$ is a vector space). If $p = \infty$ then $\|\cdot\|_p$ is simply the operator norm and we are done. In the case that $1 \leq p < \infty$, whilst some properties of the norm are easily validated it takes a considerable amount of work to show that it satisfies the triangle inequality. Before we do so, we take a look at another way of computing the Schatten p-norm, and how $\|\cdot\|_2$ is derived from an inner product for finite-dimensional H.

Proposition 3.8. *Let H be a finite-dimensional Hilbert space, $p \in \mathbb{N}$ be even and $T \in S_p(H)$. If we represent T by the matrix $A \in M_n(\mathbb{C})$ with respect to some basis, then $\|T\|_p^p = tr(|A|^p)$. In particular, if p is even then $\|T\|_p^p = tr((A^*A)^{p/2})$ and we have a significantly easier way to compute $\|T\|_p$.*

Proof. As we have seen earlier, we can decompose $|A|$ into UDU^* with U unitary and D diagonal, whereas the diagonal elements of D are the eigenvalues of $|A|$.

By Lemma 3.2 we know these to be the singular values of A as well. We now define the p-th power of $|A|$ as $|A|^p := UD^pU^*$, where D^p in turn is simply the non-negative diagonal elements exponentiated to the power p (where p may not be an integer). We note that since $U^*U = I$, should p be a natural number then this definition coincides with the natural exponentiation. Now, by this definition the eigenvalues of $|A|^p$ are the singular values of A to the power p, and thus:

$$\mathrm{tr}(|A|^p) = \mathrm{tr}(UD^pU^*) = \mathrm{tr}(D^p) = \sum_{k=1}^{n} s_k(T)^p = \|T\|_p^p.$$

Here we use that the trace is independent of the chosen basis. However, should p be an even number, then $(UD^2U^*)^{p/2} = UD^pU^*$, which yields $\mathrm{tr}((A^*A)^{p/2}) = \mathrm{tr}((UD^2U)^{p/2}) = \mathrm{tr}(UD^pU^*) = \|T\|_p^p$. □

Whenever H is finite-dimensional (say of dimension n), by Remark 3.4 we have that $S_p(H) = B(H)$. If we now choose a basis for H, we can uniquely identify each element from $B(H)$ with a matrix $M_n(\mathbb{C})$ and vice versa. Subsequently, for a finite-dimensional H we can consider the Schatten p-norm as a norm on the (complex) $n \times n$ matrices. Keeping this in mind we now look at $\|\cdot\|_2$ as a norm on $M_n(\mathbb{C})$. For $A \in M_n(\mathbb{C})$ by the previous proposition we have that $\|A\|_2^2 = \mathrm{tr}(A^*A)$. In this case it is called the *Frobenius norm* as well, and is induced by an inner product, namely the *Frobenius inner product*:

Definition 3.9. *We define the Frobenius inner product $\langle \cdot, \cdot \rangle_F : M_n(\mathbb{C}) \times M_n(\mathbb{C}) \to \mathbb{C}$ as follows:*

$$\langle A, B \rangle_F = tr(B^*A),$$

where B^* is the adjoint matrix, i.e. the conjugate transpose. Here we use the usual matrix multiplication. We verify that it is an inner product:

Proof. Let $A, B, C \in M_n(\mathbb{C})$ and $\alpha, \beta \in \mathbb{C}$. Now:

(i) For $i = 1, \ldots, n$ we have that $(A^*A)_{i,i} = \sum_{j=1}^{n} A^*_{i,j} A_{j,i} = \sum_{j=1}^{n} \overline{A}_{j,i} A_{j,i} = \sum_{j=1}^{n} |A_{j,i}|^2$. As a sum of non-negative real numbers, we find that $\langle A, A \rangle_\mathrm{F} = \mathrm{tr}(A^*A)$ is real and non-negative as well.

(ii) If $A = 0$ then clearly $\langle A, A\rangle_{\mathrm{F}} = 0$. Now if $\langle A, A\rangle_{\mathrm{F}} = 0$, and since all diagonal elements of A^*A are positive, by (i) we must have that $|A_{i,j}|^2 = 0$ for all i, j, and thus for all $i, j : A_{i,j} = 0 \Rightarrow A = 0$.

(iii) We have: $\langle \alpha A + \beta B, C\rangle_{\mathrm{F}} = \mathrm{tr}(C^*(\alpha A + \beta B)) = \alpha\mathrm{tr}(C^*A) + \beta\mathrm{tr}(C^*B) = \alpha\langle A, C\rangle_{\mathrm{F}} + \beta\langle B, C\rangle_{\mathrm{F}}$.

(iv) $\langle A, B\rangle_{\mathrm{F}} = \mathrm{tr}(B^*A) = \mathrm{tr}(\overline{B^\intercal} A) = \mathrm{tr}((\overline{B^\intercal} A)^\intercal) = \mathrm{tr}(A^\intercal \overline{B}) = \overline{\mathrm{tr}(\overline{A^\intercal} B)} = \overline{\langle B, A\rangle_{\mathrm{F}}}$ since $\mathrm{tr}(A^\intercal) = Tr(A)$ for any matrix. □

Thus, this proves that in the case $p = 2$ and finite-dimensional H that $\|\cdot\|_p$ is a norm. In particular, we have that $(M_n(\mathbb{C}), \langle\cdot,\cdot\rangle_{\mathrm{F}})$ is a Hilbert space, as any finite dimensional inner product space is complete, see [11].

At last, we will now show for any $1 \leq p < \infty$ that $\|\cdot\|_p$ satisfies the triangle inequality on $S_p(H)$ for any Hilbert space H. **Within this proof we shall assume H is a separable infinite-dimensional Hilbert space**, as for a finite-dimensional H the intermediate results follow similarly (and usually easier). To achieve this we first set up the necessary definitions and lemmas. The following proofs are based on [1] and [5].

Recall from the preliminaries that we can decompose self-adjoint compact operators as $\sum_{n=1}^{\infty} \lambda_n \langle e_n, \cdot\rangle e_n$. Elements from $S_p(H)$ we can decompose as well, but as they are not necessarily self-adjoint we instead turn to the *singular value decomposition*, a somewhat similar decomposition.

Corollary 3.10. *If T is a compact operator on H (not necessarily self-adjoint), then we can write it as*

$$T = \sum_{n=1}^{\infty} s_n(T)\langle\cdot, e_n\rangle f_n,$$

for $(e_n)_{n=1}^{\infty}, (f_n)_{n=1}^{\infty}$ orthonormal sequences in H. We call this the singular value decomposition.

Proof. We start by considering the polar decomposition of T, i.e. $T = U|T|$. By Theorem 2.11 we can write $|T| = \sum_{n=1}^{\infty} \mu_n\langle\cdot, e_n\rangle e_n$, where $(\mu_n)_{n=1}^{\infty}$ are the eigenvalues of $|T|$, $\mu_n \to 0$ as $n \to \infty$ and $(e_n)_{n=1}^{\infty}$ are the associated eigenvectors, which form an orthonormal basis. Furthermore, we know that the eigenvalues of $|T|$ are the singular values of T, so we have that $|T| = \sum_{n=1}^{\infty} s_n(T)\langle\cdot, e_n\rangle e_n$.

Now for $n \in \mathbb{N}$ define $f_n := Ue_n$. As we know that U preserves the inner product, $(f_n)_{n=1}^{\infty}$ forms an orthonormal sequence. We now define $|T|_N := \sum_{n=1}^{N} s_n(T)\langle\cdot, e_n\rangle e_n$ and $T_N := U\sum_{n=1}^{N} s_n(T)\langle\cdot, e_n\rangle e_n = \sum_{n=1}^{N} s_n(T)\langle\cdot, e_n\rangle f_n$. Lastly, we wish that $T_N \to T$, i.e. $\lim_{N\to\infty}\|T - T_N\| = 0$. Indeed, we have:

$$\|T - T_N\| = \sup_{n\geq N} s_n(T) \to 0 \text{ as } N \to \infty.$$

□

Lemma 3.11. *Let T be a compact operator. Then for $n \in \mathbb{N}$:*

$$\sum_{k=1}^{n} s_k(T) = \max\left|\sum_{k=1}^{n}\langle Tg_k, h_k\rangle\right| = \max\sum_{k=1}^{n}|\langle Tg_k, h_k\rangle|,$$

where we take the maxima over all orthonomal sequences $(g_k)_{k=1}^{n}$ and $(h_k)_{k=1}^{n}$ in H.

Proof. We start with considering the supremum, and taking the svd of T, i.e. $T = \sum_{n=1}^{\infty} s_n(T)\langle\cdot, e_n\rangle f_n$. For the second sum we now take the orthonormal sequences $(e_k)_{k=1}^{n}$ and $(f_k)_{k=1}^{n}$ and fill in T to get:

$$\sum_{k=1}^{n}\langle T_n e_k, f_k\rangle = \sum_{k=1}^{n}\langle\sum_{j=1}^{\infty} s_j(T)\langle e_j, e_k\rangle f_j, f_k\rangle = \sum_{k=1}^{n}\langle s_k(T) f_k, f_k\rangle = \sum_{k=1}^{n} s_k(T).$$

If we now take the absolute value and the supremum over all orthonormal sequences as described, and apply the triangle inequality on the sum we then obtain the following inequalities:

$$\sum_{k=1}^{n} s_k(T) \leq \sup\left|\sum_{k=1}^{n}\langle Tg_k, h_k\rangle\right| \leq \sup\sum_{k=1}^{n}|\langle Tg_k, h_k\rangle|. \tag{3.1}$$

We now wish to show that $\sup\sum_{k=1}^n |\langle Tg_k, h_k\rangle| \le \sum_{k=1}^n s_k(T)$ holds as well. Substituting the svd of T in $\sum_{k=1}^n |\langle Tg_k, h_k\rangle|$ we obtain:

$$\sum_{k=1}^n |\langle Tg_j, h_k\rangle| = \sum_{k=1}^n \left|\sum_{j=1}^\infty s_j(T)\langle g_j, e_k\rangle\langle f_k, h_j\rangle\right| \le \sum_{k=1}^\infty s_j(T)\sum_{j=1}^n |\langle g_j, e_k\rangle\langle f_k, h_j\rangle|. \tag{3.2}$$

We first look at $\sum_{j=1}^n |\langle g_j, e_k\rangle\langle f_k, h_j\rangle|$, which we note is non-negative. As the sums are (absolutely) convergent this justifies the operations we perform. If we consider it the (standard) inner product of $(|\langle g_1, e_k\rangle|, \dots, |\langle g_n, e_k\rangle|)$ and $(|\langle f_k, h_1\rangle|, \dots, |\langle f_k, h_n\rangle|)$, and we can use the Cauchy-Schwarz inequality to obtain:

$$\sum_{j=1}^n |\langle g_j, e_k\rangle\langle f_k, h_j\rangle| \overset{\text{C.S.}}{\le} \left(\sum_{j=1}^n |\langle g_j, e_k\rangle|^2\right)^{1/2} \left(\sum_{j=1}^n |\langle f_k, h_j\rangle|^2\right)^{1/2} \le \|e_k\|\|f\| = 1.$$

Applying the same trick after switching the order of summation in $\sum_{k=1}^\infty\sum_{j=1}^n |\langle e_k, g_j\rangle\langle h_j, f_k\rangle|$ (as one sum is finite), we then find

$$\begin{aligned}\sum_{k=1}^\infty\sum_{j=1}^n |\langle g_j, e_k\rangle\langle f_k, h_j\rangle| &= \sum_{j=1}^n\sum_{k=1}^\infty |\langle g_j, e_k\rangle\langle f_k, h_j\rangle| \\ &\overset{\text{C.S.}}{\le} \sum_{j=1}^n \left[\left(\sum_{k=1}^\infty |\langle g_j, e_k\rangle|^2\right)^{1/2} \left(\sum_{k=1}^\infty |\langle f_j, h_j\rangle|^2\right)^{1/2}\right] \\ &= \sum_{j=1}^n \|g_j\|\|h_j\| = \sum_{j=1}^n 1 = n.\end{aligned}$$

