Chapter 2

Normed ideals of $\mathscr{K}(\mathcal{H})$

In this chapter we review the classical theory related to compact operators: their singular values and their eigenvalues, some operator ideals, etc. We shall mainly follow [Sim] but an alternative reference is [GK]. Note that some results might be improved in the subsequent chapters. Before starting with any new material, let us emphasize one tiny but important point.

Remark 2.0.6. *Up to now, choosing $\mathbb{N} = \{0, 1, 2, \dots\}$ or $\mathbb{N} = \{1, 2, 3, \dots\}$ was not relevant and we did not impose any choice. However, in some of the subsequent formulas starting with $n = 0$ or with $n = 1$ makes a difference. So from now on we shall take the convention that $\mathbb{N} := \{1, 2, 3, \dots\}$ and stress that some formulas look different with the other convention. Thus, without further notice all sequences $\{f_n\}$ or (a_n) will start with $n = 1$. Relatedly, we shall use the convenient notation N either for a finite number or for ∞.*

2.1 Compact operators and the canonical expansion

In order to study the ideal of compact operators $\mathscr{K}(\mathcal{H})$, a standard result on analytic operator-valued functions has to be recalled. Its proof is provided for example in [RS1, Thm VI.14]. For its statement, we recall that a subset S of an open set Ω is discrete if it has no limit points in Ω.

Theorem 2.1.1 (Analytic Fredholm theorem)**.** *Let Ω be an open connected subset of $\mathbb{C}$. Let $\Psi : \Omega \to \mathscr{K}(\mathcal{H})$ be an analytic operator-valued function. Then one of the following alternative holds:*

(i) $\big(\mathbf{1} - \Psi(z)\big)^{-1}$ *exists for no $z \in \Omega$,*

(ii) $\big(\mathbf{1} - \Psi(z)\big)^{-1}$ *exists for all $z \in \Omega \setminus S$ where S is a discrete subset of Ω. In this case, $\big(\mathbf{1} - \Psi(z)\big)^{-1}$ is meromorphic in Ω, analytic in $\Omega \setminus S$, the residue at the poles are finite rank operators, and if $z \in S$ then the equation $\Psi(z)f = f$ has a nonzero solution in $\mathcal{H}$.*

This theorem has several important consequences. We state a few of them.

Corollary 2.1.2 (Fredholm alternative)**.** *If A belongs to $\mathscr{K}(\mathcal{H})$, either $(\mathbf{1}-A)^{-1}$ exists or $Af = f$ has a solution in $\mathcal{H}$.*

Proof. Set $\Psi(z) = zA$ and apply the previous theorem at $z = 1$. □

Theorem 2.1.3 (Riesz-Schauder theorem)**.** *If A belongs to $\mathscr{K}(\mathcal{H})$, then its spectrum $\sigma(A)$ is a discrete set having no limit points except perhaps 0. In addition, any non-zero $\lambda \in \sigma(A)$ is an eigenvalue of finite geometric multiplicity.*

Proof. Let us set $\Psi(z) = zA$, which makes Ψ an analytic $\mathscr{K}(\mathcal{H})$-valued function on $\mathbb{C}$. Thus from Theorem 2.1.1 one infers that the set $\{z \in \mathbb{C} \mid \Psi(z)f = f \text{ for some } f \in \mathcal{H}, f \neq 0\}$ is a discrete set. Now, if $\lambda \neq 0$ and if $\frac{1}{\lambda}$ is not in this discrete set then

$$(\lambda - A)^{-1} = \frac{1}{\lambda}\big(\mathbf{1} - \tfrac{1}{\lambda}A\big)^{-1}$$

exists, which means that $\lambda \notin \sigma(A)$. From this, one deduces that the spectrum of A consists in the discrete set mentioned above, and possibly in the value 0. Finally, the fact that the non-zero eigenvalues have finite geometric multiplicity follows directly from the compactness of A. □

The following statement is a direct consequence of Riesz-Schauder theorem together with some information deduced from the spectral theorem for self-adjoint operators, see Theorem 1.4.15.

Theorem 2.1.4 (Hilbert-Schmidt theorem)**.** *If A is self-adjoint and belongs to $\mathscr{K}(\mathcal{H})$ then there exists a complete orthonormal basis $\{f_n\}_{n\in\mathbb{N}}$ of $\mathcal{H}$ such that $Af_n = \lambda_n f_n$ and $\lim_{n\to\infty} \lambda_n = 0$.*

If A is not self-adjoint, a "canonical" description of A can still be provided. For its statement, we shall use the convenient notation $|f\rangle\langle g|$ for the rank-one operator defined by

$$|f\rangle\langle g|h := \langle g, h\rangle f, \qquad \text{for any } f, g, h \in \mathcal{H}. \tag{2.1}$$

Theorem 2.1.5. *If A belongs to $\mathscr{K}(\mathcal{H})$ then A has a norm convergent expansion*

$$A = \sum_{n=1}^{N} \mu_n(A)|g_n\rangle\langle f_n| \tag{2.2}$$

with N either a finite number or equal to ∞, with each $\mu_n(A) > 0$ and satisfying $\mu_n(A) \geq \mu_{n+1}(A)$, and with each family $\{f_n\}$ and $\{g_n\}$ orthonormal but not necessarily complete. Moreover, each $\mu_n(A)$ is uniquely determined while the f_n and g_n are usually not uniquely defined.

Proof. From the polar decomposition provided in Theorem 1.6.5 one infers that there exists a partial isometry W such that $|A| = W^*A$. Thus, $|A|$ is a compact and self-adjoint operator to which Theorem 2.1.4 applies. With the notations introduced above, this reads

$$|A| = \sum_{n=1}^{N} \mu_n(A)|f_n\rangle\langle f_n|$$

where the $\mu_n(A)$ are the eigenvalues of $|A|$ and f_n the corresponding eigenfunctions. Clearly, the family $\{f_n\}$ is orthonormal. Since W is an isometry on $\mathsf{Ran}(|A|)$ and by setting $g_n := Wf_n$ one also infers that $\{g_n\}$ is orthonormal. Since the relation $W|A| = A$ holds, one directly deduces the equality (2.2). The uniqueness follows if one observes that if (2.2) holds, then $\{\mu_n(A)^2\}$ are the eigenvalues of A^*A, $\{f_n\}$ the eigenvectors of A^*A and $\{g_n\}$ the eigenvectors of AA^*. The lack of uniqueness of f_n and g_n comes from the possible degeneracy of the eigenvalues of A^*A and AA^*. □

In the previous result, the real values $\mu_n(A)$ are called *the singular values of* A and the equality (2.2) is called *the canonical expansion of* A. Let us also emphasize that

$$\mu_n(A^*) = \mu_n(A), \tag{2.3}$$

as it can be directly deduced from (2.2) or from the fact that the spectrum of A^*A and AA^* coincide (multiplicity counted) with the possible exception of the eigenvalue 0.

Let us still add one more useful result which can be easily deduced from the construction provided in [Kat, Sec. III.6.4].

Lemma 2.1.6. *If A belongs to $\mathscr{K}(\mathcal{H})$ and $\lambda \in \sigma(A)$ is not equal to 0, then there exists a finite rank projection P_λ such that $AP_\lambda = P_\lambda A$, $\sigma\big(A \restriction P_\lambda\mathcal{H}\big) = \{\lambda\}$ and $\sigma\big(A \restriction (\mathbf{1} - P_\lambda)\mathcal{H}\big) = \sigma(A) \setminus \{\lambda\}$.*

Note that a possible expression for P_λ is provided by the formula

$$P_\lambda := -\frac{1}{2\pi i}\int_{|z-\lambda|=\varepsilon} (A - z)^{-1}\,\mathrm{d}z$$

for $\varepsilon > 0$ small enough. The dimension of $\mathsf{Ran}(P_\lambda)$ is called *the algebraic multiplicity* of λ. We still recall that the geometric and the algebraic multiplicity of an eigenvalue can be different, but the geometric multiplicity can never exceed the algebraic multiplicity.

2.2 Eigenvalues and singular values

In this section we begin the study of the singular values of any compact operator A, and then state some relations between singular values and eigenvalues. The proofs for most of these relations are not provided but references are given.

