p-Schatten Embeddings

Peter Melech

Contents

1 Introduction **3**

2 Preliminaries **5**

3 Properties of the p-Schatten classes **12**
3.1 General Properties 12
3.2 The Hilbert-Schmidt Class 17
3.3 The Trace Class 18

4 Compact embeddings **22**

5 Embeddings of Sobolev spaces **28**
5.1 Sobolev Spaces 28
5.2 Maurin's Theorem 31
5.3 Gramsch 's Theorem 32

6 Characterization of the p-Schatten classes **36**

7 Applications **45**
7.1 The spaces $s_{*,\alpha}$ 45
7.2 Discrete Sobolev spaces on weighted graphs 47
7.3 Discrete Sobolev spaces on sparse graphs 50

8 Open problems **54**

1 Introduction

The Rellich-Kondrachov compactness theorem says the embeddings

$$W^{j+m,q}(\Omega) \hookrightarrow W^{j,r}(\Omega)$$

are compact embeddings of Sobolev spaces if Ω is a bounded open subset of the space $\mathbb{R}^N$ that has the cone property and if either $mq \geq N$ or if $mq < N$ and $1 \leq r < \frac{qN}{N-mq}$ (cf. [1, Theorem 6.2]). In this case it follows for the Hilbert spaces $H^k(\Omega) = W^{k,2}(\Omega)$ and $H_0^k(\Omega) = W_0^{k,2}(\Omega)$ that the embeddings

$$H^k(\Omega) \hookrightarrow H^l(\Omega), H_0^k(\Omega) \hookrightarrow H_0^l(\Omega)$$

are compact if $k > l$. The Rellich-Kondrachov theorem is fundamental for the study of elliptic boundary value problems. The sets of non-zero singular values of these compact embeddings are discrete bounded sets. The result was improved by Maurin [17] in the sense that the embeddings are of Hilbert-Schmidt class (that means the singular values are even square summable) if $k - l > \frac{N}{2}$ and the boundary $\partial\Omega$ fulfills Sobolev conditions (see also [1]). This was useful for eigenvalue distributions, eigenfunction expansions corresponding to differential operators and integral representations of Green's functions (cf. [13],[1]). From Maurin's theorem it follows immediately that these embeddings are even of trace class if $k - l > N$, i.e. their singular values yield sequences in $\ell^1(\mathbb{N})$. The question arises if there is a way to characterize the embeddings of Sobolev spaces with singular values yielding sequences in $\ell^p(\mathbb{N})$ for $0 < p < \infty$. Classes of mappings with these properties are called p-Schatten classes. p-Schatten classes were first studied by Robert Schatten and John von Neumann as ideals of the ring of bounded linear operators on a Hilbert space $\mathcal{H}$. The characterization was achieved by Gramsch [13]: The embeddings are of p-Schatten class for $0 < p < \infty$ if and only if $k - l > \frac{N}{p}$. While Maurins proof is based on Sobolev's Lemma Gramsch uses the knowledge of an orthogonal basis of the Sobolev spaces on N-dimensional tori T^N and then factorizes Ω over T^N via a partition of one.

In [14] Hanche-Olsen and Holden achieve a very simple characterization of totally bounded subsets of the spaces $\ell^q(\mathbb{N})$ for $1 \leq q < \infty$ thus characterizing the compact embeddings

$$X \hookrightarrow \ell^q(\mathbb{N})$$

where X is some metric space. In [21] Pietsch defines generalized p-Schatten classes of mappings between Banach spaces. One would like to have a characterization of the p-Schatten embeddings

$$X \hookrightarrow \ell^q(\mathbb{N})$$

where X is some Banach space over $\mathbb{C}$ or at least a characterization of the classical p-Schatten embeddings

$$\mathcal{H} \hookrightarrow \ell^2(\mathbb{N})$$

where $\mathcal{H}$ is some complex separable Hilbert space. Our aim is to achieve this characterization. We do this by applying a very general result on p-Schatten mappings to embeddings.

In [11] Gohberg and Markus characterize the p-Schatten classes of operators on one Hilbert space $\mathcal{H}$. It is shown that a compact operator H on a Hilbert space $\mathcal{H}$ is of p-Schatten class if and only if for all orthonormal bases $(2 \leq p < \infty)$/some orthonormal basis $(0 < p \leq 2)$ $\{\phi_j\}$ of $\mathcal{H}$ one has

$$\sum_j |H\phi_j|^p < \infty. \tag{1.1}$$

Their proof is quite involved. We give a simpler proof for the characterization of p-Schatten classes of compact mappings

$$A : \mathcal{H}_1 \to \mathcal{H}_2$$

where $\mathcal{H}_1$ and $\mathcal{H}_2$ are complex separable Hilbert spaces showing that A is a p-Schatten mapping if and only if for all orthonormal bases $(2 \leq p < \infty)$/some orthonormal basis $(0 < p \leq 2)$ $\{\phi_j\}$ of $\mathcal{H}_1$ one has

$$\{|A\phi_j|_{\mathcal{H}_2}\}_{j=1}^{\infty} \in \ell^p(\mathbb{N}) \tag{1.2}$$

where $|.|_{\mathcal{H}_2}$ is the norm induced by the inner product of $\mathcal{H}_2$.

In the range $2 \leq p < \infty$ the proof involves ideas that are elaborated on by Simon in [25] and are based on a theorem that is essentially due to Markus . In the range $0 < p < 2$ an idea by Dunford and Schwartz [8] plays a major roll. This very general result can be applied to embeddings. In this way it is possible to give a characterization of the p-Schatten embeddings $\mathcal{H}_1 \hookrightarrow \ell^2(\mathbb{N})$. However it is in general not possible to describe all orthonormal bases of a space $\mathcal{H}_1 \subset \ell^2(\mathbb{N})$. If an orthonormal basis is known it turns out to be useful especially in the range $0 < p \leq 2$ for what we give several examples.

At this point I would like to thank Professor Mugnolo for giving me this interesting topic and for his idea to apply the theory to Sobolev spaces on sparse graphs (see section 7.3). I would like to thank Dr. Kerner for many suggestions for improvement. Finally I want to thank my family for their support.

2 Preliminaries

In this chapter we present the basic notions that are needed to define p-Schatten operators, p-Schatten mappings and p-Schatten embeddigs. Throughout this chapter and the following chapters $\mathcal{H}$, $\mathcal{H}_1$ and $\mathcal{H}_2$ are complex separable Hilbert spaces. Since our focus is on infinite dimensional spaces we assume these spaces to be infinite dimensional. Most results hold for finite dimensional spaces too. We assume that the inner products are linear in the first and conjugate linear in the second factor and denote them by $(.,.)$, $(.,.)_{\mathcal{H}_1}$ and $(.,.)_{\mathcal{H}_2}$ and the induced norms by $|.|,|.|_{\mathcal{H}_1}$ and $|.|_{\mathcal{H}_2}$. We start with the following definitions:

Definition 1: *We say for a bounded linear operator $H \in \mathcal{L}(\mathcal{H})$ it is positive, in signs $H \geq 0$, if $(\phi, H\phi) \geq 0$ for all $\phi \in \mathcal{H}$.*

For the next definition we recall that since $\phi \mapsto (H\phi, \psi)$ is a bounded linear functional for each ψ Riesz's representation theorem implies the existence of the operator $H^* \in \mathcal{L}(\mathcal{H})$ satisfying $(H\phi, \psi) = (\phi, H^*\psi)$ and this operator is called the adjoint operator.

Definition 2: *$H \in \mathcal{L}(\mathcal{H})$ is called self-adjoint if $H = H^*$.*

The next lemma characterizes the self-adjoint operators by their diagonal values.

Lemma 1: *H is self-adjoint if and only if $(\phi, H\phi) \in \mathbb{R}$ for all $\phi \in \mathcal{H}$.*

PROOF: If H is self-adjoint we have:

$$(\phi, H\phi) = (H\phi, \phi) = \overline{(\phi, H\phi)}$$

for all $\phi \in \mathcal{H}$ and thus $(\phi, H\phi) \in \mathbb{R}$ for all $\phi \in \mathcal{H}$. If $(\phi, H\phi) \in \mathbb{R}$ for all $\phi \in \mathcal{H}$ we have:

$$(\phi, (H - H^*)\phi) = (\phi, H\phi) - (\phi, H^*\phi) = (\phi, H\phi) - (H\phi, \phi) = (\phi, H\phi) - \overline{(\phi, H\phi)} = 0$$

and by polarization:

$$(\phi, (H - H^*)\psi) = \frac{1}{4}((\phi + \psi, (H - H^*)(\phi + \psi)) - (\phi - \psi, (H - H^*)(\phi - \psi)))$$
$$+\frac{i}{4}((\phi + i\psi, (H - H^*)(\phi + i\psi)) - (\phi - i\psi, (H - H^*)(\phi - i\psi))) = 0$$

for all $\phi, \psi \in \mathcal{H}$, e.g. $\phi = (H - H^*)\psi$, thus $|(H - H^*)\psi| = 0$ for all ψ and so $H - H^* = 0$. □

Hence every positive operator $H \in \mathcal{L}(\mathcal{H})$ is self-adjoint. Without proof we cite [23, Theorem VI.9] the so called "square root lemma":

Theorem 2: *For every positive operator $H \in \mathcal{L}(\mathcal{H})$ there is a unique positive operator $H_1 \in \mathcal{L}(\mathcal{H})$ satisfying $H_1^2 = H$ and H_1 commutes with every $H_2 \in \mathcal{L}(\mathcal{H})$, that commutes with H.*

Now let $\mathcal{H}_1, \mathcal{H}_2$ be two complex, separable Hilbert spaces and define in analogy to the case of one Hilbert space:

Definition 3: *For a bounded linear mapping $A : \mathcal{H}_1 \to \mathcal{H}_2$ its adjoint mapping $A^* : \mathcal{H}_2 \to \mathcal{H}_1$ is the bounded linear mapping satisfying*

$$(A\phi, \psi)_{\mathcal{H}_2} = (\phi, A^*\psi)_{\mathcal{H}_1}$$

for all $\phi \in \mathcal{H}_1, \psi \in \mathcal{H}_2$.

Though all separable Hilbert spaces are isomorphic, we follow Maurin [17] in distinguishing bounded linear mappings between Hilbert spaces from operators on one Hilbert space. We observe:

Lemma 3: *$A^*A \in \mathcal{L}(\mathcal{H}_1)$ is a positive operator, that is self-adjoint and has a positive and thus self-adjoint square root.*

Proof: This follows by Lemma 1 and Theorem 2.

Definition 4: *For a bounded linear mapping $A : \mathcal{H}_1 \to \mathcal{H}_2$ we define the positive mapping $|A| \in \mathcal{L}(\mathcal{H}_1)$ by $|A| = (A^*A)^{\frac{1}{2}}$.*

We observe, that for every $\phi \in \mathcal{H}_1$:

$$||A|\phi|^2_{\mathcal{H}_1} = (|A|\phi, |A|\phi)_{\mathcal{H}_1} = (A^*A\phi, \phi)_{\mathcal{H}_1} = (A\phi, A\phi)_{\mathcal{H}_2} = |A\phi|^2_{\mathcal{H}_2}. \tag{2.1}$$

In the next step we factorize A where one factor is $|A|$ and the other factor is a partial isometry. We recall that for a mapping $U : \mathcal{H}_1 \to \mathcal{H}_2$ the kernel $KernU$ is a closed subspace, thus has an orthogonal complement that we denote by $KernU^{\perp}$. We denote the range of U by $RanU$.

Definition 5: *We call $U : \mathcal{H}_1 \to \mathcal{H}_2$ a partial isometry , if $U|_{KernU^{\perp}} : KernU^{\perp} \to RanU$ is an isometry.*

In this case $RanU$ is closed:

Lemma 4: *If $U : \mathcal{H}_1 \to \mathcal{H}_2$ is a partial isometry its range $RanU$ is closed.*

Proof: If $\{\psi_n\}$ is a Cauchy sequence in $RanU$ and $\psi_n \to \psi$ in $\mathcal{H}_2$, then $\{\phi_n\}$ given by $\phi_n = (U|_{KernU^{\perp}})^{-1}\psi_n$ is a Cauchy sequence in $KernU^{\perp} \subseteq \mathcal{H}_1$, because

$$|\phi_n - \phi_m|_{\mathcal{H}_1} = |(U|_{KernU^{\perp}})^{-1}(\psi_n - \psi_m)|_{\mathcal{H}_1} = |\psi_n - \psi_m|_{\mathcal{H}_2} \to 0, n, m \to \infty.$$

Thus $\{\phi_n\}$ has a limit ϕ in $KernU^{\perp}$, which is closed. It follows:

$$|\phi - \phi_n|_{\mathcal{H}_1} = |U(\phi - \phi_n)|_{\mathcal{H}_2} = |U\phi - \psi_n|_{\mathcal{H}_2} \to 0, n \to \infty,$$

and by uniqueness of the limit $\psi = U\phi \in RanU$. □

Now we are interested in the adjoint of a partial isometry. It turns out that it is again a partial isometry.

Lemma 5: *If $U:\mathcal{H}_1\to\mathcal{H}_2$ is a partial isometry its adjoint $U^*:\mathcal{H}_2\to\mathcal{H}_1$ is a partial isometry.*

Proof: By what we have just seen:

$$\mathcal{H}_2 = RanU \oplus RanU^{\perp}.$$

So for every $\psi\in\mathcal{H}_2$ there exist unique $\phi\in KernU^{\perp}, \rho\in RanU^{\perp}$, such that

$$\psi = U\phi+\rho. \tag{2.2}$$

Now define $U':\mathcal{H}_2\to\mathcal{H}_1$ by $U'(\psi)=\phi$, then $KernU' = RanU^{\perp}$ or equivalently $RanU = KernU'^{\perp}$ and $(U'(U\phi),U'(U\phi)) = (\phi,\phi) = (U\phi,U\phi)$ for every $U\phi\in RanU$. That means $U'|_{KernU'^{\perp}}$ is an isometry. We further have for $\psi = U\phi+\rho\in\mathcal{H}_2$ and $\phi_1=\phi_1^1+\phi_1^2\in\mathcal{H}_1$, where $\phi_1^1\in KernU^{\perp}, \phi_1^2\in KernU (\mathcal{H}_1 = KernU^{\perp}\oplus KernU)$:

$$(\psi,U\phi_1)_{\mathcal{H}_2} = (U\phi+\rho,U\phi_1)_{\mathcal{H}_2} = (U\phi,U\phi_1^1)_{\mathcal{H}_2} = (\phi,\phi_1^1)_{\mathcal{H}_1} = (U'\psi,\phi_1^1)_{\mathcal{H}_1} = (U'\psi,\phi_1)_{\mathcal{H}_1},$$

where the latter identity follows from $U'\psi\in KernU^{\perp}$.
So $U'=U^*$ is a partial isometry with initial space $KernU'^{\perp} = RanU$ and final space $RanU' = KernU^{\perp}$. □

Note that U^*U is an orthogonal projection on $KernU^{\perp}\subseteq\mathcal{H}_1$ and UU^* is an orthogonal projection on $RanU\subseteq\mathcal{H}_2$. We use these results to prove the existence of a polar decomposition of a bounded linear mapping as described in the following theorem (cf. [23, Theorem VI.10]). The polar decomposition will play an important roll in our discussion.

Theorem 6: *For every bounded linear mapping $A:\mathcal{H}_1\to\mathcal{H}_2$ there exists a unique partial isometry $U:\mathcal{H}_1\to\mathcal{H}_2$, such that $A=U|A|$ and $KernU = KernA$*

Proof: We first define $U_0: Ran|A|\to RanA$ by $U_0(|A|\phi) = A\phi$. If $|A|\phi = |A|\psi$, i.e. $||A|(\phi-\psi)|_{\mathcal{H}_1}=0$, it follows by equation 2.1 $|A(\phi-\psi)|_{\mathcal{H}_2}=0$, i.e. $A\phi=A\psi$, so U_0 is well-defined. And again by equation 2.1 U_0 is an isometry.Because U_0 is an isometry, it can be extended to an isometry $U_1:\overline{Ran|A|}\to\overline{RanA}$. Now since $\mathcal{H}_1 = \overline{Ran|A|}\oplus\overline{Ran|A|}^{\perp}$the mapping $U:\mathcal{H}_1\to\mathcal{H}_2$ defined by $U|_{\overline{Ran|A|}} = U_1$ and $U|_{\overline{Ran|A|}^{\perp}}=0$ is a partial isometry on $\mathcal{H}_1$.
If $(|A|\phi,\psi)_{\mathcal{H}_1}=0$ for all $\phi\in\mathcal{H}_1$it follows $(\phi,|A|\psi)_{\mathcal{H}_1}=0$ and so $|A|\psi=0$ and similarly , if $|A|\psi=0$ it follows $(|A|\phi,\psi)_{\mathcal{H}_1}=0$ for all $\phi\in\mathcal{H}_1$. So $Ran|A|^{\perp} = Kern|A|$ and again by 2.1 $KernA = Kern|A|$.Thus $KernU = KernA$ and clearly $A=U|A|$.
$\mathcal{H}_1 = \overline{Ran|A|}\oplus\overline{Ran|A|}^{\perp}$ together with this implies, that U is unique. □

Note that $|A| = U^*A$.
We say a bounded linear mapping is of finite rank, if its image is finite dimensional.
We now consider compact mappings (cf. [7, Corollary 6.2]):

Theorem 7: *For a bounded linear mapping $A : \mathcal{H}_1 \to \mathcal{H}_2$ the following two conditions are equivalent:*
(i) A is a limit (in the operator norm) of finite rank mappings
(ii) $A(\{\phi \in \mathcal{H}_1 : |\phi|_{\mathcal{H}_1} \le 1\})$ is precompact in $\mathcal{H}_2$

PROOF: It is well known, that the unit ball in a Banach space is precompact exactly if this space is of finite dimension. Thus every finite rank mapping has property (ii). So if A has property (i), for every $\varepsilon > 0$ there exist $n(\varepsilon), m(\varepsilon) \in \mathbb{N}, A_{n(\varepsilon)} : \mathcal{H}_1 \to \mathcal{H}_2$ of finite rank and $\psi_1, ..., \psi_{m(\varepsilon)} \in \mathcal{H}_2$ such that

$$||A_{n(\varepsilon)} - A|| < \frac{\varepsilon}{2}$$

where $||.||$ denotes the operator norm and

$$A_{n(\varepsilon)}(B^1_{\mathcal{H}_1}(0)) \subset \bigcup_{j=1}^{m(\varepsilon)} B^{\frac{\varepsilon}{2}}_{\mathcal{H}_2}(\psi_j)$$

where $B^{\alpha}_{\mathcal{H}_i}(\eta)$ is the ball of radius α centered at η in $\mathcal{H}_i$. Then

$$A(B^1_{\mathcal{H}_1}(0)) \subset \bigcup_{j=1}^{m(\varepsilon)} B^{\varepsilon}_{\mathcal{H}_2}(\psi_j).$$

So A has property (ii).
Conversely if A has property (ii), i.e.

$$A(B^1_{\mathcal{H}_1}(0)) \subset \bigcup_{j=1}^{m(\varepsilon)} B^{\varepsilon}_{\mathcal{H}_2}(\psi_j)$$

for some arbitrarily small $\varepsilon > 0$ and $n < \infty$, then let P be the projection on the finite dimensional subspace spanned by $\psi_1, ..., \psi_n \in \mathcal{H}_2$, then PA is of finite rank and for any $\phi \in B^1_{\mathcal{H}_1}(0)$ there is some $j_\phi \in \{1, ..., n\}$ such that

$$|A\phi - \psi_{j_\phi}|_{\mathcal{H}_2} < \varepsilon$$

and thus

$$|PA\phi - P\psi_{j_\phi}|_{\mathcal{H}_2} = |PA\phi - \psi_{j_\phi}|_{\mathcal{H}_2} < \varepsilon,$$

because P is a contraction. This all together implies

$$||PA - A|| < 2\varepsilon,$$

i.e. A has property (i). □

Since these two properties of bounded linear mappings between Hilbert spaces are equivalent by the above theorem compactness might be defined by property (i) as well as by property (ii). However in the more general case of Banach spaces (ii) does not imply (i) and the problem is known as the "approximization problem". (For this see Remark 1 to Corollary 6.2 in [7] and Section 1.g in [15].) We choose property (i) to define compact mappings.

Definition 6: *A mapping $A : \mathcal{H}_1 \to \mathcal{H}_2$ is called compact if it is the limit of finite rank mappings.*

If $\mathcal{J}_\infty(\mathcal{H}_1, \mathcal{H}_2)$ denotes the set of compact linear mappings then the following Lemma holds.