For the sake of convenience we shall define $c_k := \sum_{j=1}^n |\langle g_j, e_k\rangle\langle f_k, h_j\rangle|$. So far we have thus shown that that $0 \le c_k \le 1$ and $\sum_{k=1}^\infty c_k \le n$. We now write:

$$\sum_{k=1}^\infty s_k(T)c_k = \sum_{k=1}^\infty\sum_{j=k}^\infty (s_j(T) - s_{j+1}(T))c_k = \sum_{j=1}^\infty\sum_{k=1}^j (s_j(T) - s_{j+1}(T))c_k,$$

which we can do as $s_j(T) \to 0$ as $j \to \infty$. The second equality is a result of enumerating the sum differently. Thus, we find that $\sum_{k=1}^\infty s_k(T)c_k = \sum_{j=1}^\infty (s_j(T) - s_{j+1}(T))\sum_{k=1}^j c_k$. But since $c_k \ge 0$ we have that $\sum_{k=1}^j c_k \le \sum_{k=1}^\infty c_k \le n$. Combined with the fact that $c_k \le 1$ we find that $\sum_{k=1}^j c_k \le \sum_{k=1}^{\min\{j,n\}} 1$. Now since $s_j(T) - s_{j+1}(T) \ge 0$, once more enumerating the sum differently we find and using that $s_j(T) \to 0$ as $j \to \infty$ we obtain:

$$\sum_{k=1}^\infty s_k(T)c_k \le \sum_{j=1}^\infty (s_j(T) - s_{j+1}(T)) \sum_{k=1}^{\min\{j,n\}} 1 = \sum_{k=1}^n\sum_{j=k}^\infty (s_j(T) - s_{j+1}(T)) = \sum_{k=1}^n s_k(T).$$

Now taking the supremum over all orthonormal sequences, the above inequality combined with (3.1) and (3.2) yields the desired inequality. Lastly, as taking $(e_k)_{k=1}^n$ and $(f_k)_{k=1}^n$ as the orthonormal sequences gave us inequality (3.1), we proves that it is in fact a maximum. □

Lemma 3.12. *For compact operators T and S and $n \in \mathbb{N}$ the following holds:*

$$\sum_{k=1}^n s_k(T+S) \le \sum_{k=1}^n s_k(T) + \sum_{k=1}^n s_k(S).$$

Proof. Using Lemma 3.11 and the triangle inequality for $|\cdot|$ we find:

$$\sum_{k=1}^n s_k(T+S) = \max\sum_{j=1}^n |\langle (T+S)g_j, h_j\rangle| \le \max\sum_{j=1}^n |\langle Tg_j, h_j\rangle| + \max\sum_{j=1}^n |\langle Sg_j, h_j\rangle|.$$

Once more using Lemma 3.11 on the last two terms yields the desired inequality. □

Lemma 3.13. *Let $f : \mathbb{R} \to \mathbb{R}$ be a convex increasing function, $n \in \mathbb{N}$ and $(a_i)_{i=1}^n, (b_i)_{i=1}^n$ be two decreasing sequences, and define the partials sums $A_k := \sum_{i=1}^k a_i$ and $B_k := \sum_{i=1}^k b_i$. If for all $k \in \{1, \dots, n\}$ we have that $A_k \leq B_k$*[1]*, then the following inequality holds:*

$$\sum_{i=1}^n f(a_i) \leq \sum_{i=1}^n f(b_i).$$

Proof. We wish to prove that $\sum_{i=1}^n (f(b_i) - f(a_i)) \geq 0$. In the case that $a_k = b_k$ for some k, the inequality resulting from the sequences with a_k and b_k removed is equivalent with the one we wish to prove. Thus, without loss of generality we can assume that $a_i \neq b_i$ for all i. But now we can define $c_i = \frac{f(b_i)-f(a_i)}{b_i-a_i}$, and then we can write

$$\sum_{i=1}^n (f(b_i) - f(a_i)) = \sum_{i=1}^n c_i(b_i - a_i).$$

As f is increasing, if $b_i > a_i$ then $f(b_i) \geq f(a_i)$ and thus $c_i \geq 0$. Likewise, if $b_i < a_i$ then $c_i \geq 0$. We now claim that $(c_i)_{i=1}^n$ is a decreasing sequence as well. Defining $R(x, y) := \frac{f(x)-f(y)}{x-y}$ for $x \neq y$, we have that f is convex if and only if R is increasing in both variables . Note that $c_i = R(a_i, b_j)$, and in the case that a_i, a_{i+1}, b_i and b_{i+1} are all different we then find that $c_i = R(a_i, b_j) \geq R(a_{i+1}, b_i) \geq R(a_{i+1}, b_{i+1}) = c_{i+1}$. In the case that two are equal, then it follows from a single step and the symmetry of R in its arguments. Now using that $a_i = A_i - A_{i-1}$ and $b_i = B_i - B_{i-1}$, we obtain the following equality:

$$\begin{aligned} \sum_{i=1}^n (f(b_i) - f(a_i)) &= \sum_{i=1}^n c_i b_i - c_i a_i = \sum_{i=1}^n c_i(B_i - B_{i-1}) - \sum_{i=1}^n c_i(A_i - A_{i-1}) \\ &= \sum_{i=1}^n c_i(B_i - A_i) - \sum_{i=1}^n c_i(B_{i-1} - A_{i-1}) \end{aligned} \tag{3.3}$$

However, since $A_0 = B_0 = 0$, we have that

$$\sum_{i=1}^n c_i(B_{i-1} - A_{i-1}) = \sum_{i=0}^{n-1} c_{i+1}(B_i - A_i) = \sum_{i=1}^{n-1} c_{i+1}(B_i - A_i).$$

Using this equality in (3.3) this yields:

$$\sum_{i=1}^n (f(b_i) - f(a_i)) = c_n(B_n - A_n) + \sum_{i=1}^{n-1} (c_i - c_{i+1})(B_i - A_i).$$

Since $(c_i)_{i=1}^n$ is decreasing, we have that $c_i - c_{i+1} \geq 0$, and we have already seen that $c_i \geq 0$ as well. But by assumption we have that $B_i - A_i \geq 0$ for any i, and thus we find that all terms are non-negative in the last expression, and thus $\sum_{i=1}^n (f(b_i) - f(a_i)) \geq 0$ as desired. □

We can now finally prove the triangle inequality for Schatten p-norm. Let $T, S \in S_p(H)$. Firstly, we apply lemma 3.12 to find that $\sum_{k=1}^n s_n(T + S) \leq \sum_{k=1}^n (s_n(T)) + s_n(S))$. We now consider the function $x \mapsto x^{1/p}$, which we note is convex and increasing for any $p \geq 1$. Now applying lemma 3.13 with $a_k = s_k(T + S)$ and $b_k = s_k(T) + s_k(S)$ and the aforementioned $x \mapsto x^p$ for any $n \in \mathbb{N}$ we find:

$$\sum_{k=1}^n (s_k(T + S))^p \leq \sum_{k=1}^n (s_k(T) + s_k(S))^p. \tag{3.4}$$

Since $x \mapsto x^{1/p}$ is increasing, we can apply this to both sides and still maintain the inequality. After raising it to the power $\frac{1}{p}$, the right sum is simply the usual p-norm for the vector $x = (s_1(T) + s_1(S), \dots, s_n(T) + s_n(S))$ in $\mathbb{C}^n$, for which we know the triangle inequality to be true. That is, we find that

$$\left(\sum_{k=1}^n (s_k(T + S))^p \right)^{1/p} \leq \left(\sum_{k=1}^n (s_k(T) + s_k(S))^p \right)^{1/p} \leq \left(\sum_{k=1}^n s_k(T)^p \right)^{1/p} + \left(\sum_{k=1}^n s_k(S)^p \right)^{1/p}. \tag{3.5}$$

[1]If we set $a = (a_1, \dots, a_n)$ and $b = (b_1, \dots, b_n)$ then it is said that b weakly majorizes a.

Lastly, by definition of $S_p(H)$ we know that $\sum_{k=1}^n s_k(T)^p < \infty$ and $\sum_{k=1}^n s_k(S)^p < \infty$. Letting $n \to \infty$ we find

$$\|T+S\|_p = \left(\sum_{k=1}^{\infty}(s_k(T+S))^p\right)^{1/p} \leq \left(\sum_{k=1}^{\infty} s_k(T)^p\right)^{1/p} + \left(\sum_{k=1}^{\infty} s_k(S)^p\right)^{1/p} = \|T\|_p + \|S\|_p.$$

□

Thus, we have shown that $\|\cdot\|_p$ satisfies the triangle inequality, and thus $S_p(H)$ is in fact a normed space. As is proven in [5], $S_p(H)$ is a Banach space as well. We will now define the Schur product for two matrices, which is essentially the entrywise product. We first consider the finite-dimensional case.

Definition 3.14. *Let $A, B \in M_n(\mathbb{C})$. We define the Schur product of A and B as an $n \times n$ matrix $A \circ B$ where*

$$A \circ B := (a_{i,j} b_{i,j})_{i,j=1}^n$$

That is, $A \circ B$ is the entry-wise multiplication of A and B.

Due to the entry-wise nature of Schur multiplication, it is easy to see that it is associative, commutative, and that it is distributive with regards to the usual addition of matrices. Furthermore, if J_n is the n matrix filled entirely with ones, we note that for any matrix $A \in M_n(\mathbb{C}) : A \circ J_n = A$. That is, J_n is identity element for Schur multiplication. Similarly, we can consider the Schur product for infinite matrices, i.e. for $A = (a_{i,j})_{i,j=1}^{\infty}$ and $B = (b_{i,j})_{i,j=1}^{\infty} : A \circ B = (a_{i,j} b_{i,j})_{i,j=1}^{\infty}$.

Example. *Let* $A = \begin{pmatrix} 1 & 1 \\ 2 & 4 \end{pmatrix}$ *and* $B = \begin{pmatrix} 2 & 5 \\ -3 & 8 \end{pmatrix}$. *Then* $A \circ B = \begin{pmatrix} 1 \cdot 2 & 1 \cdot 5 \\ 2 \cdot -3 & 4 \cdot 8 \end{pmatrix} = \begin{pmatrix} 2 & 5 \\ -6 & 32 \end{pmatrix}$.

We now define a Schur multiplier:

Definition 3.15. *Let H be a separable Hilbert space of dimension n (possibly infinite), and $A = (a_{i,j})_{i,j=1}^n$ a $n \times n$-matrix, and let $1 \leq p \leq \infty$. Now given a fixed orthonormal basis and an element $S_p(H)$, we can represent the element by a unique (possibly infinite) matrix, say $B = (b_{i,j})_{i,j=1}^n$. Subsequently, we can consider $A \circ B$ as a map from H to H. Now let us define $M_A : S_p(H) \to S_p(H)$ by $M_A(B) := A \circ B$, or in other words $M_A((b_{i,j})_{i.j=1}^n) := (a_{i,j} b_{i,j})_{i,j=1}^n$. Should M_A be well-defined, then we call M_A a* ***Schur multiplier****.*

If M_A is bounded, we can consider the operator norm, and to make clear that it is the operator norm with respect to the Schatten p-norm, we shall denote it with $\|M_A\|_p$.

Should H be finite-dimensional, we know that $S_p(H) = B(H) \cong M_n(\mathbb{C})$ by Remark 3.4, and consequently M_A is well defined for any A, and thus a Schur multiplier. Moreover, from Theorem 2.6 it follows that M_A itself is a bounded map. One should be aware that if H is infinite-dimensional, M_A will not be a Schur multiplier for every A. However, should it be a Schur multiplier, then one can easily verify that the entries of A must be bounded. Nevertheless, we claim that any Schur multiplier is bounded as a result of the closed graph theorem. Before we show this, we must first define the graph of a linear operator:

Definition 3.16. *Let X, Y be normed spaces and let $T : X \to Y$ be a linear map. We define the graph of T as:*

$$\mathcal{G}(T) := \{(x, Tx) \mid x \in X\}.$$

We note that $\mathcal{G}$ is a linear subspace of $X \times Y$. Moreover, we can equip $X \times Y$ with the norm $\|(x,y)\|_{X \times Y} := \|x\|_X + \|y\|_Y$, which then forms a Banach space. We are now ready to state the closed graph theorem:

Theorem 3.17. *(Closed graph theorem) Let X, Y be Banach spaces, and let $T : X \to Y$ be a linear map. Then T is continuous if and only if $\mathcal{G}(T)$ is closed.*

Proof. See Rynne [11]. □

Thus, to show that a linear operator between normed spaces is bounded, by 2.5 and the closed graph theorem it suffices to show that its graph is closed. Recall that in a metric space a set is closed if it contains the limit of convergent sequences within the set.

