CHAPTER 1

Preliminaries

We begin with a number of short topics. Since the first two are fairly standard, we will not give all details. The last topic is rather specialized and, in its discussion, we will use results from Chapters 2 and 3.

1.1. Absolute Value and Polar Decomposition

Throughout, all our Hilbert spaces will be *complex* and *separable* (are there any others?) and our inner product is linear in the *second* factor and conjugate linear in the first. Recall that a bounded operator, A, is called *positive* (written $A \geq 0$) if and only if $(\phi, A\phi) \geq 0$ for all ϕ; it then follows by polarization that $A^* = A$. If $A - B \geq 0$, we write $A \geq B$ or $B \leq A$. It can be proven that for any $A \geq 0$, there is a unique $B \geq 0$ with $B^2 = A$ (see [**250**, Section VI.4]); we write $B = \sqrt{A}$. For any bounded operator, A, the operator A^*A is positive and one defines

$$|A| = \sqrt{A^*A}$$

We warn the reader that

$$|A + B| \leq |A| + |B| \tag{1.1}$$

is *false* in general. For example, if

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

then

$$|A + B| = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix} \qquad |A| + |B| = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}$$

So $(\phi, |A + B|\phi) > (\phi, (|A| + |B|)\phi)$ for $\phi = \binom{1}{0}$, showing that (1.1) does not hold.

Notice that $\||A|\phi\|^2 = \|A\phi\|^2$. From this, it is easy (see [**250**, Section VI.4]) to construct uniquely an operator U so that

(i) $A = U|A|$
(ii) $\|U\psi\| = \|\psi\|$ for $\psi \in \overline{\mathrm{Ran}|A|} = (\ker A)^\perp$
(iii) $\|U\psi\| = 0$ for $\psi \in (\mathrm{Ran}|A|)^\perp = \ker A$.

Notice that $|A| = U^*A$.

The formula, $A = U|A|$, is called the *polar decomposition* of A.

1.2. Compact Operators and the Canonical Decomposition

We call a bounded operator A on a Hilbert space *finite rank* if $\dim(\mathrm{Ran}\, A) < \infty$ (the dimension of $\mathrm{Ran}\, A$ is called the *rank* of A). A bounded operator, A, is called *compact* if and only if it is a *norm limit* of finite rank operators. (For a Hilbert space, but not a general Banach space, this is equivalent to the more usual Riesz definition that $A[\{\phi \mid \|\phi\| \leq 1\}]$ has compact closure.) The following two results are standard:

THEOREM 1.1. *Let A be a compact operator on a Hilbert space, $\mathcal{H}$. Then*

(i) *$\sigma(A)$, the spectrum of A, is a set with no non-zero limit point.*

(ii) *Every non-zero $\lambda \in \sigma(A)$ is an eigenvalue of finite (geometric) multiplicity.*

(iii) *For any non-zero $\lambda \in \sigma(A)$, there is a finite rank projection P_λ so that $AP_\lambda = P_\lambda A$ with $\sigma(A \restriction P_\lambda \mathcal{H}) = \{\lambda\}$, and $\sigma(A \restriction (1 - P_\lambda)\mathcal{H}) = \sigma(A) \backslash \{\lambda\}$. $\dim P_\lambda$ is called the (algebraic) multiplicity of λ.*

THEOREM 1.2. *Let A be a self-adjoint compact operator on a Hilbert space, $\mathcal{H}$. Then $\mathcal{H}$ has an orthonormal basis of eigenvectors for A.*

Theorem 1.2 and (i) and (ii) of Theorem 1.1 are fairly standard and discussed in almost every functional analysis text (cf. [**250**, Section VI.5]). (iii) of Theorem 1.1 is a little more specialized but follows from results in Kato [**164**] or by combining results from [**250**, Section VI.5] and the appendix to [**254**, Section XII.1]. A formula for P_λ is

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

where ε is any sufficiently small positive number.

Let $\mathcal{I}_\infty$ denote the family of compact operators.

THEOREM 1.3. *$\mathcal{I}_\infty$ is a two-sided ideal closed under taking adjoints. In particular, $A \in \mathcal{I}_\infty$ if and only if $|A| \in \mathcal{I}_\infty$.*

PROOF. The finite rank operators are easily seen to be a two-sided ideal closed under taking adjoints, so $\mathcal{I}_\infty$ has the same property. □

THEOREM 1.4. *Let A be a compact operator. Then A has the norm convergent expansion*

$$A = \sum_{n=1}^{N} \mu_n(A)(\phi_n, \cdot\,)\psi_n \tag{1.2}$$

(where N is a finite non-negative integer or infinity), each $\mu_n(A) > 0$, $\mu_1(A) \geq \mu_2(A) \geq \cdots$, and the $\{\phi_n\}$ and $\{\psi_n\}$ are (not necessarily complete) orthonormal sets. Moreover, the $\mu_n(A)$ are uniquely determined and the ϕ's and ψ's are essentially uniquely determined.

REMARK. (1.2) is shorthand for

$$A\eta = \sum_{n=1}^{N} \mu_n(A)(\phi_n, \eta)\psi_n$$

PROOF. $|A| = U^*A$ is compact and self-adjoint so, by Theorem 1.2, it has a norm convergent expansion

$$|A| = \sum_{n=1}^{N} \mu_n(A)(\phi_n, \cdot\,)\phi_n$$

where the $\mu_n(A)$ are the non-zero eigenvalues of $|A|$ and ϕ_n the corresponding eigenvectors. Since U is an isometry on $\operatorname{Ran}|A|$, the $\psi_n = U\phi_n$ are orthonormal, proving (1.2). The uniqueness follows if one notes that if (1.2) holds with the stated properties, then $\mu_n(A)^2$ are the non-zero eigenvalues of A^*A, $\{\phi_n\}$ are the eigenvectors of A^*A, and $\{\psi_n\}$ are the eigenvectors of AA^*. The lack of uniqueness of the ϕ's and ψ's comes from the possibility of degenerate eigenvalues of A^*A and AA^*. □

The $\mu_n(A)$ are called the *singular values* of A. (1.2) is called the *canonical expansion* for A.

1.3. Inequalities on Singular Values, I

The basis of the inequalities on singular values we discuss now is the elementary equality

$$\mu_n(A) = \mu_n(A^*) \tag{1.3}$$

(which follows from (1.2), or alternatively from the fact that AA^* and A^*A have the same non-zero eigenvalues with the same multiplicities [**87**]) and the following:

THEOREM 1.5.

$$\mu_n(A) = \min_{\phi_1,\dots,\phi_{n-1}} \Big[\max_{\substack{\psi\in[\phi_1,\dots,\phi_{n-1}]^\perp \\ \|\psi\|=1}} \|A\psi\| \Big]$$

Theorem 1.5 follows from the min-max characterization (see [**254**, Section XIII.1]) of the eigenvalues of $-A^*A$ if we note that $\|A\psi\|^2 = (\psi, A^*A\psi)$ and that $-\mu_n(A)^2$ is the n-th eigenvalue of $-A^*A$ counting from the bottom.

THEOREM 1.6. *For any compact A and bounded B,*

$$\mu_n(AB) \le \|B\|\mu_n(A) \tag{1.4a}$$

$$\mu_n(BA) \le \|B\|\mu_n(A) \tag{1.4b}$$

PROOF. (1.4a) follows from (1.4b) and (1.3) since (1.3) implies that $\mu_n(AB) = \mu_n(B^*A^*) \le \|B^*\|\mu_n(A^*) = \|B\|\mu_n(A)$. To prove (1.4b), we need only use Theorem 1.5 and $\|BA\psi\| \le \|B\|\,\|A\psi\|$. □

REMARK AND WARNING. AB and BA have the same non-zero eigenvalues but they may not have the same singular values: For example, if $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, then $AB = 0$, but $BA = A$, so $\mu_1(AB) = \mu_2(AB) = 0$ but $\mu_1(BA) = 1$, $\mu_2(BA) = 0$.