We start with some results on singular values. Since these values can be computed by an application of the min-max principle, we first introduce this principle for positive compact operator. Note that a similar statement holds for the negative eigenvalues of any self-adjoint compact operator.

Theorem 2.2.1 (Min-max principle)**.** *Let B be a positive compact operator in $\mathcal{H}$, and let $\{\lambda_n\}$ be the set of its eigenvalues (counting multiplicity) and ordered such that $\lambda_n \geq \lambda_{n+1}$. Then*

$$\lambda_n = \min\big\{\sup\{\langle f, Bf\rangle \mid f \in \mathcal{M}_n^\perp, \|f\| = 1\} \mid \mathcal{M}_n \subset \mathcal{H}, \dim(\mathcal{M}_n) = n-1\big\}. \quad (2.4)$$

Proof. For any $n \in \mathbb{N}$ let us set $F_n := \mathsf{Span}(f_1, \dots, f_n)$ with f_j a normalized eigenvector corresponding to the eigenvalue λ_j of B. Let us also consider any subspace $\mathcal{M}_n \subset \mathcal{H}$ with $\dim(\mathcal{M}_n) = n-1$. Clearly, $F_n \cap \mathcal{M}_n^\perp \neq \{0\}$, and thus one can choose $f \in F_n \cap \mathcal{M}_n^\perp$ with $\|f\| = 1$. More precisely, $f = \sum_{j=1}^d c_j f_j$ with $\sum_{j=1}^d |c_j|^2 = 1$. It then follows that

$$\langle f, Bf\rangle = \sum_{j=1}^d \langle c_j f_j, \lambda_j c_j f_j\rangle = \sum_{j=1}^d \lambda_j |c_j|^2 \geq \lambda_n \sum_{j=1}^d |c_j|^2 = \lambda_n$$

since the eigenvalues of B are ordered. Hence we have obtained that

$$\sup\big\{\langle f, Bf\rangle \mid f \in \mathcal{M}_n^\perp, \|f\| = 1\big\} \geq \lambda_n.$$

For the converse inequality, one can choose $\mathcal{M}_n := \mathsf{Span}(f_1, \dots, f_{n-1})$ and then

$$\sup\big\{\langle f, Bf\rangle \mid f \in \mathcal{M}_n^\perp, \|f\| = 1\big\} = \lambda_n,$$

which implies the statement. □

By setting $B := A^*A$ in the previous statement and by recalling that $\mu_n(A)^2$ corresponds to the n-eigenvalue of A^*A one directly obtains a characterization of the singular values of any compact operator A, namely:

Proposition 2.2.2. *Let A belong to $\mathscr{K}(\mathcal{H})$ and let $\{\mu_n(A)\}$ denote its singular values ordered such that $\mu_n(A) \geq \mu_{n+1}(A)$. Then*

$$\mu_n(A) = \min\big\{\sup\{\|Af\| \mid f \in \mathcal{M}_n^\perp, \|f\| = 1\} \mid \mathcal{M}_n \subset \mathcal{H}, \dim(\mathcal{M}_n) = n-1\big\}. \quad (2.5)$$

As a consequence one directly infers the following estimates:

Corollary 2.2.3. *For any $A \in \mathscr{K}(\mathcal{H})$ and $B \in \mathscr{B}(\mathcal{H})$ one has*

$$\mu_n(AB) \leq \mu_n(A)\|B\| \quad \textit{and} \quad \mu_n(BA) \leq \mu_n(A)\|B\|. \quad (2.6)$$

Proof. Observe that the first inequality can be deduced from the second one and from (2.3). Indeed one has

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

For the second equality, one uses (2.5) together with the inequality $\|BAf\| \leq \|B\|\|Af\|$ for any $f \in \mathcal{H}$. □

In the next statement, we mention some generalizations of the previous result. Proofs can be found in [Fan, Thm. 2].

Proposition 2.2.4. *Let A, B belong to $\mathscr{K}(\mathcal{H})$. Then the following inequalities hold for any $m, n \in \mathbb{N}$:*

$$\mu_{m+n-1}(AB) \leq \mu_m(A)\,\mu_n(B), \tag{2.7}$$

$$\mu_{n+m-1}(A+B) \leq \mu_m(A) + \mu_n(B). \tag{2.8}$$

Note that (2.6) corresponds to the special case $m = 1$ since $\mu_1(B) = \|B\|$. One additional relation between the singular values of A, B and AB is given by:

Lemma 2.2.5. *For any A, B in $\mathscr{K}(\mathcal{H})$ and for any $n \in \mathbb{N}$ one has*

$$\prod_{j=1}^{d} \mu_j(AB) \leq \prod_{j=1}^{d} \mu_j(A)\mu_j(B). \tag{2.9}$$

Proof. See [Hor], Theorem 3 and its proof. $\square$

Let us still mention some relations linking singular values and eigenvalues. Note that the eigenvalues of a compact operator are not enumerated arbitrarily but according to the following definition.

Definition 2.2.6. *If A belongs to $\mathscr{K}(\mathcal{H})$ its eigenvalues $\lambda_1(A), \lambda_2(A), \ldots$ are ordered such that $|\lambda_j(A)| \geq |\lambda_{j+1}(A)|$ for any $j \in \mathbb{N}$, and each eigenvalue is counted up to its algebraic multiplicity.*

The following result comes from the paper [Wey].

Lemma 2.2.7. *For any A in $\mathscr{K}(\mathcal{H})$ and for any $n \in \mathbb{N}$ one has*

$$\prod_{j=1}^{d} |\lambda_j(A)| \leq \prod_{j=1}^{d} \mu_j(A).$$

As a consequence of the previous two results one has:

Proposition 2.2.8. *For any A, B in $\mathscr{K}(\mathcal{H})$ and for any monotone increasing function $\phi : [0, \infty) \to \mathbb{R}_+$ such that $t \mapsto \phi(\mathrm{e}^t)$ is convex one has*

(i)

$$\sum_j \phi\big(|\lambda_j(A)|\big) \leq \sum_j \phi\big(\mu_j(A)\big), \tag{2.10}$$

and in particular for any $p \geq 1$

$$\sum_j |\lambda_j(A)|^p \leq \sum_j \mu_j(A)^p, \tag{2.11}$$

(ii)

$$\sum_j \phi\big(\mu_j(AB)\big) \leq \sum_j \phi\big(\mu_j(A)\mu_j(B)\big). \tag{2.12}$$

Proof. These inequalities directly follow from Lemmas 2.2.5 and 2.2.7 together with Corollary 2.3.4 introduced in the next section. □

Note that the results mentioned above are usually proved for finite matrices, and then a limiting procedure is applied in order to extend the result to certain compact operators. As mentioned in the proof of Proposition 2.2.8 some additional technicalities are now required. Some of them are introduced in the next section.

2.3 Technical interlude

Let us start by introducing some ideas and results about *rearrangement* or *(sub)majorization*. This concept plays an important role when dealing with the spectrum of matrices or compact operators, and has been extensively studied in the book [MOB]. In the next definition we use the notation c_0 for the set of complex sequences $a = (a_j)_{j=1}^{\infty}$ satisfying $\lim_{j\to\infty} a_j = 0$.

Definition 2.3.1. *For any $a = (a_j)$ in $\mathbb{C}^d$ or in c_0 we denote by a^* the element of $\mathbb{R}^d$ or c_0 obtained by a non-increasing rearrangement of $\{|a_j|\}_j$.*

In other words, if means that $a_j^* \geq a_{j+1}^*$ and that the sets $\{a_j^*\}$ and $\{|a_j|\}$ are identical, counting multiplicity. For simplicity, we shall say that an element $a \in \mathbb{R}^d$ or $a \in c_0$ is *positive and ordered* if $a_j \geq 0$ and $a_j \geq a_{j+1}$ for any j. Clearly, a^* is always positive and ordered.

Now, based on the *rearrangement inequality*[1], as presented for example in [HLP, Thm 368] one infers that for two sequences a and b as in the previous definition one has

$$\sum_j |a_j b_j| \leq \sum_j a_j^* b_j^* \tag{2.13}$$

as long as the r.h.s. is meaningful (if a and b belong to $\mathbb{C}^d$ it is obviously the case). The following result, stated first in [Ma1, Lem. 1] and proved in [Ma2, Thm. 1.2], will be important later on. The version presented here is taken from [Sim, Thm. 1.9] where a proof is provided.