Proposition 3.18. *Let H be a separable Hilbert space and M_A a Schur multiplier on $S_p(H)$. Then M_A is bounded.*

Proof. By the closed graph theorem it suffices to show that $\mathcal{G}(M_A)$ is closed in $S_p(H) \times S_p(H)$.

Let $(B_n, M_A(B_n))$ be a convergent sequence in $\mathcal{G}(M_A)$ with limit (B, y). That is to say, $\lim_{n\to\infty}\|B - B_n\|_p = 0$ and $\lim_{n\to\infty}\|M_A(B_n) - y\|_p = 0$. We need to show that $y = M_A(B)$, i.e. $(B, y) \in \mathcal{G}(M_A)$. Since the entries of M_A are bounded, and convergence in the Schatten p-norm implies convergence in the operator norm, for any basis element e_i we find that $\|(M_A(B))(e_i) - (M_A(B_n))(e_i)\| = \|(M_A(B - B_n))(e_i)\| \to 0$ as $n \to \infty$. Although this in itself would not directly be enough to imply convergence in S_p in general, as we have assumed that $M_A(B_n)$ converges to something, namely y, it follows that y must be equal to $M_A(B)$. ☐

3.2. Approximating a Schur multiplier

Given a Schur multiplier as in the previous section, we are interested in its norm. As we will see in Chapter 5, whenever the underlying Hilbert space H is finite-dimensional, we can numerically approximate $\|M_A\|_p$. In this section we will show that we can approximate $\|M_A\|_p$ whenever H is infinite-dimensional with finite dimensional Schur multipliers, obtained using finite rank projections.

Throughout this section we shall assume H to be a separable infinite-dimensional Hilbert space, and $(e_i)_{i=1}^{\infty}$ to be a **fixed** orthonormal basis of H. Any elements from $S_p(H)$ we represent as a matrix with respect to the aforementioned orthonormal basis.

Definition 3.19. *For $l \in \mathbb{N}$ we define $p_l : H \to H$ as the projection on* $\operatorname{Span}\{e_1, \ldots, e_l\}$.

As $(e_i)_{i=1}^{\infty}$ is an orthonormal basis, we have that for every $x \in H : \|x - p_l x\| \to 0$ as $l \to \infty$, see [11]. Note that we can also write $x - p_l x = (I - p_l)x$, where I is the identity operator. Moreover, as $p_l \in B(H)$ as a finite rank linear map it follows that $I - p_l \in B(H)$ A similar limit can be found for compact operators, but before doing so we consider the following lemma for *arbitrary* finite rank projections. As before, by Theorem 2.6 any finite rank projection is necessarily bounded, i.e. an element of $B(H)$.

Lemma 3.20. *Let $q \in B(H)$ be a finite rank projection. Then*

$$\|q - qp_l\| = \|q(I - p_l)\| \to 0 \text{ as } l \to \infty.$$

Proof. We first consider the finite-dimensional subspace $qH := \{qx \mid x \in H\}$. Since the case $q = 0$ is trivial, we will assume $q \neq 0$. We first prove it for a projection q of rank 1. Thus, we can take $b \in qH$ such that q is the projection on H, and $\|b\| = 1$. First of all we use our orthonormal basis $(e_i)_{i=1}^{\infty}$ and write $b = \sum_{i=1}^{\infty}\langle b, e_i\rangle e_i$, where $1 = \|b\|^2 = \sum_{i=1}^{\infty}|\langle b, e_i\rangle|^2$. Subsequently we have that $(I - p_l)b = \sum_{i=l}^{\infty}\langle b, e_i\rangle e_i$, and we see that

$$\lim_{l\to\infty}\|(I - p_l)b\|^2 = \lim_{l\to\infty}\sum_{i=l+1}^{\infty}|\langle b, e_i\rangle|^2 = 0,$$

and thus $\lim_{l\to\infty}\|(I - p_l)b\| = 0$. Now let $x \in H$ be such that $\|x\| = 1$, and write $x = \sum_{i=1}^{\infty}\langle x, e_i\rangle e_i$. Furthermore, we have that $qx = \langle x, b\rangle b$. Combining the two previous statements and using the Cauchy-Schwarz inequality we find:

$$\begin{aligned}\|q(I - p_l)x\|^2 &= |\langle (I - p_l)x, b\rangle|^2 = \left|\sum_{l+1=1}^{\infty}\langle x, e_i\rangle\langle e_i, b\rangle\right|^2 \overset{\text{C.S.}}{\leq} \left(\sum_{i=l+1}^{\infty}|\langle x, e_i\rangle|^2\right)\left(\sum_{i=l+1}^{\infty}|\langle e_i, b\rangle|^2\right) \\ &\leq \|x\|^2\|(I - p_l)b\|^2 = \|(I - p_l)b\|^2.\end{aligned}$$

Since $\|(I - p_l)b\| \to 0$ as $l \to \infty$ independent of x, we find that

$$\|q(I - p_l)\| = \sup_{\substack{x\in H\\ \|x\|=1}}\|q(I - p_l)x\| \to 0 \text{ as } l \to \infty.$$

Lastly, should the rank of q be larger than 1, we can take an orthonormal basis $\{b_1, \ldots, b_n\}$ for qH and write $q = \sum_{i=1}^{n}\langle\cdot, b_i\rangle b_i$. Then, taking the norm and applying the triangle inequality and using the above result we arrive at the same conclusion. ☐

We are now ready to prove a similar statement, but for compact operators.

Proposition 3.21. *Let $T \in K(H)$ be a compact operator. Then*

$$\|T - p_l T p_l\| \to 0 \text{ as } l \to \infty.$$

Proof. First we prove it for self-adjoint T. By Theorem 2.11 we can write $T = \sum_{i=1}^{\infty} \lambda_i \langle \cdot, f_i \rangle f_i$, where $\lambda_i \to 0$ as $i \to \infty$. Now let $F \subset \mathbb{N}$ be finite and define

$$T_F = \sum_{i \in F} \lambda_i \langle \cdot, f_i \rangle f_i.$$

We note that $\|T - T_F\| = \sup_{i \in \mathbb{N} \setminus F} |\lambda_i| \to 0$ as $F \to \mathbb{N}$. We will now prove that $\|T - Tp_l\| \to 0$ as $l \to \infty$.

Let $\varepsilon > 0$. By the above we can choose a finite subset $F \subset \mathbb{N}$ such that $\|T - T_F\| < \varepsilon$. Now define the finite rank projection q_F as the projection on $\mathrm{Span}\{e_i \mid i \in F\}$ (and so $T_F q_F = T_F$). By Lemma 3.20 we can take l sufficiently large such that $\|q_F - q_F p_l\| < \frac{\varepsilon}{\|T\|}$. Now using that $T_F q_F = T_F$ and the triangle inequality we obtain:

$$\begin{aligned}
\|T - Tp_l\| &= \|T - T_F + T_F - T_F p_l + T_F p_l - T p_l\| \\
&\leq \|T - T_F\| + \|T_F - T_F p_l\| + \|T_F p_l - T p_l\| \\
&\leq \|T - T_F\| + \|T_F q_F - T_F q_F p_l\| + \|T_F - T\| \|p_l\| \\
&\leq \|T - T_F\| + \|T_F\| \|q_F - q_F p_l\| + \|T_F - T\| \cdot 1 \\
&< \varepsilon + \varepsilon + \varepsilon = 3\varepsilon.
\end{aligned}$$

Since the inequality holds for $l' \geq l$ as well, we find that $\|T - Tp_l\| \to 0$ as $l \to \infty$. Lastly, once again using the triangle inequality we find:

$$\begin{aligned}
\|T - p_l T p_l\| &= \|T - Tp_l + Tp_l - p_l T p_l\| \leq \|T - Tp_l\| + \|Tp_l - p_l T p_l\| \\
&= \|T - Tp_l\| + \|T - Tp_l\| \|p_l\| = 2\|T - Tp_l\|,
\end{aligned}$$

and so we find that $p_l T p_l \to T$ in the operator norm.

For general (not necessarily self-adjoint) compact operators on T, we can write $T = \frac{1}{2}(T + T^*) + \frac{1}{2}(T - T^*)$. Since T^* is compact it follows that $\frac{1}{2}(T + T^*)$ is compact, and we note that $\frac{1}{2}(T + T^*)$ is also self-adjoint. As for the second part, here we instead have $(\frac{1}{2}(T - T^*))^* = -\frac{1}{2}(T - T^*)$. However, if we define $S := \frac{i}{2}(T - T^*)$, we see that $S^* = S$, i.e. S is self-adjoint as well as compact. Consequently, if we have proven the proposition for self-adjoint compact operators, by the triangle inequality it follows for non-self-adjoint operators as well. □

Thus, we see that we can approximate compact operators in the operator norm using finite rank compact operators, namely $p_l T p_l$. If we consider T as an infinite matrix (indexed by $\mathbb{N} \times \mathbb{N}$), we can view this as the upper left $l \times l$ submatrix. However, for our purposes we wish to have convergence in the Schatten norm, which we shall prove below. To this end, we first consider the following lemma:

Lemma 3.22. *For all $n \in \mathbb{N}$ and $T \in K(H)$ we have that $s_n(p_l T p_l) \leq s_n(T)$.*

Proof. Let $l \in \mathbb{N}$ be fixed. Now for any $n > l$, since $p_l T p_l$ is at most of rank l, by definition of the singular values $s_n(p_l T p_l) = 0$. Now, if $n \leq l$, then we have:

$$\begin{aligned}
s_n(p_l T p_l) &= \inf\{\|p_l T p_l - S\| \mid \dim \mathrm{R}(S) < n\} \\
&= \inf\{\|p_l T p_l - p_l S p_l\| \mid \dim \mathrm{R}(S) < n\} \\
&\leq \inf\{\|p_l\| \|T - S\| \|p_l\| \mid \dim \mathrm{R}(S) < n\} \\
&= \inf\{\|T - S\| \mid \dim \mathrm{R}(S) < n\} \\
&= s_n(T).
\end{aligned}$$

To see why the second inequality holds, consider the singular value decomposition of $p_l T p_l$. □

Proposition 3.23. *Let $p \geq 1$ and $T \in S_p(H)$, then $\|T - p_l T p_l\|_p \to 0$ as $l \to \infty$.*

Proof. For clarity we define $T_l := p_l T p_l$, which is still compact. First of all, by Lemma 3.23, we find that $T_l \in S_p(H)$ and $\|T_l\|_p \leq \|T\|_p$. Now if $p = \infty$, then $\|T - T_l\|_p = \|T - T_l\|$ and we are done. Thus, assume $p < \infty$.

We start by approximating T using finite rank elements from $S_p(H)$. Consider the singular value decomposition of T, i.e. $T = \sum_{n=1}^{\infty} s_n(T)\langle \cdot, f_n\rangle g_n$, for $(f_n)_{n=1}^{\infty}, (g_n)_{n=1}^{\infty}$ orthonormal sequences in H, of which we emphasise neither are necessarily equal to the fixed $(e_n)_{n=1}^{\infty}$. Now, for $N \in \mathbb{N}$, define:

$$S_N := \sum_{n=1}^{N} s_n(T)\langle \cdot, f_n\rangle g_n.$$

As a finite sum, clearly S_N is an element of $S_p(H)$ for every N. However, as $T - S_N = \sum_{n=N+1}^{\infty} s_n(T)\langle \cdot, f_n\rangle g_n$ and $\sum_{n=1}^{\infty} s_n(T)^p < \infty$, we find that $\|T - S_N\|_p \to 0$ as $N \to \infty$.