Theorem 1.6 is the first (take $m = 0$) of a family of inequalities of Fan [**112**] which read for $n, m \ge 0$,

$$\mu_{n+m+1}(AB) \le \mu_{n+1}(A)\mu_{m+1}(B)$$

We will not need these, but the following inequalities of Fan [**112**] will be useful:

THEOREM 1.7 (Fan [**112**]). *Let A and B be compact. Then for $n \ge 0$, $m \ge 0$,*

$$\mu_{n+m+1}(A+B) \le \mu_{n+1}(A) + \mu_{m+1}(B) \tag{1.5}$$

PROOF. Let $Q_n(\phi_1,\dots,\phi_n; A) \equiv \max\{\|A\varphi\| \mid \|\psi\| = 1, \psi \in [\phi_1,\dots,\phi_n]^\perp\}$. Since $\|(A+B)\psi\| \le \|A\psi\| + \|B\psi\|$,

$$\begin{aligned} Q_{n+m}(\phi_1,\dots,\phi_{n+m}; A+B) &\le Q_{n+m}(\phi_1,\dots,\phi_{n+m}; A) + Q_{n+m}(\phi_1,\dots,\phi_{n+m}; B) \\ &\le Q_n(\phi_1,\dots,\phi_n; A) + Q_m(\phi_{n+1},\dots,\phi_{n+m}; B) \end{aligned}$$

Minimizing over $\phi_1,\dots,\phi_{n+m}$ and using Theorem 1.5, Fan's inequality (1.5) results. □

1.4. Rearrangement Inequalities and All That

DEFINITION. Let a_n be an infinite sequence of numbers with $a_n \to 0$ as $n \to \infty$. a_n^* is the sequence defined by $a_1^* = \max|a_i|$, $a_1^* + a_2^* = \max_{i \neq j}(|a_i| + |a_j|)$, etc. Thus $a_1^* \geq a_2^* \geq \cdots$ and the sets of a_i^* and $|a_i|$ are identical, counting multiplicities.

LEMMA 1.8.

$$\sum_n |a_n b_n| \leq \sum_n a_n^* b_n^*$$

PROOF. By a limiting argument, we may suppose that $a_n, b_n = 0$ for $n \geq N+1$. Without loss, we can renumber the a's so that $|a_1| \geq \cdots \geq |a_N|$. Since

$$\sum_n |a_n b_n| = |a_N| \sum_{n=1}^{N} |b_n| + (|a_{N-1}| - |a_N|) \sum_{n=1}^{N-1} |b_n| + \cdots + (|a_1| - |a_2|)|b_1|$$

and $\sum_{n=1}^{k} |b_n| \leq \sum_{n=1}^{k} b_n^*$, we clearly have the required inequality. □

The following result will be central to the treatment in Sections 1.6 and 1.7:

THEOREM 1.9 (essentially due to Markus [**213**]). *Let $a_1 \geq a_2 \geq \cdots \geq a_N \geq 0$ and let b be a point in $\mathbb{C}^N$ with*

$$\sum_{j=1}^{k} b_j^* \leq \sum_{j=1}^{k} a_j \tag{1.6}$$

for $k = 1, \ldots, N$. Then there exist points $a^{(1)}, \ldots, a^{(m)} \in \mathbb{C}^N$ with $(a^{(\ell)})^* = a$ and $0 \leq \lambda_\ell \leq 1$, $\sum_{\ell=1}^{m} \lambda_\ell = 1$ so that

$$b = \sum_{\ell=1}^{m} \lambda_\ell a^{(\ell)}$$

In particular, if (1.6) holds and Φ is a function on $([0,\infty))^N$ so that $\phi(c) = \Phi(c_1^*, \ldots, c_n^*)$ is convex on $\mathbb{C}^N$, then

$$\phi(b) \leq \phi(a)$$

PROOF DUE TO MITJAGIN [**223**]. Let H be the convex hull in $\mathbb{C}^N$ of the points c with $c^* = a$. We want to show that any b obeying (1.6) lies in H (the converse we note is easy). Let ℓ be a complex linear function on $\mathbb{C}^N$ and let

$$\beta = \max_{c \in H} \operatorname{Re}(\ell(c))$$

If we prove that $|\ell(b)| \leq \beta$ for each such ℓ, then b must lie in H by standard hyperplane results (e.g., specialize the Hahn-Banach theorem to $\mathbb{C}^N$). Suppose $\ell(c) = \sum_{n=1}^{N} \alpha_n c_n$ and let $\ell^*(c) \equiv \sum_{n=1}^{N} \alpha_n^* c_n$. By Lemma 1.8, $|\ell(b)| \leq \ell^*(b^*)$. Moreover, there exists an $a' \in \mathbb{C}^N$ with $(a')^* = a$ so that $\ell^*(a) = \ell(a')$ and thus $\ell^*(a) \leq \beta$. The proof is thus complete by noting that

$$\begin{aligned} \ell^*(b^*) &= \alpha_N^* \sum_{n=1}^{N} b_n^* + (\alpha_{N-1}^* - \alpha_N^*) \sum_{n=1}^{N-1} b_n^* + \cdots + (\alpha_1^* - \alpha_2^*) b_1^* \\ &\leq \alpha_N^* \sum_{n=1}^{N} a_n + \cdots = \ell^*(a) \end{aligned}$$

□

REMARKS. 1. Results of this genre predate Markus' work considerably. For example, Hardy, Littlewood, and Pólya [**147**] proved that if (1.6) holds with equality for $k = N$, then $b_i^* = \sum \alpha_{ij} a_j$ for a matrix α of positive numbers whose rows and columns sum to 1, and Birkhoff [**33**] has proved that any such α is a convex combination of permutation matrices $\delta_{i\pi(j)}$. Convexity results of the type quoted also go back to Hardy et al. [**147**]. See Addendum A.

2. Theorem 1.9 plays a major role in the Goh'berg-Krein [**134**] presentation of symmetric norms (see Chapter 2) but, strangely enough, they do not use it to obtain the Weyl and Horn inequalities (see Section 1.6 below).

In applications, the Φ's we will use will either be symmetric norms on $\mathbb{C}^N$ (defined below) or functions of the form $\Phi(x) = \sum_{n=1}^N f(|x_n|)$, where f is convex and monotone increasing on $[0,\infty)$. In the first case, Φ will be convex by the triangle inequality and homogeneity of norms. In the second, we need only note that for $0 \le \theta \le 1$, $c, d \in \mathbb{C}^N$,

$$\begin{aligned} \phi(\theta c + (1-\theta)d) &= \sum_{n=1}^N f(|\theta c + (1-\theta)d|_n) \\ &\le \sum_{n=1}^N \theta f(|c_n|) + (1-\theta) f(|d_n|) \\ &= \theta\phi(c) + (1-\theta)\phi(d) \end{aligned}$$

(1.6) will arise from two distinct sets of ideas. The first involves:

COROLLARY 1.10. *Let $a_1 \ge a_2 \ge \cdots \ge a_N \ge 0$, $b_1 \ge \cdots \ge b_N \ge 0$. Suppose that*

$$\prod_{j=1}^k b_j \le \prod_{j=1}^k a_j \tag{1.7}$$

for $k = 1, \dots, N$. Then for any continuous, monotone increasing function g on $[0,\infty)$ with $t \to g(e^t)$ convex, we have that

$$\sum_{j=1}^k g(b_j) \le \sum_{j=1}^k g(a_j) \tag{1.8}$$

In particular (taking $g(x) = x$), (1.6) holds.

PROOF. By a limiting argument, we may suppose that the a's and b's are all non-zero. But then replacing a_i by γa_i, b_i by γb_i, and $g(x)$ by $g(\gamma^{-1}x)$ for γ large, we can suppose all the a's and b's are bigger than 1. Taking $\tilde{a}_j = \log a_j$, $\tilde{b}_j = \log b_j$, the new variables obey (1.6). Letting $f(x) = g(e^x)$ $(x \ge 0)$ and $\Phi(x) = \sum f(x_n)$ as above, (1.8) follows from Theorem 1.9. □

The second involves the following notion:

DEFINITION. A matrix $\{\alpha_{nm}\}_{1\le n,\, m\le N}$ is called *doubly substochastic* (dss) if and only if

$$\sum_{n=1}^N |\alpha_{nm}| \le 1 \qquad m = 1, \dots, N$$

and

$$\sum_{m=1}^{N} |\alpha_{nm}| \leq 1 \qquad n = 1, \dots, N$$

The relevance of this concept to operator theory comes partly from:

PROPOSITION 1.11. *If $\{\phi_n\}, \{\psi_n\}$ are orthonormal sets, then $\alpha_{nm} = |(\phi_n, \psi_m)|^2$ is a dss matrix. More generally, if $\{\eta_n\}$, $\{\gamma_n\}$ are also orthonormal sets, then $\beta_{nm} = (\phi_n, \psi_m)(\eta_m, \gamma_n)$ is a dss matrix.*

PROOF. By Bessel's inequality, $\sum_n |\alpha_{nm}| \leq \|\psi_m\|^2 = 1$. The β result follows from the α result and the Schwarz inequality

$$\sum_n |\beta_{nm}| \leq \left(\sum_n |(\phi_n, \psi_m)|^2 \right)^{1/2} \left(\sum_n |(\eta_m, \gamma_n)|^2 \right)^{1/2} \qquad \square$$

PROPOSITION 1.12. *Let α be a doubly substochastic matrix. Let $c \in \mathbb{C}^N$ with $a = c^*$ and $b_n = \sum_{m=1}^N \alpha_{nm} c_m$. Then b and a obey* (1.6).