[1]For any $a \in \mathbb{R}^d$ let us set $a^\star$ for the non-increasing rearrangement of $\{a_j\}$ (without the absolute value). If $a, b \in \mathbb{R}^d$ the *rearrangement inequality* reads

$$\sum_{j=1}^d a_j^\star b_{n+1-j}^\star \leq \sum_{j=1}^d a_j b_j \leq \sum_{j=1}^d a_j^\star b_j^\star.$$

Theorem 2.3.2. *Let $a, b \in \mathbb{C}^d$ and assume that*

$$\sum_{j=1}^{k} b_j^* \leq \sum_{j=1}^{k} a_j^* \qquad \text{for any } k \in \{1, \ldots, n\}. \tag{2.14}$$

Then there exist m points $a^{(1)}, \ldots, a^{(m)}$ in $\mathbb{C}^d$ with $(a^{(\ell)})^* = a^*$ for $\ell \in \{1, \ldots, m\}$ and there exist $\{\lambda_\ell\} \subset [0,1]$ satisfying $\sum_{\ell=1}^m \lambda_\ell = 1$ such that

$$b = \sum_{\ell=1}^{m} \lambda_\ell \, a^{(\ell)}. \tag{2.15}$$

In addition if Φ is a positive valued function on $[0,\infty)^d$ and if the function $\phi : \mathbb{C}^d \to \mathbb{R}_+$, defined by $\phi(c) := \Phi(c_1^*, \ldots, c_n^*)$, is convex on $\mathbb{C}^d$, then

$$\phi(b) \leq \phi(a). \tag{2.16}$$

Note that condition (2.14) is often denoted by $b \lll a$ in the literature. In addition, what (2.15) really says is that b belongs to the convex hull of a family of vectors of the form $(\varepsilon_k a_{j_k})_{k=1}^d$ with $|\varepsilon_k| = 1$ and j_k is an arbitrary permutation of the numbers $1, 2, \ldots, n$. The elements $a^{(\ell)}$ are the points which define the convex hull. This, together with the fact that $\phi(a^{(\ell)}) = \phi(a)$ and the convexity of the function ϕ, directly implies the inequality (2.16).

Exercise 2.3.3. *Provide a proof of Theorem 2.3.2.*

Before mentioning two results linked to the previous statement, let us show how one can construct examples of functions Φ. Consider any function $f : [0,\infty) \to \mathbb{R}_+$ which is convex and increasing and let us set $\Phi(x) := \sum_{j=1}^d f(x_j)$ for any $x \in [0,\infty)^d$. Then one observes that for any $\theta \in [0,1]$ and $b, c \in \mathbb{C}^d$ one has

$$\begin{aligned}
\phi\big(\theta b + (1-\theta)c\big) &= \Phi\big((\theta b + (1-\theta)c)^*\big) \\
&= \sum_{j=1}^{d} f\big(\big|\theta b_j + (1-\theta)c_j\big|\big) \\
&\leq \sum_{j=1}^{d} \Big(\theta f(|b_j|) + (1-\theta) f(|c_j|)\Big) \\
&= \theta\phi(b) + (1-\theta)\phi(c)
\end{aligned}$$

which means that ϕ is convex on $\mathbb{C}^d$. As a consequence the function Φ satisfies the requirement of Theorem 2.3.2.

The next statement is an application of Theorem 2.3.2 for transforming estimates on products to estimates on sums.

Corollary 2.3.4. *Let $a, b \in \mathbb{R}^d$ be positive and ordered, and suppose that*

$$\prod_{j=1}^{k} b_j \leq \prod_{j=1}^{k} a_j \qquad \text{for any } k \in \{1, \ldots, n\}.$$

Then, for any continuous, monotone increasing function $g : [0, \infty) \to \mathbb{R}_+$ with $t \mapsto g(\mathrm{e}^t)$ convex, we have that

$$\sum_{j=1}^{k} g(b_j) \leq \sum_{j=1}^{k} g(a_j) \qquad \text{for any } k \in \{1, \ldots, n\}. \tag{2.17}$$

In particular, (2.14) *can be obtained by taking $g(x) = x$.*

Proof. Assume without loss of generality that a_j and b_j are all non-zero. By setting then $\tilde{a}_j := \gamma a_j$ and $\tilde{b}_j := \gamma b_j$ for γ large enough, we get that all $\tilde{a}_j$, $\tilde{b}_j$ are bigger than 1. By considering then $\ln(\tilde{a}_j)$ and $\ln(\tilde{b}_j)$, one observes that the condition (2.14) is satisfied for these numbers. By setting then $f(t) := g\big(\gamma^{-1} \mathrm{e}^t\big)$, the function f is convex and increasing, and by the observation made above, the function Φ defined by $\Phi(x) := \sum f(x_j)$ for any $x \in [0, \infty)^d$ satisfies the assumption of Theorem 2.3.2. The inequality (2.17) follows then directly from (2.16). □

Exercise 2.3.5. *Check the details of the previous proof.*

The second domain linked with Theorem 2.3.2 is related to the notion of doubly substochastic matrices.

Definition 2.3.6. *A matrix $\alpha = (a_{jk}) \in M_N(\mathbb{C})$ is called* doubly substochastic *(in short dss) if $\sum_{j=1}^{N} |\alpha_{jk}| \leq 1$ for all $k \in \{1, \ldots, N\}$ and $\sum_{k=1}^{N} |\alpha_{jk}| \leq 1$ for all $j \in \{1, \ldots, N\}$.*

Note that such matrices can be constructed from elements of any Hilbert space $\mathcal{H}$. Indeed, if for any $\ell \in \{1, 2, 3, 4\}$ the family $\{f_j^\ell\}_{j=1}^N \subset \mathcal{H}$ is orthonormal, then the matrix α defined by $\alpha_{jk} := |\langle f_j^1, f_k^2 \rangle|^2$ is a dss matrix, and the matrix β defined by $\beta_{jk} := \langle f_j^1, f_k^2 \rangle \langle f_k^3, f_j^4 \rangle$ is a dss matrix. The fact that these matrices are doubly substochastic can be obtained by applying Bessel and Schwartz inequalities.

The next statement is borrowed from [Sim, Prop. 1.12] to which we refer for its proof.

Proposition 2.3.7. *Let $\alpha \in M_n(\mathbb{C})$ be a dss matrix and let $c \in \mathbb{C}^d$. If one sets $a := c^*$ and $b := \alpha c$, then $a, b \in \mathbb{C}^d$ satisfy condition* (2.14) *of Theorem 2.3.2.*

We now introduce the notion of *symmetric normed spaces*. Note that a simple introduction to the subject can be found in [Sch, Sec. V.3]. For that purpose, let us denote by ℓ_∞ the set of all bounded sequences $(a_j)_{j=1}^\infty$ endowed with the sup norm (also denoted ℓ_∞-norm), and let us denote by c_c the set of complex sequences $a = (a_j)_{j=1}^\infty$

with compact support. Clearly, c_c is dense in c_0 for the ℓ_∞-norm. Recall that a norm on c_c is a map $\Phi : c_c \to \mathbb{R}_+$ which satisfies for any $a, b \in c_c$ and $\lambda \in \mathbb{C}$ the following properties: i) $\Phi(\lambda a) = |\lambda|\Phi(a)$, ii) $\Phi(a+b) \leq \Phi(a) + \Phi(b)$, iii) $\Phi(a) = 0$ if and only if $a = 0$.

Definition 2.3.8. *A norm Φ on c_c is* symmetric *if $\Phi(a) = \Phi(a^*)$ for any $a \in c_c$.*

Let us observe that a norm on c_c is symmetric if and only if it is invariant under permutations and under the map $a_j \mapsto e^{i\theta_j} a_j$ for any $\theta_j \in \mathbb{R}$.