Now let $\varepsilon > 0$, and take N sufficiently large such that $\|T - S_N\|_p < \varepsilon$. By the triangle inequality we obtain:

$$\begin{aligned}
\|T - T_l\|_p &= \|T - S_N + S_N - p_l S_N p_l + p_l S_N p_l - p_l T p_l\|_p \\
&\leq \|T - S_N\|_p + \|S_N - p_l S_N p_l\|_p + \|p_l(S_N - T)p_l\|_p \\
&\leq \|T - S_N\|_p + \|S_N - p_l S_N p_l\|_p + \|T - S_N\|_p \\
&< 2\varepsilon + \|S_N - p_l S_N p_l\|_p.
\end{aligned}$$

Here $\|p_l(S_N - T)p_l\|_p \leq \|T - S_N\|_p$ by Lemma 3.22. We now turn our attention towards $S_N - p_l S_N p_l$. We note that $S_N - p_l S_N p_l$ has at most rank $2N$, i.e. for $n > 2N : s_n(S_N - p_l S_N p_l) = 0$. Consequently:

$$\|p_l(S_N - p_l S_N p_l)p_l\|_p^p = \sum_{n=1}^{\infty} s_n(S_N - p_l S_N p_l)^p = \sum_{n=1}^{2N} s_n(S_N - p_l S_N p_l)^p.$$

Furthermore, by Proposition 3.21 we can take l large enough such that $\|S_N - p_l S_N p_l\| < \frac{\varepsilon}{2N^{1/p}}$ (as well as for $l' \geq l$). However, as $s_0(S_N - p_l S_N p_l) = \|S_N - p_l S_N p_l\|$, and $s_0(S_N - p_l S_N p_l)$ is the largest singular value, we find:

$$\|p_l(S_N - p_l S_N p_l)p_l\|_p^p = \sum_{n=1}^{2N} s_n(S_N - p_l S_N p_l)^p < \sum_{n=1}^{2N} \frac{\varepsilon^p}{2N} = \varepsilon^p.$$

That is, $\|S_N - p_l S_N p_l\|_p < \varepsilon$, and combined with the earlier found inequality we then find that $\|T - T_l\|_p < 3\varepsilon$. As this holds for larger l as well, we thus find that $\lim_{l\to\infty}\|T - T_l\|_p = \lim_{l\to\infty}\|T - p_l T p_l\|_p = 0$. $\square$

Thus, any element T from $S_p(H)$, with H infinite-dimensional, we can approximate using the finite-rank operators $p_l T p_l$, where we let $l \to \infty$. Using this, we can now prove the following proposition:

Proposition 3.24. *Let M_A be a Schur multiplier on $S_p(H)$, and define $A_l = (a_{i,j})_{i,j=1}^{l}$, i.e. the left upper $l \times l$ submatrix of A. Considering M_{A_l} as a Schur multiplier on $S_p(p_l H)$, we have:*

$$\|M_A\|_p = \sup_{l\in\mathbb{N}} \|M_{A_l}\|_p.$$

Proof. Let us define $M_A^{(l)} : S_p(H) \to S_p(H)$ by $M_A^{(l)}(T) := M_A(p_l T p_l)$. Then $\|M_{A_l}\|_p = \|M_A^{(l)}\|_p$ and $\|M_A^{(l)}\|_p \leq \|M_A\|_p$, where the latter inequality can be obtained using the definition of the operator norm and Lemma 3.22. Thus, $\|M_{A_l}\|_p \leq \|M_A\|_p$. It remains to show that $\lim_{l\to\infty}\|M_{A_l}\|_p = \|M_A\|_p$.

Recall that $\|M_A\|_p = \sup_{\|T\|_p \leq 1}\|M_A(T)\|_p$. Now let $\varepsilon > 0$ arbitrarily and take $T \neq 0$ be such that $\|M_A(T)\|_p \geq \|M_A\|_p - \varepsilon$. As M_A and the norm are continuous, and $T_l \to T$ by Proposition 3.23, we find:

$$\lim_{l\to\infty}\|M_{A_l}(T_l)\|_p = \lim_{l\to\infty}\|M_A^{(l)}(T_l)\|_p = \lim_{l\to\infty}\|M_A(T_l)\|_p = \|M_A(T)\|_p \geq \|M_A\|_p - \varepsilon,$$

where we used that $M_A^{(l)}(T_l) = M_A(T_l)$ as projections are idempotent. Since this holds for all $\varepsilon > 0$, we conclude that $\lim_{l\to\infty}\|M_{A_l}\|_p = \|M_A\|_p$. $\square$

We emphasize that for $l \in \mathbb{N}$ M_{A_l} is a finite-dimensional Schur multiplier on $S_p(p_l H) \cong M_l(\mathbb{C})$.

3.3. Amplifying Schur multipliers and Pisier's conjecture

In this section we will introduce Pisier's conjecture. We first consider the following definition:

Definition 3.25. *Let $m \in \mathbb{N}$ and $A \in M_n(\mathbb{C})$. Now define the map $\Phi^{(m)} : M_n(\mathbb{C}) \to M_{nm}(\mathbb{C})$ by*

$$(\Phi^{(m)}(A))_{i,j} = A_{\lfloor \frac{i}{m} \rfloor, \lfloor \frac{j}{m} \rfloor}.$$

That is, each entry gets replaced by an $m \times m$ block consisting of that element. We say that $\Phi^{(m)}(A)$ is the matrix A amplified m times. For shorthand we will also denote $\Phi^{(m)}(A)$ with $A^{(m)}$ and the entries with $a^{(m)}_{i,j}$. Equivalently it is defined for infinite matrices.

Example 3.26. *Let $m = 2$ and $A = \begin{pmatrix} 1 & 1 \\ 2 & 4 \end{pmatrix}$. Then*

$$\Phi^{(n)}(A) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 2 & 2 & 4 & 4 \\ 2 & 2 & 4 & 4 \end{pmatrix}$$

We use this to "amplify" a Schur multiplier, i.e. given a Schur multiplier M_A we can consider $M_{A^{(m)}}$ for $m \in \mathbb{N}$. However, what is the underlying (separable) Hilbert space in this case? If $\dim H = n < \infty$, we know that $H \cong \mathbb{C}^n$, and if $\dim H = \infty$, then $H \cong \ell^2$ [11], and so without loss of generality we can assume H to be of the latter forms. Consequently, when considering $M_{A^{(m)}}$ the underlying Hilbert space is respectively $(\mathbb{C}^n)^m = \mathbb{C}^{nm}$ and ℓ^2. To see that $M_{A^{(m)}}$ is actually a Schur multiplier as well, by the triangle inequality we have that $\|M_{A^{(m)}}\| \leq m^2 \|M_A\|$. We now consider the following definition:

Definition 3.27. *Let $p \geq 1$ and let M_A be a bounded Schur multiplier on $S_p(H)$. If $\sup_{m \in \mathbb{N}} \|M_{A^{(m)}}\| < \infty$, then we define the* ***completely bounded norm***[2] *of M_A as*

$$\|M_A\|_{cb} := \sup_{m \in \mathbb{N}} \|M_{A^{(m)}}\|.$$

At last, we now state the conjecture:

Conjecture 3.28. *Let $1 \leq p \leq \infty$ and H be a Hilbert space. We now consider the following statement:*

All Schur multipliers on $S_p(H)$ are completely bounded.

Pisier [10] conjectures that the above statement is ***false*** *for $1 < p < \infty$ and $p \neq 2$, i.e. there exists a Schur multiplier on $S_p(H)$ that is not completely bounded.*

On the other hand, as Pisier proves, for $p = 1, 2$ and ∞ the statement actually holds. Moreover, the completely bounded norm in this case is equal to the norm for any Schur multiplier. We will take a short look at $p = 2$.

Theorem 3.29. *For $p = 2$, and M_A a Schur bounded multiplier, we have that Pisier's statement holds, and $\|M_A\|_{cb} = \|M_A\|_2 = \sup |a_{i,j}|$.*

Proof. We first consider the finite-dimensional Schur multiplier M_{A_l} as in Proposition 3.24. Now $\|\cdot\|_2$ is induced by the Frobenius inner product and the matrices $e_{i,j}$ with a 1 at position (i, j) and 0 elsewhere form an orthonormal basis. Consequently, for an element of $S_p(H)$ represented by the matrix B we have:

$$\|B\|_2^2 = \sum_{i,j=1}^{n} \|b_{i,j} e_{i,j}\|^2 = \sum_{i,j=1}^{n} |b_{i,j}|^2 \|e_{i,j}\|^2 = \sum_{i,j=1}^{n} |b_{i,j}|^2. \tag{3.6}$$

Now set $\alpha_l := \sup_{1 \leq i,j \leq l} |a|_{i,j}$. Clearly, for any $m \in \mathbb{N}$ we still have that $\alpha_l = \sup_{1 \leq i,j \leq lm} |a^{(m)}_{i,j}|$. However, for any matrix $B \in M_{nm}(\mathbb{C})$, by definition of the Schur product we obtain:

$$\|M_{A_l^{(m)}}(B)\|_2^2 \overset{(3.6)}{=} \sum_{i,j=1}^{n} |a^{(m)}_{i,j} b_{i,j}|^2 \leq \sum_{i,j=1}^{n} \alpha_l^2 |b_{i,j}| \overset{(3.6)}{=} \alpha_l^2 \|B\|_2^2.$$

[2]Although not Pisier's original definition of the completely bounded norm, he shows that this definition is equivalent.

Taking the square root, we find that $\|M_{A_l^{(m)}}(B)\|_2 \leq \alpha_l \|B\|_2^2$, and thus $\|M_{A_l^{(m)}}\|_2 \leq \alpha_l$. However, since there exists an entry $a_{i',j'}^{(m)}$ such that $\alpha_l = |a_{i',j'}^{(m)}|$, we find that $\|M_{A_l^{(m)}}(e_{i',j'})\|_2 = |a_{i',j'}^{(m)}| \cdot 1 = \alpha_l \|e_{i',j'}\|_2$, and thus $\|M_{A_l^{(m)}}\|_2 \geq \alpha_l$. That is, $\|M_{A_l^{(m)}}\|_2 = \alpha_l$ for all $m \in \mathbb{N}$. Lastly, combining Proposition 3.24 with the fact that $\alpha_l \leq \sup_{i,j} |a_{i,j}|$ for any $m \in \mathbb{N}$ we obtain:

$$\|M_{A^{(m)}}\|_2 = \sup_{l \in \mathbb{N}} \|M_{A_l^{(m)}}\|_2 = \sup_{i,j \in \mathbb{N}} |a_{i,j}|.$$

$\square$

4

Inequalities for $\|M_{A^{(m)}}\|_p$.

In this chapter we will give some (in)equalities and reductions regarding the conjecture for $1 \leq p \leq \infty$. Throughout this chapter we will assume we work with finite-dimensional Hilbert spaces H, and can thus assume $H = \mathbb{C}^n$ for some $n \in \mathbb{N}$ and $S_p(H) = B(H) \cong M_n(\mathbb{C})$. However, as we can approximate M_A and its amplifications using finite-dimensional Schur multipliers, the (in)equalities found will still hold for infinite-dimensional H (but maxima turn into suprema). Furthermore, A will be an arbitrary matrix in $M_n(\mathbb{C})$, unless explicitly stated otherwise.

Recall from the previous chapter that $\|M_{A^{(m)}}\|_p = \alpha := \max_{i,j}|a_{i,j}|$ for any m if $p = 2$. However, for a non-trivial p, such as $p = 4$, this need not be the case. Firstly, observe that $\|M_{A^{(m)}}\|_p \geq \alpha$, which follows easily from considering $B = e_{i',j'}$ where $\alpha = |a^{(m)}_{i',j'}|$. To show this can also be a strict inequality, we give the following example for $p = 4$ and $n = 2$:

Example 4.1. Let $A = \begin{pmatrix} 1 & 1 \\ 1 & \frac{1}{2} \end{pmatrix}$ and $B = \begin{pmatrix} -3 & 1 \\ 1 & 1 \end{pmatrix}$. Using Proposition 3.8 one can calculate easily that $\|B\|_4^4 = 112$ and $\|M_A(B)\|_4^4 = 114\frac{1}{16}$. That is, we find that $\|M_A\|_4 > 1$, whereas 1 is the maximal element of A in modulus.

The above example shows that $p = 4$ is indeed a non-trivial case. So far we have that the maximal absolute entry of A is a lower bound for $\|M_{A^{(m)}}\|$. However, we can give a stronger lower bound for $\|M_{A^{(m)}}\|$, for which we consider the following proposition:

Proposition 4.2. *For any $m \in \mathbb{N}$ and $1 \leq p \leq \infty$ we have that $\|M_A\|_p \leq \|M_{A^{(m)}}\|_p$.*

Proof. Recall that $\|M_A\|_p = \sup_{\|B\|_p=1}\|M_A(B)\|_p$, and recall from 2.4 that in this case the unit sphere is compact, and thus we can take $B \in M_n(\mathbb{C})$ such that $\|M_A(B)\|_p = \|M_A\|_p$. We shall now construct a $B' \in M_{nm}(\mathbb{C})$ from B.