PROOF. By relabelling rows and columns of α, we can suppose that $|b_1| \geq \cdots \geq |b_N|$ and $|c_1| \geq \cdots \geq |c_N|$. Then

$$\sum_{n=1}^{k} b_n^* \leq \sum_{n=1}^{k} \sum_{m=1}^{N} |\alpha_{nm}|\, |a_m| = \sum_{m=1}^{N} \gamma_n |a_m|$$

where the γ's obey $0 \leq \gamma_m \leq 1$, $\sum_{m=1}^N \gamma_m \leq k$. From these properties of the γ's and $a_1 \geq \cdots \geq a_N$, one easily sees that $\sum_{m=1}^N \gamma_m |a_m| \leq \sum_{m=1}^k a_m$. $\square$

REMARK. The point of this proposition is the following: α is dss if and only if α is a contraction on $\mathbb{C}^N$ in both the ℓ_1 and ℓ_∞ norm. This proposition and Theorem 1.9 imply that α is then a contraction in all symmetric norms (see Theorem 1.16 below).

1.5. Antisymmetric Tensor Products

We review the basic ideas; see [**250**, Sections II.4 and VIII.10] and [**254**, Section XIII.17] for further details. Given Hilbert spaces $\mathcal{H}_1, \dots, \mathcal{H}_n$, we use $\hom(\mathcal{H}_1, \dots, \mathcal{H}_n)$ to denote the maps $\ell : \mathcal{H}_1 \times \cdots \times \mathcal{H}_n \to \mathbb{C}$ which are multilinear. Given $\phi_1 \in \mathcal{H}_1$, $\phi_1 \otimes \cdots \otimes \phi_n$ denotes the multilinear function

$$\phi_1 \otimes \cdots \otimes \phi_n : (\psi_1, \dots, \psi_n) \mapsto \prod_{i=1}^{n} (\phi_i, \psi_i)$$

$\hom_f(\mathcal{H}_1, \dots, \mathcal{H}_n)$ denotes the algebraic span of the $\phi_1 \otimes \cdots \otimes \phi_n$ in $\hom(\mathcal{H}_1, \dots, \mathcal{H}_n)$. $\hom_f$ supports a unique inner product with

$$\langle \ell, \phi_1 \otimes \cdots \otimes \phi_n \rangle = \ell(\phi_1, \dots, \phi_n) \tag{1.9}$$

Using (1.9), the abstract completion of $\hom_f$ can be realized as a subset of $\hom(\mathcal{H}_1, \dots, \mathcal{H}_n)$ denoted $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$. Given maps $A_i : \mathcal{H}_i \to \mathcal{H}_i$, there is a map of $\hom(\mathcal{H}_1, \dots, \mathcal{H}_n)$ into itself defined by

$$[(A_1 \otimes \cdots \otimes A_n)(\ell)](\psi_1, \dots, \psi_n) = \ell(A_1^* \psi_1, \dots, A_n^* \phi_n)$$

which takes $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ into itself and obeys

$$(A_1 \otimes \cdots \otimes A_n)(\phi_1 \otimes \cdots \otimes \phi_n) = (A_1\phi_1) \otimes \cdots \otimes (A_n\phi_n)$$

$\otimes_n\mathcal{H}$ is just $\mathcal{H} \otimes \cdots \otimes \mathcal{H}$ (n times). If $\{\phi_i\}$ is an orthonormal basis for $\mathcal{H}$, then $\{\phi_{i_1} \otimes \cdots \otimes \phi_{i_n}\}_{i_1,\dots,i_n}$ is an orthonormal basis for $\otimes_n\mathcal{H}$.

Given $\psi_1, \dots, \psi_n \in \mathcal{H}$, $\psi_1 \wedge \cdots \wedge \psi_n$ is defined by

$$\psi_1 \wedge \cdots \wedge \psi_n = \frac{1}{\sqrt{n!}} \sum_{\pi\in\sigma_n} (-1)^\pi \psi_{\pi(1)} \otimes \cdots \otimes \psi_{\pi(n)}$$

where σ_n is the family of all permutations on $\{1, \dots, n\}$ and $(-1)^\pi$ is the sign of the permutation π. $\Lambda^n\mathcal{H}$ is the Hilbert-span of the $\psi_1 \wedge \cdots \wedge \psi_n$. $\Lambda^0\mathcal{H} = \mathbb{C}$. Now, a simple calculation shows that

$$(\phi_1 \wedge \cdots \wedge \phi_n,\, \psi_1 \wedge \cdots \wedge \psi_n) = \det((\phi_i, \psi_j)_{1\le i,\, j\le n}) \tag{1.10}$$

so that if $\{\phi_i\}$ is an orthonormal basis for $\mathcal{H}$, then $\{\phi_{i_1} \wedge \cdots \wedge \phi_{i_n}\}_{i_1<\cdots<i_n}$ is an orthonormal basis for $\Lambda^n\mathcal{H}$. $A \otimes \cdots \otimes A$ (n times) leaves $\Lambda^n\mathcal{H}$ invariant and is denoted $\Lambda^n(A)$. Note that $\Lambda^n(AB) = \Lambda^n(A)\Lambda^n(B)$.

If $\dim\mathcal{H} = n < \infty$, then $\Lambda^n\mathcal{H} = \mathbb{C}$, and

$$\Lambda^n(A) = \det(A) \tag{1.11}$$

which leads to an easy (and quite standard) proof that $\det(AB) = \det(A)\det(B)$.

This machinery will play a major role in the definition of determinants. It is also the natural framework for the following two results:

THEOREM 1.13 (Horn [**150**]). *For any compact operators A and B and all N,*

$$\prod_{n=1}^N \mu_n(AB) \le \prod_{n=1}^N \mu_n(A)\mu_n(B) \tag{1.12}$$

DEFINITION. Given any compact operator, A, we will denote the eigenvalues by $\lambda_1(A), \lambda_2(A), \dots, \lambda_n(A), \dots$ ordered so that $|\lambda_1| \ge |\lambda_2| \ge \cdots$ and so that each eigenvalue is counted up to its algebraic multiplicity.

THEOREM 1.14 (Weyl [**340**]). *For any compact operator A and all N,*

$$\prod_{n=1}^N |\lambda_n(A)| \le \prod_{n=1}^N \mu_n(A) \tag{1.13}$$

PROOF OF THEOREMS 1.13 AND 1.14. Let C be a positive compact self-adjoint operator. Let $\phi_1, \dots, \phi_n, \dots$ be a complete orthonormal set of eigenvectors for C. Then $\phi_{i_1} \wedge \cdots \wedge \phi_{i_m}$ ($i_1 < \cdots < i_m$) is a complete set of eigenvectors for $\Lambda^m(C)$ so that

$$\|\Lambda^m(C)\| = \prod_{n=1}^m \mu_n(C) \tag{1.14}$$

Since $|\Lambda^m(C)| = \Lambda^m(|C|)$ for any C, (1.14) holds for any C. (1.12) now follows from $\|\Lambda^m(AB)\| \le \|\Lambda^m(A)\|\,\|\Lambda^m(B)\|$. To prove (1.13), one notices that by using a Jordan normal form in $\operatorname{Ran} P_\lambda$ for $\lambda = \lambda_1, \dots, \lambda_N$, we find a set $\{\eta_n\}_{n=1}^N$ of independent vectors so that $A\eta_n = \lambda_n\eta_n + x_n\eta_{n-1}$ with $x_n = 0$ or 1. Clearly,

$$\Lambda^N(A)(\eta_1 \wedge \cdots \wedge \eta_N) = \left(\prod_{n=1}^N \lambda_n\right)\eta_1 \wedge \cdots \wedge \eta_N$$

so $\prod_{n=1}^N \lambda_n$ is an eigenvalue of $\Lambda^N(A)$ (since $\eta_1 \wedge \cdots \wedge \eta_N \neq 0$ on account of the independence of the η's). (1.13) now follows from (1.14) and the trivial fact that the norm must dominate any eigenvalue. □

1.6. Inequalities on Eigenvalues, I

In 1909, Schur [**281**] proved that

$$\sum_{i=1}^N |\lambda_i(A)|^2 \leq \sum_{1 \leq i,\, j \leq N} |a_{ij}|^2$$

for an $N \times N$ matrix. The right side of this inequality is precisely $\sum_{n=1}^N |\mu_n(A)|^2$ which we will eventually denote by $\|A\|_2^2$. Not long thereafter, Lalesco [**186**] proved that $\sum_{i=1}^N |\lambda_i(BC)| \leq \|B\|_2 \|C\|_2$. For any A, we can choose $B = U|A|^{1/2}$, $C = |A|^{1/2}$ and thus obtain

$$\sum_{i=1}^N |\lambda_i(A)|^p \leq \sum |\mu_i(A)|^p \tag{1.15}$$

for $p = 1$. (Lalesco dealt with infinite matrices, but, as trace class and singular values had not been invented, he discussed products of Hilbert-Schmidt operators.) Much later, Weyl made a systematic study of this question in terms of relations of eigenvalues and singular values.