Definition 2.3.9. *Let Φ be a norm on c_c. The* maximal space *J_Φ consists in the set of sequence $a = (a_j)_{j=1}^\infty$ such that the limit $\lim_{n\to\infty} \Phi\big((a_1, a_2, \ldots, a_n, 0, 0 \ldots)\big)$ exists (we denote it then by $\Phi(a)$). The* minimal space *$J_\Phi^{(0)}$ consists in the closure of c_c with the norm Φ. If $J_\Phi = J_\Phi^{(0)}$, that is if c_c is dense in J_Φ the norm Φ is called* regular *(or* mononormalizing *in some references).*

Examples 2.3.10. *1) For $p \geq 1$, if $\Phi(a) \equiv \|a\|_p := \big(\sum_j |a_j|^p\big)^{1/p}$ then J_Φ corresponds to the usual ℓ_p space. We also set $\|a\|_\infty := \sup_j |a_j|$. Note that if $p < \infty$ the norm $\|\cdot\|_p$ is regular.*

2) For $p > 1$ let us set

$$\|a\|_{p,w} := \sup_n \Big(n^{-1+\frac{1}{p}} \sum_{j=1}^{d} a_j^*\Big), \tag{2.18}$$

which is a symmetric norm, called Calderón norm. The maximal space associated with this norm is denoted by $\ell_{p,w}$ and called weak *ℓ_p-space. The minimal space corresponds to the elements $a \in \ell_{p,w}$ satisfying the additional condition $\lim_{j\to\infty} j^{\frac{1}{p}} a_j^* = 0$, which means that the Calderón norms are not regular. Note also that the following inequalities hold*

$$\|a\|'_{p,w} \leq \|a\|_{p,w} \leq \frac{p}{p-1}\|a\|'_{p,w}$$

with $\|a\|'_{p,w} := \sup_j(j^{\frac{1}{p}} a_j^*)$. Clearly, this expression is simpler than (2.18) but $\|\cdot\|'_{p,w}$ does not define a norm. However, the set of $a \in \ell_\infty$ satisfying $\|a\|'_{p,w} < \infty$ corresponds to $\ell_{p,w}$, and this expression can also be used for $p = 1$.

In the following statement, several properties of maximal and minimal spaces are summarized. Note that Theorem 2.3.2 and Proposition 2.3.7 play an important role in the proof, and that condition (2.14) is explicitly mentioned in the point (b). For the proof of these statements, we refer to [Sim, Thm. 1.16].

Theorem 2.3.11. *Let Φ be a symmetric norm on c_c, then*

(a) If $a \in J_\Phi$ and $\lim_{j\to\infty} a_j = 0$, then $\Phi(a) = \Phi(a^*)$,

(b) If $a, b \in J_\Phi$ with $\lim_{j\to\infty} a_j = 0$ and $\lim_{j\to\infty} b_j = 0$, and if $\sum_{j=1}^d b_j^* \leq \sum_{j=1}^d a_j^*$ for any $n \in \mathbb{N}$, then $\Phi(b) \leq \Phi(a)$,

(c) *If* $\Phi\big((1,0,0,\dots)\big) = \mathrm{c}$, *then* $\mathrm{c}\|a\|_\infty \leq \Phi(a) \leq \mathrm{c}\|a\|_1$ *for any* $a \in J_\Phi$,

(d) *Both* J_Φ *and* $J_\Phi^{(0)}$ *are Banach spaces,*

(e) *If* α *is a substochastic matrix and* $a \in J_\Phi$, *resp.* $a \in J_\Phi^{(0)}$, *then* αa *is in* J_Φ, *resp. in* $J_\Phi^{(0)}$, *and* $\Phi(\alpha a) \leq \Phi(a)$,

(f) *If* Φ *is inequivalent to* $\|\cdot\|_\infty$, *then* J_Φ *consists only of sequences which vanish at infinity,*

(g) *If* $J_\Phi = J_\Psi$, *then* Φ *and* Ψ *are equivalent norms.*

For each symmetric norm Φ on c_c one can define a conjugate norm Φ' on c_c by the following construction: For any $b \in c_c$ one sets

$$\Phi'(b) := \sup\big\{\big|\sum_j a_j b_j\big| \mid a \in c_c, \Phi(a) \leq 1\big\}. \tag{2.19}$$

As a consequence of (2.13) one easily infers that for $b, c \in c_c$ with $c = c^*$

$$\sup\big\{\big|\sum_j a_j b_j\big| \mid a^* = c\big\} = \sum_j b_j^* c_j$$

and then that Φ' is a symmetric norm on c_c. Some additional standard duality results are gathered in the following statement.

Theorem 2.3.12. *Let* Φ *be a symmetric norm on* c_c. *Then*

(a) $\sum_j |a_j b_j| \leq \Phi(a)\Phi'(b)$,

(b) $\big(J_\Phi^{(0)}\big)^* = J_{\Phi'}$ *in the sense that any continuous linear functional on* $J_\Phi^{(0)}$ *has the form* $a \mapsto \sum_j a_j b_j$ *for some* $b \in J_{\Phi'}$,

(c) $J_\Phi^{(0)}$, *resp.* J_Φ, *is reflexive if and only if both* Φ *and* Φ' *are regular.*

The proof of the above statement is provided in [Sim, Thm. 1.17] and is based on standard duality arguments.

Exercise 2.3.13. *Provide the proofs of Theorems 2.3.11 and 2.3.12.*

We close this section with a few results related to singular values of pairs of compact operators. Proofs can be found in [Sim, Sec. 1.8 & 1.9].

Proposition 2.3.14. *a) For any pair of compact operators* A *and* B *one has*

$$\mu_n(A) - \mu_n(B) = \sum_m \alpha_{nm}\mu_m(A - B) \tag{2.20}$$

for a dss matrix α.

b) For any pair of finite dimensional self-adjoint matrices A, B, let $\lambda_n^\star(A)$ denote the eigenvalues of A listed in decreasing order. Then

$$\lambda_n^\star(A) - \lambda_n^\star(B) = \sum_m \beta_{nm}\lambda_m^\star(A-B) \tag{2.21}$$

for a dss matrix β.

For any $A \in \mathscr{K}(\mathcal{H})$ we set $\|A\|_p := \big(\sum_n \mu_n(A)^p\big)^{1/p}$ whenever the summation is meaningful.

Proposition 2.3.15 (Clarkson-McCarthy inequalities). *a) For $2 \leq p < \infty$ one has*

$$\|A+B\|_p^p + \|A-B\|_p^p \leq 2^{p-1}\big(\|A\|_p^p + \|B\|_p^p\big). \tag{2.22}$$

b) For $1 < p \leq 2$ and for $p' = p/(p-1)$ one has

$$\|A+B\|_p^{p'} + \|A-B\|_p^{p'} \leq 2\big(\|A\|_p^p + \|B\|_p^p\big)^{p'/p}. \tag{2.23}$$

Note that for A, B positive an additional relation holds:

Proposition 2.3.16. *For $p \geq 1$ and for A, B positive compact operators, one has*

$$2^{1-p}\|A+B\|_p^p \leq \|A\|_p^p + \|B\|_p^p \leq \|A+B\|_p^p. \tag{2.24}$$

Exercise 2.3.17. *Consider the matrices $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & i \\ -i & 0 \end{smallmatrix}\right)$ and compute $\|A\|_1$, $\|B\|_1$ and $\|A+iB\|_1$. What do you observe and can you compare this result with the commutative case ?*

2.4 Normed ideals of $\mathscr{B}(\mathcal{H})$

In this section we begin the study of two-sided ideals in $\mathscr{B}(\mathcal{H})$. By definition, as linear subspace $\mathscr{J}$ of $\mathscr{B}(\mathcal{H})$ is *a two-sided ideal of $\mathscr{B}(\mathcal{H})$* if AB and BA belong to $\mathscr{J}$ whenever $A \in \mathscr{J}$ and $B \in \mathscr{B}(\mathcal{H})$. Some of these spaces will be linked to sequences introduced in the previous sections. We begin with two standard results about operator ideals. The first one state that the biggest ideal of $\mathscr{B}(\mathcal{H})$ is $\mathscr{K}(\mathcal{H})$.