Lemma 8: *For Hilbert spaces $\mathcal{H}_0, \mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ and bounded linear mappings $B : \mathcal{H}_0 \to \mathcal{H}_1$, $A \in \mathcal{J}_\infty(\mathcal{H}_1, \mathcal{H}_2)$ and a bounded linear mapping $C : \mathcal{H}_2 \to \mathcal{H}_3$ we have $AB \in \mathcal{J}_\infty(\mathcal{H}_0, \mathcal{H}_2)$, $CA \in \mathcal{J}_\infty(\mathcal{H}_1, \mathcal{H}_3)$ and $CAB \in \mathcal{J}_\infty(\mathcal{H}_0, \mathcal{H}_3)$. .*

Proof: If A is compact there exist finite rank operators $A_n : \mathcal{H}_1 \to \mathcal{H}_2$ so that

$$||A - A_n|| \to 0, n \to \infty.$$

Clearly $A_n B : \mathcal{H}_0 \to \mathcal{H}_2$ are finite rank operators and

$$||AB - A_n B|| \leq ||A - A_n|| ||B|| \to 0, n \to \infty.$$

The rest is shown in the same way. □

So in the case of compact operators $\mathcal{J}_\infty(\mathcal{H}, \mathcal{H}) = \mathcal{J}_\infty(\mathcal{H})$ is a two sided ideal in the ring of bounded linear operators. This explains the notation. From Lemma 8 we see that if $A : \mathcal{H}_1 \to \mathcal{H}_2$ is compact then $|A| = U^* A \in \mathcal{L}(\mathcal{H}_1)$ is compact and if $|A|$ is compact then $A = U|A|$ is compact. Moreover if A is compact $|A|$ is a compact, positive and thus self-adjoint linear operator. To such operators applies the following Theorem (cf. [7, Theorem 6.8 and Theorem 6.11]).We recall that the spectrum of an operator $H \in \mathcal{L}(\mathcal{H})$ is defined as the complement of its resolvent set, i.e. $\sigma(H) = \mathbb{C} \backslash \rho(H)$ where

$$\rho(H) = \{\lambda \in \mathbb{C} : Kern(\lambda id_{\mathcal{H}} - H) = \{0\} \text{ and } Ran(\lambda id_{\mathcal{H}} - H) = \mathcal{H}\}.$$

Note that by the open mapping theorem the operators $\lambda id_{\mathcal{H}} - H$ have a bounded inverse for $\lambda \in \rho(H)$.

Theorem 9: *If $H \in \mathcal{J}_\infty(\mathcal{H})$ is self-adjoint then there exists an orthonormal basis of $\mathcal{H}$ consisting of eigenvectors of H and the spectrum consists of 0 and the non-zero eigenvalues of H and either*

$$\sigma(H) = \{0\} \text{ i.e. } H = 0$$

or

$$\sigma(H) \backslash \{0\} \ is \ finite$$

or

$$\sigma(H) \backslash \{0\} \text{ is a sequence converging to } 0.$$

If $\sigma(H) \backslash \{0\}$ is finite there exists a basis consisting of eigenvectors where only a finite number of these eigenvectors correspond to non-zero eigenvalues. By applying the operator H to this basis one sees H is a finite rank operator.

We apply this theorem to the operator $|A| \in \mathcal{L}(\mathcal{H}_1)$ connected to the compact linear mapping $A : \mathcal{H}_1 \to \mathcal{H}_2$ as defined in Definition 4. Proofs can be found in almost every book on functional analysis like in [7] as cited above . The following definition and the following theorem are important for the discussion of p-Schatten mappings.

Definition 7: *If $A : \mathcal{H}_1 \to \mathcal{H}_2$ is a compact linear mapping then the elements of the spectrum of $|A| \in \mathcal{L}(\mathcal{H}_1)$ are denoted $\sigma_j(A)$ and are called the **singular values** of A.*

In Definition 7 the index j is assumed to run through some set J. As we have seen $|A|$ is compact if A is compact and in this case the spectrum is discrete by Theorem 9. So the index can be assumed to run through $\mathbb{N}$. We will consider the sequence $\{\sigma_j(A)\}_{j=1}^{\infty}$ as the sequence of the singular values in non-increasing order and that implies we consider the sequence of the **non-zero** singular values if $\sigma(|A|)\backslash\{0\}$ is infinite which in this case is a sequence converging to 0. Note that since $|A|$ is positive the singular values are non-negative real numbers. Now we are able to expand A as described below (cf. [25, Theorem 1.4]).

Theorem 10: *If $A \in \mathcal{J}_\infty(\mathcal{H}_1, \mathcal{H}_2)$ then A has a norm convergent expansion*

$$A = \sum_{j=1}^{\omega} \sigma_j(A)(., \phi_j)_{\mathcal{H}_1} \psi_j \tag{2.3}$$

where $\omega \in \mathbb{N} \cup \{0, \infty\}$, $\sigma_j(A)$ are the non-zero singular values of A in non-increasing order counting multiplicities, $\sigma_j^2(A)$ are the non-zero eigenvalues of A^*A and AA^* in non-increasing order counting multiplicities, $\{\phi_j\}_{j=1}^{\omega}$ is an orthonormal family in $\mathcal{H}_1$ consisting of the corresponding eigenvectors of A^*A and $\{\psi_j\}_{j=1}^{\infty}$ is an orthonormal family in $\mathcal{H}_2$ consisting of the corresponding eigenvectors of AA^*.

Proof: If $A = 0$ we choose $\omega = 0$ and the right hand side is understood as an empty sum. If $A \neq 0$ then it follows from Theorem 9 that there is an expansion

$$|A| = \sum_{j=1}^{\omega} \sigma_j(A)(., \phi_j)_{\mathcal{H}_1} \phi_j$$

where $1 \leq \omega \leq \infty$, $\sigma_j(A)$ are the non-zero singular values of A in non-increasing order and ϕ_j the corresponding orthonormal eigenvectors. Applying the partial isometry $U : \mathcal{H}_1 \to \mathcal{H}_2$ derived from the polar decomposition of A to the equation yields

$$A = U|A| = \sum_{j=1}^{\omega} \sigma_j(A)(., \phi_j)_{\mathcal{H}_1} U\phi_j$$

and because U is an isometry on $Ran|A|$ and clearly $\phi_j \in Ran|A|$ it follows that $\{\psi_j\}_{j=1}^{\omega}$ where $\psi_j = U\phi_j$ defines an orthonormal family. Furthermore $A^*A\phi_j = |A|^2\phi_j = \sigma_j(A)^2\phi_j$ for $j = 1, ..., \omega$ shows $\{\phi_j\}$ are the orthonormal eigenvectors of A^*A to the eigenvalues $\sigma_j(A)^2$ and

$$AA^*\psi_j = U|A| \underbrace{A^*U}_{=(U^*A)^*=|A|^*=|A|} \phi_j = U|A|^2\phi_j = \sigma_j(A)^2 U\phi_j = \sigma_j(A)^2\psi_j$$

for $j = 1, ..., \omega$ shows $\{\psi_j\}_{j=1}^{\omega}$ are the orthonormal eigenvectors of AA^* to the eigenvalues $\sigma_j(A)^2$. □

We are now able to define the p-Schatten classes for $0 < p < \infty$.

Definition 8: *A compact linear mapping $A : \mathcal{H}_1 \to \mathcal{H}_2$ is a p-Schatten mapping if*

$$\sum_{j=1}^{\infty} \sigma_j(A)^p < \infty,$$

i.e. if $\{\sigma_j(A)\}_{j=1}^{\infty} \in \ell^p(\mathbb{N})$.
The p-Schatten class is the set of all p-Schatten mappings:

$$\mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2) = \{A : \mathcal{H}_1 \to \mathcal{H}_2 : A \textit{ is compact, linear and } \{\sigma_j(A)\}_{j=1}^{\infty} \in \ell^p(\mathbb{N})\}.$$

An operator $A \in \mathcal{J}_p(\mathcal{H})$ is called a p-Schatten operator and an embedding $A : \mathcal{H}_1 \hookrightarrow \mathcal{H}_2$, i.e. an injective map where we identify the elements of $\mathcal{H}_1$ with their images under A is called a p-Schatten embedding if $A \in \mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2)$. Finally $\mathcal{J}_1(\mathcal{H}_1, \mathcal{H}_2)$ is called the trace class and $\mathcal{J}_2(\mathcal{H}_1, \mathcal{H}_2)$ is called the Hilbert-Schmidt class.

One can define more generally for two Banach spaces X and Y and a bounded linear mapping $A : X \to Y$ the sets $\mathcal{L}_j(X, Y)$ of all bounded linear mappings with a j-dimensional range and the numbers

$$\alpha_j(A) = \inf_{B \in \mathcal{L}_j(X,Y)} ||A - B||, j = 0, 1, 2, ...$$

and then define the classes $\ell^p(X, Y)$ as the sets of all bounded linear mappings A that satisfy $\sum_{j=1}^{\infty} \alpha_j(A)^p < \infty$ for $0 < p < \infty$. Only if X and Y are Hilbert spaces and $1 \leq p < \infty$ one can show in general that these classes are Banach spaces. In this case $\alpha_j(A) = \sigma_{j+1}(A)$ for $j = 0, 1, 2, ...$ where the singular values are numbered in non-increasing order and $\ell^p(X, Y) = \mathcal{J}_p(X, Y)$. For this see [13] , [21, Theorem 5] and [2].

3 Properties of the p-Schatten classes

3.1 General Properties

In this chapter we prove some results that describe the classes $\mathcal{J}_p$. Many of the properties of these classes mirror properties of the well known sequence spaces $\ell^p(\mathbb{N})$. By the following definition the classes $\mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2)$ are equipped with a " norm ":

Definition 9: *For $0 < p < \infty$ define the map $|.|_p : \mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2) \to [0, \infty)$ by*

$$|A|_p = \left(\sum_{j=1}^{\infty} \sigma_j(A)^p \right)^{\frac{1}{p}} .$$

We need to characterize the singular numbers using the minimax theorem which we cite here for this purpose (cf. [8, Theorem X.4.3]):

Theorem 11: *For a compact self-adjoint operator $H \in \mathcal{L}(\mathcal{H})$ its eigenvalues λ_j in non-increasing order are given by*

$$\lambda_{j+1} = \min_{\phi_1,\ldots,\phi_j \in \mathcal{H}} \max_{\substack{|\phi|=1 \\ (\phi,\phi_1)=\ldots=(\phi,\phi_j)=0}} (H\phi, \phi), j \geq 0.$$

Because $A^*A \in \mathcal{J}_\infty(\mathcal{H}_1)$ if A is compact this theorem can be applied to $H = A^*A$ (cf. [8, Lemma XI.9.2]and [25, Theorem 1.5]).

Lemma 12: *The singular values $\sigma_j(A)$ of a compact linear mapping $A : \mathcal{H}_1 \to \mathcal{H}_2$ in non-increasing order are given by*

$$\sigma_{j+1}(A) = \min_{\phi_1,\ldots,\phi_j \in \mathcal{H}_1} \max_{\substack{|\phi|=1 \\ (\phi,\phi_1)=\ldots=(\phi,\phi_j)=0}} |A\phi|_{\mathcal{H}_2} \tag{3.1}$$

$j \geq 0$.

Proof: Equation 3.1 is clearly equivalent to

$$\sigma_{j+1}^2(A) = \min_{\phi_1,\ldots,\phi_j \in \mathcal{H}_1} \max_{\substack{|\phi|=1 \\ (\phi,\phi_1)=\ldots=(\phi,\phi_j)=0}} |A\phi|_{\mathcal{H}_2}^2$$

and because $|A\phi|_{\mathcal{H}_2}^2 = (A^*A\phi, \phi)_{\mathcal{H}_1}$ this is exactly the minimax theorem applied to A^*A. □

This characterization is used to show the following lemma which in turn will be used to prove a Hölder inequality, a not yet perfect " triangle inequality " and an " ideal property " (cf. [8, Corollary XI.9.3] and [25, Theorem 1.6 and Theorem 1.7]). The inequalities that are stated in the lemma go back to Fan (cf. [10, Theorem 2]).

Lemma 13: *If $H \in \mathcal{J}_\infty(\mathcal{H}_1)$ and $A \in \mathcal{J}_\infty(\mathcal{H}_1, \mathcal{H}_2)$ then*

$$\sigma_{k+l+1}(AH) \leq \sigma_{k+1}(A)\sigma_{l+1}(H). \tag{3.2}$$

If $A_1, A_2 \in \mathcal{J}_\infty(\mathcal{H}_1, \mathcal{H}_2)$ then

$$\sigma_{k+l+1}(A_1 + A_2) \leq \sigma_{k+1}(A_1) + \sigma_{l+1}(A_2). \tag{3.3}$$

If $A \in \mathcal{J}_\infty(\mathcal{H}_1, \mathcal{H}_2)$ and $B : \mathcal{H}_0 \to \mathcal{H}_1$ and $C : \mathcal{H}_2 \to \mathcal{H}_3$ are bounded then

$$\sigma_k(AB) \leq \sigma_k(A)||B|| \text{ and } \sigma_k(CA) \leq ||C||\sigma_k(A), \tag{3.4}$$

where $||.||$ denotes the appropriate operator norms.

PROOF: By equation 3.1 the inequalities

$$\begin{aligned}
\sigma_{k+l+1}(AH) &= \min_{\phi_1,\ldots,\phi_{k+l}\in\mathcal{H}_1} \max_{\substack{|\phi|=1\\(\phi,\phi_1)=\ldots=(\phi,\phi_{k+l})=0}} |AH\phi|_{\mathcal{H}_2} \\
&\leq \min_{\phi_1,\ldots,\phi_{k+l}} \max_{\substack{|\phi|=1\\(\phi,H^*\phi_1)=\ldots=(\phi,H^*\phi_k)=\ldots=\\(\phi,\phi_{k+1})=\ldots=(\phi,\phi_{k+l})=0}} |AH\phi|_{\mathcal{H}_2} \\
&= \min_{\phi_1,\ldots,\phi_{k+l}} \max_{\substack{(H\phi,\phi_1)=\ldots=(H\phi,\phi_k)=0|\\(\phi,\phi_{k+1})=\ldots=(\phi,\phi_{k+l})=0}} \frac{|A(H\phi)|_{\mathcal{H}_2}}{|H\phi|_{\mathcal{H}_1}} \frac{|H\phi|_{\mathcal{H}_1}}{|\phi|_{\mathcal{H}_1}} \\
&\leq (\min_{\phi_1,\ldots,\phi_k} \max_{\substack{|\psi|=1\\(\psi,\phi_1)=\ldots=(\psi,\phi_k)=0}} |A\psi|_{\mathcal{H}_2})(\min_{\phi_{k+1},\ldots,\phi_{k+l}} \max_{\substack{|\phi|=1\\(\phi,\phi_{k+1})=\ldots=(\phi,\phi_{k+l})=0}} |H\phi|_{\mathcal{H}_1}) \\
&= \sigma_{k+1}(A)\sigma_{l+1}(H)
\end{aligned}$$

prove assertion 3.2.
Again by 3.1 the inequalities

$$\begin{aligned}
\sigma_{k+l+1}(A_1 + A_2) &= \min_{\phi_1,\ldots,\phi_{k+l}\in\mathcal{H}_1} \max_{\substack{|\phi|=1\\(\phi,\phi_1)=\ldots=(\phi,\phi_{k+l})=0}} |(A_1 + A_2)\phi|_{\mathcal{H}_2} \\
&\leq \min_{\phi_1,\ldots,\phi_{k+l}\in\mathcal{H}_1} \max_{\substack{|\phi|=1\\(\phi,\phi_1)=\ldots=(\phi,\phi_{k+l})=0}} (|A_1\phi|_{\mathcal{H}_2} + |A_2\phi|_{\mathcal{H}_2}) \\
&\leq \min_{\phi_1,\ldots,\phi_k\in\mathcal{H}_1} \max_{\substack{|\phi|=1\\(\phi,\phi_1)=\ldots=(\phi,\phi_k)=0}} |A_1\phi|_{\mathcal{H}_2} + \min_{\phi_{k+1},\ldots,\phi_{k+l}\in\mathcal{H}_1} \max_{\substack{|\phi|=1\\(\phi,\phi_{k+1})=\ldots=(\phi,\phi_{k+l})=0}} |A_2\phi|_{\mathcal{H}_2} \\
&= \sigma_{k+1}(A_1) + \sigma_{l+1}(A_2)
\end{aligned}$$

prove assertion 3.3.
Finally $\sigma_k(CA) \leq ||C||\sigma_k(A)$ for bounded C and compact A is clear from 3.1 and $|CA\phi|_{\mathcal{H}_3} \leq ||C|||A\phi|_{\mathcal{H}_2}$ for every $\phi \in \mathcal{H}_1$.
From Theorem 10 it follows that A^*A and AA^* have the same eigenvalues counting multiplicities. This implies $\sigma_k(A) = \sigma_k(A^*)$ and thus

$$\sigma_k(AB) = \sigma_k(B^*A^*) \leq ||B^*||\sigma_k(A^*) = ||B||\sigma_k(A)$$

where we used the facts that A is compact if and only if A^* is compact (a theorem due to Schauder) and that $||A|| = ||A^*||$ (cf. [3, Theorem 10.1 and Theorem 10.6] and [7, Theorem 6.4]).

□

One could have also used formula 3.1 to define singular values for non-compact mappings. In this case 3.4 would have been a simple consequence of 3.2 as from 3.1 one sees immediately $\sigma_1(A) = ||A||$ (cf. [8, Corollary XI.9.4]).
From the inequalities given in Lemma 13 we can deduce the following theorem which already gives some results concerning the algebraic structure of the p-Schatten classes (cf. [8, Lemma XI.9.9])

Theorem 14: *Let $0 < p_1, p_2 < \infty$. If $T \in \mathcal{J}_{p_1}(\mathcal{H}_1)$ and $B \in \mathcal{J}_{p_2}(\mathcal{H}_1, \mathcal{H}_2)$ then $A = BT \in \mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2)$ where $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$ and*

$$|A|_p \leq 2^{\frac{1}{p}}|B|_{p_1}|T|_{p_2}. \tag{3.5}$$

If $0 < p < \infty$ and $A_1, A_2 \in \mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2)$ then $A_1 + A_2 \in \mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2)$ and

$$|A_1 + A_2|_p \leq 2^{\frac{1}{p}}|A_1|_p + 2^{\frac{1}{p}}|A_2|_p \text{ for } p \geq 1 \tag{3.6}$$

and

$$|A_1 + A_2|_p^p \leq 2|A_1|_p^p + 2|A_2|_p^p \text{ for } 0 < p < 1. \tag{3.7}$$

If $0 < p < \infty$, $A \in \mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2)$ and $B : \mathcal{H}_0 \to \mathcal{H}_1$ and $C : \mathcal{H}_2 \to \mathcal{H}_3$ are bounded then $AB \in \mathcal{J}_p(\mathcal{H}_0, \mathcal{H}_2)$ and $CA \in \mathcal{J}(\mathcal{H}_1, \mathcal{H}_3)$ and

$$|AB|_p \leq |A|_p||B|| \text{ and } |CA|_p \leq ||C|||A|_p. \tag{3.8}$$

Proof: By Lemma 13 3.2 and the general Hölder inequality for sequences

$$\left(\sum_{j=0}^{\infty}\sigma_{2j+1}(A)^p\right)^{\frac{1}{p}} \leq \left(\sum_{j=0}^{\infty}(\sigma_{j+1}(B)\sigma_{j+1}(T))^p\right)^{\frac{1}{p}}$$

$$\leq \left(\sum_{j=0}^{\infty}\sigma_{j+1}(B)^{p_1}\right)^{\frac{1}{p_1}} \left(\sum_{j=0}^{\infty}\sigma_{j+1}(T)^{p_2}\right)^{\frac{1}{p_2}}$$

and

$$\left(\sum_{j=0}^{\infty}\sigma_{2j+2}(A)^p\right)^{\frac{1}{p}} \le \left(\sum_{j=0}^{\infty}\sigma_{j+1}(B)^{p_1}\right)^{\frac{1}{p_1}}\left(\sum_{j=0}^{\infty}\sigma_{j+1}(T)^{p_2}\right)^{\frac{1}{p_2}}.$$

This implies

$$\sum_{j=0}^{\infty}\sigma_{2j+1}(A)^p+\sum_{j=0}^{\infty}\sigma_{2j+2}(A)^p \le 2\left(\sum_{j=0}^{\infty}\sigma_{j+1}(B)^{p_1}\right)^{\frac{p}{p_1}}\left(\sum_{j=0}^{\infty}\sigma_{j+1}(T)^{p_2}\right)^{\frac{p}{p_2}}.$$

Taking the $\frac{1}{p}$ 'th power completes the proof of inequality 3.5.
By Lemma 133.3 and Minkowski's inequality we get

$$\left(\sum_{j=0}^{\infty}\sigma_{2j+1}(A_1+A_2)^p\right)^{\frac{1}{p}} \le \left(\sum_{j=0}^{\infty}\sigma_{j+1}(A_1)^p\right)^{\frac{1}{p}}+\left(\sum_{j=0}^{\infty}\sigma_{j+1}(A_2)^p\right)^{\frac{1}{p}}$$

and

$$\left(\sum_{j=0}^{\infty}\sigma_{2j+2}(A_1+A_2)^p\right)^{\frac{1}{p}} \le \left(\sum_{j=0}^{\infty}\sigma_{j+1}(A_1)^p\right)^{\frac{1}{p}}+\left(\sum_{j=0}^{\infty}\sigma_{j+1}(A_2)^p\right)^{\frac{1}{p}}$$

and so

$$\sum_{j=1}^{\infty}\sigma_j(A_1+A_2)^p \le 2\left(\left(\sum_{j=0}^{\infty}\sigma_{j+1}(A_1)^p\right)^{\frac{1}{p}}+\left(\sum_{j=0}^{\infty}\sigma_{j+1}(A_2)^p\right)^{\frac{1}{p}}\right)^p.$$

Taking the $\frac{1}{p}$ 'th power completes the proof of inequality 3.6.
In the case $0<p<1$ the same argument and the inequality $|x+y|^p_{\ell^p(\mathbb{N})} \le |x|^p_{\ell^p(\mathbb{N})}+|y|^p_{\ell^p(\mathbb{N})}$ valid in the metric spaces $\ell^p(\mathbb{N})$ prove 3.7.
Finally by Lemma 133.3

$$\sum_{j=1}^{\infty}\sigma_j(AB)^p \le \sum_{j=1}^{\infty}\sigma_j(A)^p||B||^p = ||B||^p\sum_{j=1}^{\infty}\sigma_j(A)^p$$

and taking the $\frac{1}{p}$ 'th power proves the first inequality in 3.8.The second is shown in the same way. □

From Theorem 14 it is easy to deduce that the p-Schatten classes $\mathcal{J}_p(\mathcal{H}_1,\mathcal{H}_2)$ are $\mathbb{C}$-vector spaces and that the operator classes $\mathcal{J}_p(\mathcal{H})$ are two sided ideals in the ring of bounded operators $\mathcal{L}(\mathcal{H})$. For $1\le p<\infty$ the p-Schatten classes are even Banach spaces with respect to the norms $|.|_p$. The difficulty here is to prove the triangle inequality. In chapter 6 we will prove that the mappings $|.|_p$ are obtained by maximizing ($p\ge 2$) or minimizing (p<2) the expressions $(\sum_{j=1}^{\infty}|A\phi_j|^p_{\mathcal{H}_2})^{\frac{1}{p}}$ where the maximum or minimum is taken over all orthonormal systems $\{\phi_j\}$. But this means the $\ell^p(\mathbb{N})$-norms of the sequences $\{A\phi_j\}$ are maximized or minimized. For $p\ge 2$ the triangle inequality follows immediately from this and Minkowski's inequality. This idea doesn't work for $1\le p<2$. Since we don't need the Banach space property in our further discussion we will omit the proof of the triangle inequality for $1 \leq p < 2$. We will present the proof for the case $p \geq 2$ because it shows a nice application of Corollary 7 in chapter 6 which is one of the main results of our discussion. For proofs of the triangle inequality for operators on one Hilbert space see [8, Lemma XI.9.14] and [24, Satz 4.3.6].