Theorem 1.15 (Weyl [**340**], Horn [**150**]). *Let ϕ be a non-negative monotone increasing function on $[0, \infty)$ so that $t \mapsto \phi(e^t)$ is convex. Then for any compact A,*

$$\sum \phi(|\lambda_n(A)|) \leq \sum \phi(\mu_n(A))$$

In particular, (1.15) *holds. Moreover, for any compact A and B,*

$$\sum \phi(\mu_n(AB)) \leq \sum \phi(\mu_n(A)\mu_n(B))$$

Proof. For the first part, it clearly suffices to prove that

$$\sum_{n=1}^N \phi(|\lambda_n(A)|) \leq \sum_{n=1}^N \phi(\mu_n(A))$$

for each N. This follows from Corollary 1.10 and Theorem 1.14. The second part follows similarly, using Theorem 1.13. □

Remarks. 1. The first inequality is due to Weyl, the second to Horn.

2. If one just wants (1.15), a proof exists which avoids the use of rearrangement-type ideas [**300, 254**].

1.7. Symmetrically Normed Spaces

Let f denote the infinite sequences with only finitely many non-zero elements. Following Schatten [**278**], a norm Φ on f is called *symmetric* if and only if $\Phi(a) = \Phi(a^*)$ for any $a \in f$, that is, if and only if Φ is invariant under permutations and the maps $a_n \to e^{i\theta_n} a_n$. A sequence a_n is said to lie in the *maximal space* s_Φ if and only if $\lim_{n\to\infty} \Phi(a_1, \dots, a_n, 0, 0, \dots) \equiv \Phi(a)$ exists and is finite. It is then easy to see that Φ is a norm on s_Φ. The *minimal space* $s_\Phi^{(0)}$ is defined to be the closure (in s_Φ) of f. If $s_\Phi = s_\Phi^{(0)}$, that is, if f is dense in s_Φ, we call Φ *regular* (called mononormalizing in [**134**]).

EXAMPLE. If $\Phi(a) = (\sum |a_n|^p)^{1/p} = \|a\|_p$, then s_Φ is just the usual ℓ_p space. For $p < \infty$, Φ is regular. The space $\ell_{p,\infty}$ (weak-ℓ_p) is those sequences a with

$$\|a\|_{p,w}^* \equiv \sup_n \, (n^{1/p} a_n^*) < \infty$$

$\|\cdot\|_{p,w}^*$ is not a norm (e.g., if $a = (1, 2^{-1/p}, 0, 0, \dots)$ and $b = (2^{-1/p}, 1, 0, 0, \dots)$, then $\|a\|_{p,w}^* = \|b\|_{p,w}^* = 1$, but $\|a+b\|_{p,w}^* = 1 + 2^{1/p}$). However, for $p \neq 1$, one has the Calderón norm:

$$\|a\|_{p,w} = \sup_n \biggl(n^{-1+\frac{1}{p}} \sum_{j=1}^n a_j^* \biggr)$$

which is a symmetric *norm* (as $\sup\{|S|^{1+\frac{1}{p}} \sum_{j\in S} |a_j| \mid S \text{ a finite subset}\}$) obeying

$$\|a\|_{p,w}^* \le \|a\|_{p,w} \le \frac{p}{p-1} \, \|a\|_{p,w}^*$$

$\ell_{p,w}$ is the maximal space associated to $\|\cdot\|_{p,w}$. The minimal space is just those $a \in \ell_{p,w}$ with $\lim_{n\to\infty} n^{1/p} a_n^* = 0$. Thus the Calderón norms are not regular for any p.

The properties of minimal and maximal spaces are summarized in:

THEOREM 1.16. *Let Φ be a symmetric norm. Then*

(a) *If $a_n \to 0$ as $n \to \infty$, then $\Phi(a) = \Phi(a^*)$.*

(b) *If $a_n, b_n \to 0$ as $n \to \infty$ and if $\sum_{n=1}^N b_n^* \le \sum_{n=1}^N a_n^*$ (in particular if $b_n^* \le a_n^*$), then $\Phi(b) \le \Phi(a)$.*

(c) *If $\Phi(1, 0, 0, \dots) = c$, then*

$$c\|a\|_\infty \le \Phi(a) \le c\|a\|_1$$

for any $a \in s_\Phi$.

(d) *Both s_Φ and $s_\Phi^{(0)}$ are Banach spaces.*

(e) *If α is a doubly substochastic matrix and $a \in s_\Phi$ (resp. in $s_\Phi^{(0)}$), then αa, defined by $(\alpha a)_n = \sum_m \alpha_{nm} a_m$ (the sum being absolutely convergent by* (c)) *is in s_Φ (resp. in $s_\Phi^{(0)}$) and $\Phi(\alpha a) \le \Phi(a)$.*

(f) *If Φ is inequivalent to $\|\cdot\|_\infty$, then s_Φ only consists of sequences a_n with $a_n \to 0$ at ∞.*

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

PROOF. (a),(b) Part (b) for finite sequences follows from Theorem 1.9. If we prove (a), then (b) extends to all sequences. Using (b) for finite sequences, we see that $\Phi(a_1, \dots, a_n, 0, 0, \dots)$ is monotone in n and that $\Phi(a_1, \dots, a_n, 0, 0 \dots) \le \Phi(a_1^*, \dots, a_n^*, 0, 0 \dots)$ so that $\lim \Phi(a_1, \dots, a_n, 0, 0, \dots)$ always exists and $\Phi(a) \le \Phi(a^*)$. Given n, find N so that $a_1^*, \dots, a_n^*$ are among $|a_1|, \dots, |a_N|$. Then, by (b) for finite sequences, $\Phi(a_1^*, \dots, a_n^*, 0, 0, \dots) \le \Phi(a_1, \dots, a_N, 0, 0, \dots)$, proving $\Phi(a^*) \le \Phi(a)$.

(c) Clearly, $\Phi(0, \dots, a_n, 0, \dots) \le \Phi(a) \le \sum_n \Phi(0, \dots, a_n, \dots, 0)$ (use (b) for the first inequality, the triangle inequality for the second).

(d) Since $s_\Phi^{(0)}$ is closed in s_Φ, we need only consider the case of s_Φ. Let $a^{(m)} \in s_\Phi$ be Cauchy. By (c), $a_n^{(m)}$ is Cauchy for each n, so there exists a so that $a_n^{(m)} \to a_n$ for each n. Notice that

$$\begin{aligned}\Phi(a_1 - a_1^{(m)}, \dots, a_N - a_N^{(m)}, 0, \dots) &= \lim_{n\to\infty} \Phi(a_1^{(n)} - a_1^{(m)}, \dots, a_N^{(n)} - a_N^{(m)}, 0, \dots) \\ &\le \lim_{n\to\infty} \Phi(a^{(n)} - a^{(m)})\end{aligned}$$

where the inequality comes from (b). Taking N to ∞, it follows that $a \in s_\Phi$ and $\Phi(a - a^{(m)}) \to 0$.