Proposition 2.4.1. *Let $\mathscr{J}$ be a two-sided ideal of $\mathscr{B}(\mathcal{H})$ containing an element A which is not compact. Then $\mathscr{J} = \mathscr{B}(\mathcal{H})$.*

Proof. By the polar decomposition of Theorem 1.6.5, there exists a partial isometry W such that $|A| = W^*A$. If follows that $\mathscr{J}$ contains the positive self-adjoint operator $|A|$ which is not compact. By the spectral theorem, for any $a > 0$ let us set $P_a := \chi_{[a,\infty)}(|A|) \equiv E\big([a,\infty)\big)$ where E denotes the spectral measure associated with the operator $|A|$. If each P_a is a finite dimensional projection, then $|A| = u{-}\lim_{a\to 0}|A|P_a$

would be compact (as a norm limit of finite dimensional operator) which is a contradiction with the fact that $|A|$ is not compact. Thus there exists at least one P_a which is not finite dimensional. In addition, since $|A|^{-1}P_a$ is bounded (by functional calculus), then $P_a = |A|\,(|A|^{-1}P_a)$ is an element of $\mathscr{J}$ (as a product of an element in $\mathscr{J}$ and a bounded operator). Thus there exists an infinite dimensional projection P_a which belongs to $\mathscr{J}$. Then, by a general argument there exists an isometry V from $\mathcal{H}$ to $P_a\mathcal{H}$, and then $V^*P_aV = \mathbf{1} \in \mathscr{J}$, since $P_a \in \mathscr{J}$. The fact that $\mathbf{1} \in \mathscr{J}$ automatically implies that $\mathscr{J} = \mathscr{B}(\mathcal{H})$. □

As a consequence of the previous statement, any two-sided ideal in $\mathscr{B}(\mathcal{H})$ consists in compact operators. On the other hand, one easily shows that if this ideal $\mathscr{J}$ contains at least one rank one projection and is norm closed, then it is automatically equal to $\mathscr{K}(\mathcal{H})$. However, without this assumption of closeness, more possibilities exist. Let us first add a short but rather astonishing lemma of comparison between elements of $\mathscr{K}(\mathcal{H})$.

Lemma 2.4.2. *Let $\mathscr{J}$ be a two-sided proper ideal in $\mathscr{B}(\mathcal{H})$, and let $A, B \in \mathscr{K}(\mathcal{H})$ with $\mu_n(B) \leq \mu_n(A)$ for any $n \in \mathbb{N}$. If $A \in \mathscr{J}$, then $B \in \mathscr{J}$.*

Proof. Let $A = \sum_n \mu_n(A)|g_n\rangle\langle f_n|$ and $B = \sum_n \mu_n(B)|k_n\rangle\langle h_n|$ be the canonical expansion of A and B, as introduced in (2.2). Since these respective families of vectors are orthonormal there exist a partial isometry D with $D^*f_n = h_n$ and a contraction C with $Cg_h = \mu_n(B)\mu_n(A)^{-1}k_n$. Since $B = CAD$ it follows that B belongs to A, as stated. □

Corollary 2.4.3. *Every two-sided ideal $\mathscr{J}$ of $\mathscr{B}(\mathcal{H})$ is invariant under taking the adjoint,* i.e. *if $A \in \mathscr{J}$ then $A^* \in \mathscr{J}$.*

Proof. Since $\mu_n(A^*) = \mu_n(A)$ for any $n \in \mathbb{N}$, the statement follows from the previous lemma. □

Another consequence of the previous lemma is that two-sided ideals of $\mathscr{B}(\mathcal{H})$ are completely described by a set of sequences. Let us be more precise about such a statement, by following the adaptation of the main result of [Cal, Sec. 1] provided in [Sim, Chap. 2].

Definition 2.4.4. *A vector subspace J of c_0 is called* a Calkin space *if it possesses the* Calkin property, *namely whenever $a \in J$ and $b \in c_0$ with $b_j^* \leq a_j^*$ for any $j \in \mathbb{N}$, then $b \in J$.*

Theorem 2.4.5 (Calkin correspondence)**.** *There exists a bijective relation between the set of Calkin spaces and the set of two-sided ideals of $\mathscr{B}(\mathcal{H})$.*

Exercise 2.4.6. *Provide a proof of this theorem, and provide a construction for this correspondence as explicitly as possible.*

The previous result together with Theorem 2.3.11 establish a relation between symmetric norms discussed in Section 2.3 and two-sided ideals. Indeed, if Φ is a symmetric norm on c_c which is not equivalent to ℓ_∞ then the maximal space J_Φ and the minimal space $J_\Phi^{(0)}$ are Calkin spaces. The corresponding two-sided ideals of $\mathscr{B}(\mathcal{H})$ will be denoted respectively by $\mathscr{J}_\Phi$ and $\mathscr{J}_\Phi^{(0)}$. More precisely $\mathscr{J}_\Phi$, resp. $\mathscr{J}_\Phi^{(0)}$, are defined by the set of compact operators whose singular values belong to J_Φ, resp. $J_\Phi^{(0)}$. Then, for any $A \in \mathscr{J}_\Phi$ we set

$$\Phi(A) := \Phi\big((\mu_1(A), \mu_2(A), \dots)\big). \tag{2.25}$$

Let us mention a different way of computing this number, see [Sim, Prop. 2.6] for its proof. For that purpose we let $\mathcal{L}$ represent the set of all orthonormal sets $\{f_n\} \subset \mathcal{H}$.

Proposition 2.4.7. *If $A \in \mathscr{J}_\Phi$, then*

$$\Phi(A) = \sup_{\{f_n\},\{g_n\} \in \mathcal{L}} \Phi\big((\langle g_n, Af_n\rangle)\big).$$

The links between Φ and $\mathscr{J}_\Phi$ are summarized in the following statement.

Theorem 2.4.8. *Let Φ be a symmetric norm on c_c, and let $\mathscr{J}_\Phi$ be the corresponding two-sided ideal of $\mathscr{B}(\mathcal{H})$.*

(a) Φ defines a norm on $\mathscr{J}_\Phi$ by the relation (2.25) and satisfies for all $B \in \mathscr{J}_\Phi$ and $A, C \in \mathscr{B}(\mathcal{H})$:

$$\Phi(ABC) \le \|A\|\,\|C\|\,\Phi(B) \tag{2.26}$$

$$\Phi(B) \ge \|B\|\,\Phi\big((1,0,\dots)\big). \tag{2.27}$$

(b) $\mathscr{J}_\Phi$ and $\mathscr{J}_\Phi^{(0)}$ are Banach spaces with the norm Φ, and $\mathscr{J}_\Phi^{(0)}$ is the closure in $\mathscr{J}_\Phi$ of the finite rank operators. For any $A \in \mathscr{J}_\Phi^{(0)}$ the canonical decomposition provided in (2.2) converges in the Φ-norm.

(c) Any norm on a two-sided ideal $\mathscr{J}$ obeying (2.26) agrees, on the finite rank operators, with a norm $\tilde{\Phi}$ defined by a symmetric norm on c_c. In addition one has $\mathscr{J} \subset \mathscr{J}_{\tilde{\Phi}}$, and if $\mathscr{J}$ is a Banach space with its norm then $\mathscr{J}_{\tilde{\Phi}}^{(0)} \subset \mathscr{J}$.

(d) (non-commutative Fatou Lemma) If $A_m \in \mathscr{J}_\Phi$ with $w\!-\!\lim_{m\to\infty} A_m = A \in \mathscr{K}(\mathcal{H})$ and if $\sup_m \Phi(A_m) < \infty$, then $A \in \mathscr{J}_\Phi$ and $\Phi(A) \le \sup_m \Phi(A_m)$. If Φ is not equivalent to ℓ_∞, then A need not be assumed to be compact a priori.

As a consequence of the point (a) we shall call $\mathscr{J}_\Phi$ *a normed ideal of $\mathscr{B}(\mathcal{H})$*.

Exercise 2.4.9. *Provide a proof of the above statement. Note that the material introduced in Section 2.3 and in particular Theorem 2.3.11 are extensively used for this proof.*

Let us still mention and prove some general results which apply to arbitrary norms Φ. More detailed investigations on certain normed ideal will be realized in the next section. For the time being, let us just mention the spaces $\mathscr{I}_p$ and $\mathscr{I}_{p,w}$ which correspond to the normed ideals constructed from the symmetric norms $\|\cdot\|_p$ and $\|\cdot\|_{p,w}$ exhibited in Examples 2.3.10.