Theorem 15: *For $p \geq 2$ the p-Schatten classes $\mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2)$ are Banach spaces.*

PROOF: By Corollary 7 in chapter 6 for $p \geq 2$ we have

$$\begin{aligned} |A_1 + A_2|_p = \max |\{(A_1 + A_2)\phi_j\}_{j=1}^{\infty}|_{\ell^p(\mathbb{N})} &= \max |\{A_1\phi_j\}_{j=1}^{\infty} + \{A_2\phi_j\}_{j=1}^{\infty}|_{\ell^p(\mathbb{N})} \\ &\leq \max(|\{A_1\phi_j\}_{j=1}^{\infty}|_{\ell^p(\mathbb{N})} + |\{A_2\phi_j\}_{j=1}^{\infty}|_{\ell^p(\mathbb{N})}) \qquad (3.9) \\ \leq \max |\{A_1\phi_j\}_{j=1}^{\infty}|_{\ell^p(\mathbb{N})} &+ \max |\{A_2\phi_j\}_{j=1}^{\infty}|_{\ell^p(\mathbb{N})} = |A_1|_p + |A_2|_p \end{aligned}$$

where 3.9 is Minkowski 's inequality and the maximum is taken over all orthonormal systems of $\mathcal{H}_1$.
Clearly $\alpha A \in \mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2)$ and $|\alpha A|_p = |\alpha||A|_p$ for $\alpha \in \mathbb{C}$ and $A \in \mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2)$ because

$$|\alpha A| = ((\alpha A)^*(\alpha A))^{\frac{1}{2}} = (\overline{\alpha}\alpha A^* A)^{\frac{1}{2}} = |\alpha||A|$$

and thus $\sigma_j(\alpha A) = |\alpha|\sigma_j(A)$.
Hence $\mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2)$ is equipped with a topology induced by the norm $|.|_p$. It follows addition $+ : \mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2) \times \mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2) \to \mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2)$ and scalar multiplication $: \mathbb{C} \times \mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2) \to \mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2)$ are continuous mappings with respect to this topology and the norm is continuous.
We have to show $\mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_1)$ is complete (cf. [8, Corollary XI.9.4 and Lemma XI.9.10]). By 3.1 $||A|| = \sigma_1(A) \leq |A|_p$ for $1 \leq p < \infty$ and so every Cauchy sequence in $\mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2)$ is a Cauchy sequence in $\mathcal{J}_\infty(\mathcal{H}_1, \mathcal{H}_2)$ which is complete (cf. [8, Corollary VI.5.5]). Thus for every Cauchy sequence $\{A_n\}_{n=1}^{\infty}$in $\mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2)$ there exists a mapping $A \in \mathcal{J}_\infty(\mathcal{H}_1, \mathcal{H}_2)$ such that

$$||A - A_n|| \to 0, n \to \infty,$$

i.e. A is a compact limit in the uniform topology. By 3.3 for compact linear mappings A_1, A_2 we have

$$\sigma_{k+j+1}(A_2) = \sigma_{k+j+1}(A_1 + A_2 - A_1) \leq \sigma_{k+1}(A_1) + \sigma_{j+1}(A_2 - A_1).$$

For $j = 0$ and by symmetry in A_1, A_2 this yields

$$|\sigma_{k+1}(A_1) - \sigma_{k+1}(A_2)| \leq ||A_1 - A_2||$$

for $k \geq 0$. It follows for $j, n, m \geq 1$ that

$$|\sigma_j(A_n - A_m) - \sigma_j(A_n - A)| \leq ||A - A_m|| \to 0, m \to \infty.$$

Thus $\lim_{m\to\infty} \sigma_j(A_n - A_m) = \sigma_j(A_n - A)$ for $j, n \geq 1$ and this implies for $N \geq 1$

$$(\sum_{j=1}^{N} |\sigma_j(A_n - A)|^p)^{\frac{1}{p}} \leq \limsup_{m\to\infty}(\sum_{j=1}^{\infty} |\sigma_j(A_n - A_m)|^p)^{\frac{1}{p}} = \limsup_{m\to\infty} |A_n - A_m|_p.$$

So for $N \to \infty$ it follows

$$|A_n - A|_p \leq \limsup_{m\to\infty} |A_n - A_m|_p$$

and this implies

$$\lim_{n\to\infty} |A_n - A|_p \leq \lim_{n,m\to\infty} |A_n - A_m|_p = 0.$$

So A is a limit in the topology induced by $|.|_p$ on $\mathcal{J}_p$ and by the triangle inequality

$$|A|_p \leq |A_n - A|_p + |A_n|_p \tag{3.10}$$

for $n \geq 1$. Because $\{A_n\}$ is a Cauchy sequence it is bounded,say by C>0.Thus taking the limit $n \to \infty$ in 3.10 yields

$$|A|_p \leq \limsup_{n\to\infty} |A_n|_p \leq C < \infty.$$

This means $A \in \mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2)$. □

As already mentioned for $1 \leq p \leq 2$ the classes $\mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2)$ are Banach spaces too. Another analogue to the sequence spaces constitutes the following lemma.

Lemma 16: *For $0 < p \leq q < \infty$ we have the following inclusion of spaces*

$$\mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2) \subseteq \mathcal{J}_q(\mathcal{H}_1, \mathcal{H}_2).$$

PROOF: This is a direct consequence of the well known inclusion

$$\ell^p(\mathbb{N}) \subseteq \ell^q(\mathbb{N}).$$

3.2 The Hilbert-Schmidt Class

For a compact linear mapping $A : \mathcal{H}_1 \to \mathcal{H}_2$ and two orthonormal bases $\{\phi_j\}_{j=1}^\infty$ and $\{\psi_k\}_{k=1}^\infty$ of $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively the equation (cf. [17, Lemma 1])

$$\sum_{j=1}^\infty |A\phi_j|_{\mathcal{H}_2}^2 = \sum_{j=1}^\infty \sum_{k=1}^\infty |(A\phi_j, \psi_k)|_{\mathcal{H}_2}^2 = \sum_{k=1}^\infty \sum_{j=1}^\infty |(\phi_j, A^*\psi_k)^2|_{\mathcal{H}_1} = \sum_{k=1}^\infty |A^*\psi_k|_{\mathcal{H}_1}^2 \tag{3.11}$$

shows that the expression yields the same value for any orthonormal basis since the right hand side does not depend on the choice of the basis $\{\phi_j\}_{j=1}^\infty$ of $\mathcal{H}_1$ and the left hand side does not depend on the choice of the basis $\{\psi_j\}_{j=1}^\infty$ of $\mathcal{H}_2$. In particular it takes a finite value for one orthonormal basis if and only if it takes a finite value for all orthonormal bases. Its square root is called the Hilbert-Schmidt norm of A and A is called a Hilbert-Schmidt mapping if this norm is finite. By taking for $\{\phi_j\}_{j=1}^\infty$ the completion of the orthonormal eigenvectors of A^*A the equation

$$\sum_{j=1}^\infty |A\phi_j|_{\mathcal{H}_2}^2 = \sum_{j=1}^\infty (A\phi_j, A\phi_j)_{\mathcal{H}_2} = \sum_{j=1}^\infty (A^*A\phi_j, \phi_j)_{\mathcal{H}_1} = \sum_{j=1}^\infty \sigma_j^2(A)$$

shows that the Hilbert-Schmidt mappings are exactly the compact linear mappings of 2-Schatten class and the Hilbert-Schmidt norm is exactly the norm $|.|_2$ defined in the above section. As we will see later in our discussion the Hilbert-Schmidt mappings constitute a special case because they fit especially well to the Hilbert space structure. We see from equation 3.11 that

$$|A|_2 = \left(\sum_{j=1}^{\infty} |A\phi_j|^2_{\mathcal{H}_2} \right)^{\frac{1}{2}} = |A^*|_2 = \left(\sum_{j=1}^{\infty} |A^*\psi_j|^2_{\mathcal{H}_2} \right)^{\frac{1}{2}}$$

for an arbitrary basis $\{\phi_j\}_{j=1}^{\infty}$ of $\mathcal{H}_1$ and an arbitrary basis $\{\psi_j\}_{j=1}^{\infty}$ of $\mathcal{H}_2$.

In [22] Pietsch defines absolutely p-summable mappings $A : X \to Y$ between normed spaces X, Y as follows:
For $1 \leq p < \infty$ a linear mapping A is absolutely p-summable if there exists a number $\rho \geq 0$ such that for every finite sytem $\{\phi_1, ..., \phi_k\} \subset X$ the inequality

$$\left(\sum_{j=1}^{k} |A\phi_j|^p_Y \right)^{\frac{1}{p}} \leq \rho \sup_{|\psi|_X \leq 1} \left(\sum_{j=1}^{k} |(\phi_j, \psi)^p_X \right)^{\frac{1}{p}}$$

holds. Then $\pi_p(A)$ is defined as the smallest possible number ρ and $\Pi_p(X, Y)$ as the set of all absolutely p-summable mappings. Pietsch proves (cf. [22, Theorem 1]) that if X and Y are Hilbert spaces then

$$\Pi_2(X, Y) = \mathcal{J}_2(X, Y) \text{ and } \pi_2(A) = |A|_2$$

for every $A \in \mathcal{J}_2(X, Y)$, mentions that even

$$\Pi_p(X, Y) = \mathcal{J}_2(X, Y) \text{ for } 1 \leq p \leq 2$$

and asks if this even holds for $1 \leq p < \infty$.
For the following space there is a special characterization of the Hilbert-Schmidt mappings (cf. [7, Theorem 6.12]). Let $\mathcal{H} = L^2(\Omega)$ be the Hilbert space of square integrable functions on a domain $\Omega \subset \mathbb{R}^N$. Then $A \in \mathcal{J}_2(\mathcal{H})$ if and only if there exists a function $\kappa \in L^2(\Omega \times \Omega)$ such that

$$A\phi(x) = \int_{\Omega} \kappa(x, y)\phi(y)dy.$$

for all $x \in \Omega$.
The product of two Hilbert-Schmidt mappings belongs to an even “better” class that we discuss in the next section.

3.3 The Trace Class

We recall that $\mathcal{J}_1(\mathcal{H}_1, \mathcal{H}_2)$ is called the trace class.

Definition 10: *A is called nuclear if there exist sequences* $\{\gamma_j\}_{j=1}^{\infty}$ *and* $\{\eta_j\}_{j=1}^{\infty}$ *in* $\mathcal{H}_1$ *and* $\mathcal{H}_2$ *respectively such that*

$$A = \sum_{j=1}^{\infty} (.,\gamma_j)_{\mathcal{H}_1}\eta_j \text{ and } \sum_{j=1}^{\infty} |\gamma_j|_{\mathcal{H}_1}|\eta_j|_{\mathcal{H}_2} < \infty.$$

We have the following characterization of the trace class by nuclear operators (cf. [17]):

Lemma 17: *A mapping* $A \in \mathcal{J}_\infty(\mathcal{H}_1, \mathcal{H}_2)$ *is of trace class if and only if it is nuclear.*

PROOF: If A is of trace class, i.e. in the 1-Schatten class, its polar decomposition

$$A = \sum_{j=1}^{\infty} \sigma_j(A)(.,\phi_j)_{\mathcal{H}_1}\psi_j$$

shows A is nuclear because if $\gamma_j = \sigma_j(A)\phi_j$ and $\eta_j = \psi_j$ then

$$\sum_{j=1}^{\infty} |\gamma_j|_{\mathcal{H}_1}|\eta_j|_{\mathcal{H}_2} = \sum_{j=1}^{\infty} \sigma_j(A) < \infty.$$

If A is nuclear it is compact, hence admits a polar decomposition $A = \sum_{j=1}^{\infty} \sigma_j(A)(.,\phi_j)\psi_j$. Then we simply compute

$$\sigma_k(A) = (A\phi_k, \psi_k)_{\mathcal{H}_2} = \sum_{j=1}^{\infty} (\phi_k, \gamma_j)_{\mathcal{H}_1}(\eta_j, \psi_k)_{\mathcal{H}_2}$$

to see

$$\sum_{k=1}^{\infty} \sigma_k(A) = \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} (\phi_k, \gamma_j)_{\mathcal{H}_1}(\eta_j, \psi_k)_{\mathcal{H}_2}$$

$$\leq \sum_{j=1}^{\infty} \left(\sum_{k=1}^{\infty} |(\phi_k, \gamma_j)_{\mathcal{H}_1}|^2\right)^{\frac{1}{2}} \left(\sum_{k=1}^{\infty} |(\eta_j, \psi_k)_{\mathcal{H}_2}|^2\right)^{\frac{1}{2}} = \sum_{j=1}^{\infty} |\gamma_j|_{\mathcal{H}_1}|\eta_j|_{\mathcal{H}_2} < \infty$$

that A is in the trace class. So nuclear compact linear mappings are exactly those of trace class. □

We can characterize the trace class mappings by products of Hilbert-Schmidt mappings (cf. [17, Satz 3] and [24, Remark 1. to 4.3.6.]).

Lemma 18: *Let* $\mathcal{H}_3$ *be another infinite dimensional separable Hilbert space and* $B : \mathcal{H}_2 \to \mathcal{H}_3$ *another compact linear mapping. Then* $BA : \mathcal{H}_1 \to \mathcal{H}_3$ *is of trace class if A and B are Hilbert Schmidt mappings. If conversely A is of trace class it is a product of two Hilbert-Schmidt mappings.*

PROOF: Let $\{\phi_j\}_{j=1}^{\infty}$ and $\{\psi_j\}_{j=1}^{\infty}$ be orthonormal bases of $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. Then

$$A\phi = \sum_{j=1}^{\infty} (A\phi, \psi_j)_{\mathcal{H}_2}\psi_j$$

for all $\phi \in \mathcal{H}_1$ implies

$$(BA)\phi = \sum_{j=1}^{\infty}(A\phi, \psi_j)_{\mathcal{H}_2} B\psi_j = \sum_{j=1}^{\infty}(\phi, A^*\psi_j)_{\mathcal{H}_1} B\psi_j$$

and

$$\sum_{j=1}^{\infty} |A^*\psi_j|_{\mathcal{H}_1} |B\psi_j|_{\mathcal{H}_3} \leq \left(\sum_{j=1}^{\infty} |A^*\psi_j|^2_{\mathcal{H}_1}\right)^{\frac{1}{2}} \left(\sum_{j=1}^{\infty} |B\psi_j|^2_{\mathcal{H}_3}\right)^{\frac{1}{2}} < \infty$$

since with A and B Hilbert-Schmidt A^* and B are Hilbert-Schmidt. So BA is nuclear and thus by Lemma 17 of trace class.
If $A \in \mathcal{J}_1(\mathcal{H}_1, \mathcal{H}_2)$ consider its polar decomposition

$$A = U|A| = \underbrace{U|A|^{\frac{1}{2}}}_{=H_1} \underbrace{|A|^{\frac{1}{2}}}_{=H_2}$$

where $H_2 = |A|^{\frac{1}{2}} \in \mathcal{J}_2(\mathcal{H}_1)$ because $\sigma_j(|A|^{\frac{1}{2}}) = \sigma_j(|A|)^{\frac{1}{2}}$ and so

$$\sum_{j=1}^{\infty} \sigma_j(|A|^{\frac{1}{2}})^2 = \sum_{j=1}^{\infty} \sigma_j(|A|) = \sum_{j=1}^{\infty} \sigma_j(A) < \infty$$

and thus $H_1 \in \mathcal{J}_2(\mathcal{H}_1, \mathcal{H}_2)$ because of the ideal property. □

To understand why $\mathcal{J}_1(\mathcal{H}_1, \mathcal{H}_1)$ is called the trace class we define for a positive operator $H \in \mathcal{L}(\mathcal{H})$ its trace:

Definition 11: *For $H \in \mathcal{L}(\mathcal{H})$ with $H \geq 0$ we define its trace tr with values in $\mathbb{R} \cup \{\infty\}$ by*

$$trH = \sum_{j=1}^{\infty}(H\phi_j, \phi_j) \tag{3.12}$$

where $\{\phi_j\}_{j=1}^{\infty}$ is an orthonormal basis of $\mathcal{H}$.

To see that the trace is well defined we have to show that 3.12 is independent of the chosen basis. To see this let $\{\psi_j\}_{j=1}^{\infty}$ be another basis and recall that since $H \geq 0$ it has a positive square root:

$$\sum_{j=1}^{\infty}(H\phi_j, \phi_j) = \sum_{j=1}^{\infty} |H^{\frac{1}{2}}\phi_j|^2 = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} |(H^{\frac{1}{2}}\phi_j, \psi_k)|^2$$

$$= \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} |(\phi_j, H^{\frac{1}{2}}\psi_k)|^2 = \sum_{k=1}^{\infty} |H^{\frac{1}{2}}\psi_k|^2 = \sum_{k=1}^{\infty}(H\psi_k, \psi_k).$$

We note some properties of the trace map that are easy to prove. Recall an operator $U \in \mathcal{L}(\mathcal{H})$ is called unitary if $U^*U = UU^* = id_{\mathcal{H}}$ (cf. [24, 4.3.4]).

Lemma 19: *For positive bounded operators H_1, H_2,$\lambda \geq 0$ and a unitary operator U we have*

$$tr(H_1 + H_2) = tr(H_1) + tr(H_2).$$

and

$$tr(\lambda H_1) = \lambda tr(H_1).$$

and

$$tr(UH_1U^*) = tr(H_1).$$

A mapping $A \in \mathcal{J}_\infty(\mathcal{H}_1, \mathcal{H}_2)$ is in the trace class if and only if $tr|A| < \infty$. If $A \in \mathcal{J}_1(\mathcal{H}_1, \mathcal{H}_1)$ let $\{\phi_j\}_{j=1}^\infty$ be an orthonormal basis of $\mathcal{H}_1$ consisting of eigenvectors of $|A|$ then

$$tr|A| = \sum_{j=1}^{\infty}(|A|\phi_j, \phi_j)_{\mathcal{H}_1} = \sum_{j=1}^{\infty}\sigma_j(A) < \infty \tag{3.13}$$

If for $A \in \mathcal{J}_\infty(\mathcal{H}_1, \mathcal{H}_2)$ the trace is finite, equation 3.13 shows $A \in \mathcal{J}_1(\mathcal{H}_1, \mathcal{H}_2)$. For the class $\mathcal{J}_1(\mathcal{H})$ one can show that for all $A \in \mathcal{J}_1(\mathcal{H})$ the expression

$$trA = \sum_{j=1}^{\infty}(A\phi_j, \phi_j)$$

is independent of the choice of the basis $\{\phi_j\}$ and the series converges absolutely. Then this expression is called the trace and the function tr is linear and continuous on the space $\mathcal{J}_1(\mathcal{H})$. Because our aim is to study Schatten class embeddings that do not contain operators on one Hilbert space we give no proof for this and just cite [8, Lemma XI.9.13].

A characterization of the trace class operators on $L^2(\Omega)$ as in the case of Hilbert-Schmidt operators is not possible (cf. [24, Remark 3. to 4.3.6]).