(e) Let $b = \alpha a$ and let $b^{(N)} = (b_1, \dots, b_N, 0, \dots)$. By the argument in Proposition 1.12, $\sum_{j=1}^k b_j^{(N)} \le \sum_{j=1}^k a_j^*$ whence $\Phi(b^{(N)}) \le \Phi(a^*)$ by (b). All that remains is to show that α takes $s_\Phi^{(0)}$ into $s_\Phi^{(0)}$. It clearly suffices to show that $\alpha a \in s_\Phi^{(0)}$ for $a \in f$. Let $a = (a_1, \dots, a_N, 0, \dots)$. Only the first N columns of α count for αa. Given ε, pick M so that $\sum_{1>M} |\alpha_{ij}| \le \varepsilon$ for $j = 1, \dots, N$. Then $\Phi(0, \dots, 0, (\alpha a)_{M+1}, \dots) \le \varepsilon \Phi(a)$ so $\alpha a \in s_\Phi^{(0)}$.

(f) If $a_n \in s_\Phi$ and $a_n \nrightarrow 0$, then (b) and a simple argument show that $(1, 1, \dots) \in s_\Phi$ whence $\|a\|_\infty \Phi(1, 1, 1, \dots) \ge \Phi(a)$. This inequality and (c) show that Φ is equivalent to $\|\cdot\|_\infty$.

(g) By (c), the identity map from s_Φ to s_Ψ is closed, so Φ is equivalent to Ψ by the closed graph theorem. □

The very simple duality theory for the s_Φ's is due to Schatten [**278**]. For any symmetric norm Φ, we define the conjugate norm Φ' on f by

$$\Phi'(b) = \sup\left\{ \left| \sum a_n b_n \right| \;\middle|\; a \in f,\, \Phi(a) \le 1 \right\}$$

Clearly, using Theorem 1.16(b), the sup is taken by an a which has no more non-zero elements than b. Moreover, using Lemma 1.8, it is easy to see that for b, c fixed with $c = c^*$,

$$\sup\left\{ \left| \sum a_n b_n \right| \;\middle|\; a^* = c \right\} = \sum b_n^* c_n$$

so that Φ' is a symmetric norm.

Theorem 1.17 (Schatten [**278**]). *Fix a symmetric norm Φ. Then,*

(a) $\sum |a_n b_n| \le \Phi(a)\Phi'(b)$

(b) $(s_\Phi^{(0)})^* = s_{\Phi'}$ *(including norm) in the sense that any continuous linear functional on $s_\Phi^{(0)}$ has the form $a \mapsto \sum a_n b_n$ for some $b \in s_{\Phi'}$.*

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

Proof. (a) By Lemma 1.8 and Theorem 1.16(b),

$$\begin{aligned}\sum_1^N |a_n b_n| &\le \sum_1^N a_n^* b_n \\ &\le \Phi(a_1, \dots, a_N, 0, \dots)\Phi'(b_1, \dots, b_N, 0, \dots) \\ &\le \Phi(a)\Phi'(b)\end{aligned}$$

(b) By (a), every b in $s_{\Phi'}$ defines a linear functional. For each N, there is an $a \in f$ so that $\Phi(a) = 1$ and $|\sum a_n b_n| = \Phi'(b_1, \dots, b_N, 0, 0, \dots)$ so that $\Phi'(b)$ is the norm of b as a linear functional. Thus, we need only show that any $\ell \in s_\Phi^{(0)}$ is of the requisite form. Since $(\mathbb{C}^n)^* = \mathbb{C}^n$, we can find a sequence b_n with $\ell(a) = \sum a_n b_n$ for any $a \in f$. Moreover, by definition of Φ', $\Phi'(b_1, \dots, b_n, 0, \dots) = \sup\{\ell(a) \mid a = (a_1, \dots, a_n, 0, \dots),\ \Phi(a) = 1\} \leq \|\ell\|$ so $b \in s_{\Phi'}$. Since b is in $s_{\Phi'}$ and $\ell = b$ on f, $\ell = b$ on $s_\Phi^{(0)}$.

(c) If Φ and Φ' are regular, $(s_\Phi^{(0)})^{**} = s_\Phi^{(0)}$ by (b). If Φ is not regular, $s_\Phi \subset (s_\Phi^{(0)})^{**}$, so $s_\Phi^{(0)}$ is not reflexive. And if Φ' is not regular, $(s_\Phi^{(0)})^{**} = (s_{\Phi'})^*$ is strictly bigger than $s_\Phi = (s_{\Phi'}^{(0)})^*$ since, by the Hahn-Banach theorem, there is an element of $s_{\Phi'}$ vanishing on $s_{\Phi'}^{(0)}$. Finally, if $s_\Psi^{(0)}$ is *not* reflexive, then $s_\Phi = (s_{\Phi'}^{(0)})^*$ cannot be reflexive. □

Note. For additional discussion of sequence spaces, see Köthe [**177**] and Lindenstrauss-Tzafriri [**204**]. For additional discussion of weak-L^p spaces, see Stein-Weiss [**319**].

1.8. Inequalities on Singular Values and Eigenvalues, II

We collect here a few additional inequalities which follow from Theorem 1.9 and Proposition 1.12 and their infinite-dimensional consequence, Theorem 1.16(e).

THEOREM 1.18. *Let A be a compact operator and let $\{\eta_m\}_{m=1}^\infty$ be an orthonormal family. Then*

$$\|A\eta_m\|^2 = \sum_n \alpha_{mn} \mu_n(A)^2 \tag{1.16}$$

for a dss matrix α. Thus, for any symmetric norm, Φ, with $\mu_n(A)^2 \in s_\Phi$, $\Phi(\|A\eta_m\|^2) \leq \Phi(\mu_n(A)^2)$ and, in particular, for any $p \geq 2$,

$$\sum_{m=1}^\infty \|A\eta_m\|^p \leq \sum_{m=1}^\infty (\mu_n(A))^p \tag{1.17}$$

PROOF. By the canonical decomposition (1.2), equation (1.6) holds with

$$\alpha_{mn} = |(\Phi_n, \eta_m)|^2$$

α is dss by Proposition 1.11. The Φ-inequality follows from Theorem 1.16(e) and (1.17) is just the $\frac{1}{2}p$ power of the Φ-inequality for the symmetric norm $\Phi(a) = (\sum |a_n|^{p/2})^{2/p}$. ($p \geq 2$ is necessary for this to be a norm.) □

REMARKS. 1. The related inequality $\Phi(\langle \eta_m, A\eta_m \rangle) \leq \Phi(\mu_n(A))$ is a special case of Proposition 2.6 below.

2. The requirement $p \geq 2$ is critical for (1.17) to hold. Indeed, if η_m is complete, we have, using (1.2), that

$$\begin{aligned}\sum_{n=1}^\infty \mu_n(A) &= \sum_{m=1}^\infty \langle \eta_m, |A|\eta_m \rangle \\ &\leq \sum_{m=1}^\infty \||A|\eta_m\| \\ &= \sum_{m=1}^\infty \|A\eta_m\|\end{aligned}$$

Unless the η_m are all eigenvectors of A, the inequality will be strict. More generally, for $p \le 2$,

$$\sum_{n=1}^{\infty} |\mu_n(A)|^p \le \sum_{m=1}^{\infty} \|A\eta_n\|^p$$

(See [**134**, p. 95 and references therein] for this result and for the history of Theorem 1.18; see also McCarthy [**219**].)

3. Note that for $p = 2$ and η_m, a basis equality holds in (1.17) since, in that case, $\sum_m \alpha_{mn} = 1$.

THEOREM 1.19 (Goh'berg-Krein [**134**]). *Let A be a compact operator and let $\{P_j\}_{j=1}^J$ be some family of mutually orthogonal self-adjoint projections. Let $B = \sum_{j=1}^J P_j A P_j$. Let Φ be a symmetric norm. Then $\mu_n(B) \in s_\Phi$ whenever $\mu_n(A) \in s_\Phi$ and*

$$\Phi(\mu_n(B)) \le \Phi(\mu_n(A)) \tag{1.18}$$

PROOF. By Theorem 1.16(e), it suffices to prove that

$$\mu_m(B) = \sum_{m=1}^{\infty} \alpha_{mn}\mu_n(A) \tag{1.19}$$

for a dss matrix α. Let (1.2) be the canonical decomposition for A and let $B = \sum_{m=1}^{\infty} \mu_m(B)(\gamma_m, \cdot\,)\eta_m$ be the canonical decomposition for B. Then

$$\mu_M(B) = (\eta_m, B\gamma_m) = \sum_{j=1}^{J} (P_j\eta_m, AP_j\gamma_m)$$

so that (1.19) holds with

$$\alpha_{mn} = \sum_{j=1}^{J} (P_j\eta_m, \psi_n)(\phi_n, P_j\gamma_m)$$

This α is dss as in the proof of Proposition 1.11, for example,

$$\begin{aligned}\sum_n |\alpha_{mn}| &\le \Big(\sum_{j,n} |(P_j\eta_m, \psi_n)|^2\Big)^{1/2} \Big(\sum_{j,n} |(\phi_n, P_j\gamma_m)|^2\Big)^{1/2} \\ &\le \Big(\sum_j \|P_j\eta_m\|^2\Big)^{1/2} \Big(\sum_j \|P_j\gamma_m\|^2\Big)^{1/2} \le \|\eta_m\|\,\|\gamma_m\| \qquad \square\end{aligned}$$

REMARK. Notice that $\{\mu_n(B)\}$ is the symmetric rearrangement of $\cup_j\{\mu_n(P_jAP_j)\}$. Thus, a special case of (1.18) is the fact that for P a self-adjoint projection and $Q = 1 - P$,

$$\sum_{n=1}^{\infty} |\mu_n(PAP)|^p + |\mu_n(QAQ)|^p \le \sum_{n=1}^{\infty} |\mu_n(A)|^p \tag{1.20}$$

for any $1 \le p < \infty$.