We partition B' into $m \times m$ index by $B'_{i,j}$, where $1 \leq i, j \leq n$. Consider an entry $b_{i,j}$ of B and the associated $m \times m$ block $B'_{i,j}$ in B'. Now set the entry at $(1, 1)$ of $B'_{i,j}$ to $b_{i,j}$, and the rest 0. For example, for $m = 2$ and B as below we would have:

$$B = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}, \quad B' = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Claim: $\|B'\|_p = 1$ and $\|M_A(B)\|_p = \|M_{A^{(m)}}(B')\|_p = \|M_A(B)\|_p$.
To show the first part, we once more turn to Proposition 3.8. Looking at the $m \times m$ block of $(B')^*B'$ at position (i, j), we have that this is equal to $\sum_{k=1}^n (B')^*_{i,k}B'_{k,j}$. But since all of the blocks are diagonal matrices with only one non-zero entry at $(1, 1)$, we see that the result is a block with $\sum_{k=1}^n \overline{b_{k,i}}b_{k,j}$ at position $(1, 1)$, and 0 elsewhere. However, these are also equal to the entries of B^*B at position (i, j). To

illustrate this with the prior example, we would have:

$$B^*B = \begin{pmatrix} 10 & 5 \\ 5 & 5 \end{pmatrix} \quad \text{and} \quad (B')^*B' = \begin{pmatrix} 10 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 \\ 5 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{4.1}$$

If we diagonalize B' for instance, then we would find that for any $B \in M_n(\mathbb{C})$, the non-zero singular values of B and B' are in fact equal. Consequently, $\|B'\|_p^p = \|B\|_p^p = 1$. Furthermore, if we were to construct a matrix in $M_{nm}(\mathbb{C})$ from $M_A(B)$ similarly as to how we constructed B' from B, one observes that this matrix is equal to $M_{A^{(n)}}(B')$. By the previous equality we then find that $\|M_{A^{(m)}}(B')\| = \|M_A(B)\|$, and thus $\|M_A\|_p \leq \|M_{A^{(m)}}\|_p$. $\square$

Corollary 4.3. For arbitrary $m, m' \in \mathbb{N}$ such that $m \leq m'$ we can similarly embed a (unit) $B \in M_{nm}(\mathbb{C})$ into $B' \in M_{nm'}((C))$ and repeat the previous proof to find $\|M_{A^{(m)}}\|_p \leq \|M_{A^{(m')}}\|_p$ as well.

Thus, we find that $\|M_{A^{(m)}}\|_p$ can only increase with respect to m. Next we look at how we can use unitary transformations to reduce the problem of finding $\|M_{A^{(m)}}\|_p$ for $m \geq 2$. Recall that a matrix $U \in M_n(\mathbb{C})$ is unitary if and only if $U^*U = UU^* = I$. This brings us to the following lemma:

Lemma 4.4. *Let $B, U \in M_n(\mathbb{C})$, where U is a unitary matrix. Then $\|UB\|_p = \|BU\|_p = \|B\|_p$.*

Proof. This is easily verified using the singular values. First of all, we have:

$$\begin{aligned} s_n(UB) &= \inf\{\|UB - S\| \mid S \in K(H) \text{ and } \operatorname{rank}(S) < n\} \\ &= \inf\{\|UB - US\| \mid S \in K(H) \text{ and } \operatorname{rank}(S) < n\} \\ &= \inf\{\|U(B - S)\| \mid S \in K(H) \text{ and } \operatorname{rank}(S) < n\} \\ &= \inf\{\|B - S\| \mid S \in K(H) \text{ and } \operatorname{rank}(S) < n\} \\ &= s_n(B), \end{aligned} \tag{4.2}$$

where we use that U is unitary and $\{S \in K(H,H') \mid \text{ and } \operatorname{rank}(S) < n\} = \{US \in K(H,H') \mid \text{ and } \operatorname{rank}(S) < n\}$ as a result of U being bijective. Consequently, we find that $\|UB\|_p = \|B\|_p$. That $\|BU\|_p = \|B\|_p$ holds as well follows similarly. $\square$

Remark 4.5. Since U^* is a unitary matrix, the same holds for U^*.

However, the same will not be true in general when we consider $\|M_{A^{(m)}}(UB)\|_p^p$. Instead, there is an extra requirement for U:

Lemma 4.6. *Let $m \in \mathbb{N}$ and $B, U \in M_{nm}(\mathbb{C})$, where $B \neq 0$. If we partition U into $m \times m$ blocks, and the diagonal blocks of U are unitary, then we have*

$$\frac{\|M_{A^{(m)}}(UB)\|_p}{\|UB\|_p} = \frac{\|M_{A^{(m)}}(BU)\|_p}{\|BU\|_p} = \frac{\|M_{A^{(m)}}(B)\|_p}{\|B\|_p}.$$

Proof. Note that U is unitary, and thus from lemma 4.4 it directly follows that the denominators are equal, and thus it only remains to show that $\|M_{A^{(m)}}(UB)\|_p = \|M_{A^{(m)}}(BU)\|_p = \|M_{A^{(m)}}(B)\|_p$. Claim: we have that $M_{A^{(m)}}(UB) = UM_{A^{(m)}}(B)$. We first partition all $nm \times nm$ matrices into $m \times m$ blocks, and look at $M_{A^{(m)}}(UB) = A^{(m)} \circ (UB)$. We write U and B as

$$U = \begin{pmatrix} U_1 & & \\ & \ddots & \\ & & U_n \end{pmatrix}, \quad B = \begin{pmatrix} B_{1,1} & \dots & B_{1,n} \\ \vdots & \ddots & \vdots \\ B_{n,1} & \dots & B_{n,n} \end{pmatrix},$$

where the U_i are unitary. Subsequently, for the product UB we have:

$$UB = \begin{pmatrix} U_1B_{1,1} & \dots & U_1B_{1,n} \\ \vdots & \ddots & \vdots \\ U_nB_{n,1} & \dots & U_nB_{n,n} \end{pmatrix}.$$

We look at $A^{(m)}$ in more detail. By definition of $A^{(m)}$, the blocks are filled entirely with an entry of A. That is, we have:

$$A^{(m)} = \begin{pmatrix} A^{(m)}_{1,1} & \dots & A^{(m)}_{1,n} \\ \vdots & \ddots & \vdots \\ A^{(m)}_{n,1} & \dots & A^{(m)}_{n,n} \end{pmatrix}, \quad \text{where} \quad A^{(m)}_{i,j} = \begin{pmatrix} a_{i,j} & \dots & a_{i,j} \\ \vdots & \ddots & \vdots \\ a_{i,j} & \dots & a_{i,j} \end{pmatrix} = a_{i,j} J_m.$$

It is important to note the distinction that $a_{i,j}$ is a scalar here, whereas $A^{(m)}_{i,j}$ is in fact an $m \times m$ block. Recall that J_m is the $m \times m$ matrix filled entirely with ones. We now look at the Schur product of $A^{(m)}$ and UB. Since this now comes down to taking the Schur product of the blocks at the same position, we finally obtain:

$$\begin{aligned} A^{(m)} \circ (UB) &= \begin{pmatrix} A^{(m)}_{1,1} \circ (U_1 B_{1,1}) & \dots & A^{(m)}_{1,n} \circ (U_1 B_{1,n}) \\ \vdots & \ddots & \vdots \\ A^{(m)}_{n,1} \circ (U_n B_{n,1}) & \dots & A^{(m)}_{n,n} \circ (U_n B_{n,n}) \end{pmatrix} & (4.3) \\ &= \begin{pmatrix} U_1(a_{1,1} B_{1,1}) & \dots & U_1(a_{1,n} B_{1,n}) \\ \vdots & \ddots & \vdots \\ U_n(a_{n,1} B_{n,1}) & \dots & U_n(a_{n,n} B_{n,n}) \end{pmatrix} = U(A^{(m)} \circ B). & (4.4) \end{aligned}$$

Here $A^{(m)}_{i,j} \circ (U_i B_{i,j}) = a_{i,j} J_m \circ (U_i B_{i,j}) = a_{i,j}(U_i B_{i,j}) = U_i(A_{i,j} B_{i,j})$, as $a_{i,j}$ is a scalar and J_m was the identity element under Schur multiplication. Thus, we have found that $A^{(m)} \circ (UB) = U(A^{(m)} \circ B) = UM_{A^{(m)}}(B)$. As $M_{A^{(m)}}(B)$ is simply another matrix, by Lemma 4.4 we have that $\|M_{A^{(m)}}(UB)\|_p = \|UM_{A^{(m)}}(B)\|_p = \|M_{A^{(m)}}(B)\|_p$. For $\|M_{A^{(m)}}(BU)\|_p$ the result follows analogously. □

Remark 4.7. In particular, if B is such that $\|M_{A^{(m)}}\|_p = \frac{\|M_{A^{(m)}}(B)\|_p}{\|B\|_p}$, i.e. a maximum, then we can multiply it with unitary matrices as above and still maintain a maximum.

We now combine Lemma 4.6 and Proposition 2.13 to prove the following theorem:

Theorem 4.8. *Let $m \in \mathbb{N}$ and $B \in M_{nm}(\mathbb{C})$. Then we can transform B into $B' \in M_{nm}(\mathbb{C})$ such that $\frac{\|M_{A^{(m)}}(B)\|_p}{\|B\|_p} = \frac{\|M_{A^{(m)}}(B')\|_p}{\|B'\|_p}$, and if we partition B' into $m \times m$ blocks, then in each row and column at least one block will be a diagonal matrix with real entries.*

Proof. Throughout this proof we shall assume all $nm \times nm$ matrices to be partitioned into $m \times m$ blocks. Now consider a block $B_{i,j}$ of B at position (i, j). Using the polar decomposition we can write $B_{i,j} = U_i S_i$, for U_i a unitary $m \times m$ matrix, and S_i a self-adjoint $m \times m$ matrix. Now set

$$U = \begin{pmatrix} I_m & & & & & \\ & \ddots & & & & \\ & & U_i^* & & & \\ & & & I_m & & \\ & & & & \ddots & \\ & & & & & I_m \end{pmatrix} \in M_{nm}(\mathbb{C}), \tag{4.5}$$

where I_m is the identity $m \times m$ block and U_i^* is the i-th diagonal block. Now set $B' := UB$. By Lemma 4.6 we have that $\frac{\|M_{A^{(m)}}(B)\|_p}{\|B\|_p} = \frac{\|M_{A^{(m)}}(B')\|_p}{\|B'\|_p}$. Furthermore, we have:

$$B' = \begin{pmatrix} B_{1,1} & \dots & & \dots & B_{1,n} \\ \vdots & & & & \vdots \\ U_i^* B_{i,1} & \dots & U_i^* B_{i,j} & \dots & U_i^* B_{i,n} \\ \vdots & & & & \vdots \\ B_{n,1} & \dots & & \dots & B_{n,n} \end{pmatrix} = \begin{pmatrix} B_{1,1} & \dots & & \dots & B_{1,n} \\ \vdots & & & & \vdots \\ U_i^* B_{i,1} & \dots & S_i & \dots & U_i^* B_{i,n} \\ \vdots & & & & \vdots \\ B_{n,1} & \dots & & \dots & B_{n,n} \end{pmatrix}.$$

That is, $B'_{i,j}$ has now become a self-adjoint $m \times m$ matrix. Note that since we multiply the entire i-th row with U_i^*, the other blocks there change as well. We consider B' and this block. By the spectral theorem we can decompose $B'_{i,j}$ into $U'_{i,j} D_{i,j} (U'_{i,j})^*$, for $U'_{i,j}$ a unitary $m \times m$ matrix and $D_{i,j}$ a diagonal $m \times m$ matrix. Similar as in (4.5) we construct U', i.e. where $U'_{i,j}$ is the i-th diagonal element. Lastly we set $B'' = U'B'(U')^*$. Then by Lemma 4.6 twice we see that $\frac{\|M_{A^{(m)}}(B)\|_p}{\|B\|_p} = \frac{\|M_{A^{(m)}}(B'')\|_p}{\|B''\|_p}$, by writing out multiplication we find that $B''_{i,j}$ is now a diagonal block. Since we change the other blocks in row i and column j we can't do this for all blocks simultaneously, but doing it for exactly one block in each column and row (e.g. the diagonal blocks), we can guarantee that in the resulting matrix B'' in each row and column at least one block is a diagonal matrix. ☐

Remark 4.9. In particular, when looking for a $B \in M_{nm}(\mathbb{C})$ such that $\frac{\|M_{A^{(m)}}(B)\|_p}{\|B\|_p} = \|M_{A^{(m)}}\|_p$, we can assume without loss of generality that the blocks on the diagonal of B are diagonal matrices themselves.