4 Compact embeddings

As mentioned in our introduction Hanche-Olsen and Holden characterized the totally bounded sets in $\ell^q(\mathbb{N})$ for $1 \leq q < \infty$ (cf. [14]). In this way they characterized the compact embeddings

$$X \hookrightarrow \ell^q(\mathbb{N})$$

where X is some metric space. Since the characterization of the p-Schatten embeddings into $\ell^2(\mathbb{N})$ that we want to achieve is modeled on this result we present it here.
We recall the following notion:

Definition 12: *A subset M of a metric space X is called **totally bounded** if it can be covered by a finite number of balls with radii $< \varepsilon$ and centers in M for every $\varepsilon > 0$.*

Without proof we cite (cf. [26]):

Theorem 20: *A subset M of a complete metric space is precompact if and only if it is totally bounded.*

To see the importance of this chapter for our discussion we state:

Theorem 21: *An embedding $A : \mathcal{H}_1 \to \mathcal{H}_2$ of Hilbert spaces is compact if and only if $A(\{\phi \in \mathcal{H}_1 : |\phi|_{\mathcal{H}_1} \leq 1\})$ is totally bounded.*

Proof: This follows from Theorem 20 and 7. □

Hanche-Olsen and Holden use a very simple lemma to prove several compactness results (cf. [14, Lemma 1]):

Lemma 22: *Let X be a metric space with metric d_X. If for every $\varepsilon > 0$ there exists $\delta(\varepsilon) > 0$ and a metric space $W(\varepsilon)$ with metric $d_{W(\varepsilon)}$ and a mapping $\Phi : X \to W(\varepsilon)$ such that $\Phi(X)$ is a totally bounded subset of $W(\varepsilon)$ and $d_{W(\varepsilon)}(\Phi(x), \Phi(y)) < \delta(\varepsilon)$ implies $d_X(x, y) < \varepsilon$ for all $x, y \in X$ then X is totally bounded.*

Proof: For $\varepsilon > 0$ and $\Phi(X) \subset W(\varepsilon)$ choose a finite covering by balls with radii $\delta(\varepsilon)$ and centers $m_1, ..., m_n \in \Phi(X)$ and denote them by $B^{\delta(\varepsilon)}(m_1), ..., B^{\delta(\varepsilon)}(m_n)$. This is possible since $\Phi(X)$ is a totally bounded subset of $W(\varepsilon)$. Then by the assumption $\Phi^{-1}(B^{\delta(\varepsilon)}(m_j)) \subset B^{\varepsilon}(x_j)$ for some $x_j \in X$ and $j = 1, ..., n$ where $B^{\varepsilon}(x_j)$ are balls with radii ε and centers x_j. It follows $X \subset \bigcup_{j=1}^{n} B^{\varepsilon}(x_j)$ i.e. X is covered by a finite number of balls with radii $\leq \varepsilon$. This implies X is totally bounded. □

Lemma 4 can be used to show the Arzelá-Ascoli theorem (cf. [14, Theorem 2]). We will use it to prove the following characterization of the totally bounded subsets of $\ell^q(\mathbb{N})$, i.e. of the precompact subsets for $1 \le q < \infty$.

Theorem 23: *For $1 \le q < \infty$ a subset $M \subset \ell^q(\mathbb{N})$ is totally bounded if and only if*

(i) for every $j \in \mathbb{N}$ there exists a number $C_j > 0$ so that $|x_j| < C_j$ for all $x \in M$

i.e. M is pointwise bounded, and
(ii) for every $\varepsilon > 0$ there exists a number $n(\varepsilon) \in \mathbb{N}$ so that for all $x \in M$

$$\left(\sum_{j=n(\varepsilon)+1}^{\infty} |x_j|^q \right)^{\frac{1}{q}} < \varepsilon.$$

Proof: If $M \subset \ell^q(\mathbb{N})$ satisfies (i) and (ii) then for every $\varepsilon > 0$ and $n(\varepsilon)$ as in (ii) define $\Phi : M \to \mathbb{R}^{n(\varepsilon)}$ by

$$\Phi(x) = (x_1, ..., x_{n(\varepsilon)})^T, x \in M.$$

Clearly $\mathbb{R}^{n(\varepsilon)}$ is a metric space with the metric induced by the q-norm and by (i) for all $x \in M$:

$$|\Phi(x)|_q = (|x_1|^q + ... + |x_{n(\varepsilon)}|^q)^{\frac{1}{q}} < \max\{C_1, ..., C_{n(\varepsilon)}\}.$$

Thus as a bounded subset of a finite dimensional space $\Phi(M)$ is totally bounded. Furthermore $|\Phi(x) - \Phi(y)|_q = (\sum_{j=1}^{n(\varepsilon)} |x_j - y_j|^q)^{\frac{1}{q}} < \varepsilon$ implies

$$|x-y|_q = \left(\sum_{j=1}^{\infty} |x_j - y_j|^q \right)^{\frac{1}{q}} \le \left(\sum_{j=1}^{n(\varepsilon)} |x_j - y_j|^q \right)^{\frac{1}{q}} + \left(\sum_{n(\varepsilon)+1}^{\infty} |x_i - y_j|^q \right)^{\frac{1}{q}}$$

$$\le \left(\sum_{j=1}^{n(\varepsilon)} |x_j - y_j|^q \right)^{\frac{1}{q}} + \left(\sum_{n(\varepsilon)+1}^{\infty} |x_i|^q \right)^{\frac{1}{q}} + \left(\sum_{n(\varepsilon)+1}^{\infty} |y_j|^q \right)^{\frac{1}{q}} < \varepsilon + \varepsilon + \varepsilon = 3\varepsilon$$

where we used Minkowski's inequality. Thus the conditions of Lemma 4 are fulfilled and so M is totally bounded.
For the converse let $\varepsilon > 0$ and assume M is totally bounded. Then choose a finite $\frac{\varepsilon}{4}$-cover by balls $B^{\frac{\varepsilon}{4}}(x^1), ..., B^{\frac{\varepsilon}{4}}(x^n)$. For any sequences $x, y \in M$ we have

$$|x-y|_q \le n\frac{\varepsilon}{2}.$$

Fix $y \in M$ and put $|y|_q = C$ to see

$$|x|_q \le |x-y|_q + |y|_q \le n\frac{\varepsilon}{2} + C$$

for every $x \in M$. So M is bounded and thus pointwise bounded. For every $k \in \{1, ..., n\}$ there exists a number N_k so that

$$\left(\sum_{j=N_k+1}^{\infty} |x_j^k|^q \right)^{\frac{1}{q}} < \frac{\varepsilon}{2}.$$

Put $N = \max\{N_1, ..., N_k\}$. Then for any $x \in M$ one has $x \in B^{\frac{\varepsilon}{2}}(x^k)$ for some k and so

$$\left(\sum_{j=N+1}^{\infty} |x_j|^q\right)^{\frac{1}{q}} \leq |x - x^k|_q + \left(\sum_{j=N+1}^{\infty} |x^k|^q\right)^{\frac{1}{q}} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus condition (ii) in Lemma 4 holds. □

Since we are especially interested in the compact embeddings $\mathcal{H} \hookrightarrow \ell^2(\mathbb{N})$ we conclude:

Corollary 1: *An embedding $\mathcal{H} \hookrightarrow \ell^2(\mathbb{N})$ of Hilbert spaces is compact if and only if*

(i) for every $j \in \mathbb{N}$ there exists $C_j > 0$ so that $|x_j| < C_j$ for all $x \in \{x \in \mathcal{H} : |x|_{\mathcal{H}} \leq 1\}$

and
(ii) for every $\varepsilon > 0$ there exist a number $n(\varepsilon) \in \mathbb{N}$ so that for all $x \in \{x \in \mathcal{H} : |x|_{\mathcal{H}} \leq 1\}$

$$\left(\sum_{j=n(\varepsilon)+1}^{\infty} |x_j|^2\right)^{\frac{1}{2}} < \varepsilon.$$

Proof: This follows from Theorem 23 and Theorem 21. □

Note that we identify the elements of $\mathcal{H}$ with their images in $\ell^2(\mathbb{N})$ under the embedding. Theorem 25 below shows another application of Lemma 4. We present its proof at this point since it is of interest to our discussion that the same techniques are applied to the Lebesgue spaces as to the sequence spaces. We will need Jensen's inequality (cf. [9, Appendix B.1,Theorem 2]).

Theorem 24 (Jensen's inequality)**:** *Let $\Omega \subset \mathbb{R}^N$ be an open, bounded subset of the N-dimensional Euclidean space. For any convex function $\phi : \mathbb{R} \to \mathbb{R}$ and every summable function $f : \Omega \to \mathbb{R}$ we have*

$$\phi\left(\frac{1}{|\Omega|}\int_\Omega f(x)d\lambda\right) \leq \frac{1}{|\Omega|}\int_\Omega \phi(f(x))d\lambda.$$

Here λ is the Lebesgue measure and $|\Omega| = \lambda(\Omega) = \int_\Omega d\lambda$. The summability of f means $\int_\Omega f(x)d\lambda$ exists and is finite. The following Theorem 25 is known as the Kolmogorov -Riesz theorem. For its history see [14, 4].

Theorem 25 (Kolmogorov Riesz)**:** *For $1 \leq q < \infty$ a subset M of $L^q(\mathbb{R}^N)$ is totally bounded if and only if*
(i)M is bounded
(ii)for every $\epsilon > 0$ there exists a radius $R > 0$ so that for every $f \in M$

$$\left(\int_{\mathbb{R}^N \setminus B^R(0)} |f(x)|^q d\lambda\right)^{\frac{1}{q}} < \varepsilon,$$

(iii)for every $\varepsilon > 0$ there exists a radius $\rho > 0$ so that for every $f \in M$ and $y \in B^\rho(0)$

$$\left(\int_{\mathbb{R}^N} |f(x+y) - f(x)| d\lambda(x)\right)^{\frac{1}{q}} < \varepsilon.$$

Proof: Consider a subset $M \subset L^q(\mathbb{R}^N)$ satisfying conditions (i),(ii),(iii) and a given number $\varepsilon > 0$ and choose R and ρ as in (ii) and (iii).
Let$Q$ be an open cube that is centered at the origin and is contained in the sphere $B^{\frac{\rho}{2}}(0)$ and let $Q_j = Q + v_j$ for $v_j \in \mathbb{R}^N$ and $j = 1, ..., n$ be mutually non-overlapping translates of Q so that

$$B^R(0) \subset \overline{\bigcup_{j=1}^{n} Q_j}. \tag{4.1}$$

We define the projection $P : L^q(\mathbb{R}^N) \to span\{\chi_{Q_1}, ..., \chi_{Q_n}\}$ where χ_{Q_j} is the characteristic function of the cube Q_j by

$$Pf(x) = \begin{cases} \frac{1}{Q_j} \int_{Q_j} f(z) d\lambda & \text{if } x \in Q_j, j = 1, ..., n \\ 0 & \text{otherwise.} \end{cases}$$

For all $f \in M$ it follows from (ii) and 4.1 that

$$|f - Pf|_q^q < \varepsilon^q + \sum_{j=1}^{n} \int_{Q_j} |f(x) - Pf(x)|^q d\lambda$$

$$= \epsilon^q + \sum_{j=1}^{n} \int_{Q_j} |\frac{1}{|Q_j|} \int_{Q_j} (f(x) - f(z)) d\lambda(z)|^q d\lambda(x)$$

by the definition of P and $\frac{1}{|Q_j|} \int_{Q_j} f(x) d\lambda(z) = f(x)$.
Now we observe that whenever two points are in the same cube Q_j their difference is in the cube obtained from Q by doubling its side length. Indeed if $x, z \in Q_j$ there exist $q_x, q_z \in Q$ such that $x = q_x + v_j$ and $z = q_z + v_j$ and so $x - z = q_x - q_z \in Q - Q$ and because Q is centered at the origin $Q - Q = 2Q$. A change of a variable of integration by $z = x + y$, the fact that $|.|^q$ is convex for $q \geq 1$, Jensen's inequality 24 and (iii) yield

$$|f - Pf|_q^q < \epsilon^q + \sum_{j=1}^{n} \int_{Q_j} \frac{1}{|Q_j|} \int_{Q_j} |f(x) - f(z)|^q d\lambda(z) d\lambda(x)$$

$$\leq \epsilon^q + \sum_{j=1}^{n} \int_{Q_j} |\frac{1}{|Q_j|} \int_{2Q} |f(x) - f(x + y)|^q d\lambda(y)| d\lambda(x)$$

$$\leq \epsilon^q + \frac{1}{|Q|} \int_{2Q} \int_{\mathbb{R}^N} |f(x) - f(x + y)|^q d\lambda(x) d\lambda(y)$$

$$< \varepsilon^q + \frac{1}{|Q|} \int_{2Q} \varepsilon^q d\lambda(y) = (2^N + 1)\varepsilon^q$$

because clearly $|2Q| = 2^N |Q|$. It follows $|f - Pf|_q < (2^N + 1)^{\frac{1}{q}} \varepsilon$ and by the triangle inequality $|f|_q < (2^N + 1)^{\frac{1}{q}} \varepsilon + |Pf|_q$. By the linearity of the integral P is linear and so $|Pf - Pg|_q < \varepsilon$ implies $|f - g| < ((2^N + 1)^{\frac{1}{q}} + 1)\varepsilon$.
As a projection on a finite dimensional subspace P is bounded and because M is bounded by (i) this implies $P(M)$ is bounded in a finite dimensional space and thus totally

bounded. Hence it follows by Lemma 4 that M is totally bounded.
Now we assume M to be totally bounded. Then M is bounded. So condition (i) holds. For $\varepsilon > 0$ choose $g_1, ..., g_n \in M$ so that

$$M \subset \bigcup_{j=1}^{n} B^{\varepsilon}(g_j)$$

and $R > 0$ so that

$$\left(\int_{\mathbb{R}^N \setminus B^R(0)} |g_j(x)|^q d\lambda \right)^{\frac{1}{q}} < \varepsilon, j = 1, ..., n.$$

This is possible because $g_j \in L^q(\mathbb{R}^N)$ for $j = 1, ..., n$. For every $f \in M$ there exists $j \in \{1, ..., n\}$ so that $f \in B^{\varepsilon}(g_j)$ and thus

$$\left(\int_{\mathbb{R}^N \setminus B^R(0)} |f(x)|^q d\lambda \right)^{\frac{1}{q}} \leq \left(\int_{\mathbb{R}^N \setminus B^R(0)} |f(x) - g_j(x)|^q d\lambda \right)^{\frac{1}{q}} + \left(\int_{\mathbb{R}^N \setminus B^R(0)} |g_j(x)|^q d\lambda \right)^{\frac{1}{q}}$$

$$\leq |f - g_j|_q + \left(\int_{\mathbb{R}^N \setminus B^R(0)} |g_j(x)|^q d\lambda \right)^{\frac{1}{q}} < 2\varepsilon.$$

So condition (ii) holds.
The set of smooth functions with compact support , denoted by $C_c^{\infty}(\mathbb{R}^N)$ is known to be dense in $L^q(\mathbb{R}^N)$. So for every single function $f \in M$ there is a function $\phi \in C_c^{\infty}(\mathbb{R}^N)$ so that $|f - \phi|_q < \varepsilon$ for given ε. For smooth functions with compact support a radius $\rho > 0$ establishing the inequality of condition (iii) can clearly be found and so

$$\left(\int_{\mathbb{R}^N} |f(x+y) - f(x)|^q d\lambda(x) \right)^{\frac{1}{q}} \leq \left(\int_{\mathbb{R}^N} |f(x+y) - \phi(x+y) - (f(x) - \phi(x))|^q d\lambda(x) \right)^{\frac{1}{q}}$$

$$+ \left(\int_{\mathbb{R}^N} |\phi(x+y) - \phi(x)|^q d\lambda(x) \right)^{\frac{1}{q}}$$

$$\leq 2|f - \phi|_q + \left(\int_{\mathbb{R}^N} |\phi(x+y) - \phi(x)|^q d\lambda(x) \right)^{\frac{1}{q}} \leq 3\varepsilon$$

for every $y \in B^{\rho}(0)$. It follows that there exists a radius $\rho > 0$ so that

$$\left(\int_{\mathbb{R}^N} |g_j(x+y) - g_j(x)|^q d\lambda(x) \right)^{\frac{1}{q}} < \varepsilon \text{ for } y \in B^{\rho}(0), j = 1, ..., n$$

where g_j are the centers of the ε-cover given above. Then for every $f \in M$ there is $j \in \{1, ..., n\}$ so that $f \in B^{\varepsilon}(g_j)$ and so

$$\left(\int_{\mathbb{R}^N} |f(x+y) - f(x)|^q d\lambda(x) \right)^{\frac{1}{q}} \leq \left(\int_{\mathbb{R}^N} |f(x+y) - g_j(x+y)|^q d\lambda(x) \right)^{\frac{1}{q}}$$

$$+ \left(\int_{\mathbb{R}^N} |g_j(x+y) - g_j(x)|^q d\lambda(x) \right)^{\frac{1}{q}} + \left(\int_{\mathbb{R}^N} |g_j(x) - f(x)|^q d\lambda(x) \right)^{\frac{1}{q}} < 3\varepsilon$$

for every $y \in B^{\rho}(0)$. So condition (iii) holds. □

The Kolmogorov-Riesz theorem can be used to show several corollaries and a simpler version of the Rellich-Kondrachov theorem. For this see [14, Corollary 7,8,9 and Theorem 10].
We observe that the proofs of Theorem 23 and of Theorem 25 use the same techniques. The compact embeddings of metric spaces X in $\ell^q(\mathbb{N})$ or $L^q(\mathbb{R}^N)$ are characterized by very similar summability conditions that the elements of images of bounded sets of X must fulfill. Of course all elements of an orthonormal basis of a Hilbert space that is embedded in $\ell^2(\mathbb{N})$ or $L^2(\mathbb{R}^N)$ must fulfill these conditions to yield a compact embedding. We will find summability conditions that orthonormal bases of the embedded Hilbert spaces must fulfill to yield p-Schatten embeddings and the techniques we use can be applied to any separable Hilbert space. The disadvantage is that it is in general not possible to describe all orthonormal bases. In the next chapter we will present results by Maurin and Gramsch that allow to characterize certain embeddings of Sobolev spaces. The techniques used there don't seem to apply to the discrete setting as described in [19, chapter 3.1].

5 Embeddings of Sobolev spaces

In this chapter we will present results by Maurin [17] and Gramsch [13] concerning the embeddings of Sobolev spaces. Gramsch's result is a direct generalization of Maurin's result but we will see that Maurin's proof is very short whilst Gramsch's proof is quite involved . A reason for this is that the sum

$$\sum_{j=1}^{\infty} |A\phi_j|^2_{\mathcal{H}_2}$$

for a compact linear mapping $A : \mathcal{H}_1 \to \mathcal{H}_2$ does not depend on the orthonormal basis of $\mathcal{H}_1$ while this is in general the case for $p \neq 2$. Another reason is that the Hilbert-Schmidt class fits especially well to the Hilbert space structure which allows a very convenient use of Riesz's representation theorem. In the first section we will recall the definition of Sobolev spaces and some facts concerning them. In the second and third section we will present Maurin's and Gramsch's results respectively.

5.1 Sobolev Spaces

We recall the important notion of a weak derivative (cf. [9, 5.2.1]). Let $\Omega \subseteq \mathbb{R}^N$ be open. As a motivation we observe that for a function $f \in C^1(\Omega)$ and a test function $\phi \in C_c^\infty(\Omega)$, i.e. a smooth function with compact support in Ω one has by integration by parts

$$\int_\Omega f(x)\frac{\partial \phi(x)}{\partial x_j} d\lambda = -\int_\Omega \frac{\partial f(x)}{\partial x_j}\phi(x) d\lambda, j = 1, ..., N$$

where there is no boundary term because ϕ has compact support in Ω. For a multiindex $\alpha = (\alpha_1, ..., \alpha_N) \in \mathbb{N}_0^N$ we say it is of order k if $|\alpha| = \alpha_1 + ... + \alpha_N = k$ and we write for $f \in C^k(\Omega)$:

$$D^\alpha f = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} ... \frac{\partial^{\alpha_N}}{\partial x_N^{\alpha_N}} f.$$

We observe that for $f \in C^k(\Omega)$ and $\phi \in C_c^\infty(\Omega)$:

$$\int_\Omega f(x) D^\alpha \phi(x) d\lambda = (-1)^{|\alpha|} \int_\Omega D^\alpha f(x)\phi(x) d\lambda. \tag{5.1}$$

We recall that locally integrable functions are functions that are integrable on compact sets contained in Ω and the set of all locally integrable functions is denoted $L^1_{loc}(\Omega)$. Now we can define weak derivatives:

Definition 13: *Assume $f \in L^1_{loc}(\Omega)$ and a multiindex α are given. Then $g \in L^1_{loc}(\Omega)$ is called the α'th weak derivate of f if*

$$\int_\Omega f(x)D^\alpha \phi(x)d\lambda = (-1)^{|\alpha|}\int_\Omega g(x)\phi(x)d\lambda$$

for all $\phi \in C_c^\infty(\Omega)$.

It is easy to prove that weak partial derivatives are unique up to a set of measure zero (cf. [9, 5.2.1 Lemma]). We denote the α'th partial derivative of a locally integrable function f by $D^\alpha f$. We define the Sobolev spaces :

Definition 14: *Let $\Omega \subseteq \mathbb{R}^N$ be open. For $k \in \mathbb{N}_0$ and $1 \leq q < \infty$ the Sobolev space*

$$W^{k,q}(\Omega)$$

is defined as the space of all functions $f \in L^1_{loc}(\Omega)$ so that $D^\alpha f$ exists for all multiindices α with $|\alpha| \leq k$ and belongs to $L^q(\Omega)$. We define the norm

$$|f|_{W^{k,q}(\Omega)} = \left(\sum_{|\alpha|\leq k}\int_\Omega |D^\alpha f(x)|^q d\lambda\right)^{\frac{1}{q}}.$$

Without proof we cite (cf. [9, 5.2.3 Theorem 2]):

Theorem 26: *The mapping $|.|_{W^{k,q}(\Omega)}$ is a norm on $W^{k,q}(\Omega)$ and $W^{k,q}(\Omega)$ is a Banach space with this norm.*

We denote by $W_0^{k,q}$ the closure of $C_c^\infty(\Omega)$ in $W^{k,q}(\Omega)$ and write $H^k(\Omega) = W^{k,2}(\Omega)$ and $H_0^k(\Omega) = W_0^{k,2}(\Omega)$. We note:

Corollary 2: *For $k \in \mathbb{N}_0$ the Sobolev space $H^k(\Omega)$ is a Hilbert space with the inner product*

$$(f,g)_{H^k(\Omega)} = \sum_{|\alpha|\leq k}\int_\Omega (D^\alpha f(x))\overline{D^\alpha g(x)}d\lambda, f,g \in H^k(\Omega).$$

$H_0^k(\Omega)$ is a Hilbert space where the inner product $(.,.)_{H_0^k(\Omega)}$ is obtained by restricting $(.,.)_{H^k(\Omega)}$ to $H_0^k(\Omega) \times H_0^k(\Omega)$.