Part (b) of the following result is due to Lidskii [**193**].

THEOREM 1.20. (a) *For any pair of compact operators, A, B,*

$$\mu_n(A) - \mu_n(B) = \sum_{m=1}^{\infty} \alpha_{nm}\mu_m(A-B) \tag{1.21}$$

for a dss matrix α.

(b) *For any pair of finite-dimensional self-adjoint matrices, A, B, let $\tilde{\lambda}_n(A)$ denote the eigenvalues listed in decreasing order. Then*

$$\tilde{\lambda}_n(A) - \tilde{\lambda}_n(B) = \sum_{m=1}^{N} \beta_{nm}\tilde{\lambda}_m(A-B)$$

for a dss matrix β.

PROOF. We will need some facts from elementary eigenvalue perturbation theory [**164, 254**]. Let $Q(x)$ be a polynomial in x with coefficients which are compact operators, so that for x real and near zero, $Q(x)$ is positive. Then a given non-zero eigenvalue, $\lambda_n(Q(x))$, is either degenerate at $x=0$ in such a way that the degeneracy is removed to some extent for x small, or else $\lambda_n(Q(x))$ is analytic near $x=0$, and in that case,

$$\left.\frac{d\lambda_n(Qx))}{dx}\right|_{x=0} = \left(\phi_n, \frac{dQ}{dx}\,(0)\phi_n\right)$$

where ϕ_n is any normalized vector with $Q(0)\phi_n = \lambda_n(Q(0))\phi_n$. In the finite-dimensional case, the same is true for $\tilde{\lambda}_n$ if $Q(x)$ is only assumed self-adjoint. Now fix C and D and let $Q(x) = (C+xD)^*(C+xD)$. Write $C = U|C|$ and pick ϕ_n so that $|C|\phi_n = \mu_n(C)\phi_n$. Then, unless degeneracy is removed,

$$\begin{aligned}
\left.\frac{d\mu_n(C+xD)}{dx}\right|_{x=0} &= \tfrac{1}{2}\,\mu_n(C)^{-1}\,\frac{d[\mu_n(C+xD)]^2}{dx} \\
&= \tfrac{1}{2}\,\mu_n(C)^{-1}\,\frac{d}{dx}\,\lambda_n(Q(x)) \\
&= \tfrac{1}{2}\,\mu_n(C)^{-1}(\phi_n, (C^*D + D^*C)\phi_n) \\
&= \mathrm{Re}(U\phi_n, D\phi_n) = \sum_{m=1}^{\infty} a_{nm}\mu_m(D)
\end{aligned}$$

where $a_{nm} = \mathrm{Re}[(U\phi_n, \eta_m)(\gamma_m, \phi_n)]$ if $D = \sum \mu_m(D)(\gamma_m, \cdot\,)\eta_m$ is the canonical decomposition for D. As in Proposition 1.11, A is a dss matrix. Next, let $A(x) = B + x(A-B)$. Then $\mu_n(A(x))$ is continuous for all x and differentiable for all n except for a countable set. For x not in this exceptional set, we have that

$$\frac{d\mu_n(A(x))}{dx} = \sum_{m=1}^{\infty} a_{nm}(x)\mu_m(A-B)$$

by the above calculation. Thus (1.21) holds with $\alpha_{nm} = \int_0^1 a_{nm}(x)\,dx$. Since each a is dss, so is α.

The proof of (b) is similar, except that we use

$$\frac{d\tilde{\lambda}_n(A(x))}{dx} = (\phi_n(x), (A-B)\phi_n(x)) = \sum b_{nm}(x)\tilde{\lambda}_m(A-B)$$

where $b_{nm} = |(\phi_n(x), \eta_m)|^2$ and

$$A - B = \sum_{m=1}^{\infty} \tilde{\lambda}_m (A-B)(\eta_m, \,\cdot\,)\eta_m$$

is the eigenexpansion for $A - B$. □

Remark. We have not been explicit about the usual kind of consequences of these dss relations; we note, however, the inequality,

$$\sum_{n=1}^{\infty} |\mu_n(A) - \mu_n(B)|^p \leq \sum_{n=1}^{\infty} |\mu_n(A-B)|^p \tag{1.22}$$

1.9. Clarkson-McCarthy Inequalities

We want to note two remarkable inequalities expressed in terms of $\|A\|_p = (\sum |\mu_n(A)|^p)^{1/p}$:

Theorem 1.21. (a) *For $2 \leq p < \infty$,*

$$\|A+B\|_p^p + \|A-B\|_p^p \leq 2^{p-1}(\|A\|_p^p + \|B\|_p^p) \tag{1.23}$$

(b) *For $1 < p \leq 2$ and $p' = p/(p-1)$,*

$$\|A+B\|_p^{p'} + \|A-B\|_p^{p'} \leq 2(\|A\|_p^p + \|B\|_p^p)^{p'/p} \tag{1.24}$$

These inequalities, which we call *Clarkson-McCarthy inequalities*, will only be used once again in these notes, namely, to mildly improve a convergence theorem in Chapter 2. They are non-commutative analogs, due to McCarthy [**219**], of some celebrated estimates of Clarkson [**74**]. We will only prove (1.23), which is the simpler of the two. We use some technical results from Chapters 2 and 3. Our proof of (1.23) depends on the following result of McCarthy [**219**] of independent interest:

Theorem 1.22. *Fix $1 \leq p$. Let A and B be two non-negative self-adjoint operators with finite $\|\cdot\|_p$. Then*

$$2^{1-p}\|A+B\|_p^p \leq \|A\|_p^p + \|B\|_p^p \leq \|A+B\|_p^p \tag{1.25}$$

Proof. We begin with two preliminaries. Let X be a positive self-adjoint operator and let $\|\phi\| \leq 1$. Then for $p \geq 1$,

$$|(\phi, X\phi)|^p \leq (\phi, X^p\phi) \tag{1.26}$$

Since the spectral theorem asserts that $(\phi, X^a\phi) = \int_0^\infty x^a \, d\mu(x)$ for a suitable measure $d\mu$, (1.26) follows from Hölder's inequality

$$\left(\int_0^\infty x \, d\mu(x)\right) \leq \left(\int_0^\infty x^p \, d\mu(x)\right)^{1/p} \left(\int_0^\infty d\mu(x)\right)^{(p-1)/p}$$

For positive numbers a and b,

$$(a+b)^p \leq 2^{p-1}(a^p + b^p) \tag{1.27}$$

(1.27) is equivalent to $(\frac{1}{2}(a+b))^p \leq \frac{1}{2}a^p + \frac{1}{2}b^p$, which just expresses the fact that x^p is a convex function.

To prove the leftmost inequality in (1.25), pick η_n a unit vector with $(A + B)\eta_n = \mu_n(A+B)\eta_n$. Then

$$2^{1-p}\mu_n(A+B)^p = 2^{1-p}[(\eta_n, A\eta_n) + (\eta_n, B\eta_n)]^p$$

$$\begin{aligned}&\le (\eta_n, A\eta_n)^p + (\eta_n, B\eta_n)^p \qquad &\text{(by (1.27))}\\ &\le (\eta_n, A^p\eta_n) + (\eta_n, B^p\eta_n) \qquad &\text{(by (1.26))}\end{aligned}$$

Summing over n and using $\sum_n(\eta_n, C^p\eta_n) = \text{Tr}(C^p) = \|C\|_p^p$, we have the first inequality in (1.25).