Lastly, we give a proposition that gives a relation between $\|M_A(B)\|_p$ and $\|M_{A^{(m)}}(B)\|_p$ for particular matrices B.

Proposition 4.10. *Let $m \in \mathbb{N}$ and let $B \in M_{nm}(\mathbb{C})$ with $B \neq 0$ have the property that every block is a multiple of I_m, i.e. $B_{i,j} = \beta_{i,j} I_m$ for $\beta_{i,j} \in \mathbb{C}$. If we define $C := (\beta_{i,j})_{i,j=1}^n \in M_n(\mathbb{C})$, then*

$$\frac{\|M_{A^{(m)}}(B)\|_p}{\|B\|_p} = \frac{\|M_A(C)\|_p}{\|C\|_p}.$$

Proof. The reasoning is similar to that used in the proof of Proposition 4.2. If we now consider B^*B, we have that $(B^*B)_{i,j} = \sum_{k=1}^n \overline{c_{k,i}} c_{k,j} I_m$, i.e. each block at position (i,j) of B^*B is the identity matrix multiplied by the entry of C^*C at position (i,j). Consequently, we find that the singular values of B are those of C, but now repeated m times. For $1 \leq p < \infty$:

$$\|B\|_p^p = \sum_{k=1}^{\infty} s_k^p(B) = m \sum_{k=1}^{\infty} s_k^p(C) = \|C\|_p^p.$$

Thus, $\|B\|_p = m^{1/p}\|C\|_p$. For $p = \infty$ we simply find $\|B\|_p = \|C\|_p$. Similarly, for $1 \leq p < \infty$ one finds that $\|M_{A^{(m)}}(B)\|_p = m^{1/p}\|M_A(C)\|_p$, and $\|M_{A^{(m)}}(B)\|_p = \|M_A(C)\|_p$ if $p = \infty$. Lastly, substitution these equalities in $\frac{\|M_{A^{(m)}}(B)\|_p}{\|B\|_p}$ we find the desired result. ☐

We note that if we can find $B \in M_{nm}(\mathbb{C})$ in the above form such that $\|M_{A^{(m)}}\|_p = \|M_{A^{(m)}}(B)\|_p$, then by the prior proposition and Proposition 4.2 it would follow that $\|M_A\|_p = \|M_{A^{(m)}}\|_p$. If we are able to do this for all m, this would mean that $\|M_A\|_p = \|M_A\|_{cb}$ and thus disprove Pisier's conjecture.

5

Computational results

5.1. Approximating the operator norm

In this chapter we will look at the method for numerically approximating the operator norm of a finite dimensional Schur multiplier, given a matrix $A \in M_n(\mathbb{C})$. Recall that in this case for Schur multiplier M_A we have $\|M_A\|_p = \max_{B \neq 0} \frac{\|M_A(B)\|_p}{\|B\|_p}$. Thus, we have arrived at an optimization problem, where we wish to maximize the quotient $\frac{\|M_A(B)\|_p}{\|B\|_p}$ in the domain $M_n(\mathbb{C}) \cong \mathbb{C}^{n^2}$.

A first concern is however, how do we calculate $s_n(B)$ for an arbitrary matrix B in order to calculate $\|B\|_p$? As we have seen in Chapter 3 as well, the singular values coincide with the square roots of the eigenvalues of B^*B. Thus, one possibility is computing the eigenvalues of B^*B, for which many algorithms such as the QR algorithm [8] exist. However, this is generally computationally expensive. If we instead only consider even p, then by Proposition 3.8 we have an easy way of computing $\|B\|_p$, namely $\|B\|_p = (\mathrm{tr}((B^*B)^{p/2}))^{1/p}$, where computing $\mathrm{tr}((B^*B)^{p/2})$ requires only elementary matrix operations, and taking the p-th root is a fast operation as well. Thus, we will only consider even p, i.e. $p = 2k \in \mathbb{N}$, and we will mostly focus on $p = 4$ during the computations.

5.1.1. A function to maximize

Next we have to choose a method to maximize $\frac{\|M_A(B)\|_p}{\|B\|_p}$ on $M_{nm}(\mathbb{C})$ for even p. Note that maximizing $\frac{\|M_A(B)\|_p}{\|B\|_p}$ is equivalent with maximizing $\frac{\|M_A(B)\|_p^p}{\|B\|_p^p}$, as $x \mapsto x^p$ is increasing, and we choose the latter to save on computing the p-th root until the end. For a given $A \in M_n(\mathbb{C})$ and $m \in \mathbb{N}$, for $B \in M_{nm}(\mathbb{C})$ we now define:

$$f_{A^{(m)}}(B) := \frac{\|M_A(B)\|_p^p}{\|B\|_p^p} = \frac{\mathrm{tr}((M_A(B)^* M_A(B))^{p/2})}{\mathrm{tr}((B^*B)^{p/2})}.$$

We now make a couple of observations. First of all, we know that $f_{A^{(m)}}$ is bounded, since $f_{A^{(m)}} \geq 0$ and $f_{A^{(m)}}(B) \leq \|M_{A^{(m)}}\| < \infty$, where $\|M_{A^{(m)}}\| < \infty$ as any Schur multiplier is bounded.
Secondly, we have that $\mathrm{tr}((M_A(B)^* M_A(B))^{p/2}$ and $\mathrm{tr}((B^*B)^{p/2})$ are polynomials with respect to the entries of B. Since polynomials are smooth, and we know that $\|B\|_p \neq 0$ for $B \neq 0$, outside of a neighborhood of 0, we have that $f_{A^{(m)}}$ is smooth, and the denominator is non-zero.

5.1.2. Methods for maximization

We now choose a method for maximizing $f_{A^{(m)}}$. If we exclude a small neighborhood of 0 from our domain, the gradient of f exists as f is then smooth. A common method for maximizing a differentiable function is the iterative method gradient ascent, where given some initial vector B we move into the direction of the gradient and obtain a higher value for $f_{A^{(m)}}$. However, since higher derivatives of $f_{A^{(m)}}$ exist as well, we may also employ more advanced algorithms, such as the Broyden–Fletcher–Goldfarb–Shanno algorithm [8] (BFGS). The BFGS algorithm is a minimization algorithm that minimizes a (possibly) non-linear function f on an unconstrained real domain, which performs well if f is at least twice differentiable. As f takes complex arguments, we split each entry into its real and imaginary part, i.e. $z = x + \mathrm{i}y$, which doubles the amount of variables. As it is a minimization algorithm, we shall instead apply it to $-f_{A^{(m)}}$. Similar to gradient descent, the BFGS algorithm requires an initial guess, and hopefully converges to a local minimum. In order to obtain a global minimum, we make a number of initial random guesses and apply the BFGS algorithm to obtain a local minimum. Lastly, we take the minimum of all the local minima produced.

5.1.3. Implementation

We now construct the pseudocode for the implementation. Here we assume $A \in M_n(\mathbb{C})$, and p is even. Furthermore, we assume that we have an implementation of the BFGS algorithm. The pseudocode will now be as follows:

Algorithm 1 Operator norm approximation

```
procedure NORM(B)                                  ▷ p-th power of the norm
    Bstar ← conjugate(transpose(B))
    C ← Bstar · B
    return trace(C^{p/2})

procedure SCHURPRODUCT(B)
    C ← zeros(n, n)                                ▷ n × n matrix filled with zeros
    for i, j = 1, ..., n do
        C_{i,j} ← A_{i,j} · B_{i,j}
    return C

procedure f(B)
    a ← norm(schurproduct(B))
    b ← norm(B)
    return a/b

procedure OPERATORNORM                             ▷ Approximate the operator norm
    result ← 0                                     ▷ n × n matrix filled with zeros
    for i = 1, ..., n^2 do
        B ← random(n, n)                           ▷ Random complex n × n matrix
        temp = −BFGS(−f, B)
        if temp > result then
            result ← temp
    return result^{(1/p)}
```

We wish to approximate $\|M_{A^{(m)}}\|$ as well for $m \in \mathbb{N}$. To that extent, $A^{(n)}$ is easily constructed as follows:

Constructing $A^{(m)}$

```
procedure AMPLIFY(m)
    A_new = zeros(nm, nm)
    for i, j = 1, ..., nm do
        (A_new)_{i,j} ← A_{⌊i/m⌋,⌊j/m⌋}
    n ← n · m
    A ← A_new
```

After amplifying A we can approximate the norm of the Schur multiplier with Algorithm 1 as well. However, as this increases the number of variables we can expect to need more iterations to reach similar results as in the case of smaller matrices. Hence, in algorithm 1 we perform the BFGS algortihm n^2 times, as this scales with the number of variables.

To actually implement the program, we turn to the programming language Python 3. To work more easily with matrices we use the package NumPy [9], and from the package SciPy [6] we employ the use

of the BFGS algorithm. A full implementation in Python 3.6.3 can be found in the appendix.
However, up until now, we have not used the reductions we made in Chapter 4 when approximating $\|M_{A^{(m)}}\|_p$, and in particular Theorem 4.8. By Remark 4.9, we can at the least assume that the maximum is assumed in a matrix where the diagonal blocks (with the size of these blocks depending on the amplification) are diagonal blocks. Thus, instead of taking $M_{nm}(\mathbb{C})$ as the domain, we can instead only consider the matrices in the aforementioned form. That is, if we again consider the matrix entries are variables themselves, we need not consider those outside of the diagonal in the diagonal blocks (since these are zero). This lowers the amount of variables by $nm^2 - nm = nm(m-1)$, which presumably increases the accuracy of the approximation, whereas the new variable count is thus $n^2m^2 - nm(m-1)$. To implement this, we could for instance let $f_{A^{(m)}}$ take a vector in $\mathbb{C}^{n^2m^2-nm(m-1)}$ and transform it into a matrix of the previously mentioned form. As the pseudocode for this transformation would not be insightful we have chosen to omit it.

5.1.4. Time Complexity

We take a quick look at the time complexity when computing $\|M_{A^{(n)}}\|_p$ for $A \in M_n(\mathbb{C})$ and $m \in \mathbb{N}$. Firstly, we look at the time complexity of evaluating $f_{A^{(m)}}(B)$, for some $B \in M_{nm}(\mathbb{C})$. Whereas taking the Schur product is $\mathcal{O}((nm)^2)$, multiplying the matrices takes $\mathcal{O}((nm)^3)$ time. Thus, we find that in total computing $f_{A^{(m)}}(B)$ takes $\mathcal{O}((nm)^3)$ time. Secondly, when minimizing a function of k variables the BFGS algorithm has a time complexity of $\mathcal{O}(k^2)$. In both the initial and reduced amount of variables case this is equal to $\mathcal{O}((nm)^2)$. This does not take into account the time complexity of the function evaluation, but as an iterative algorithm with a maximum amount of iterations this adds $\mathcal{O}((nm)^3)$ by the previous calculation. Thus, we find that executing the BFGS algorithm a single time altogether takes $\mathcal{O}((nm)^3)$ time as well. Lastly, since we perform the BFGS algorithm $(nm)^2$ times, we find that the total time complexity of the algorithm is $\mathcal{O}((nm)^5)$. Thus, it would be infeasible to employ this implementation for large n or m, whereas the most time is spent on computing $f_{A^{(m)}}(B)$.

5.2. Results

In this section we present and interpret the results produced by the implementation from the previous section.