Because it can be considered as a starting point for the discussion in this chapter we will cite the Rellich-Kondrachov compactness theorem. We will need the Sobolev embedding theorem in the next section too. To achieve these embeddings it is necessary that the boundary $\partial\Omega$ fulfills certain conditions. The most important regularity property that we want the domain Ω to have is the cone property (cf. [1, 4.1]).

Definition 15: *For a point $x \in \mathbb{R}^N$, a ball $B^r(x)$ and a ball B_0 that does not contain x we define the finite cone C with vertex x by*

$$C = B^r(x) \cap \{x + \lambda(y - x) : y \in B_0, \lambda > 0\}$$

For an open domain $\Omega \subseteq \mathbb{R}^N$ we say it has the cone property if there exists a finite cone C so that for every $x \in \Omega$ there is a cone C_x with vertex x that is contained in Ω and congruent to C.

It is clear that the boundary $\partial\Omega$ must have some regularity properties so that Ω has the cone property. It is not clear how this concept could be transferred in a reasonable way to the discrete setting as described in [19, chapter 3.1]. The Sobolev embedding theorem includes three cases. We will only consider the case where $m > \frac{N}{2}$ and we embed $H^{m+k}(\Omega)$ in $C_b^k(\Omega) = \{f \in C^k(\Omega) : D^\alpha f \text{ is bounded on } \Omega \text{ for} |\alpha| \leq k\}$(cf.[1, 5.3 Part I Case C]).

Theorem 27: *For a bounded domain that has the cone property and $m > \frac{N}{2}$ there is a continuous embedding*

$$H^{m+k}(\Omega) \hookrightarrow C_b^k(\Omega).$$

For the full Sobolev embedding theorem and a proof see [1, 5.3]. It follows by the definition of the space $C_b^k(\Omega)$ (cf. [17]):

Lemma 28: *Let Ω be a bounded domain that has the cone property and $m > \frac{N}{2}$. Then the mappings*

$$H^{m+k}(\Omega) \stackrel{T_y^\alpha}{\to} \mathbb{C}$$
$$u \mapsto D^\alpha u(y)$$

are bounded linear functionals satisfying

$$|T_y^\alpha|_{(H^{m+k}(\Omega))^*} \leq C \tag{5.2}$$

for all α with $|\alpha| \leq k$ and all $y \in \Omega$.

Here $|.|_{(H^{m+k}(\Omega))^*}$ denotes the operator norm on the dual space of $H^{m+k}(\Omega)$ and C does depend on k and Ω but not on the multiindex α if $|\alpha| \leq k$ and not on y. The same lemma holds for $H_0^{m+k}(\Omega)$ instead of $H^{m+k}(\Omega)$. In this case one doesn't need cone property of Ω. We don't cite the Rellich-Kondrachov theorem in its most general form since we are primarily interested in its applications to the Hilbert spaces. For a more general form see [1, Theorem 6.2].

Theorem 29: *Let Ω be a bounded subdomain of $\mathbb{R}^N$ that has the cone property. Then the embeddings*

$$W^{j+m,q}(\Omega) \hookrightarrow W^{j,r}(\Omega)$$

are compact if either $mq \geq N$ or if $mq < N$ and $1 \leq r < \frac{qN}{N-mq}$.

It follows

Lemma 30: *Let Ω be a bounded subdomain of $\mathbb{R}^N$ that has the cone property. Then the embeddings*

$$H^k(\Omega) \hookrightarrow H^l(\Omega)$$

are compact if $k - l > 0$.

PROOF: We have by the Rellich-Kondrachov compactness theorem that for $q = r = 2$ the embedding

$$H^{j+m}(\Omega) \hookrightarrow H^j(\Omega)$$

is compact if $2m \geq N$ or if $2m < N$ and $1 \leq 2 < \frac{2N}{N-2m}$ where the latter inequality is trivial . Thus all embeddings are compact for $m > 0$. Put $k = j + m$ and $l = j$. $\square$

In the next section this result will be improved.

5.2 Maurin's Theorem

For the first time in our discussion we encounter Schatten class embeddings (cf. [17, Satz 4]):

Theorem 31: *Let $\Omega \subset \mathbb{R}^N$ be a bounded domain that has the cone property , k a non-negative integer and $m > \frac{N}{2}$. Then the embeddings*

$$H^{m+k}(\Omega) \hookrightarrow H^k(\Omega)$$

$$H_0^{m+k}(\Omega) \hookrightarrow H_0^k(\Omega)$$

are Hilbert-Schmidt embeddings

PROOF: By Lemma 28 and Riesz's representation theorem for any point $y \in \Omega$ and every multiindex α there exists an element $g_y^\alpha \in H^{m+k}(\Omega)$ such that

$$T_y^\alpha(u) = (u, g_y^\alpha)_{H^{m+k}(\Omega)}$$

for all $u \in H^{m+k}(\Omega)$ and

$$|T_y^\alpha|_{(H^{m+k}(\Omega))^*} = |g_y^\alpha|_{H^{m+k}(\Omega)}.$$

For an orthonormal basis $\{\phi_j\}_{j=1}^\infty$ of $H^{m+k}(\Omega)$ we thus have

$$|T_y^\alpha|^2_{(H^{m+k}(\Omega))^*} = |g_y^\alpha|^2_{H^{m+k}(\Omega)} = \sum_{j=1}^\infty |(\phi_j, g_y^\alpha)_{H^{m+k}(\Omega)}|^2 = \sum_{j=1}^\infty |D^\alpha \phi_j(y)|^2 \leq C^2$$

by inequality 5.2.
Since C does not depend on $|\alpha| \leq k$ and $y \in \Omega$ we can sum up over $|\alpha| \leq k$ and integrate over Ω to get

$$\int_\Omega \left(\sum_{|\alpha| \leq k} \sum_{j=1}^\infty |D^\alpha \phi_j(y)|^2 \right) dy = \sum_{j=1}^\infty \sum_{|\alpha| \leq k} \int_\Omega |D^\alpha \phi_j(y)|^2 = \sum_{j=1}^\infty |\phi_j|^2_{H^k(\Omega)} \leq |\Omega| \nu(k) C^2 < \infty$$

since Ω is bounded and $\nu(k)$ (- the number of multi-indices of order $\leq k$) is finite. This shows the assertion. The proof is the same for $H_0^{m+k}(\Omega) \hookrightarrow H_0^k(\Omega)$ where we don 't need the cone property condition. □

By what we have seen in 18 the composition of two Hilbert-Schmidt embeddings is of trace class i.e. nuclear. Applying this to the composite embeddings

$$H^{2m+k}(\Omega) \hookrightarrow H^{m+k}(\Omega) \hookrightarrow H^k(\Omega)$$

$$H_0^{2m+k}(\Omega) \hookrightarrow H_0^{m+k}(\Omega) \hookrightarrow H_0^k(\Omega)$$

yields (cf. [17, Korollar])

Corollary 3: *Let $\Omega \subset \mathbb{R}^N$ a bounded domain,that has the cone property k a non-negative integer and $m > \frac{N}{2}$ Then the embeddings*

$$H^{2m+k}(\Omega) \hookrightarrow H^k(\Omega)$$

$$H_0^{2m+k}(\Omega) \hookrightarrow H_0^k(\Omega)$$

are of trace class.

Again one does not need the cone property condition in the second case.
It follows from Theorem 31 and its Corollary 3 that the embeddings

$$H^k(\Omega) \hookrightarrow H^l(\Omega)$$

are of Hilbert-Schmidt class i.e. of 2-Schatten class if $k - l > \frac{N}{2}$ and of trace class i.e. 1-Schatten class if $k - l > \frac{N}{1}$. Now one might conjecture that the embeddings are of p-Schatten class if $k - l > \frac{N}{p}$. This will be achieved in the next section. We see that Maurin's proof is remarkably short. The result had far reaching consequences. For its applications to elliptic boundary value problems and eigenfunction expansions see [17, sections 3 and 4].

5.3 Gramsch 's Theorem

In his proof of the characterization of p-Schatten embeddings of the Sobolev spaces on tori $\mathcal{H}^m(T^n) \hookrightarrow \mathcal{H}^l(T^n)$ in [13] Gramsch uses a method that can be generalized in the following way:

Theorem 32: *Let $\epsilon : \mathcal{H}_1 \hookrightarrow \mathcal{H}_2$ be a compact embedding. If there exists an orthogonal basis $\{\phi_j\}_{j=1}^{\infty}$ of $\mathcal{H}_1$ **and** $\mathcal{H}_2$ then ϵ is a p-Schatten embedding if and only if*

$$\sum_{j=1}^{\infty} \left(\frac{|\phi_j|_{\mathcal{H}_2}}{|\phi_j|_{\mathcal{H}_1}} \right)^p < \infty \tag{5.3}$$

PROOF: Define $\epsilon^* : \mathcal{H}_2 \to \mathcal{H}_1$ by

$$\epsilon^*(\phi_j) = \frac{{|\phi_j|_{\mathcal{H}_2}}^2}{|\phi_j|_{\mathcal{H}_1}^2} \phi_j, j = 1, 2, ...$$

Then because $\epsilon\phi_j = \phi_j$ and $\{\phi_j\}_{j=1}^{\infty}$ is orthogonal we have on the one hand

$$(\epsilon\phi_j, \phi_k)_{\mathcal{H}_2} = \delta_{jk}(\phi_j, \phi_k)_{\mathcal{H}_2} \tag{5.4}$$

where δ_{jk} is the Kronecker symbol and on the other hand

$$(\phi_j, \epsilon^*\phi_k)_{\mathcal{H}_1} = (\phi_j, \frac{|\phi_k|_{\mathcal{H}_2}^2}{|\phi_k|_{\mathcal{H}_1}^2} \phi_k)_{\mathcal{H}_1} = \underbrace{\frac{(\phi_j, \phi_k)_{\mathcal{H}_1}}{|\phi_k|_{\mathcal{H}_1}^2}}_{=\delta_{jk}} |\phi_k|_{\mathcal{H}_2}^2 = \delta_{jk}(\phi_j, \phi_k)_{\mathcal{H}_2}. \tag{5.5}$$

Comparing 5.4 and 5.5 shows that ϵ^* is the adjoint of the embedding ϵ. Furthermore

$$\epsilon^*\epsilon\frac{\phi_j}{|\phi_j|_{\mathcal{H}_1}} = \epsilon^*\frac{\phi_j}{|\phi_j|_{\mathcal{H}_1}} = \frac{{|\phi_j|_{\mathcal{H}_2}}^2}{|\phi_j|^2_{\mathcal{H}_1}}\frac{\phi_j}{|\phi_j|_{\mathcal{H}_1}}$$

shows that $\{\frac{\phi_j}{|\phi_j|_{\mathcal{H}_1}}\}_{j=1}^{\infty}$ is an orthonormal basis of eigenvectors of $\epsilon^*\epsilon \in \mathcal{J}_\infty(\mathcal{H}_1)$ and

$$\epsilon\epsilon^*\frac{\phi_j}{|\phi_j|_{\mathcal{H}_2}} = \frac{{|\phi_j|_{\mathcal{H}_2}}^2}{|\phi_j|^2_{\mathcal{H}_1}}\frac{\phi_j}{|\phi|_{\mathcal{H}_2}}$$

shows that $\{\frac{\phi_j}{|\phi_j|_{\mathcal{H}_2}}\}_{j=1}^{\infty}$ is an orthonormal basis of eigenvectors of $\epsilon\epsilon^* \in \mathcal{J}_\infty(\mathcal{H}_2)$ .Thus

$$\epsilon = \sum_{j=1}^{\infty}\frac{|\phi_j|_{\mathcal{H}_2}}{|\phi_j|_{\mathcal{H}_1}}(., \frac{\phi_j}{|\phi_j|_{\mathcal{H}_1}})\frac{\phi_j}{|\phi_j|_{\mathcal{H}_2}}$$

is a polar decomposition of ϵ and the singular numbers are given by

$$\sigma_j(\epsilon) = \frac{|\phi_j|_{\mathcal{H}_2}}{|\phi_j|_{\mathcal{H}_1}}.$$

This completes the proof. □

In the case of the Hilbert spaces $\mathcal{H}_1 = H^k(T^N)$ and $\mathcal{H}_2 = H^l(T^N)$ where T^N is the N-dimensional torus we are in the lucky situation that the assumption of Theorem 32 is given, i.e. that one orthogonal basis of both spaces exists and is known. To achieve a result for Sobolev spaces on certain manifolds Ω Gramsch uses a partition of one to factorize Ω over T^N. The method requires quite complicated techniques that are described in detail in [20, pp. 162-172]. We define the N-dimensional torus T^N (cf. [20, p.156]).

Definition 16: *For $N \in \mathbb{N}$ the N-dimensional torus T^N is defined by*

$$T^N = \mathbb{R}^N/2\pi\mathbb{Z}^N.$$

The set $\{x \in \mathbb{R}^N : -\pi \leq x_j < \pi, j = 1, ..., N\}$ is called the fundamental cube.

Clearly every element of T^N has a representative in the fundamental cube. Sobolev functions on T^N resemble functions with a period in $2\pi\mathbb{Z}$ in every variable and integrals over T^N should be thought of as integrals over the fundamental cube (cf. [20, p.156]). We consider the following set of functions:

Definition 17: *For $\nu \in \mathbb{Z}^N$ we define $\phi_\nu : \mathbb{R}^\mathbb{N} \to \mathbb{C}$ by*

$$\phi_\nu(t) = (2\pi)^{-\frac{N}{2}}exp(i(\nu_1 t_1 + ... + \nu_N t_N)), t \in \mathbb{R}^\mathbb{N}.$$

Clearly the functions ϕ_ν are $2\pi\mathbb{Z}$-periodic in every variable. We define a new norm on the closure of their span:

Definition 18: *For $k \geq 1$ and*

$$f = \sum_{\nu \in \mathbb{Z}^N} c_\nu \phi_\nu \tag{5.6}$$

put

$$|f|_k = \left(\sum_{\nu \in \mathbb{Z}^N} (1 + (\nu, \nu)_2)^k |c_\nu|^2 \right)^{\frac{1}{2}}$$

where $(.,.)_2$ is the usual euclidean product in $\mathbb{Z}^N$ and the sum 5.6 is understood to converge with respect to the Sobolev norm $|.|_{H^k(T^N)}$.

Now we see that the norms $|.|_k$ are Hilbert norms that are equivalent to $|.|_{H^k(T^N)}$ (cf. [5, p.166]) , that

$$H^k(\Omega) = \overline{span\{\phi_\nu : \nu \in \mathbb{Z}^N\}}$$

and $\{\phi_\nu\}_{\nu \in \mathbb{Z}^N}$ is an orthogonal basis for $H^k(T^N)$ with respect to $|.|_k$ and to $|.|_{H^k(T^N)}$ for every $k \geq 1$ and

$$|\phi_\nu|_k = (1 + (\nu, \nu)_2)^{\frac{k}{2}}.$$

Now Theorem 32 applies: The embedding $H^k(T^N) \hookrightarrow H^l(T^N)$ is of p-Schatten class if and only if

$$\sum_{\nu \in \mathbb{Z}^N} \left(\frac{|\phi_\nu|_l}{|\phi_\nu|_k} \right)^p = \sum_{\nu \in \mathbb{Z}^N} \left(\frac{1}{(1 + \nu_1^2 + ... + \nu_N^2)^{\frac{k-l}{2}}} \right)^p < \infty.$$

Let $V_N = \{\nu \in \mathbb{Z}^N : \nu_j \neq 0, j = 1, ..., N\}$. By the inequality of arithmetic and geometric means:

$$\sum_{\nu \in V_N} \left(\frac{1}{(1 + \nu_1^2 + ... + \nu_N^2)^{\frac{k-l}{2}}} \right)^p \leq \sum_{\nu \in \mathbb{Z}^N} \left(\frac{1}{(\nu_1^2 + ... + \nu_N^2)^{\frac{k-l}{2}}} \right)^p$$

$$\leq \sum_{\nu \in V_N} \frac{1}{\nu_1^2 ... \nu_N^2}^{\frac{k-l}{2N} p} = \sum_{\nu \in V_N} \frac{1}{|\nu_1 ... \nu_N|}^{\frac{k-l}{N} p}$$

$$\leq \left(\sum_{\nu_1 \in \mathbb{Z} \setminus \{0\}} (\frac{1}{\nu_1})^{\frac{k-l}{N} p} \right) ... \left(\sum_{\nu_N \in \mathbb{Z} \setminus \{0\}} (\frac{1}{\nu_N})^{\frac{k-l}{N} p} \right)$$

and these sums converge for $k - l > \frac{N}{p}$.
For $V_m = \{\nu \in \mathbb{Z}^N : |\{j : \nu_j \neq 0\}| = m\}$ where $0 \leq m < N$ the same argument yields convergence for $k - l > \frac{m}{p}$. So the sum over $\mathbb{Z}^N = \bigcup_{m=0}^N V_m$ converges for $k - l > \frac{N}{p}$. This implies the embedding of the Hilbert spaces $H^k(T^N) \hookrightarrow H^l(T^N)$ equipped with the norms $|.|_k, |.|_l$ are of p-Schatten class and because the norms are equivalent to the Sobolev norms we have

$$\sum_{\nu \in \mathbb{Z}^N} \left(\frac{|\phi_\nu|_{H^l(T^N)}}{|\phi_\nu|_{H^k(T^N)}} \right)^p < \infty.$$

Because $\{\phi_\nu\}$ is an orthogonal basis of $H^k(T^N)$ and $H^l(T^N)$ with respect to $|.|_{H^k(T^N)}$ and $|.|_{H^l(T^N)}$ Theorem 32 applies again. It follows the embedding of the Sobolev spaces $H^k(T^N) \hookrightarrow H^l(T^N)$ is of p-Schatten class.

For $k-l \leq \frac{N}{p}$ the embedding $H^K(\Omega) \hookrightarrow H^l(\Omega)$ is not of p-Schatten class: By what we have seen it is clear that for the sum of the p-th powers of the singular numbers σ_ν of this embedding we have

$$\sum_{\nu \in \mathbb{Z}^N} \sigma_\nu^p = \sum_{m=0}^{\infty} \sum_{\nu \in \mathbb{Z}^N, |\nu|_1 = m} \left(\frac{1}{(1+\nu_1^2+...+\nu_N^2)^{\frac{k-l}{2}}} \right)^p \tag{5.7}$$

where $|\nu|_1 = |\nu_1| + ...|\nu_N|$. For $|\nu|_1 = m$ we have

$$\sigma_\nu \geq (\frac{1}{2Nm^2})^{\frac{k-l}{2}}, m > 0$$

and for m sufficiently large $|\{\nu \in \mathbb{Z}^N : |\nu|_1 = m\}| \geq (\frac{m}{N^3})^{N-1}$. We get

$$\sum_{\nu \in \mathbb{Z}^N} \sigma_\nu^p \geq \sum_{m=1}^{\infty} (\frac{m}{N^3})^{N-1} (\frac{1}{2Nm^2})^{\frac{k-l}{2}p} \geq C(N) \sum_{m=1}^{\infty} m^{N-1-(k-l)p}$$

where $C(N)$ is some constant dependent on N. For $k-l \leq \frac{N}{p}$ it follows $N-1-(k-l)p \geq -1$ and thus that the sum 5.7 does not converge. It follows:

Lemma 33: *The embedding $H^k(T^N) \hookrightarrow H^l(T^N)$ is of p-Schatten class if and only if $k-l > \frac{N}{p}$.*

As we already mentioned this result can be generalized to certain compact manifolds, for example to the closure of bounded domains with a smooth boundary. Since the proof and even the proper formulation of this requires some effort and the involved techniques and concepts are really marginal for our further conclusions we just refer to [13, pp.84-85] and [20, pp.162-172]. Like Maurin's theorem Gram's theorem has applications to elliptic differential operators, see [13, pp.85-86].

6 Characterization of the p-Schatten classes

In this chapter we will characterize the p-Schatten classes and thus the p-Schatten embeddings. As mentioned in the introduction Gohberg and Markus (cf. [11]) achieved a characterization of the p-Schatten classes $\mathcal{J}_p(\mathcal{H})$ on one Hilbert space $\mathcal{H}$. Their proof involves some complicated arguments, e.g. transfinite induction. For two separable Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ we characterize the classes $\mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2)$. We will see that there are essentially two cases to consider. In the case $0 < p < 2$ the set

$$\left\{ \left(\sum_{j=1}^{\infty} |A\phi_j|_{\mathcal{H}_2}^p \right)^{\frac{1}{p}} : \{\phi_j\}_{j=1}^{\infty} \text{ an orthonormal basis} \right\}$$

turns out to be unbounded and to contain ∞ for all $A \in \mathcal{J}_\infty(\mathcal{H}_1, \mathcal{H}_2)\backslash\{0\}$. It is necessary and sufficient for A to be a p-Schatten mapping that this set is not identical to $\{\infty\}$ i.e. that there exists an orthonormal basis so that the sum converges. In the case $2 \leq p < \infty$ the set may be bounded and it is necessary and sufficient for A to be a p-Schatten mapping that this set does not contain ∞ i.e. that the sum converges for all orthonormal bases. We will see that this implies its boundedness by the p-Schatten norm of A. In the case $p = 2$ we know that the set shrinks to a singleton set that contains either ∞ or the Hilbert-Schmidt norm of the mapping. We start with the case $0 < p < 2$. The proof of the following theorem is due to Dunford and Schwartz (see [8, XI.9.32 Lemma]). See also [25, remarks to Theorem 1.18] and [12, p.95].