Theorem 2.4.10 (Abstract Hölder inequality)**.** *Let Φ_1, Φ_2 and Φ_3 be three symmetric norms on c_c and let J_{Φ_1}, J_{Φ_2} and J_{Φ_3} denote the corresponding maximal spaces. If for any $a \in J_{\Phi_2}$ and $b \in J_{\Phi_3}$ one has $ab \in J_{\Phi_1}$ (pointwise product) and*

$$\Phi_1(ab) \leq \Phi_2(a)\,\Phi_3(b),$$

then if $A \in \mathscr{I}_{\Phi_2}$ and $B \in \mathscr{I}_{\Phi_3}$ it follows that $AB \in \mathscr{I}_{\Phi_1}$ and

$$\Phi_1(AB) \leq \Phi_2(A)\,\Phi_3(B).$$

If either $A \in \mathscr{I}^{(0)}_{\Phi_2}$ or $B \in \mathscr{I}^{(0)}_{\Phi_3}$, then $AB \in \mathscr{I}^{(0)}_{\Phi_1}$.

Proof. By the inequality (2.9) together with Corollary 2.3.4 one infers that

$$\sum_{j=1}^{d} \mu_j(AB) \leq \sum_{j=1}^{d} \mu_j(A)\,\mu_j(B).$$

Then, by Theorem 2.3.11.(b) one deduces that

$$\begin{aligned}\Phi_1(AB) = \Phi_1\Big(\big(\mu_n(AB)\big)\Big) &\leq \Phi_1\Big(\big(\mu_n(A)\,\mu_n(B)\big)\Big)\\ &\leq \Phi_2\Big(\big(\mu_n(A)\big)\Big)\,\Phi_3\Big(\big(\mu_n(B)\big)\Big) = \Phi_2(A)\Phi_3(B).\end{aligned}$$

The second part of the statement is straightforward. □

Corollary 2.4.11. *Let $p, q, r \geq 1$ satisfy $p^{-1} = q^{-1} + r^{-1}$. If $A \in \mathscr{I}_q$ and $B \in \mathscr{I}_r$, then $AB \in \mathscr{I}_p$ with*

$$\|AB\|_p \leq \|A\|_q\,\|B\|_r.$$

For $p > 1$, if $A \in \mathscr{I}_{q,w}$ and $B \in \mathscr{I}_{r,w}$, then $AB \in \mathscr{I}_{p,w}$ with

$$\|AB\|_{p,w} \leq \frac{p}{p-1}\|A\|_{q,w}\,\|B\|_{r,w}.$$

Proof. The first part of the statement is a direct application of the previous theorem together with Hölder inequality while the second one follows from the inequality

$$\frac{(p-1)}{p}\|ab\|_{p,w} \leq \|a\|'_{q,w}\,\|b\|'_{r,w} \leq \|a\|_{q,w}\,\|b\|_{r,w}$$

with the notations introduced in Examples 2.3.10. □

Extension 2.4.12. *Study the complex interpolation in this general framework as introduced in Theorem 2.9 and 2.10 of [Sim]. Provide some applications of these abstract results.*

2.5 The Schatten ideals $\mathscr{J}_p$

In this section we focus on the normed ideals $\mathscr{J}_p$ which are closely related to the commutative ℓ_p-spaces. As mentioned earlier, the norm on $\mathscr{J}_p$ is constructed from the usual ℓ_p-norm. This material is very classical and can be found in any textbook on operator theory. Note that $\mathscr{J}_2$ is usually called *the set of Hilbert-Schmidt operators* while $\mathscr{J}_1$ corresponds to the set of *trace class operators*. More generally, the space $\mathscr{J}_p$ is called the p-Schatten ideal.

The first result about Hilbert-Schmidt operators is useful in applications. We refer for example to [Sim, Thm. 2.11] or [Amr, Prop. 2.15] for its proof.

Theorem 2.5.1. *Let (Ω, μ) be a measure space such that $\mathcal{H} := L^2(\Omega, \mu)$ is separable. An operator A belongs to $\mathscr{J}_2$ if and only if there exists a measurable function $k \in L^2(\Omega \times \Omega, \mu \otimes \mu)$ such that*

$$[Af](x) = \int_\Omega k(x,y)f(y)\mu(\mathrm{d}y). \tag{2.28}$$

In addition the following relation holds

$$\|A\|_{HS} := \|A\|_2 = \|k\|_{L^2(\Omega\times\Omega)}. \tag{2.29}$$

Note that we have used the convenient notation $\|\cdot\|_{HS}$ for the norm $\|\cdot\|_2$ which is often used for Hilbert-Schmidt operators. Now, such a simple characterization of trace class operators does not exist, and this is quite unfortunate since trace class operators often play an important role. Nevertheless, some partial results exist, as presented in the next statement for positive operators.

Theorem 2.5.2. *Let μ be a Baire measure*[2] *on a locally compact Hausdorff space Ω. Let $\mathcal{H} := L^2(\Omega, \mu)$ and let k be a continuous function on $\Omega \times \Omega$. Assume that the following two conditions hold:*

(i) For any $f \in C_c(\Omega)$ one has

$$\iint_{\Omega\times\Omega} \bar{f}(x)f(y)k(x,y)\mu(\mathrm{d}x)\mu(\mathrm{d}y) \geq 0,$$

(ii) $\int_\Omega k(x,x)\mu(\mathrm{d}x) < \infty$.

Then there exists a positive operator A defined by (2.28) which belongs to $\mathscr{J}_1$ and the following relation holds:

$$\|A\|_1 = \int_\Omega k(x,x)\mu(\mathrm{d}x). \tag{2.30}$$

[2]Recall that the Baire sets form a σ-algebra of a topological space that avoids some of the pathological properties of Borel sets. However, in Euclidean spaces the concept of a Baire set coincides with that of a Borel set.

For the proof, we refer to page 65 of [RS3, Sec. XI.4], or to [GK, Sec. III.10] for a more comprehensive approach to the subject.

Extension 2.5.3. *Study the more recent results obtained by C. Brislawn in [Bri] and mentioned in the Addendum D of [Sim, page 128].*

Let us add two results which are often used as a definition of trace class and Hilbert-Schmidt operators. Since we have introduced $\mathscr{J}_1$ and $\mathscr{J}_2$ through a different approach, one has to show that the two definitions coincide. We refer for example to [Mur, Sec. 2.4] for this alternative approach. Note that in the approach used for example in [Mur] various properties of the norms $\|\cdot\|_1$ and $\|\cdot\|_2$ have to be shown independently, while in our approach all these results follow from the general theory of symmetric norms on c_c.

Proposition 2.5.4. *1) Let $A \in \mathscr{B}(\mathcal{H})$ be positive and let $\{f_n\}$ be an orthonormal basis of $\mathcal{H}$. Then $\sum_n \langle f_n, Af_n\rangle$ is independent of the choice of basis, and it is finite if and only if $A \in \mathscr{J}_1$, with $\sum_n \langle f_n, Af_n\rangle = \|A\|_1$.*

2) Let $B \in \mathscr{B}(\mathcal{H})$ and let $\{f_n\}, \{g_n\}$ be orthonormal bases of $\mathcal{H}$. Then $\sum_n \|Bf_n\|^2$ and $\sum_{n,m} |\langle g_m, Bf_n\rangle|^2$ are independent of the choice of bases and equal. They are finite if and only if $B \in \mathscr{J}_2$, and in this case are equal to $\|B\|_2^2$.

Proof. Let us first observe that 1) follows from 2). Indeed, by setting $B := A^{1/2}$, one infers from 2) that for any orthonormal basis $\{f_n\}$

$$\sum_n \langle f_n, Af_n\rangle = \sum_n \|Bf_n\|^2 = \|B\|_2^2 = \sum_j \mu_j(B)^2 = \sum_j \mu_j(A) = \|A\|_1$$

where we have used that $\mu_j(B)^2 = \mu_j(A)$ which is a direct consequence of the spectral theorem for self-adjoint operators.