5.2.1. A fixed matrix

We begin with the same 2×2 matrix from Example 4.1, i.e.

$$A = \begin{pmatrix} 1 & 1 \\ 1 & \frac{1}{2} \end{pmatrix}.$$

In the example we have already seen that the norm will at least be $\left(\frac{114.25}{112}\right)^{1/4} \approx 1.0045$, which the implementation will presumably confirm as well. Although we approximate the operator norm using complex matrices, on top of this we can also try to approximate it with real matrices. Whereas the operator norm can only be equal or less when we restrict the domain to real matrices, should it be equal in both cases then the approximation using real matrices might be better due to there being half the variables. We then of course take the largest value of both approximations. As A is a real matrix we speculate that $f_{A^{(m)}}$ assumes its maximum for a real matrix.

In the table below the results of running the algorithm can be seen for $m = 1 \ldots 4$. In Proposition 4.2 and the subsequent remark we saw that $\|M_{A^{(m)}}\|_4$ can only increase with respect to m, and to that end we consider the difference $\|M_{A^{(m)}}\|_4 - \|M_A\|_4$ as well. Furthermore, since the amount of variables is smallest for $m = 1$, we can expect this approximation to be the best. Lastly, the total run time was approximately 4 minutes[1], whereas the majority of the time is spent on the calculations for larger m.

[1]These computations were performed on a Intel Core i7-7700 @ 3.60GHz

Table 5.1: Results for the default algorithm. Run time: 4 minutes.

m	$\|M_{A^{(m)}}\|_4$	$\|M_{A^{(m)}}\|_4$ - $\|M_A\|_4$
1	1.01311512	0
2	1.01311512	$-1.9984 \cdot 10^{-15}$
3	1.01311512	$-7.3053 \cdot 10^{-14}$
4	1.01311512	$-2.1116 \cdot 10^{-13}$
5	1.01311512	$-1.3327 \cdot 10^{-12}$

The first and foremost observation is that $\|M_{A^{(m)}}\|_4$ seems to be equal for $m = 1 \dots 5$. Whilst it is important to remember that we have only approximated the operator norms (from below), with reasonable certainty we can say the first nine digits are correct. Furthermore, as we will investigate later, it might also only be the case for this particular matrix A.
From the second column we see that our approximations for the norms are in fact decreasing. As Remark 4.3 gives us that the operator norm cannot decrease, we can deduce that this is the results of numerical errors or a (slightly) worse approximation. Here the latter is easily attributed due to working with more variables, and the former due to the amount operations increasing. Lastly, although not presented in the table, $f_{A^{(m)}}$ achieves its maximum for real matrices, as well as for complex matrices.

We now do the same using the algorithm with the reduced variable count as discussed in Subsection 5.1.3, i.e. we only look at matrices where the $m\ timesm$ diagonal blocks are diagonal matrices, reducing the variable count and hopefully improving the algorithm. The results can be seen in the tables below.

Table 5.2: Results for the algorithm with the reduction in variables. Run time: 4 minutes.

m	$\|M_{A^{(m)}}\|_4$	$\|M_{A^{(m)}}\|_4$ - $\|M_A\|_4$
1	1.01311512	0
2	1.01311512	$-6.6613 \cdot 10^{-16}$
3	1.01311512	$-7.3275 \cdot 10^{-15}$
4	1.01311512	$-7.9936 \cdot 10^{-14}$
5	1.01311512	$-2.8422 \cdot 10^{-14}$

Although the approximations seem to be identical with those of table 5.1, the second column reveals this not to be the case. On the other hand, as both algorithms are identical for $m = 1$ and only 4 variables are involved, it is no surprise that the approximations for $\|M_A\|_4$ are identical. However, in the second column we see that the difference with $\|M_A\|_4$ grows less as m increases. If we are once more willing to accept for now that the operator norm does not change, then we see that the small errors grow slower than in the previous case, which too can be attributed to the amount of variables present. As the run times are similar, we prefer this algorithm over the first one.

Seeing the previous results, one may wonder if this is due to A being a self-adjoint matrix. To this end, we consider the following (arbitrarily chosen) non-self-adjoint matrix:

$$A = \begin{pmatrix} 1 & 2 & 2 \\ -2 & 1 & 3 \\ 0 & 2 & -2 \end{pmatrix}$$

This in turns yields the following results:

Table 5.3: Results for the algorithm with the reduction in variables. Run time: 20 minutes.

m	$\|M_{A^{(m)}}\|_4$	$\|M_{A^{(m)}}\|_4$ - $\|M_A\|_4$
1	3.04915498	0
2	3.04915498	$-8.0028 \cdot 10^{-11}$
3	3.04915498	$-1.5293 \cdot 10^{-10}$
4	3.04915498	$-2.8443 \cdot 10^{-10}$
5	3.04915498	$-4.8709 \cdot 10^{-10}$

Here too we see that the approximations are equal in the first 9 digits, and worsen as we increase m. That is, for this non-self-adjoint matrix we see that $\|M_{A^{(m)}}\|_4$ does not grow either with m. However, compared to the results of Table 5.2 the errors are worse, which we can attribute to working with larger matrices. Lastly, one may also note that the run time is 5 times as large, which is in line with what we found in Subsection 5.1.4.

5.2.2. RANDOM MATRICES

To gain some more insight as to whether $\|M_{A^{(m)}}\|_4$ does not grow when m increases, we repeat the above procedure for a number of random **complex** matrices of varying sizes. As mentioned earlier, the algorithm scales poorly in terms of n and m, and therefore we will only consider some smaller n and m. For each $n = 2, \dots, 4$ we take N random complex matrices A and for each of these we approximate $\|M_{A^{(m)}}\|_4$ for some m. In the table below we list the averages of $\|M_{A^{(m)}}\|_4$ for these matrices and if for any particular matrices $\|M_{A^{(m)}}\|_4$ is larger $\|M_{A^{(m')}}\|_4$ for any $m \geq m'$, i.e. it increases. As for some combinations of n and m the run times become unreasonable, we skip these.

Table 5.4: Average of $\|M_{A^{(m)}}\|_4$ for random complex matrices. Run time: 3 hours.

n	N	$\|M_A\|_4$	$\|M_{A^{(2)}}\|_4$	$\|M_{A^{(3)}}\|_4$
2	50	1.080694501430	1.080694501430	1.080694501423
3	50	1.148449704359	1.148449704313	-
4	10	0.255923647109	0.255923647073	-

In particular, we have that $\|M_{A^{(m)}}\|_4$ does not grow when m grows **for any** $\|M_A\|$ that we used, and instead stay constant (up to a small numerical error). This strengthens our suspicions that for for any Schur multiplier and $p = 4$ the operator norm does not grow when we amplify the Schur multiplier. Lastly, we can do the same for other even p a well, for instance $p = 6$ and $p = 8$. The results can be seen below:

Table 5.5: Average of $\|M_{A^{(m)}}\|_6$ and $\|M_{A^{(m)}}\|_8$ for random complex matrices. Run time: 2 hours.

n	N	$\|M_A\|_6$	$\|M_{A^{(2)}}\|_6$	$\|M_{A^{(3)}}\|_6$	$\|M_A\|_8$	$\|M_{A^{(2)}}\|_8$	$\|M_{A^{(3)}}\|_8$
2	20	1.04849944	1.04849944	1.04849944	1.05123843	1.05123843	1.05123843
3	20	1.13260095	1.13260095	-	1.13673182	1.13673049	-
4	5	1.28387169	1.28387169	-	1.29026375	1.29026375	-

For no combination of Schur multiplier and p did the norm increase when amplified.

5.2.3. COMPUTATIONS FOR ARBITRARY p

Before we chose to approximate $\|M_A\|_4$ for even $p \in \mathbb{N}$, as the elementary matrix operations involved allowed somewhat reasonable run times. However, as mentioned in the beginning of the chapter, algorithms exist that can directly approximate the singular values of a matrix. Moreover, NumPy contains an implementation of such an algorithm. Although the time complexity involved for computing the singular values is the same as for matrix multiplication, namely $\mathcal{O}(n^3)$ (for square matrices), the computational cost is generally higher and therefore we can expect the run times to be larger.
Now, for $B \in M_n(\mathbb{C})$ we can compute $\|B\|_p$ by definition for any $1 \leq p \leq \infty$. As done earlier in this chapter, we can use the BFGS algorithm in an attempt to approximate $\|M_{A^{(m)}}\|_p$, where we now instead define $f_{A^{(m)}}(B) := \frac{\|M_A(B)\|_p^p}{\|B\|_p^p}$. One problem that might arise is that $f_{A^{(m)}}$ might not be twice differentiable outside of zero for all p, which in turn could lead to a worse performance of the BFGS algorithm. Nevertheless, we performed a small series of computations with the same matrices used in tables 5.2 and 5.3. The results can be seen in Table 5.6 and 5.7 respectively.

Table 5.6: Results for the algorithm where the singular values where calculated. Run time: 20 minutes.

m	$\|M_{A^{(m)}}\|_3$	$\|M_{A^{(m)}}\|_{3.5}$	$\|M_{A^{(m)}}\|_{4.5}$	$\|M_{A^{(m)}}\|_5$
1	1.00555816	1.00941559	1.01653545	1.01965451
2	1.00555816	1.00941559	1.01653545	1.01965451
3	1.00555816	1.00941559	1.01653545	1.01965451
4	1.00555816	1.00941559	1.01653545	1.01965450

Table 5.7: Results for the algorithm where the singular values where calculated. Run time: 40 minutes.

m	$\|M_{A^{(m)}}\|_3$	$\|M_{A^{(m)}}\|_{3.5}$	$\|M_{A^{(m)}}\|_{4.5}$	$\|M_{A^{(m)}}\|_5$
1	3.00876850	3.02718405	3.07097511	3.09124261
2	3.00876849	3.02718405	3.07097511	3.09124260
3	3.00876849	3.02718405	3.07097511	3.09124260

To our surprise, here too we see that $\|M_{A^{(m)}}\|_p$ does not seem to grow as m grows. In fact, much like before, the approximations only worsen. What we do see, is that $\|M_{A^{(m)}}\|_p$ grows with respect to p. Furthermore, the run times are significantly longer as well, as was to be expected.
Similar to what we did in Subsection 5.2.2, we can also perform the above computations for a number of random matrices. However, due to the added computational cost, we are more restricted in terms of the sizes of matrices we work with. The results can be seen in Table 5.8.

Table 5.8: Average of $\|M_{A^{(m)}}\|_p$ for random complex matrices with various p. Run time: 2.5 hours.

n	N	$\|M_A\|_3$	$\|M_{A^{(2)}}\|_3$	$\|M_{A^{(3)}}\|_3$	$\|M_A\|_{3.5}$	$\|M_{A^{(2)}}\|_{3.5}$	$\|M_{A^{(3)}}\|_{3.5}$
2	10	1.02609799	1.02609799	1.02609799	1.02748576	1.02748576	1.02748576
3	10	1.15783284	1.15783284	-	1.15800175	1.15800175	-

Table 5.8 continued

$\|M_A\|_{4.5}$	$\|M_{A^{(2)}}\|_{4.5}$	$\|M_{A^{(3)}}\|_{4.5}$	$\|M_A\|_5$	$\|M_{A^{(2)}}\|_5$	$\|M_{A^{(3)}}\|_5$
1.03092012	1.03092012	1.03092012	1.03260235	1.03260235	1.03260235
1.15856526	1.15856526	-	1.15990010	1.15982128	-

Once more, for not a single Schur multiplier we have that $\|M_{A^{(m)}}\|_p$ increases in m.

5.3. New conjecture and further studies

We look back to the results of the previous section. In none of the cases has the norm of a Schur multiplier increased when amplified for small m, and instead stayed constant. Based on these results, extending n and m to all natural numbers, we pose the following conjecture:

Conjecture 5.1. *Given a Schur multiplier the Schatten p-norm for $1 \leq p \leq \infty$ does not change when it is amplified, i.e.* $\|M_A\|_p = \|M_A\|_{cb}$.

That is to say, we conjecture that Pisier's conjecture does not hold. We emphasize that Conjecture 5.1 is more strongly supported for even p, as we have performed more computations for this case, and in particular $p = 4$. Although the results are only for finite-dimensional Schur multipliers, this directly extends to infinite-dimensional Schur multipliers as well. By Proposition 3.24 we can approximate $\|M_{A^{(m)}}\|_p$ using finite-dimensional Schur multipliers, and if the latter do not depend on m, neither will the limit.