Theorem 34: *If $0 < p < 2$ and $A : \mathcal{H}_1 \to \mathcal{H}_2$ is a compact mapping then $\sum_{j=1}^{\infty} \sigma_j(A)^p < \infty$, i.e. A is of p-Schatten class if and only if*

$$\sum_{j=1}^{\infty} |A\phi_j|_{\mathcal{H}_2}^p < \infty$$

for some orthonormal basis $\{\phi_j\}_{j=1}^{\infty}$ of $\mathcal{H}_1$.

Proof: If A is of p-Schatten class then the eigenvectors of A^*A yield such an orthonormal basis. To see this let $\{\phi_j\}_{j=1}^{\infty}$ be an orthonormal basis consisting of eigenvectors and simply compute

$$\sum_{j=1}^{\infty} |A\phi_j|_{\mathcal{H}_2}^p = \sum_{j=1}^{\infty} (A\phi_j, A\phi_j)_{\mathcal{H}_2}^{\frac{p}{2}}$$

$$= \sum_{j=1}^{\infty} (A^*A\phi_j, \phi_j)_{\mathcal{H}_1}^{\frac{p}{2}} = \sum_{j=1}^{\infty} \sigma_j(A)^p < \infty.$$

Now suppose $\sum_{j=1}^{\infty} |A\phi_j|_{\mathcal{H}_2}^p < \infty$ for some orthonormal basis $\{\phi_j\}_{j=1}^{\infty}$ of $\mathcal{H}_1$. Define the mapping $B : \mathcal{H}_1 \to \mathcal{H}_2$ by $B\phi_j = |A\phi_j|_{\mathcal{H}_2}^{\frac{p}{2}-1} A\phi_j$ for j if $A\phi_j \neq 0$ and $B\phi_j = 0$ otherwise. Then

$$\sum_{j=1}^{\infty} |B\phi_j|_{\mathcal{H}_2}^2 = \sum_{j=1}^{\infty} |A\phi_j|^p < \infty$$

i.e. B is a Hilbert-Schmidt operator and thus of 2-Schatten class. Now define $T : \mathcal{H}_1 \to \mathcal{H}_1$ by $T\phi_j = |A\phi_j|^{1-\frac{p}{2}}\phi_j$ for $j = 1, 2, \ldots.$ Then T is obviously self-adjoint and of q-Schatten class where $q(1 - \frac{p}{2}) = p$ since $\sigma_j(T) = |A\phi_j|^{1-\frac{p}{2}}$ and

$$\sum_{j=1}^{\infty} |A\phi_j|_{\mathcal{H}_2}^{q(1-\frac{p}{2})} = \sum_{j=1}^{\infty} |A\phi_j|_{\mathcal{H}_2}^p < \infty.$$

By Theorem 14 it follows that $A = BT$ belongs to the r-Schatten class where $\frac{1}{r} = \frac{1}{2} + \frac{1}{q} = \frac{1}{2} + \frac{1-\frac{p}{2}}{p} = \frac{1}{p}$, i.e. A is of p-Schatten class. □

We note that the positivity of $1 - \frac{p}{2}$ is essential for the existence of $q > 0$. Thus the existence of an orthonormal basis $\{\phi_j\}$ of $\mathcal{H}_1$ so that $\{|A\phi_j|_{\mathcal{H}_2}\} \in \ell^p(\mathbb{N})$ is sufficient for $A \in \mathcal{J}_p(\mathcal{H}_1, \mathcal{H}_2)$. If we think the other way around the question arises if for $0 < p < 2$ there always exists an orthonormal basis $\{\gamma_j\}_{j=1}^{\infty}$ so that $\{|A\gamma_j|_{\mathcal{H}_2}\}_{j=1}^{\infty} \notin \ell^p(\mathbb{N})$. The answer is positive if $A \neq 0$ (cf. [11, Theorem 3]).

Lemma 35: *For $0 < p < 2$ and a non-zero bounded linear mapping $A : \mathcal{H}_1 \to \mathcal{H}_2$ there always exists an orthonormal basis $\{\gamma_j\}_{j=1}^{\infty}$ so that*

$$\sum_{j=1}^{\infty} |A\gamma_j|_{\mathcal{H}_2}^p = \infty.$$

Proof: If A is not compact then for all orthonormal bases $\{\eta_j\}_{j=1}^{\infty}$ we have

$$\sum_{j=1}^{\infty} |A\eta_j|_{\mathcal{H}_2}^2 = \infty \tag{6.1}$$

because otherwise for $\gamma \in \mathcal{H}_1$

$$\sum_{j=1}^{\infty} |(\gamma, \eta_j)_{\mathcal{H}_1}||A\eta_j|_{\mathcal{H}_2} \leq \underbrace{(\sum_{j=1}^{\infty} |(\gamma, \eta_j)_{\mathcal{H}_1}|^2)^{\frac{1}{2}}}_{=|\gamma|} \underbrace{(\sum_{j=1}^{\infty} |A\eta_j|_{\mathcal{H}_2}^2)^{\frac{1}{2}}}_{<\infty} < \infty$$

implied that A is the limit of the finite rank mappings $\sum_{j=1}^{N} |(., \eta_j)_{\mathcal{H}_1}||A\eta_j|_{\mathcal{H}_2}$ for $N \to \infty$, i.e. A is compact. Thus for all non-compact A:

$$\sum_{j=1}^{\infty} |A\eta_j|_{\mathcal{H}_2}^p = \infty$$

for all $0 < p < 2$ and all orthonormal bases $\{\eta\}_{j=1}^{\infty}$ of $\mathcal{H}_1$.
If A is compact it admits a polar decomposition:

$$A = \sum_{j=1}^{\infty} \sigma_j(A)(., \phi_j)_{\mathcal{H}_1} \psi_j.$$

Now we fix some orthonormal basis $\{\eta_j\}_{j=1}^{\infty}$ of $\mathcal{H}_1$. Then $s_1 = \{(\eta_j, \phi_1)_{\mathcal{H}_1}\}_{j=1}^{\infty}$ is a sequence of length 1 in $\ell^2(\mathbb{N})$. The sequence $s = \{\frac{1}{i^{\frac{1}{p}}}\}$ is in $\ell^2(\mathbb{N})$. Thus $s_2 = \frac{1}{|s|_{\ell^2}} s$ is a sequence of length 1 in $\ell^2(\mathbb{N})$. Whence there exists a map $\alpha : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ that respects the inner product on $\ell^2(\mathbb{N})$ and maps s_1 to s_2. It can be represented by an orthogonal matrix $\{\alpha_{ij}\}_{i,j=1}^{\infty}$ so that

$$\sum_{j=1}^{\infty} \alpha_{ij} \overline{\alpha_{jk}} = \delta_{ik} \tag{6.2}$$

and

$$\sum_{j=1}^{\infty} \alpha_{ij} (\eta_j, \phi_1)_{\mathcal{H}_1} = \frac{1}{|s|_{\ell^2}} \frac{1}{i^{\frac{1}{p}}}. \tag{6.3}$$

Now define $\gamma_i = \sum_{j=1}^{\infty} \alpha_{ij} \eta_j$. Then by 6.2 it follows that $\{\gamma_i\}_{i=1}^{\infty}$ is an orthonormal system and by 6.3 for $i = 1, 2, ..$ one has $(\gamma_i, \phi_1)_{\mathcal{H}_1} = \frac{1}{|s|_{\ell^2}} \frac{1}{i^{\frac{1}{p}}}$. For every i using the polar decomposition

$$|A\gamma_i|_{\mathcal{H}_2}^2 = \sum_{j=1}^{\infty} \sigma_j^2(A) |(\gamma_i, \phi_j)_{\mathcal{H}_1}|^2 \geq \sigma_1^2(A) |(\gamma_i, \phi_1)_{\mathcal{H}_1}|^2$$

so $|A\gamma_i|_{\mathcal{H}_2} \geq \sigma_1(A) |(\gamma_i, \phi_1)_{\mathcal{H}_1}|$ and so for $0 < p < 2$

$$\sum_{i=1}^{\infty} |A\gamma_i|_{\mathcal{H}_2}^p \geq \sum_{i=1}^{\infty} \sigma_1^p(A) |(\gamma_i, \phi_1)_{\mathcal{H}_1}|^p = \underbrace{\frac{\sigma_1^p(A)}{|s|_{\ell^2}^p}}_{>0} \sum_{i=1}^{\infty} \frac{1}{i} = \infty$$

This completes the proof. □

We see that by Theorem 34 we obtain a characterization of the p-Schatten embeddings for $0 < p < 2$:

Corollary 4: *For $0 < p < 2$ a compact embedding $\mathcal{H}_1 \hookrightarrow \mathcal{H}_2$ is a p-Schatten embedding if and only if $\{|\phi_j|_{\mathcal{H}_2}\}_{j=1}^{\infty} \in \ell^p(\mathbb{N})$ for some orthonormal basis $\{\phi_j\}_{j=1}^{\infty}$ of $\mathcal{H}_1$.*

The case $2 \leq p$ needs some more effort. In the border case $p = 2$ we know that $\{|A\phi_j|_{\mathcal{H}_2}\} \in \ell^2(\mathbb{N})$ for an orthonormal basis $\{\phi_j\}$ of $\mathcal{H}_1$ if and only if $\{|A\phi_j|_{\mathcal{H}_2}\} \in \ell^2(\mathbb{N})$ for all orthonormal bases $\{\phi_j\}$ of $\mathcal{H}_1$. For $2 \leq p$ the function $x \mapsto x^{\frac{p}{2}}$ is convex and we will use the convexity of this function. We define doubly substochastic matrices.

Definition 19: *$\{\alpha_{jk}\}_{j,k=1}^{\infty}$ is called a doubly substochastic matrix if $\alpha_{jk} \geq 0$ and*

$$\sum_{j=1}^{\infty} \alpha_{jk} \leq 1, k = 1, 2, ... \text{ and } \sum_{k=1}^{\infty} \alpha_{jk} \leq 1, j = 1, 2, ...$$

We use the following technical lemma due to Fan (cf. [10, Lemma 1A] and [25, Proposition 1.12]):

Lemma 36: *Let* $\{\alpha_{jk}\}_{j,k=1}^{\infty}$ *be a doubly substochastic matrix and* $b = \{b_j\}_{j=1}^{\infty}$ *a non-negative, non-increasing sequence,such that* $\{\alpha_{jk}b_k\}_{k=1}^{\infty} \in \ell^1(\mathbb{N})$ *for each* $j \in \mathbb{N}$ *then*

$$\sum_{j=1}^{m}\sum_{k=1}^{\infty} \alpha_{jk}b_k \leq \sum_{j=1}^{m} b_j$$

for any $m \in \mathbb{N}$.

PROOF: Let $c_j = \sum_{k=1}^{\infty} \alpha_{jk}b_k$ for $j \in \mathbb{N}$ and $d_k = \sum_{j=1}^{m} \alpha_{jk}$ for $k \in \mathbb{N}$ and for some fixed integer m. Then $0 \leq d_k \leq 1$ and

$$\sum_{k=1}^{\infty} d_k = \sum_{j=1}^{m} \underbrace{\sum_{k=1}^{\infty} \alpha_{jk}}_{\leq 1} \leq m.$$

This and the fact that b is non-increasing imply

$$\sum_{k=1}^{\infty} d_k b_k \leq \sum_{k=1}^{m-1} d_k b_k + (m - \sum_{k=1}^{m-1} d_k) b_m \leq \sum_{k=1}^{m-1} d_k(b_k - b_m) + m b_m \leq \sum_{k=1}^{m-1}(b_k - b_m) + m b_m$$

since $d_k \leq 1$ and so

$$\sum_{k=1}^{\infty} d_k b_k = \sum_{j=1}^{m}\sum_{k=1}^{\infty} \alpha_{jk} b_k \leq \sum_{k=1}^{m-1}(b_k - b_m) + m b_m = \sum_{j=1}^{m} b_j.$$

We use this lemma to prove (cf. [10, Theorem 1]):

Theorem 37: *For* $H = H^*$ *a compact, non-negative operator on* $\mathcal{H}_1$,$\{\lambda_j\}_{j=1}^{\infty}$ *the sequence of its eigenvalues in non-increasing order we have:*

$$\max \sum_{j=1}^{m} (H\psi_j, \psi_j) = \sum_{j=1}^{m} \lambda_j \text{ for any } m \in \mathbb{N} \tag{6.4}$$

where the maximum is taken over all finite orthonormal sets $\{\psi\}_{j=1}^{m}$ *of order* m *in* $\mathcal{H}_1$.

PROOF: Let $\{\psi_j\}$ be an arbitrary orthonormal set in $\mathcal{H}_1$. By the spectral theorem there exists an orthonormal family of eigenvectors $\{\phi_j\}_{j=1}^{\infty}$, i.e. $H\phi_j = \lambda_j \phi_j$ for $j = 1, 2, ...$ ($\{\phi_j\}_{j=1}^{\infty}$ is not necessarily a basis, because we leave out the eigenvectors in the kernel, if there are infinitely many non-zero eigenvectors.) We have $\sum_{j=1}^{m} (H\phi_j, \phi_j) = \sum_{j=1}^{m} \lambda_j$ for every $m \in \mathbb{N}$.
We have for $1 \leq j \leq m$:

$$(H\psi_j, \psi_j) = \sum_{k=1}^{\infty} \lambda_k |(\psi_j, \phi_k)|^2.$$

By Bessel's inequality $\{|(\psi_j,\phi_k)|^2\}_{j,k}$ is a doubly substochastic matrix and so Lemma 36 implies:

$$\sum_{j=1}^{m}(H\psi_j,\psi_j) \le \sum_{j=1}^{m}\lambda_j \text{ for any } m\in\mathbb{N}.$$

□

In our application of this theorem we will put $H = A^*A$ and use $(A^*A\psi,\psi) = |A\psi|^2$. For a sequence $\alpha \in c_0$, where c_0 is the space of sequences of complex numbers tending to zero let α^* be the sequence of its absolute values in non-increasing order. We define (cf. [11]) the relations $\prec\prec$ and $\prec$ on the spaces c_0 and $\ell^1(\mathbb{N})$ respectively.

Definition 20: *We write $\beta \prec\prec \alpha$ for $\alpha,\beta\in c_0$ if*

$$\sum_{j=1}^{k}\beta_j^* \le \sum_{j=1}^{k}\alpha_j^* \textit{ for } k=1,2,... \tag{6.5}$$

We write $\beta\prec\alpha$ for $\alpha,\beta\in\ell^1(\mathbb{N})$ if

$$\beta\prec\prec\alpha \textit{ and } \sum_{j=1}^{\infty}\beta_j^* = \sum_{j=1}^{\infty}\alpha_j^*. \tag{6.6}$$

For H as above Theorem 37 shows that for any orthonormal basis $\{\phi_j\}_{j=1}^\infty$:

$$\{(H\phi_j,\phi_j)\}_{j=1}^\infty \prec\prec \{\lambda_j\}_{j=1}^\infty$$

and if $\{\lambda_j\}_{j=1}^\infty \in \ell^1(\mathbb{N})$

$$\{(H\phi_j,\phi_j)\}_{j=1}^\infty \prec \{\lambda_j\}_{j=1}^\infty$$

because $tr(H) = \sum_{j=1}^\infty (H\phi_j,\phi_j)$ does not depend on the basis for a non-negative operator. Note that the sets $\{|\alpha_j|\}$ and $\{\alpha_j^*\}$ need not be identical, as the example $\alpha = \{1,0,\frac{1}{2},0,\frac{1}{3},0,..\}$, $\alpha^* = \{1,\frac{1}{2},\frac{1}{3},...\}$ shows, whereas the sets $\{|\alpha_j|\}-\{0\}$ and $\{\alpha_j^*\}-\{0\}$ are identical, even counting multiplicities. Also note that for α,β non-increasing sequences of non-negative numbers $\alpha\prec\beta$ implies $\inf\sum_{m=1}^k \alpha_{j_m} \ge \inf\sum_{m=1}^k \beta_{j_m}$, where the infimum is taken over all sets of k different integers $\{j_1,...,j_k\}$. When we say for a non-negative compact operator H, let $\{\lambda_j(H)\}$ be its non-increasing sequence of eigenvalues, we actually mean $\{\lambda_j^*(H)\}$, for some sequence $\{\lambda_j(H)\}$ of its eigenvalues, counting multiplicities, but then 0 appears in this sequence, only if there are finitely many non-zero eigenvalues, counting multiplicities. We want to show that if $\beta\prec\prec\alpha$ for non-negative sequences of real numbers and a convex function $f:[0,\infty)\to[0,\infty)$ we have $\sum_{j=1}^k f(\beta_j) \le \sum_{j=1}^k f(\alpha_j)$ for $k=1,2,...$, i.e. $\{f(\beta_j)\}_{j=1}^\infty \prec\prec \{f(\alpha_j)\}_{j=1}^\infty$, so that convex f respect the relation $\prec\prec$. This needs some preparation.
We fix an integer N and for a vector $\alpha\in\mathbb{C}^N$ we define the vector $\alpha^*\in[0,\infty)^N$ as the vector of absolute values $|\alpha_j|$ renumbered, so that $\alpha_1^*\ge\alpha_2^*\ge...\ge\alpha_N^*$. Then the following lemma holds (cf. [25, Lemma 1.8]).

Lemma 38: *For $\alpha, \beta \in \mathbb{C}^N$*

$$\sum_{j=1}^{N} |\alpha_j \beta_j| \leq \sum_{j=1}^{N} \alpha_j^* \beta_j^*.$$

Proof: Without loss of generality we can renumber the coordinates of α and β so that $|\alpha_1| \geq |\alpha_2| \geq ...|\alpha_N|$. We compute

$$\sum_{j=1}^{N} |\alpha_j \beta_j| = |\alpha_N| \sum_{j=1}^{N} |\beta_j| + (|\alpha_{N-1}| - |\alpha_N|) \sum_{j=1}^{N-1} |\beta_j| + ... + (|\alpha_1| - |\alpha_2|)|\beta_1|. \tag{6.7}$$

By definition of the $*$-operation we have

$$\sum_{j=1}^{k} |\beta_j| \leq \sum_{j=1}^{k} \beta_j^* \text{ k=1,2,...,N.} \tag{6.8}$$

6.7 and 6.8 imply

$$\sum_{j=1}^{N} |\alpha_j \beta_j| \leq \sum_{j=1}^{N} \alpha_j^* \beta_j^*.$$

We will use this lemma in the proof of the following theorem that according to Simon is essentially due to Markus while the proof is due to Mityagin. Our classification in the case $p \geq 2$ is primarily based on this theorem (cf. [25, Theorem 1.9] and [16] and[18]).