For the proof of 2), observe first by Parseval's identity one has

$$\sum_n \|Bf_n\|^2 = \sum_{n,m} |\langle g_m, Bf_n\rangle|^2 = \sum_{n,m} |\langle B^* g_m, f_n\rangle|^2 = \sum_m \|B^* g_m\|^2.$$

By symmetry, one directly gets the required equality and the independence with respect to the choice of a basis. Now, if $B \in \mathscr{J}_2$, *i.e.* if $\sum_n \mu_n(B)^2 = \|B\|_2^2 < \infty$, one easily gets from the canonical decomposition of B provided in (2.2) that $\sum_n \|Bg_n\|^2$ is finite and equal to $\sum_n \mu_n(B)^2$. Conversely, if $\sum_{n,m} |\langle g_m, Bf_n\rangle|^2 < \infty$ one has

$$\sum_n |\langle g_n, Bf_n\rangle|^2 \le \sum_{n,m} |\langle g_m, Bf_n\rangle|^2 < \infty$$

and $B \in \mathscr{J}_2$ by Proposition 2.4.7. □

In the next statement we emphasize once more the relation between Hilbert-Schmidt operators and trace class operators. Its proof follows easily from what has been introduced so far, see also [Mur, Thm. 2.4.13].

Proposition 2.5.5. *Let A be an element of $\mathscr{B}(\mathcal{H})$. The following conditions are equivalent:*

(i) A is a trace class operator,

(ii) $|A|$ is a trace class operator,

(iii) $|A|^{1/2}$ is a Hilbert-Schmidt operator,

(iv) There exists Hilbert-Schmidt operators B_1, B_2 such that $A = B_1 B_2$.

We close this section with several results on convergence in $\mathscr{I}_p$. The first one is clearly an analog of the dominated convergence theorem. Note that we shall use the convenient notation $A^{(*)}$ for A and for its adjoint A^*. For example, the condition $|A^{(*)}| \leq B$ means $|A| \leq B$ and $|A^*| \leq B$.

Theorem 2.5.6. *Let $A_m, A, B \in \mathscr{B}(\mathcal{H})$ with B self-adjoint. Assume that $|A_m^{(*)}| \leq B$, $|A^{(*)}| \leq B$ and that $w - \lim_{m\to\infty} A_m = A$. Then, if $B \in \mathscr{I}_p$ for some $p < \infty$, then $\|A_m - A\|_p \to 0$ as $m \to \infty$.*

A proof of this statement is provided in [Sim, Thm. 2.16]. The following result is also proved at the end of chapter 2 of [Sim].

Theorem 2.5.7. *Let p belongs to $[1, \infty)$, and let $\{A_n\} \subset \mathscr{I}_p$ and $A \in \mathscr{I}_p$. If $w - \lim_{n\to\infty} A_n = A$ and $\lim_{n\to\infty} \|A_n\|_p = \|A\|_p$, then $\lim_{n\to\infty} \|A_n - A\|_p = 0$.*

Let us mention that more generally, results like the previous one are a consequence of the uniform convexity of some Banach spaces. We shall not go further in this direction here.

Extension 2.5.8. *Study the notion of uniform convexity for Banach spaces and deduce from this notion the content of the previous theorem.*

2.6 Usual trace

We can finally define the notion of trace, which extends the usual one on matrices. Based on Proposition 2.5.4.(1) one infers that the domain for the "trace" which is closed under $A \mapsto |A|$ can only be $\mathscr{I}_1$. More precisely one has:

Theorem 2.6.1. *Let $A \in \mathscr{I}_1$ with its canonical decomposition $A = \sum_n \mu_n(A)|g_n\rangle\langle f_n|$. Then for any orthonormal basis $\{h_m\}$ of $\mathcal{H}$ one has $\sum_m |\langle h_m, Ah_m\rangle| < \infty$ and*

$$\sum_m \langle h_m, Ah_m\rangle = \sum_n \mu_n(A)\langle f_n, g_n\rangle =: \mathrm{Tr}(A) \tag{2.31}$$

is independent of this basis. Moreover

$$|\mathrm{Tr}(A)| \leq \|A\|_1, \tag{2.32}$$

the map $A \mapsto \mathrm{Tr}(A)$ is a bounded linear functional on $\mathscr{I}_1$, and for any $A \in \mathscr{I}_1$ and $B \in \mathscr{B}(\mathcal{H})$ one has $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

Proof. Let us set $\alpha_{nm} := \langle f_n, h_m\rangle\langle h_m, g_n\rangle$ which is a dss matrix. Then $\sum_m |\alpha_{nm}| \leq 1$ for each n, and one has

$$\sum_m |\langle h_m, Ah_m\rangle| = \sum_{m,n} |\alpha_{nm}|\mu_n(A) \leq \|A\|_1$$

which directly proves (2.32). In addition, the absolute convergence of the last double sum justifies an interchange in

$$\sum_m \langle h_m, Ah_m\rangle = \sum_{m,n} \alpha_{nm}\mu_n(A) = \sum_n \Big(\mu_n(A)\sum_m \alpha_{nm}\Big) = \sum_n \mu_n(A)\langle f_n, g_n\rangle.$$

The linearity of the map $A \mapsto \mathrm{Tr}(A)$ follows again from the absolute convergence of the sums. Finally, if $B \in \mathscr{B}(\mathcal{H})$ one has

$$\mathrm{Tr}(AB) = \sum_m \langle h_m, ABh_m\rangle = \sum_n \mu_n(A)\langle B^* f_n, g_n\rangle = \sum_m \langle h_m, BAh_m\rangle = \mathrm{Tr}(BA). \quad \square$$

Corollary 2.6.2. *If $A \in \mathscr{J}_1$ and $B \in \mathscr{B}(\mathcal{H})$, then one has*

$$|\mathrm{Tr}(AB)| \leq \|B\|\mathrm{Tr}(|A|).$$

Proof. By the previous theorem one has

$$|\mathrm{Tr}(AB)| \leq \|AB\|_1 \leq \|B\|\,\|A\|_1 = \|B\|\,\mathrm{Tr}(|A|). \qquad \square$$

From the duality theory for symmetric norm introduced just before Theorem 2.3.12 and from the results contained in this statement one easily gets:

Theorem 2.6.3. *Let Φ and Φ' be conjugate symmetric norms on c_c. If $A \in \mathscr{J}_\Phi$ and $B \in \mathscr{J}_{\Phi'}$, then $AB \in \mathscr{J}_1$. Moreover, for each fixed $B \in \mathscr{J}_{\Phi'}$ the map $A \mapsto \mathrm{Tr}(AB)$ is a bounded linear functional in $\mathscr{J}_\Phi$ with norm $\Phi'(B)$. If Φ' is not equivalent to the ℓ_∞-norm, then every functional on $\mathscr{J}_\Phi^{(0)}$ is of this form, that is $\big(\mathscr{J}_\Phi^{(0)}\big)^* = \mathscr{J}_{\Phi'}$. In particular, $\mathscr{J}_\Phi$ is a reflexive space if and only if both Φ and Φ' are regular. If Φ is the ℓ_1-norm, then $\mathscr{J}_\Phi^* = \mathscr{B}(\mathcal{H})$.*

Since the trace on elements of $\mathscr{J}_1$ has now been defined in (2.31), a natural question is about the equality

$$\mathrm{Tr}(A) = \sum_n \lambda_n(A) \tag{2.33}$$

where $\{\lambda_n(A)\}$ corresponds to the set of eigenvalues of A, multiplicity counted. This equality is indeed correct, but as emphasized in any textbook on the subject its proof is surprisingly difficult. It has only been proved in 1959 by Lidskii. Note that the main difficulty comes from nilpotent or quasinilpotent operators (an operator A satisfying respectively $A^d = \mathbf{0}$ for some $n \in \mathbb{N}$ or $\sigma(A) = \{0\}$). We do not provide the proof of the equality (2.33) but suggest to study it as an extension:

Extension 2.6.4. *Study the proof the Lidskii's theorem, namely the equality* (2.33)*, either from the information provided in [Sim, Chap. 3] or from any other reference.*

One direct consequence of the equality (2.33) is contained in the following statement.

Corollary 2.6.5. *If $A, B \in \mathscr{B}(\mathcal{H})$ have the property that both AB and BA belong to $\mathscr{J}_1$ (as for example if $A \in \mathscr{J}_\Phi$ and $B \in \mathscr{J}_{\Phi'}$ for any conjugate symmetric norms on c_c), then*

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA). \tag{2.34}$$

Proof. As well known, and shown for example in [Sak, Prop. 1.1.8], the operators AB and BA share the same spectrum, including the algebraic multiplicity, with the only possible exception of 0. Thus, the equality (2.34) follows directly from this fact and from (2.33). □

Let us add one more result related to integral operators which are trace class. Note that the following statement does not contradict Theorem 2.5.2 since it is assumed from the beginning that the operator is trace class.