Firstly, to strengthen the conjecture, more computations like those in Section 5.2 should be performed, preferably with larger matrices and amplifications. To achieve the latter in a reasonable amount of time improvements to the algorithm might need to be made (or ran on a more powerful computer).
Secondly, to prove this conjecture it would suffice to find a map which sends a maximum of $f_{A^{(m)}}$ to a maximum of f_A, as from this and Proposition 4.2 it would follow that the norms are actually equal.

In fact, it would suffice to find a map which sends a maximum of B of $f_{A^{(m)}}$ to a matrix B' where the $m \times m$ blocks are all multiples of I_m, as we have seen in Proposition 4.10. Experimentally, using the reduced variable count algorithm (i.e. we have a maximum of $f_{A^{(m)}}$ in B with B having diagonal blocks on the diagonal), and assuming the diagonal blocks are multiplies of the identity matrix, we have always found a maximum in this form. Moreover, the remaining blocks of B we found to be either Hermitian or anti-Hermitian, i.e. one could diagonalize such a block using unitary matrices. Now suppose such an $m \times m$ block is diagonalized using U (an $m \times m$ matrix as well). If we then conjugate B with the matrix

$$U' = \begin{pmatrix} U & & \\ & \ddots & \\ & & U \end{pmatrix},$$

we would retain the maximum and gain one block in diagonal form. More so, the blocks that were already a multiple of the identity matrix remain so as well. Should this new diagonal block also be a multiple of the identity matrix, and should another non-diagonal (anti-) Hermitian block be found, the procedure can be repeated, until hopefully B is in the desired form. However, the writer's attempts in proving the existence of matrices on which $f_{A^{(m)}}$ attains a maximum and meet the aforementioned conditions have not proven successful. Explicit derivatives can in theory be calculated, but these would form large expressions.

6

Conclusion

In this thesis we have looked at a subspace of the compact operators on a separable Hilbert space, namely the Schatten classes $S_p(H)$, which consists of compact operators of which the singular values are p-summable. On a Schatten class we can define a Schur multiplier which sends bounded linear maps from $S_p(H)$ to $S_p(H)$, which "multiplies" elements from $S_p(H)$ with with a matrix A. The latter is achieved by considering the linear maps as matrices with respect to a basis. An important result is that we can approximate Schatten class operators using projections on the first n elements of an orthonormal basis, not only in the operator norm, but in the Schatten p-norm as well.

Following this, we introduced the concept of amplifying Schur multipliers by a natural number m. Furthermore, should the norm of the amplified Schur multiplier be bounded for all m, then the Schur multiplier was said to be completely bounded.
Subsequently, we were able to state a conjecture posed by Pisier [10], in which he conjectures that for $p \notin \{1, 2, \infty\}$ there exist bounded Schur multipliers that are not completely bounded, i.e. there exists Schur multipliers of which the norm grows when amplified.
From here on we proved some relations and reductions between a Schur multiplier an its amplified equivalent for $1 \leq p \leq \infty$, which focused around unitary matrices. Here for instance we found that the operator norm can only increase when a Schur multiplier is amplified. Moreover, for finite-dimensional Schur multipliers we can always find the maximum in a matrix a number of diagonal blocks, i.e. fewer non-zero entries.

These results we were in turn able to use in numerically approximating both fixed and random Schur multipliers using the Broyden–Fletcher–Goldfarb–Shanno algorithm. The majority of the computations were performed for $p = 4, 6$, and 8, and we performed a smaller number of computations for several non-odd p. Although the problematic scaling of the run-time prevented us from amplifying Schur multipliers by larger numbers, surprisingly, the computations showed that for the random Schur multipliers we had taken, the operator norm did not increase, and instead stayed equal. Based on these results we posed a new conjecture that for even p bounded Schur multipliers are completely bounded as well, and moreover, the completely bounded norm is equal to the operator norm. That is, we conjecture that for these p Pisier's conjecture does not hold. Lastly, we discussed a possible way to prove the aforementioned new conjecture.

Bibliography

[1] J. Bell. The singular value decomposition of compact operators on Hilbert spaces, 2014.

[2] M. Caspers and M. De La Salle. Schur and Fourier multipliers of an amenable group acting on non-commutative L^p-spaces. *Transactions of the American Mathematical Society*, 2015.

[3] J. B. Conway. *A course in functional analysis*, volume 96. Springer Science & Business Media, 2013.

[4] A. Harcharras. Fourier analysis, Schur multipliers on S_p and non-commutative $\Lambda(p)$-sets. Citeseer, 1999.

[5] T. Hytönen, J. van Neerven, M. Veraar, and L. Weis. *Analysis in Banach spaces*. Springer, 2016.

[6] E. Jones, T. Oliphant, P. Peterson, et al. SciPy: Open source scientific tools for Python, 2001. [Online; accessed 2018-05-26].

[7] S. Neuwirth and É. Ricard. Transfer of fourier multipliers into schur multipliers and sumsets in a discrete group. 2010.

[8] J. Nocedal and W. S. J. *Numerical Optimization*. Springer, 2006.

[9] T. E. Oliphant. *A guide to NumPy*, volume 1. Trelgol Publishing USA, 2006.

[10] G. Pisier. Non-commutative vector valued L_p-spaces and completely p-summing maps. *Asterisque-Societe Mathematique de France*, 247, 1998.

[11] B. P. Rynne and M. A. Youngson. *Linear functional analysis*. Springer Science & Business Media, 2000.

[12] N. Yu, R. Duan, and Q. Xu. Bounds on the distance between a unital quantum channel and the convex hull of unitary channels, with applications to the asymptotic quantum Birkhoff conjecture. 2012.

A

A.1. Python code

```
import numpy as np
import scipy.optimize
from datetime import datetime
from math import sqrt


class schur():
    """A class for finite dimensional Schur Multipliers"""

    def __init__(self, A, p=2, complex_valued=False):
        """Initialises a Schur Multiplier using A, where the p-norm is used.
        If complex_valued = True then the operator norm is approached using
        complex matrices, and otherwise only real matrices."""
        self.A_orig = A
        self.A = A
        self.n_orig = A.shape[0]
        self.n = A.shape[0]
        self.p = p
        self.k = 0
        self.reduc_insert = []
        self.init_basis()
        self.amp(1)
        self.cvals = complex_valued

    def amp(self, m):
        """Amplifies the matrix by a factor m. Resets the matrix with m = 1 and
        initializes self.basis_vectors."""
        self.A = np.repeat(self.A_orig, m, axis=0)
        self.A = np.repeat(self.A, m, axis=1)
        self.n = m * self.n_orig
        self.init_basis()

        self.reduc_insert = []
        self.k = self.n**2 - self.n**2//self.n_orig + self.n
        curr_idx = self.k - 1
        for block_row in range(self.n_orig - 1, -1, -1):
            curr_l = m - 1
            for row in range(m - 1, -1, -1):
                self.reduc_insert.extend([curr_idx + 1] * (m-1-curr_l))
                self.reduc_insert.extend([curr_idx] * curr_l)
                curr_l -= 1
                curr_idx -= (self.n - (m - 1))
            curr_idx -= m
        self.reduc_insert.reverse()

    def p_norm_power(self, B):
        "Calulates the p-th power Schatten p-norm for even integer p."
        if type(self.p) != int:
```

```
            svs = np.linalg.svd(B)[1]
            if self.p == float("inf"):
                return svs.max()
            else:
                svs = svs**self.p
                return svs.sum()
        B_adj = B.conjugate().transpose()
        temp = B_adj.dot(B)
        res = temp
        for i in range(1, self.p//2):
            res = res.dot(temp)
        tr = res.trace()
        return abs(tr)

    def p_norm(self, B):
        "Calculates the Schatten-p-norm."
        if self.p == float("inf"):
            return self.p_norm_power(B)
        return self.p_norm_power(B)**(1/self.p)

    def init_basis(self):
        "Creates the standard basis for M_n(C)"
        self.basis_vectors = {}
        n = self.n
        for i in range(n):
            self.basis_vectors[i] = {}
            for j in range(n):
                E_ij = np.zeros((n, n), dtype=np.complex128)
                E_ij[i, j] = 1
                self.basis_vectors[i][j] = E_ij

    def gradient(self, B, delta=0.001):
        "Returns the (numerically calculated) gradient of schur.fun in B."
        grad = np.zeros((self.n, self.n))
        for i in range(self.n):
            for j in range(self.n):
                B1 = B - delta*self.basis_vectors[i][j]
                B2 = B + delta*self.basis_vectors[i][j]
                B1i = B - 1j*delta*self.basis_vectors[i][j]
                B2i = B + 1j*delta*self.basis_vectors[i][j]
                grad[i][j] = (self.fun(B2) - self.fun(B1))/(2*delta)
                grad[i][j] += (self.fun(B2i) - self.fun(B1i))/(2*delta)
        return grad

    def fun(self, B):
        "Calculates ||AB||/||B||, i.e. a lower bound for the operator norm."
        return self.p_norm_power(self.A*B)/self.p_norm_power(B)

    def f_negative(self, B):
        "Returns -f and reshapes the argument if necessary."
        if len(B.shape) == 1 and self.cvals:
            x1 = B[:len(B)//2]
            x2 = B[len(B)//2:]
            B = x1.reshape(self.A.shape) + 1j*x2.reshape(self.A.shape)
        elif len(B.shape) == 1:
            B = B.reshape(self.A.shape)
        return -self.fun(B)

    def operator_norm(self):
        """Approximates the operator norm of the Schur multiplier (from below).
        Using the BFGS method from scipy. Returns the norm and the associated
        element."""
        norm = 0.0
        for i in range(self.n**2):
            if self.cvals:
                B = np.random.rand(2*self.n**2)
            else:
                B = np.random.rand(self.n**2)
            t = scipy.optimize.minimize(self.f_negative, B, tol=10**-15,
                                        method='BFGS')
            val = -t.fun
```

```python
            if val > norm:
                norm = val
                out = t.x
        if self.cvals:
            x1 = out[:len(out)//2]
            x2 = out[len(out)//2:]
            out = x1.reshape(self.A.shape) + 1j*x2.reshape(self.A.shape)
        else:
            out = np.reshape(out, self.A.shape)
        return norm**(1/self.p), out/self.p_norm(out)

    def f_reduc(self, x):
        if self.cvals:
            x1 = x[:len(x)//2]
            x2 = x[len(x)//2:]
            B = np.insert(x1, self.reduc_insert, 0) \
            + 1j*np.insert(x2, self.reduc_insert, 0)
        else:
            B = np.insert(x, self.reduc_insert, 0)
        B = B.reshape(self.A.shape)
        return self.fun(B)

    def f_neg_reduc(self, B):
        return -self.f_reduc(B)

    def operator_norm_reduced(self):
        """Approximates the operator norm of the Schur multiplier (from below).
        Using the BFGS method from scipy using some digonal blocks."""
        norm = 0.0
        for i in range(self.k):
            if self.cvals:
                B = np.random.rand(2*self.k)
            else:
                B = np.random.rand(self.k)
            t = scipy.optimize.minimize(self.f_neg_reduc, B, tol=10**-15,
                                        method='BFGS')
            val = -t.fun
            if val > norm:
                norm = val
                out = t.x
        if self.cvals:
            x1 = np.insert(out[:len(out)//2], self.reduc_insert, 0)
            x2 = np.insert(out[len(out)//2:], self.reduc_insert, 0)
            out = x1.reshape(self.A.shape) + 1j*x2.reshape(self.A.shape)
        else:
            out = np.insert(out, self.reduc_insert, 0)
        out = np.reshape(out, self.A.shape)
        return norm**(1/self.p), out/self.p_norm(out)

    def __str__(self):
        return str(self.A)


if __name__ == "__main__":
    np.warnings.filterwarnings('ignore')
    start = datetime.now()
    A = [[1, 1],
         [1, 1.0/2]    ]
    A = np.array(A, dtype=np.float64)  #operator norm: 1.0131151203640094
    A_mult = schur(A, p=4, complex_valued=False)
    for k in range(1, 4):
        A_mult.amp(k)
        A_norm, out = A_mult.operator_norm_reduced()
        print ("Matrix amplfied by {}, norm: {}".format(k, A_norm))
    print ("Elapsed time:", datetime.now() - start)
```