Theorem 39: *If $\alpha \in [0, \infty)^N$ with $\alpha_1 \geq \alpha_2 \geq .... \geq \alpha_N$ and $\beta \in \mathbb{C}^N$ satisfy*

$$\sum_{j=1}^{k} \beta^* \leq \sum_{j=1}^{k} \alpha_j \textit{ for k=1,2,...,N} \tag{6.9}$$

then there exist points $\alpha^{(1)}, ..., \alpha^{(m)} \in \mathbb{C}^N$,$m \in \mathbb{N}$ with $\alpha^{(l)*} = \alpha$ for $l = 1, ..., m$ and $0 \leq \lambda_l \leq 1$ with $\sum_{l=1}^{m} \lambda_l = 1$ such that

$$\beta = \sum_{l=1}^{m} \lambda_l \alpha^{(l)}.$$

Proof: Let $A = \{\gamma \in \mathbb{C}^N : \gamma^* = \alpha\}$ and $\mathrm{conv}(A)$ its convex hull. The statement of the theorem is then: If $\beta \in \mathbb{C}^N$ obeys 6.9, then $\beta \in \mathrm{conv}(A)$. We prove that for each complex linear function ℓ the following holds

$$|\ell(\beta)| \leq \max_{\gamma \in \mathrm{conv}(A)} \mathrm{Re}(\ell(\gamma)). \tag{6.10}$$

If we define $\eta_j = (\delta_{jk})_{k=1}^N$ for $j = 1, ..., N$, where δ_{jk} is the Kronecker symbol and $\ell_j = \ell(\eta_j)$ for $j = 1, ..., N$ then $\ell(\beta) = \sum_{j=1}^{N} \ell_j \beta_j$ and

$$|\ell(\beta)| \leq \sum_{j=1}^{N} |\ell_j \beta_j| \leq \sum_{j=1}^{N} \ell_j^* \beta_j^* \tag{6.11}$$

by Lemma 38. Furthermore

$$\sum_{j=1}^{N} \ell_j^* \beta_j^* = \ell_N^* \sum_{j=1}^{N} \beta_j^* + (\ell_{N-1}^* - \ell_N^*) \sum_{j=1}^{N-1} \beta_j^* + ... + (\ell_1^* - \ell_2^*)\beta_1^* \tag{6.12}$$

$$\leq \ell_N^* \sum_{j=1}^{N} \alpha_j + (\ell_{N-1}^* - \ell_N^*) \sum_{j=1}^{N-1} \alpha_j + ... + (\ell_1^* - \ell_2^*)\alpha_1 = \sum_{j=1}^{N} \ell_j^* \alpha_j$$

by 6.9. But there exists an element $\hat{\alpha} \in A$ so that $\ell(\hat{\alpha}) = \sum_{j=1}^{N} \ell_j^* \alpha_j$. This together with 6.11 and 6.12 imply 6.10. To see the existence of $\hat{\alpha}$ put $\ell_j^* = |\ell_{k_j}|$ and define $\hat{\alpha}$ by $\hat{\alpha}_{k_j} = \alpha_j$ if $|\ell_{k_j}| = 0$ and $\hat{\alpha}_{k_j} = \ell_{k_j}^{-1} |\ell_{k_j}| \alpha_j$ otherwise. Then $\hat{\alpha}^* = \alpha$, i.e. $\hat{\alpha} \in A$ and

$$\ell(\hat{\alpha}) = \sum_{k_j} \ell_{k_j} \hat{\alpha}_{k_j} = \sum_{k_j, |\ell_{k_j}| \neq 0} \ell_{k_j} \ell_{k_j}^{-1} |\ell_{k_j}| \alpha_j = \sum_{k_j, |\ell_{k_j}| \neq 0} |\ell_{k_j}| \alpha_j = \sum_{j=1}^{N} \ell_j^* \alpha_j.$$

Now suppose that $\beta \notin \mathrm{conv}(A)$. Since $\mathrm{conv}(A)$ is closed, convex and bounded, there is a hyperplane $E = \{\gamma \in \mathbb{C}^N : \ell(\gamma) = s\}$ separating $\mathrm{conv}(A)$ and β:

$$\ell(\gamma) < s \text{ for } \gamma \in \mathrm{conv}(A) \text{ and } \ell(\beta) > s \tag{6.13}$$

for a real valued functional ℓ and $s > 0$ that can be chosen so that 6.13 contradicts 6.10. It follows $\beta \in \mathrm{conv}(A)$. □

The importance of Theorem 39 for our purpose consists in the following two corollaries.

Corollary 5: *If $\hat{F} : [0, \infty)^N \to \mathbb{R}$ is a function so that $F : \mathbb{C}^N \to \mathbb{R}, \gamma \mapsto \hat{F}(\gamma^*)$ is convex and $\beta \in \mathbb{C}^N$ and $\alpha \in [0, \infty)^N$ obey 6.9 then*

$$F(\beta) \leq F(\alpha).$$

Proof: We have

$$F(\beta) = \hat{F}((\sum_{l=1}^{m} \lambda_l \alpha^{(l)})^*) \leq \sum_{l=1}^{m} \lambda_l \hat{F}(\alpha^{(l)^*}) = \sum_{l=1}^{m} \lambda_l F(\alpha) = F(\alpha).$$

□

Corollary 6: *If $f : [0, \infty) \to [0, \infty)$ is a monotone and convex function then $F : \mathbb{C}^N \to [0, \infty), \beta \mapsto \sum_{j=1}^{N} f(\beta_j^*)$ is a convex function and if $\alpha = \{\alpha_j\}_{j=1}^{\infty}$ and $\beta = \{\beta_j\}_{j=1}^{\infty}$ are non-increasing sequences of non-negative real numbers, satisfying $\beta \prec\prec \alpha$ then*

$$\sum_{j=1}^{k} f(\beta_j) \leq \sum_{j=1}^{k} f(\alpha_j) \text{ for } k = 1, 2, ...$$

, i.e. $\{f(\beta_j)\}_{j=1}^{\infty} \prec\prec \{f(\alpha_j)\}_{j=1}^{\infty}$.

PROOF: For $\alpha, \beta \in \mathbb{C}^N$ and $0 \le \theta \le 1$ we have

$$F(\theta\alpha + (1-\theta)\beta) = \sum_{j=1}^{N} f((\theta\alpha + (1-\theta)\beta)_j^*) = \sum_{j=1}^{N} f(|\theta\alpha_j + (1-\theta)\beta_j|)$$

$$\le \sum_{j=1}^{N} f(\theta|\alpha_j| + (1-\theta)|\beta_j|) \le \sum_{j=1}^{N} \theta f(|\alpha_j|) + (1-\theta) f(|\beta_j|) = \theta F(\alpha) + (1-\theta)F(\beta). \quad \square$$

Now consider a mapping $A : \mathcal{H}_1 \to \mathcal{H}_2$ as mentioned at the beginning and its canonical decomposition. Let $p \ge 2$. Then the map $f : [0,\infty) \to [0,\infty), x \mapsto x^{\frac{p}{2}}$ is convex.The following holds:

Theorem 40: *If $A : \mathcal{H}_1 \to \mathcal{H}_2$ is a compact mapping and $1 \le N < \infty$ then*

$$\max \sum_{j=1}^{N} |A\phi_j|_{\mathcal{H}_2}^p = \sum_{j=1}^{N} \sigma_j^p(A)$$

where the maximum is taken over all orthonormal systems $\{\phi_j\}_{j=1}^N$.

PROOF: Clearly $H = A^*A$ is self-adjoint and non-negative. By Theorem 37 we have for any orthonormal system $\{\phi_j\}_{j=1}^N$ in $\mathcal{H}_1$:

$$\sum_{j=1}^{k} (H\phi_j, \phi_j)_{\mathcal{H}_1} \le \sum_{j=1}^{k} \lambda_j(H) \text{ for } k = 1, 2, ..., N \tag{6.14}$$

but since $(H\phi_j, \phi_j)_{\mathcal{H}_1} = |A\phi_j|_{\mathcal{H}_2}^2$ and $\lambda_j(H) = \sigma_j^2(A)$ where $\sigma_j(A)$ are the singular values of A relation 6.14 becomes

$$\sum_{j=1}^{k} |A\phi_j|_{\mathcal{H}_2}^2 \le \sum_{j=1}^{k} \sigma_j^2(A) \text{ for } k = 1, 2, ..., N \tag{6.15}$$

and then the convexity of $f(x) = x^{\frac{p}{2}}$ $(p \ge 2)$ and Corollary 6 imply

$$\sum_{j=1}^{k} |A\phi_j|_{\mathcal{H}_2}^p \le \sum_{j=1}^{k} \sigma_j^p(A) \text{ for } k = 1, 2, ..., N \tag{6.16}$$

and we get equality in 6.15 and thus in 6.16 , if we take for $\{\phi_j\}_{j=1}^N$ the eigenvectors of A^*A in (1). $\square$

As a direct consequence of Theorem 40 we get:

Corollary 7: *If $A : \mathcal{H}_1 \to \mathcal{H}_2$ is a compact mapping then*

$$\max \sum_{j=1}^{\infty} |A\phi_j|_{\mathcal{H}_2}^p = \sum_{j=1}^{\infty} \sigma_j^p(A)$$

where the maximum is taken over all orthonormal bases $\{\phi_j\}_{j=1}^\infty$ of $\mathcal{H}_1$.

PROOF: By Theorem 40 for all orthonormal bases $\{\phi_j\}_{j=1}^{\infty}$ we have

$$\sum_{j=1}^{\infty} |A\phi_j|_{\mathcal{H}_2}^p \leq \sum_{j=1}^{\infty} \sigma_j^p(A) \tag{6.17}$$

and so taking a completion of the eigenvectors for A^*A , equation 6.17 becomes an equality.

Finally we apply this result to embeddings:

Corollary 8: *For $2 \leq p$ a compact embedding $\mathcal{H}_1 \hookrightarrow \mathcal{H}_2$ is a p-Schatten embedding if and only if $\{|\phi_j|_{\mathcal{H}_2}\}_{j=1}^{\infty} \in \ell^p(\mathbb{N})$ for every orthonormal basis $\{\phi_j\}_{j=1}^{\infty}$ of $\mathcal{H}_1$.*

In this Corollary we assumed $\mathcal{H}$ to be infinite dimensional. Of course every finite dimensional space can be embedded via a p-Schatten embedding. It catches one's eye that the methods used for $0 < p < 2$ and $2 \leq p$ differ completely. This is not the case in the proof given by Gohberg and Markus but they use a theorem (cf. [11, Theorem 1]) whose proof is not easy to understand. The characterization of the p-Schatten embeddings bases upon summability conditions on orthonormal bases whilst the characterization of compact embeddings in $\ell^2(\mathbb{N})$ given by Hanche-Olsen and Holden bases upon summability conditions on elements of bounded sets. Furhtermore one has to know at least one orthonormal basis to get a vivid description. We will encounter such applications in the next chapter. In respect of our aim to characterize the p-Schatten embeddings $\mathcal{H} \hookrightarrow \ell^2(\mathbb{N})$ we just apply our general result:

Corollary 9: *A compact embedding $\mathcal{H} \hookrightarrow \ell^2(\mathbb{N})$ is a p-Schatten embedding if and only if for every ($2 \leq p$)/for some (0<p<2) orthonormal basis $\{f^k\}_{k=1}^{\infty}$ of $\mathcal{H}$:*

$$\sum_{k=1}^{\infty} \left(\sum_{j=1}^{\infty} |f_j^k|^2 \right)^{\frac{p}{2}} < \infty.$$

7 Applications

In this chapter we apply the result of the preceding chapter to several spaces. As we already mentioned we need to know an orthonormal basis. In the first section we study spaces that we construct by imitating Sobolev spaces on the discrete set $\mathbb{N}$. In the second section we study the embeddings of discrete Sobolev spaces on weighted graphs. In the third section we embed Sobolev spaces on sparse graphs.

7.1 The spaces $s_{*,\alpha}$

For the construction of the following spaces we recall the definition of the right shift operator and the left shift operator on $\ell^2(\mathbb{N})$:

Definition 21: *For a sequence $a = \{a_j\}_{j=1}^{\infty} \in \ell^2(\mathbb{N})$ the right shift operator S is defined by*

$$Sa = \{a_{j-1}\}_{j=1}^{\infty} \text{ where } a_0 = 0$$

and the left shift operator T is defined by

$$Ta = \{a_{j+1}\}_{j=1}^{\infty}.$$

Recall that the left shift operator is the adjoint of the right shift operator: $T = S^*$. Note that the right and left shift operator resemble creation and annihilation operators in quantum physics that are weighted forward and backward shifts defined on an orthonormal basis of a Hilbert space. For this see [4, 2.1.2.10 and 2.2.1.2.].

We consider a Sobolev like norm on the space $\ell^2(\mathbb{N})$, where the left shift operator acts similar to a derivative. To "approach the boundary of $\mathbb{N}$" we apply it infinitely many times:

$$|a|_* = \left(\sum_{n=0}^{\infty} |(S^*)^n a|_2^2\right)^{\frac{1}{2}}$$

for $a \in \ell^2(\mathbb{N})$, where $|.|_2$ is the usual ℓ^2-norm and S^* is the adjoint of the shift operator, the left shift operator.
It is associated with the inner product

$$(a, b)_* = \sum_{n=0}^{\infty} ((S^*)^n a, (S^*)^n b)_2$$

where $a, b \in \ell^2(\mathbb{N})$ and $(.,.)_2$ is the usual inner product on ℓ^2.
We compute:

$$|a|_*^2 = \sum_{j=1}^{\infty} |a_j|^2 + \sum_{j=1}^{\infty} |a_{j+1}|^2 + ... = \sum_{j=1}^{\infty} j|a_j|^2 = |\{j^{\frac{1}{2}} a_j\}_{j=1}^{\infty}|_2^2$$

and define more generally for $\alpha \geq 0$ and $a, b \in \ell^2(\mathbb{N})$:

Definition 22:

$$|a|_{*,\alpha} = |\{j^\alpha a_j\}_{j=1}^\infty|_2,$$
$$(a,b)_{*,\alpha} = (\{j^\alpha a_j\}_{j=1}^\infty, \{j^\alpha b_j\}_{j=1}^\infty)_2 \text{ and}$$
$$s_{*,\alpha} = \{a \in \ell^2(\mathbb{N}) | |a|_{*,\alpha} < \infty\}.$$

We have $|.|_* = |.|_{*,\frac{1}{2}}$ and $|.|_2 = |.|_{*,0}$. The following theorem holds:

Theorem 41: *$s_{*,\alpha}$ are Hilbert spaces and*
$s_{*,\alpha} \hookrightarrow \ell^2(\mathbb{N})$ are compact embeddings if $\alpha > 0$ and
$s_{*,\alpha} \hookrightarrow \ell^2(\mathbb{N})$ are p-Schatten embeddings for $0 < p \leq 2$ if $p\alpha > 1$.

PROOF: $|.|_{*,\alpha}$ is a norm because $|.|_2$ is a norm and thus $s_{*,\alpha}$ are normed linear spaces. The norm is induced by the inner product $(.,.)_{*,\alpha}$ that is an inner product because $(.,.)_2$ is an inner product. Thus $s_{*,\alpha}$ are pre-Hilbert spaces. If $\{a^n\}_{n=1}^\infty$ is a Cauchy sequence in $s_{*,\alpha}$, there exists a sequence $\hat{a}$ in $\ell^2(\mathbb{N})$ such that

$$|\hat{a} - \{j^\alpha a_j^n\}_{j=1}^\infty|_2 \to 0, n \to \infty, |\hat{a}|_2 < \infty$$

because $\{j^\alpha a_j^n\}_{j=1}^\infty$ for $n = 1, 2, ...$ constitute a Cauchy sequence with respect to $|.|_2$ and $\ell^2(\mathbb{N})$ is a Hilbert space. Define a by $a_j = j^{-\alpha}\hat{a}_j$ then,

$$|a - a^n|_{*,\alpha} \to 0, n \to \infty,$$

and $|a|_{*,\alpha} < \infty$ by a standard argument and thus $s_{*,\alpha}$ are Hilbert spaces.
To show the second assertion we apply Corollary 1 in chapter 4. Let $\alpha > 0$. We denote by $B^1_{*,\alpha}(0)$ the unit ball $\{a \in s_{*,\alpha} : |a|_{*,\alpha} \leq 1\}$. Clearly all elements of $B^1_{*,\alpha}(0)$ are pointwise bounded. We have to show that for every $\varepsilon > 0$ there exists a number $n(\varepsilon) \in \mathbb{N}$ so that all $a \in B^1_{*,\alpha}(0)$ satisfy

$$\sum_{j=n(\varepsilon)+1}^{\infty} |a_j|^2 < \varepsilon^2.$$

Assume on the contrary that there exists a number $\varepsilon > 0$ so that for all $n \in \mathbb{N}$ there exists a sequence $a \in B^1_{*,\alpha}(0)$ so that

$$\sum_{j=n+1}^{\infty} |a_j|^2 \geq \varepsilon^2.$$

Choose $N = \lfloor \varepsilon^{-\frac{1}{\alpha}} \rfloor + 1$ then by our assumption there exists a sequence $a \in B^1_{*,\alpha}(0)$ so that

$$|a|^2_{*,\alpha} = \sum_{j=1}^{\infty} j^{2\alpha}|a_j|^2 \geq \sum_{N+1}^{\infty} j^{2\alpha}|a_j|^2$$

$$\underbrace{>}_{\alpha>0} N^{2\alpha} \sum_{j=N+1}^{\infty} |a_j|^2 > \varepsilon^{-2}\varepsilon^2 = 1$$

a contradiction. Thus by Corollary 1 $s_{*,\alpha} \hookrightarrow \ell^2(\mathbb{N})$ is a compact embedding for $\alpha > 0$.

If we define $e^{m,\alpha}$ by $e_j^{m,\alpha} = m^{-\alpha}\delta_{m,j}$ for $j = 1, 2, ...$ then $\{e^{m,\alpha}\}_{m=1}^{\infty}$ is an orthonormal basis for $s_{*,\alpha}$. Thus by Corollary 4 the equation

$$\sum_{m=1}^{\infty} |e^{m,\alpha}|_2^p = \sum_{m=1}^{\infty} m^{-p\alpha},$$

shows the third assertion since the above sum converges if $p\alpha > 1$. □

In particular: $s_{*,\alpha} \hookrightarrow \ell^2(\mathbb{N})$ is a Hilbert-Schmidt embedding for $\alpha > \frac{1}{2}$ and $s_{*,\alpha} \hookrightarrow \ell^2(\mathbb{N})$ is of trace class for $\alpha > 1$. If for some $p \geq 2$ one has $p\alpha \leq 1$ then it follows by Corollary 8 that the embedding $s_{*,\alpha} \hookrightarrow \ell^2(\mathbb{N})$ is no p-Schatten embedding and thus by Lemma 16 no p_1-Schatten embedding for all $0 < p_1 \leq p$.
In exactly the same way one proves that if $s_{*,\alpha} \hookrightarrow s_{*,\beta}$ are compact for $\alpha - \beta > 0$ then $s_{*,\alpha} \hookrightarrow s_{*,\beta}$ are p-Schatten embeddings for $0 < p \leq 2$ if $(\alpha - \beta)p > 1$.

Note the analogy to the result by Gramsch where the condition for the embeddings $H^k(\Omega) \to H^l(\Omega)$ to be p-Schatten embeddings was $(k - l)p > N$.

7.2 Discrete Sobolev spaces on weighted graphs

There are other spaces where an orthonormal basis is not easily obtained. As an example we consider discrete Sobolev spaces on graphs. We will need some preparation to understand this setting. We define a directed graph (cf. [19, Definition A.1]):

Definition 23: *For a finite or countable set V and a subset $E \subset V \times V$ the set $G = (V, E)$ is called a directed graph, the elements of V are called the vertices of the graph and the elements of E its edges. G is called simple if for all $v, w \in V$*

at most one of the pairs (v, w) and (w, v) is an element of E,

$$(v, v) \notin E.$$

For an edge $e = (v, w)$ we call $e_{init} = v$ its initial endpoint and $e_{term} = w$ its terminal endpoint.

Now we consider weights on vertices and edges and define function spaces on V and E (cf. [19, Definition 3.1 and 3.2]):

Definition 24: *For $\nu : V \to (0, \infty)$ and $\rho : E \to (0, \infty)$ and $1 \leq q < \infty$ we define the spaces $\ell^q_\nu(V)$ and $\ell^q_\rho(E)$ by*

$$\ell^q_\nu(V) = \{f : V \to \mathbb{C} : \left(\sum_{v \in V} |f(v)|^q \nu(v)\right)^{\frac{1}{q}} < \infty\} \text{ and}$$

$$\ell^q_\rho(E) = \{u : E \to \mathbb{C} : \left(\sum_{e \in E} |u(e)|^q \rho(e)\right)^{\frac{1}{q}} < \infty\}$$

respectively.