Theorem 2.6.6. *Let $\mathcal{H} := L^2([a, b])$ and let $A \in \mathscr{J}_1$ be of the form $[Af](x) = \int_a^b k(x, y)f(y)\,\mathrm{d}y$ for some continuous function $k : [a, b] \times [a, b] \to \mathbb{C}$ and all $f \in \mathcal{H}$. Then*

$$\mathrm{Tr}(A) = \int_a^b k(x, x)\,\mathrm{d}x.$$

The proof of this statement is provided in [Sim, Thm. 3.9] and is based on the construction of an explicit basis for $\mathcal{H} := L^2([a, b])$. Many applications of the theory developed so far could be presented. Quite a lot of them are presented in the subsequent chapters of [Sim].

Up to this point, the uniqueness of the above trace has not been discussed. In fact, this uniqueness holds under an additional condition which we are going to introduce. The following material is borrowed from [Les], and we start by recalling an extension of the notion of trace. Recall that if V is a real vector space, then a map $\Phi : V \to [0, \infty]$ is *positive homogeneous* if $\Phi(\lambda v) = \lambda\Phi(v)$ for any $\lambda \geq 0$ and $v \in V$, and is *additive* if $\Phi(v + w) = \Phi(v) + \Phi(w)$ for any $v, w \in V$.

Definition 2.6.7. *A* weight *on $\mathscr{B}(\mathcal{H})$ is a map $\tau : \mathscr{B}(\mathcal{H})_+ \to [0, \infty]$ which is positive homogeneous and additive. Such a weight is* tracial *if $\tau(BB^*) = \tau(B^*B)$ for any $B \in \mathscr{B}(\mathcal{H})$.*

Note that in some references a tracial weight is simply called a trace. However, let us emphasize that a weight is only defined on the positive cone of $\mathscr{B}(\mathcal{H})$, and it can take the value ∞. Now, the trace Tr defined in (2.31) for any $A \in \mathscr{B}(\mathcal{H})_+$ is clearly a tracial weight on $\mathscr{B}(\mathcal{H})$, see also Proposition 3.4.3 and Corollary 3.4.4 in [Ped] for the proof of this statement. In addition, if we define the subset of $\mathscr{B}(\mathcal{H})_+$ on which Tr is finite, one gets $(\mathscr{J}_1)_+$, and then its linear span leads to $\mathscr{J}_1$, as introduced in the previous section.

Let us now show that up to a normalization constant this trace is the unique one on the set $\mathscr{F}(\mathcal{H})$ of finite rank operators in $\mathcal{H}$. In the present context, a *trace* τ on a complex algebra $\mathscr{A}$ is a linear functional $\mathscr{A} \to \mathbb{C}$ satisfying $\tau(AB) = \tau(BA)$ for any $A, B \in \mathscr{A}$.

Lemma 2.6.8. *Any trace on* $\mathscr{F}(\mathcal{H})$ *is proportional to* Tr.

Proof. Let $P, Q \in \mathscr{F}(\mathcal{H})$ be rank one orthogonal projections, or in other words $P = |f\rangle\langle f|$ and $Q = |g\rangle\langle g|$ for some $f, g \in \mathcal{H}$ with $\|f\| = \|g\| = 1$. We now set $T := |g\rangle\langle f|$ which is still a finite rank operator and satisfies $TT^* = |g\rangle\langle g| = Q$ and $T^*T = |f\rangle\langle f| = P$. Thus, if τ is a trace on $\mathscr{F}(\mathcal{H})$ one has

$$\tau(P) = \tau(T^*T) = \tau(TT^*) = \tau(Q),$$

which means that τ takes the same value $\lambda_\tau \in \mathbb{C}$ on all rank one orthogonal projections. Thus for any rank one orthogonal projection P one has

$$\tau(P) = \lambda_\tau = \lambda_\tau \mathrm{Tr}(P).$$

Since any $T \in \mathscr{F}(\mathcal{H})$ is a linear combination of rank one orthogonal projections, the result follows by linearity of τ and Tr. □

The properties of Tr mentioned so far are not sufficient for showing that any tracial weight on $\mathscr{B}(\mathcal{H})$ is proportional to Tr. The necessary additional property is *normality*, as introduced below. Note that we shall also impose that $\tau(B) \geq 0$ if $B \geq 0$, which is a natural requirement.

Definition 2.6.9. *A tracial weight* τ *on* $\mathscr{B}(\mathcal{H})$ *is* normal *if for any increasing sequence* $\{B_n\} \subset \mathscr{B}(\mathcal{H})_+$ *such that* $s - \lim_{n\to\infty} B_n = B \in \mathscr{B}(\mathcal{H})_+$ *one has* $\tau(B) = \sup_n \tau(B_n)$.

One can now prove the following statement:

Theorem 2.6.10. *(i) The usual trace* Tr *on* $\mathscr{B}(\mathcal{H})$ *is normal,*

(ii) If τ *is any normal tracial weight on* $\mathscr{B}(\mathcal{H})$ *then there exists a constant* $\lambda_\tau \in [0, \infty)$ *such that* $\tau(B) = \lambda_\tau \mathrm{Tr}(B)$ *for any* $B \in \mathscr{B}(\mathcal{H})_+$.

Note that an additional pathological case also exists: The tracial weight τ_∞ is defined by $\tau_\infty(B) = \infty$ for any $B \in \mathscr{B}(\mathcal{H})_+ \setminus \{\mathbf{0}\}$ and $\tau_\infty(\mathbf{0}) = 0$. In such a case one has $\lambda_{\tau_\infty} = \infty$. We shall not consider this case subsequently.

Proof. i) Let $\{f_m\}$ be an orthonormal basis of $\mathcal{H}$ and let $s - \lim_{n\to\infty} B_n = B$ in $\mathscr{B}(\mathcal{H})_+$ be an increasing sequence. Then one has $\langle f_m B_n f_m\rangle \nearrow \langle f_m, B f_m\rangle$ for any m, and therefore $\sup_n \langle f_m B_n f_m\rangle = \langle f_m, B f_m\rangle$. It follows then from the monotone convergence theorem for the discrete measure on $\mathbb{N}$ that

$$\mathrm{Tr}(B) = \sum_m \langle f_m, B f_m\rangle = \sum_m \sup_n \langle f_m, B_n f_m\rangle = \sup_n \sum_m \langle f_m, B_n f_m\rangle = \sup_n \mathrm{Tr}(B_n).$$

ii) As in the proof of Lemma 2.6.8 one observes that $\tau \upharpoonright \mathscr{F}(\mathcal{H}) = \lambda_\tau \mathrm{Tr} \upharpoonright \mathscr{F}(\mathcal{H})$ for some $\lambda_\tau \in [0, \infty)$. Let us now choose any increasing sequence $\{P_n\}_{n\in\mathbb{N}}$ of orthogonal projections with the dimension of $\mathsf{Ran}(P_n)$ equal to n. Then, given any $B \in \mathscr{B}(\mathcal{H})_+$ one can consider the increasing sequence $\{B^{1/2}P_nB^{1/2}\}_{n\in\mathbb{N}} \subset \mathscr{F}(\mathcal{H})$ which converges strongly to B. Since τ is assumed to be normal we get

$$\tau(B) = \sup_n \tau\big(B^{1/2}P_nB^{1/2}\big) = \sup_n \lambda_\tau \mathrm{Tr}\big(B^{1/2}P_nB^{1/2}\big) = \lambda_\tau \mathrm{Tr}(B). \qquad \square$$

The conclusion of the previous construction is that on $\mathscr{B}(\mathcal{H})$ and up to a multiplicative constant, the only tracial normal weight is Tr. If we drop the condition of normality, this is not the case, as shown in the next chapter. Note finally that the previous proof is based on the fact that the strong closure of $\mathscr{K}(\mathcal{H})$ is $\mathscr{B}(\mathcal{H})$ itself. In other contexts, such an approximation argument might not be available.