For the second, we may as well suppose that $A > 0$, since the inequality then holds for general A by limiting arguments (using result of Chapter 2). Let $C = A^{1/2}(A+B)^{-1/2}$, $D = B^{1/2}(A-B)^{-1/2}$. Then $C^*C + D^*D = 1$ so, in particular, $\|C\| \le 1$. Moreover, $A = C(A+B)C^*$. Now pick ϕ_n orthonormal with $A\phi_n = \mu_n(A)\phi_n$. Then, since $\|C^*\phi_n\| \le 1$, we have, by (1.26), that

$$\begin{aligned}\mu_n(A)^p &= (\phi_n, A\phi_n)^p\\ &= (C^*\phi_n, (A+B)C^*\phi_n)^p\\ &\le (C^*\phi_n, (A+B)^pC^*\phi_n)\end{aligned}$$

Therefore, using the cyclicity of the trace,

$$\begin{aligned}\|A\|_p^P + \|B\|_p^p &\le \text{Tr}(C(A+B)^pC^* + D(A+B)^pD^*)\\ &= \text{Tr}((C^*C + D^*D)(A+B)^p) = \|A+B\|_p^p \qquad \square\end{aligned}$$

Warning. (1.25) is trivial in the commutative case, but there are trivial results in the commutative case which do not extend. For example, if f and g are real-valued functions so that $\|f + ig\|_1 = \|f\|_1$, then $\|g\|_1 = 0$. But if

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$$

then $\|A + iB\|_1 = \|A\|_1$.

Proof of (1.23). We have that, since $p/2 \ge 1$,

$$\begin{aligned}&2^{p-1}(\|A\|_p^p + \|B\|_p^p)\\ &\quad= 2^{p/2}[2^{(p/2)-1}(\|A^*A\|_{p/2}^{p/2} + \|B^*B\|_{p/2}^{p/2})]\\ &\quad\ge 2^{p/2}\|A^*A + B^*B)\|_{p/2}^{p/2} \qquad &\text{(by (1.25))}\\ &\quad= \|(A+B)^*(A+B) + (A-B)^*(A-B)\|_{p/2}^{p/2}\\ &\quad\ge \|(A+B)^*(A+B)\|_{p/2}^{p/2} + \|(A-B)^*(A-B)\|_{p/2}^{p/2} \qquad &\text{(by (1.25))}\\ &\quad= \|A+B\|_p^p + \|A-B\|_p^p \quad \square\end{aligned}$$

M. Klaus [**170**] has remarked that there is a simple proof of the Clarkson-McCarthy inequalities which results from mimicking the elegant proof that Boas [**49**] gave of Clarkson's original L^p inequalities. Namely, let $\mathcal{I}_p^q$ consist of pairs of operators A, B in $\mathcal{I}_p$ with norm

$$\|(A,B)\|_{p;q} = [\|A\|_p^q + \|B\|_p^q]^{1/q}$$

Then (1.23) asserts that the map

$$T : (A, B) \to (A + B, A - B)$$

is bounded from $\mathcal{I}_p^p$ to $\mathcal{I}_p^p$ with norm $2^{1-p^{-1}}$ for $2 \le p \le \infty$, and (1.24) asserts that T is bounded from $\mathcal{I}_p^{p'}$ to $\mathcal{I}_p^{p'}$ with norm $2^{1/p'}$ for $1 \le p \le 2$. The cases $p = 1, 2, \infty$

of these inequalities are easy. The full inequalities then follow from the Calderón-Lions interpolation theory [**65, 205, 206, 251**] and the calculation that in their framework, $[\mathcal{I}_{p_0}^{q_0}, \mathcal{I}_{p_1}^{q_1}]_t = \mathcal{I}_{p_t}^{q_t}$ with $p_t^{-1} = tp_1^{-1} + (1-t)p_0^{-1}$ and $q_t^{-1} = tq_1^{-1} + (1-t)q_0^{-1}$.

In Chapter 2, we will define $\mathcal{I}_p$ to be those compact operators A with $\|A\|_p < \infty$. We will see that $\mathcal{I}_p$ is a Banach space with $\|\cdot\|_p$ as norm. Recall:

Definition. A Banach space X is called *uniformly convex* if and only if for all $0 < \varepsilon < 2$, the *modulus of convexity*,

$$\delta(\varepsilon) \equiv \inf\left\{1 - \left\|\frac{x+y}{2}\right\| \;\middle|\; \|x\| = 1,\, \|y\| = 1,\, \|x-y\| \geq \varepsilon\right\}$$

is strictly positive.

If one notes that equality holds in (1.23)/(1.24) when A and B are self-adjoint operators with $\ker A = (\ker B)^\perp$, we see that, by these inequalities,

$$\begin{aligned} \delta(\varepsilon) &= 1 - \left[1 - \left(\frac{1}{2}\,\varepsilon\right)^p\right]^{1/p} && \text{for } \mathcal{I}_p,\ p \geq 2 \\ &= 1 - \left[1 - \left(\frac{1}{2}\,\varepsilon\right)^{p'}\right]^{1/p'} && \text{for } \mathcal{I}_p,\ p \geq 2;\ p' = \frac{p}{(p-1)} \end{aligned}$$

In particular,

Theorem 1.23 (Dixmier [**96**] for $p \geq 2$; McCarthy [**219**] for general p). *For $1 < p < \infty$, $\mathcal{I}_p$ is uniformly convex.*

Uniform convexity in general spaces is important for two reasons: First, uniformly convex spaces are reflexive — since one can compute $\mathcal{I}_p^*$ fairly easily (see Chapter 2), this is of minor interest here. Second, we have the following standard result:

Theorem 1.24. *Let X be a uniformly convex Banach space and suppose that x_n converges to x weakly, that is, $\ell(x_n) \to \ell(x)$ for all $\ell \in X^*$ and that $\|x_n\| \to \|x\|$. Then $\|x_n - x\| \to 0$.*

Proof. If $\|x\| = 0$, we are done. If not, we can replace x_n by $x_n/\|x_n\|$ so we are reduced to the case $\|x_n\| = \|x\| = 1$. By the uniform convexity, it suffices to prove that $\|(x_n+x)/2\| \to 1$. Clearly, $\|(x_n+x)/2\| \leq 1$. Now, by the Hahn-Banach theorem, we can find $\ell \in X^*$ so that $\|\ell\| = \ell(x) = 1$. Then

$$\left\|\frac{(x_n+x)}{2}\right\| \geq \frac{1}{2}\,|\ell(x_n+x)| \to 1 \qquad \square$$

We discuss the consequences of this theorem for $\mathcal{I}_p$ in the next chapter. See also Addendum H.

CHAPTER 2

Calkin's Theory of Operator Ideals and Symmetrically Normed Ideals; Convergence Theorems for $\mathcal{I}_P$

In this chapter, we will begin by classifying all two-sided ideals of operators (not necessarily norm closed) in $\mathcal{L}(\mathcal{H})$ following Calkin [**66**]. We will set up a one-one correspondence between ideals and certain sequence spaces. Among these sequence spaces are the symmetrically normed spaces s_Φ and $s_\Phi^{(0)}$ discussed in Chapter 1; the corresponding ideals will be denoted $\mathcal{I}_\Phi$ and $\mathcal{I}_\Phi^{(0)}$. Taking $\Phi = \ell_p$ or $\ell_{p,w}$, we obtain the trace ideals $\mathcal{I}_p$ and $\mathcal{I}_{p,w}$ which play a major role in later developments. We then discuss criteria for integral operators to lie in the *trace class*, $\mathcal{I}_1$, and the Hilbert-Schmidt class, $\mathcal{I}_2$. Finally, we discuss some convergence theorems for $\mathcal{I}_p$.

We begin our presentation of Calkin's theory with two elementary propositions:

Proposition 2.1. *Let $\mathcal{I}$ be a two-sided ideal in $\mathcal{L}(\mathcal{H})$ containing an operator A which is not compact. Then $\mathcal{I} = \mathcal{L}(\mathcal{H})$.*

Remark. We emphasize our standing convention that $\mathcal{H}$ is separable since it is essential for the truth of this proposition.

Proof. Since $|A| = U^*A$, $\mathcal{I}$ contains a positive self-adjoint operator which is not compact. Let $P_a = P_{[a,\infty)}(|A|)$ be the spectral projection for $|A|$. If each P_a ($a > 0$) is finite-dimensional, then $|A| = \text{norm-lim}_{a\downarrow 0}|A|P_a$ is compact, so some P_a is infinite-dimensional. Since $|A|^{-1}P_a$ is bounded for $a > 0$, each P_a is in $\mathcal{I}$. If P_a is infinite-dimensional, there exists an isometry V from $\mathcal{H}$ onto $\operatorname{Ran} P_a$. Then $V^*P_aV = 1 \in \mathcal{I}$, so $\mathcal{I} = \mathcal{L}(\mathcal{H})$. □

Henceforth, we suppose $\mathcal{I} \neq \mathcal{L}(\mathcal{H})$ so, by Proposition 2.1, it must consist of compact operators.