We are now able to define discrete Sobolev spaces of first order on directed graphs (cf. [19],):

Definition 25: *For a function* $f \in \ell^q_\nu(V)$ *we define its derivative* $f^{'}$ *on* E *by*

$$f^{'}(e) = f(e_{term}) - f(e_{init}).$$

The Sobolev space $w^{1,q}_{\rho,\nu}$ *is defined by*

$$w^{1,q}_{\rho,\nu}(V) = \{f \in \ell^q_\nu(V) : f^{'} \in \ell^q_\rho(E)\}.$$

Without proof we cite [19, Lemma 3.6]:

Lemma 42: *For* $1 \leq q < \infty$ *the space* $w^{1,q}_{\rho,\nu}(V)$ *is a Banach space with respect to the norm defined by*

$$|f|^q_{w^{1,q}_{\rho,\nu}(V)} = \sum_{v \in V} |f(v)|^q \nu(v) + \sum_{e \in E} |f^{'}(e)|^q \rho(e).$$

From 42 it follows immediately that $w^{1,2}_{\rho,\nu}(V)$ is a Hilbert space with respect to the inner product $(.,.)_{w^{1,2}_{\rho,\nu}(V)}$ defined by

$$(f,g)_{w^{1,2}_{\rho,\nu}(V)} = \sum_{v \in V} f(v)\overline{g(v)}\nu(v) + \sum_{e \in E} f^{'}(e)\overline{g^{'}(e)}\rho(e).$$

We need to define the distance of two vertices (cf. [19, Definition 3.8]) :

Definition 26: *A path from the vertex* $v \in V$ *to the vertex* $w \in V$ *is a sequence of edges* $e^1, ..., e^k$ *where* $e^1_{init} = v, e^j_{term} = e^{j+1}_{init}$ *for* $j = 1, ..., k-1$ *and* $e^k_{term} = w$*. Its length is defined by* $\rho(e^1) + ... + \rho(e^k)$ *and the distance of two vertices* $dist_\rho(v,w)$ *is defined as the infimum of the lenghts of all paths from* v *to* w*. A graph is connected if for any pair of vertices* $v \neq w$ *there exists a path from* v *to* w*. Finally the ball of radius* $r > 0$ *centered at* $v \in V$ *is defined by*

$$B^r_\rho(v) = \{w \in V : dist_\rho(v,w) < r\}.$$

The conditions of Hanche-Olsen and Holden lead to a compactness criterion (cf. [19, Proposition 3.8]):

Theorem 43: *If* G *is a connected directed graph and* $1 \leq q < \infty$ *the embedding* $w^{1,q}_{\nu,\rho}(V) \hookrightarrow \ell^q_\nu(V)$ *is compact if for every* $\varepsilon > 0$ *there exist a vertex* $v \in V$ *and a radius* $r > 0$ *so that*

$$(i) B^r_\rho(v) \text{ is finite and}$$

$$(ii) \left(\sum_{w \notin B^r_\rho(v)} |f(w)|^q \nu(w) \right)^{\frac{1}{q}} < \varepsilon$$

for all f *in the unit ball of* $w^{1,q}_{\nu,\rho}(V)$*.*

Using our results achieved in chapter 6 we want to give summability conditions for the embedding $w^{1,2}_{\rho,\nu}(V) \hookrightarrow \ell^2_\nu(V)$ to be a p-Schatten embedding where $0 < p \le 2$ for some simple graphs. From the preceding chapter we get:

Corollary 10: *A compact embedding $w^{1,2}_{\rho,\nu}(V) \hookrightarrow \ell^2_\nu(V)$ is a p-Schatten embedding if for every $(2 \le p)$/for some (0<p<2) orthonormal basis $\{f^k\}_{k=1}^\infty$ with respect to $(.,.)_{w^{1,2}_{\rho,\nu}(V)}$:*

$$\sum_{k=1}^{\infty} \left(\sum_{v \in V} |f^k(v)|^2 \nu(v) \right)^{\frac{p}{2}} < \infty.$$

Since in general there is no obvious orthonormal basis of the space $w^{1,q}_{\rho,\nu}(V)$ our strategy will be to apply the Gram-Schmidt orthogonalization method to a standard basis. Consider the following weighted path Γ:

$$\underset{v_1}{\bullet} \xrightarrow{\rho(e_1)} \underset{v_2}{\bullet} \xrightarrow{\rho(e_2)} \underset{v_3}{\bullet} \longrightarrow \ldots$$

For a further simplification assume $\nu \equiv 1$. We write the Sobolev functions $f \in w^{1,2}_\rho(V)$ as vectors :

$$f = (f(v_1), ...), f_j = f(v_j).$$

By applying the Gram-Schmidt orthogonalization method to the basis $\{e^k\}_{k=1}^\infty$ where $e^k = \{\delta_{kj}\}_{j=1}^\infty$ one obtains the following recursive definition of an orthogonal basis of $w^{1,2}_\rho(\Gamma)$:

$$f^1 = (1, 0, ...),$$

$$f^k = \left(\frac{\rho(e_1)...\rho(e_{k-1})}{|f^1|^2_{w^{1,2}_\rho(V)}...|f^{k-1}|^2_{w^{1,2}_\rho(V)}}, \frac{\rho(e_2)...\rho(e_{k-1})}{|f^2|^2_{w^{1,2}_\rho(V)}...|f^{k-1}|^2_{w^{1,2}_\rho(V)}}, ..., \frac{\rho(e_{k-1})}{|f^{k-1}|^2_{w^{1,2}_\rho(V)}}, 1, 0, ...\right) \; k = 2, ...$$

We see that:

$$|f^k|^2_{w^{1,2}_\rho(V)} = \sum_{j=1}^{k-1} \left(\frac{\rho(e_j)...\rho(e_{k-1})}{|f^j|^2_{w^{1,2}_\rho(V)}...|f^{k-1}|^2_{w^{1,2}_\rho(V)}} \right)^2 + 1 \tag{7.1}$$

$$+ \sum_{j=1}^{k-1} \left(\frac{\rho(e_j)...\rho(e_{k-1})}{|f^j|^2_{w^{1,2}_\rho(V)}...|f^{k-1}|^2_{w^{1,2}_\rho(V)}} - \frac{\rho(e_{j-1})...\rho(e_{k-1})}{|f^j|^2_{w^{1,2}_\rho(V)}...|f^{k-1}|^2_{w^{1,2}_\rho(V)}} \right)^2 \rho(e_{j-1})$$

$$+ \left(1 - \frac{\rho(e_{k-1})}{|f^{k-1}|^2_{w^{1,2}_\rho(V)}} \right)^2 \rho(e_{k-1}) + 1\rho(e_k).$$

Assume $\omega : \mathbb{N} \to \mathbb{R}$ is a lower bound for this expression:

$$|f^k|^2_{w^{1,2}_\rho} \ge \omega(k) > 1 \; , \; k = 1, 2, ...$$

Then if for $0 < p \le 2$

$$\sum_{k=1}^{\infty} \left| \frac{f^k}{|f^k|_{w^{1,2}_\rho(V)}} \right|^p_{\ell^2(V)} \le \sum_{k=1}^{\infty} \omega(k)^{-\frac{p}{2}} \left(\sum_{j=1}^{k-1} \left(\frac{\rho(e_j)...\rho(e_{k-1})}{|f^j|^2_{w^{1,2}_\rho(V)}...|f^{k-1}|^2_{w^{1,2}_\rho(V)}} \right)^2 + 1 \right)^{\frac{p}{2}} < \infty \tag{7.2}$$

it follows that $w_\rho^{1,2}(V) \hookrightarrow \ell^2(V)$ is a p-Schatten embedding.
For $p \geq 2$

$$\sum_{k=1}^{\infty} \left| \frac{f^k}{|f^k|_{w_\rho^{1,2}(V)}} \right|_{\ell^2(V)}^p < \infty$$

is necessary for $w_\rho^{1,2}(V) \hookrightarrow \ell^2(V)$ to be a p_1-Schatten embedding for all $0 < p_1 \leq p$. Even if no suitable lower bound is known 7.1 and 7.2 (put $\omega(k) = |f^k|^2_{w_\rho^{1,2}}$) provide a method that yields a summability condition that only depends on the weight ρ.
The border case $p = 2$ yields:

Corollary 11: *The embedding $w_\rho^{1,2}(V) \hookrightarrow \ell^2(V)$ is a Hilbert-Schmidt embedding if and only if*

$$\sum_{k=1}^{\infty} |f^k|^{-2}_{w_\rho^{1,2}(V)} \left(\sum_{j=1}^{k-1} \left(\frac{\rho(e_j)...\rho(e_{k-1})}{|f^j|^2_{w_\rho^{1,2}(V)}...|f^{k-1}|^2_{w_\rho^{1,2}(V)}} \right)^2 + 1 \right) < \infty.$$

7.3 Discrete Sobolev spaces on sparse graphs

For this section we adopt the definition of a graph in the preceding section but we assume the set of edges E to be symmetric, i.e. we have $(v, w) \in E$ if and only if $(w, v) \in E$ and in this case we say that v and w are adjacent and write $v \sim w$. We also assume the graph $G = (V, E)$ to contain no loops, i.e. $(v, v) \notin E$ for all $v \in V$ and to be locally finite, i.e. $|\{w \in V : w \sim v\}| < \infty$ for every $v \in V$. We define:

Definition 27: *For a graph $G = (V, E)$ where E is symmetric and that contains no loops the degree function* $\deg : V \to \mathbb{N} \cup \{0\}$ *is defined by*

$$\deg(v) = |\{w \in V : w \sim v\}|.$$

For a subset $W \subseteq V$ we denote the induced subgraph by $G_W = (W, E_W)$ where $E_W = E \cap (W \times W)$. In this section we consider unweighted graphs, i.e. we put $\nu \equiv 1$ and $\rho \equiv 1$ and we denote the spaces $\ell_1^2(V)$ and $w_{1,1}^{1,2}(V)$ as defined in the preceding section by $\ell^2(V)$ and $w^{1,2}(V)$ respectively. For a function $g : V \to \mathbb{C}$ we denote the multiplication operator on $\ell^2(V)$ given by $f \mapsto gf$ i.e. by pointwise multiplication again by g, e.g. in the case $g = \deg$.
We define the discrete Schrödinger operator $\Delta + q$ where $q : V \to [0, \infty)$ is a non-negative function (cf. [6]):

Definition 28: *On the domain*

$$\mathcal{D}(\Delta + q) = \left\{ f \in \ell^2(V) : \left(v \mapsto \sum_{w \sim v} (f(v) - f(w)) + q(v)f(v) \right) \in \ell^2(v) \right\}$$

we define the Schrödinger operator $\Delta + q : \ell^2(V) \to \ell^2(V)$ by

$$(\Delta + q)f(v) = \sum_{w \sim v} (f(v) - f(w)) + q(v)f(v)$$

for all $f \in \ell^2(V), v \in V$.

In their paper [6] Bonnefont, Golénia and Keller show for sparse graphs a functional inequality for the Schrödinger operator. Intuitively a graph is understood to be sparse if it has few edges. Bonnefont, Golénia and Keller give the following definition:

Definition 29: *A graph $G = (V, E)$ is k-sparse if for finite $W \subseteq V$*

$$|E_W| \leq k|W|.$$

By $T_1 \leq T_2$ for two operators T_1, T_2 on a Hilbert space we mean that the operator $T_2 - T_1$ is positive. For a real valued function $g : V \to \mathbb{R}$ we define

$$\lim\inf_{|v|\to\infty} g(v) = \sup_{\substack{W\subseteq V\\ |W|<\infty}} \inf_{v\in V\setminus W} g(v), \lim\sup_{|v|\to\infty} g(v) = \inf_{\substack{W\subseteq V\\ |W|<\infty}} \sup_{v\in V\setminus W} g(v).$$

We cite [6, Theorem 1.1]:

Theorem 44: *If $G = (V, E)$ is a k-sparse graph, we have*
(i) for all $0 < \varepsilon \leq 1$

$$(1-\varepsilon)(\deg + q) - \frac{k}{2}(\frac{1}{\varepsilon} - \varepsilon) \leq \Delta + q \leq (1+\varepsilon)(\deg + q) + \frac{k}{2}(\frac{1}{\varepsilon} - \varepsilon)$$

on $C_c(V)$, the space of real valued functions with finite support on V,

(ii) $\Delta + q$ has purely discrete spectrum if and only if

$$\lim\inf_{|v|\to\infty} (\deg(v) + q(v)) = \infty.$$

We remark that purely discrete means that the operator has a discrete spectrum consisting of eigenvalues of finite multiplicity.
We see that for a function $f \in \ell^2(V)$ one has

$$((\Delta + q)f, f)_{\ell^2(V)} = \sum_{v\in V} \sum_{w\sim v} (f(v) - f(w))\overline{f(v)} + \sum_{v\in V} q(v)f(v)\overline{f(v)}$$

and rearranging the terms in the first sum and noting that

$$(f(v) - f(w))\overline{f(v)} + (f(w) - f(v))\overline{f(w)} = (f(v) - f(w))\overline{(f(v) - f(w))} = |f(v) - f(w)|^2$$

we see that for $q \equiv 1$:

$$((\Delta + 1)f, f)_{\ell^2(V)} = \frac{1}{2} \sum_{(v,w)\in E} |f(v) - f(w)|^2 + \sum_{v\in V} |f(v)|^2$$

and thus by

$$\frac{1}{2}|f|^2_{w^{1,2}(V)} \leq ((\Delta + 1)f, f)_{\ell^2(V)} \leq |f|^2_{w^{1,2}(V)}$$

that $((\Delta + 1).,.)_{\ell^2(V)}$ induces a norm that is equivalent to $|.|_{w^{1,2}(V)}$. We define another space:

Definition 30: *For a locally finite graph $G = (V, E)$ where E is symmetric and that contains no loops we define for $f \in \ell^2(V)$:*

$$|f|_{\deg} = \left(\sum_{v \in V} |f(v)|^2 (\deg(v) + 1) \right)^{\frac{1}{2}}$$

and

$$\ell^2_{\deg}(V) = \{f \in \ell^2(V) : |f|_{\deg} < \infty\}.$$

It follows by Theorem 44 (i) (cf. [6, Theorem 1.1 (a)]) that for finitely supported real valued functions f the norm $|f|_{\deg}$ is finite if and only if $|f|_{w^{1,2}(V)}$ is finite. Noting that this property for norms $|.|_1, |.|_2$ on a real valued function space is inherited by the norms $\sqrt{|Re(.)|_1^2 + |Im(.)|_1^2}, \sqrt{|Re(.)|_2^2 + |Im(.)|_2^2}$ on the corresponding complex valued function space and by extending the supports we can conclude that

$$\ell^2_{\deg}(V) = w^{1,2}(V). \tag{7.3}$$

as sets. In the same way as for the spaces $s_{*,\alpha}$ one can show that $\ell^2_{\deg}(V)$ are Hilbert spaces endowed with the inner product

$$(f, g) = \sum_{v \in V} (\deg(v) + 1) f(v) \overline{g(v)}$$

for $f, g \in \ell^2_{\deg}(V)$ and we know that $w^{1,2}(V)$ are Hilbert-spaces. By 44 (i) (cf. [6, Theorem 1.1 (a)]) one has $\Delta + 1 \leq 2(\deg + 1)$ and arguing as above we see that the embedding

$$\ell^2_{\deg}(V) \hookrightarrow w^{1,2}(V)$$

is continuous. Since the sets are identical it follows this embedding is surjective and thus by the open mapping theorem it is an isomorphism of Hilbert spaces. Now for convenience we suppose that the vertices are numbered in such a way that $\deg(j) = \deg(v_j)$ (we denote the function on $\mathbb{N}$ by deg again) is non-decreasing. We are now able to apply Theorem 41:

Corollary 12: *(i) If there exists $\alpha > 0$ so that $j^{2\alpha} \leq \deg(j) + 1$ for all but finitely many j the embedding $w^{1,2}(V) \cong \ell^2_{\deg}(V) \hookrightarrow \ell^2(V)$ is compact, and:*
(ii) If $0 < p \leq 2$ and $j^{2\alpha} \leq \deg(j) + 1$ for all but finitely many j and $p\alpha > 1$ the embedding $w^{1,2}(V) \cong \ell^2_{\deg}(V) \hookrightarrow \ell^2(V)$ is a p-Schatten embedding.

Proof: Suppose there exists a number $\alpha > 0$ so that $j^{2\alpha} \leq \deg(j) + 1$ for all but finitely many j. If we denote for $f \in \ell^2(V)$ the sequence $\{f(v_j)\}_{j=1}^{\infty}$ by $\hat{f}$ there clearly exists a constant $C > 0$ so that

$$|\hat{f}|^2_{*,\alpha} = \sum_{j=1}^{\infty} j^{2\alpha} |f(v_j)|^2 \leq C \sum_{j=1}^{\infty} (\deg(j) + 1) |f(v_j)|^2 \tag{7.4}$$

$$= C \sum_{v \in V} (\deg(v) + 1) |f(v)|^2 = C |f|^2_{\deg}.$$

We have the following composition of embeddings:

$$\ell^2_{\deg}(V) \hookrightarrow s_{*,\alpha} \hookrightarrow \ell^2(\mathbb{N}) \cong \ell^2(V)$$

where the first one is bounded by 7.4 and the second is compact by Theorem 41. By the ideal property of the classes $\mathcal{J}_\infty$ as shown in Lemma 8 it follows that

$$w^{1,2}(V) \cong \ell^2_{\deg}(V) \hookrightarrow \ell^2(V)$$

is compact. If $0 < p \leq 2$ and $p\alpha > 1$ the second embedding is of p-Schatten class by Theorem 41 and by Theorem 143.8 it follows that

$$w^{1,2}(V) \cong \ell^2_{\deg}(V) \hookrightarrow \ell^2(V)$$

is of p-Schatten class. $\square$

8 Open problems

The notion of an orthonormal basis was crucial to our discussion. It is not clear if there is a similar characterization of p-Schatten classes for mappings between Banach spaces $X_1 \to X_2$ in the sense of [21] for example via Schauder bases or more specifically a characterization of p-Schatten embeddings $X \hookrightarrow \ell^q(\mathbb{N})$ for $q \in [1, \infty)\backslash\{2\}$.
In [13] Gramsch defines generalized p-Schatten embeddings but as we saw in section 5.3 his discussion is restricted to the Hilbert spaces $H^k(\Omega)$. One might ask if there is a characterization of the generalized p-Schatten embeddings $W^{k,q}(\Omega) \hookrightarrow W^{l,q}(\Omega)$ for $q \in [1, \infty)\backslash\{2\}$.
The question asked by Pietsch in [22] that we mentioned in section 3.2 seems to be unanswered: Is the set of p-summable mappings $\Pi_p(X, Y)$ identical to the Hilbert-Schmidt class for $p > 2$.
In [11] Gohberg and Markus prove that if $H = H^*$ is a non-negative compact operator on $\mathcal{H}_1$ with $\{\lambda_j(H)\}_{j=1}^{\infty} \in \ell^1(\mathbb{N})$ its sequence of eigenvalues in non-increasing order and $\{\alpha_j\}_{j=1}^{\infty}$ a non-increasing sequence of non-negative numbers satisfying the condition

$$\{\alpha_j\}_{j=1}^{\infty} \prec \{\lambda_j(H)\}_{j=1}^{\infty}. \tag{8.1}$$

there is a number m $(0 \leq m \leq \infty)$ and an orthonormal basis, consisting of orthonormal systems $\{\phi_j\}_{j=1}^{\infty}$ and $\{f_j\}_{j=1}^{m}$ such that

$$(H\phi_j, \phi_j) = \alpha_j \;\; j = 1, 2, ...$$

$$(Hf_j, f_j) = 0 \;\; j = 1, ..., m.$$

One would like to have a simple proof for this theorem because it would provide another proof of both Theorem 34 and Theorem 40 and thus Corollary 7 (see [11]).
For a simple path Γ we found sufficient $(0 < p \leq 2)$ and necessary $(2 \leq p < \infty)$ summability conditions for the embeddings of the discrete Sobolev spaces $w_{\nu,\rho}^{1,2} \hookrightarrow \ell_\nu^2(V)$ to be p-Schatten embeddings, that only depend on the weights , but could not write them down explicitly because they are bound to a recursive definition of the underlying basis. The question is still open if there is a simple characterization of the p-Schatten embeddings $w_{\nu,\rho}^{1,2}(V) \hookrightarrow \ell_\nu^2(V)$ for certain broader classes of graphs or even a characterization of the generalized p-Schatten embeddings $w_{\nu,\rho}^{1,q}(V) \hookrightarrow \ell_\nu^q(V)$ where $q \in [1, \infty)$. Finally we conjecture that the embeddings $s_{*,\alpha} \hookrightarrow s_{*,\beta}$ are compact for $\alpha - \beta > 0$ and of p-Schatten class for $0 < p < \infty$ if and only if $(\alpha - \beta)p > 1$.

Bibliography

[1] Adams, R. A., and Fournier, J. J. F. *Sobolev Spaces.* Academic Press, 1975.

[2] Almira, J., and Perez-Acosta, F. Convergence in generalized Schatten Classes of Operators. *Le Matematiche* (1998), 305–317.

[3] Alt, W. *Lineare Funktionalanalysis.* Springer, 1992.

[4] Bagarello, F., Gazeau, J. P., and Szafraniec, F. H. *Non-Selfadjoint Operators in Quantum Physics.* Wiley, 2015.

[5] Bers, L., John, F., and Schechter, M. *Partial Differential Equations.* Interscience Publisher, 1964.

[6] Bonnefont, M., Golénia, S., and Keller, M. Eigenvalue Asymptotics for Schrödinger Operators on sparse Graphs. *preprint* (2014). arxiv.org.

[7] Brezis, H. *Functional Analysis,Sobolev spaces and Partial Differential Equations.* Springer, New York, 2011.

[8] Dunford, N., and Schwartz, J. T. *Linear Operators Part II Spectral Theory.* Wiley, 1988.

[9] Evans, L. C. *Partial Differential Equations.* American Mathematical Society, 1998.

[10] Fan, K. Maximum properties and inequalities for the eigenvalues of completely continuous operators. *Proceedings of the National Academy of Sciences* (1951), 760–766.

[11] Gohberg, I., and Markus, A. Some relations between eigenvalues and matrix elements of linear operators. *Matematicheskii Sbornik* (1964), 481–496. Russian.

[12] Gohberg, I., and M.G.Krein. *Introduction to the theory of linear non-seldadjoint operators.* Translations of mathematical monographs, 1969.

[13] Gramsch, B. Zum Einbettungssatz von Rellich bei Sobolevräumen. *Mathematische Zeitschriften* (1967), 81–87.

[14] Hanche-Olsen, H., and Holden, H. The Kolmogorov-Riesz Compactness Theorem. *Expos. Mathematicae* (2010), 385–394.

[15] Lindenstrauss, J., and Tzafriri, L. *Classical Banach Spaces*, vol. 2. Springer, 1996.

[16] Markus, A. S. The eigen- and singular values of the sum and product of linear operators. *Uspekhi Mat. Nauk.* (1964), 93–123.

[17] Maurin, K. Abbildungen vom Hilbert-Schmidtschen Typus und ihre Anwendungen. *Mathematica Scandinavica* (1961), 359–371.

[18] Mityagin, B. S. Normed ideals of intermediate type. *American Mathematical Society Translations* (1967), 180–194.

[19] Mugnolo, D. *Semigroup Methods for Evolution Equations on Networks.* Springer, 2014.

[20] Palais, R. S. *Seminar on the Atiyah-Singer index theorem.* Princeton University Press, 1965.

[21] Pietsch, A. Einige neue Klassen von kompakten linearen Abbildungen. *Revue Roumaine de Mathematiques pures et appliques* (1963), 427–447.

[22] Pietsch, A. Absolut p-summierbare Abbildungen in normierten Räumen. *Studia mathematica* (1967), 333–353.

[23] Reed, M., and Simon, B. *Methods of Modern Mathematical Physics*, vol. 1. Academic Press, 1980.

[24] Schröder, H. *Funktionalanalysis.* Harri Deutsch, 2000.

[25] Simon, B. *Trace Ideals and their Applications.* American Mathematical Society, 2005.

[26] Yosida, K. *Functional Analysis.* Springer, 1980.