Proposition 2.2. *Let $\mathcal{I}$ be a two-sided proper ideal in $\mathcal{L}(\mathcal{H})$. Let A and B be compact operators with $\mu_n(B) \leq \mu_n(A)$ for all n. If $A \in \mathcal{I}$, then $B \in \mathcal{I}$.*

Proof. Let $A = \sum \mu_n(A)(\phi_n, \cdot)\psi_n$, $B = \sum \mu_n(B)(\eta_n, \cdot)\gamma_n$ be the canonical decompositions of A and B. Since ϕ, ψ, η, γ are orthonormal and $\mu_n(B) \leq \mu_n(A)$, there exists a partial isometry D with $D^*\phi_n = \eta_n$ and a contraction C with $C\psi_n = \mu_n(B)\mu_n(A)^{-1}\gamma_n$. Since $B = CAD$, the proposition is proven. □

This proposition immediately implies:

Corollary 2.3. *If $\mathcal{I} \neq 0$, then every finite rank operator lies in $\mathcal{I}$. In particular, the only norm closed, two-sided ideals are $\{0\}$, $\mathcal{I}_\infty$, and $\mathcal{L}(\mathcal{H})$.*

Since $\mu_n(A) = \mu_n(A^*)$, Proposition 2.2 implies that

COROLLARY 2.4. *Every two-sided ideal in* $\mathcal{L}(\mathcal{H})$ *is left invariant by the adjoint operation.*

The most important consequence of Proposition 2.2 is that ideals are completely described by a set of sequences. Rather than associate $\mathcal{I}$ just with its family of possible singular values, we take a definition allowing enough sequences to form a vector space:

DEFINITION. Fix an orthonormal set ϕ_n in $\mathcal{H}$. Given an ideal, $\mathcal{I} \neq \mathcal{L}(\mathcal{H})$, let $\mathcal{S}(\mathcal{I}) = \{a = (a_1, a_2, \dots) \mid \sum a_n(\phi_n, \cdot)\phi_n \in \mathcal{I}\}$.

Thus $a \in \mathcal{S}(\mathcal{I})$ if and only if a^* is a possible set of singular values for an operator in $\mathcal{I}$.

DEFINITION. Given a sequence space s, $I(s)$ is the family of compact operators, A, with $(\mu_1(A), \dots) \in s$.

DEFINITION. A *Calkin space* is a vector space, s, of sequences a_n with $a_n \to 0$ at infinity, with the *Calkin property*: $a \in s$ and $b_n^* \leq a_n^*$ implies $b \in s$.

The following sets up a one-one correspondence between Calkin spaces and two-sided ideals:

THEOREM 2.5 (Calkin [**66**]). *If* s *is a Calkin space, then* $I(s)$ *is a two-sided ideal of operators and* $\mathcal{S}(I(s)) = s$. *If* $\mathcal{I}$ *is a two-sided ideal, then* $\mathcal{S}(\mathcal{I})$ *is a Calkin space and* $I(\mathcal{S}(\mathcal{I})) = \mathcal{I}$.

PROOF. If $\mathcal{I}$ is a two-sided ideal, $\mathcal{S}(\mathcal{I})$ is a vector space since $\mathcal{I}$ is a vector space and it has the Calkin property by Proposition 2.2. Moreover, by construction, $I(\mathcal{S}(\mathcal{I})) = \mathcal{I}$.

Now let s be a Calkin space. If we can show that $I(s)$ is an ideal, then $\mathcal{S}(I(s)) = s$ is evident. By the inequalities (1.4), if $A \in I(s)$ and $B \in \mathcal{L}(\mathcal{H})$, then AB and $BA \in I(s)$ by the Calkin property. All that remains is to show that if $A, B \in I(s)$, then so is $A + B$. By Fan's inequalities, (1.5), we have that

$$\begin{aligned} \mu_{2n+1}(A+B) &\leq \mu_{n+1}(A) + \mu_{n+1}(B) \\ \mu_{2n}(A+B) &\leq \mu_{2n-1}(A+B) \leq \mu_n(A) + \mu_n(B) \end{aligned}$$

As a result, $\mu_n(A+B) \leq (\alpha_n + \beta_n)$ where

$$\begin{aligned} \alpha &= (\mu_1(A), \mu_1(A), \mu_2(A), \mu_2(A), \dots) \\ \beta &= (\mu_1(B), \mu_1(B), \mu_2(B), \mu_2(B), \dots) \end{aligned}$$

so it suffices that $\alpha, \beta \in s$, by the Calkin property. But by the Calkin property again, $(\mu_1(A), 0, \mu_2(A), 0, \dots)$ and $(0, \mu_1(A), 0, \mu_2(A), \dots)$ are in s. □

Calkin's theorem clearly sets up some relation between symmetric norms on sequences and ideals since every s_Φ (Φ is not equivalent to ℓ_∞) is a Calkin space, as is every $s_\Phi^{(0)}$. We let $\mathcal{I}_\Phi \equiv I(s_\Phi)$ and $\mathcal{I}_\Phi^{(0)} \equiv I(s_\Phi^{(0)})$. (*Exception and warning*: By this criterion, we should denote the compacts by $\mathcal{I}_\infty^{(0)}$, but we will use $\mathcal{I}_\infty$.) For $A \in \mathcal{I}_\Phi$, we let

$$\Phi(A) = \Phi(\mu_1(A), \mu_2(A), \dots) \tag{2.1}$$

PROPOSITION 2.6 (Simon [**296**]). *Let $\mathcal{B}$ denote the set of pairs of orthonormal sets $\{\phi_n\}, \{\psi_n\}$. If $A \in \mathcal{I}_\Phi$, then*

$$\Phi(A) = \sup_{\phi,\psi\in\mathcal{B}} \Phi((\phi_n, A\psi_n)) \tag{2.2}$$

Conversely, if A is compact and the right side of (2.2) is finite, then $A \in \mathcal{I}_\Phi$. If Φ is inequivalent to ℓ_∞, then "A is compact" may be replaced by "A is bounded" in the last statement.

PROOF. Let A be compact and let $A = \sum \mu_n(A)(\eta_n, \cdot\,)\gamma_n$ be its canonical expansion. Choosing $\psi_n = \eta_n$ and $\phi_n = \gamma_n$, we see that $(\phi_n, A\psi_n) = \mu_n(A)$, so if the right side of (2.2) is finite, then $A \in \mathcal{I}_\Phi$ and $\Phi(A) \leq \sup \ldots$. On the other hand, for any $\phi, \psi \in \mathcal{B}$, we have that $(\phi_n, A\psi_n) = \sum a_{nm}\mu_m(A)$ where $a_{nm} \equiv (\phi_n, \gamma_m)(\eta_m, \psi_n)$ is dss by Proposition 1.11. Thus, by Theorem 1.16(e), $\Phi((\phi_n, A\psi_n)) \leq \Phi(A)$, proving (2.2).

Finally, suppose that Φ is inequivalent to ℓ_∞, that $A \in \mathcal{L}(\mathcal{H})$, and that the right side of (2.2) is finite. Thus, by Theorem 1.16(f), $(\phi_n, A\psi_n) \to 0$ as $n \to \infty$ for all $\phi, \psi \in \mathcal{B}$. It follows that A is compact (completing the proof) by the following argument: Let ψ_n be an orthonormal set in $(\ker(A))^\perp$ and let $U^*\psi_n = \phi_n$ where $A = U|A|$. Then $\||A|^{1/2}\psi_n\| \to 0$. Choose ψ_n inductively by requiring

$$\||A|^{1/2}\psi_n\| \geq \tfrac{1}{2} \sup_{\substack{\phi\in[\psi_1,\ldots,\psi_{n-1}]\\ \|\phi\|=1}} \||A|^{1/2}\phi\| \equiv \tfrac{1}{2}\,\beta_n$$

and we conclude that $\beta_n \to 0$. But then

$$\left\| |A|^{1/2} - \sum_{n=1}^{N} (\psi_n, \cdot\,)|A|^{1/2}\psi_n \right\| \to 0$$

as $N \to \infty$. □

The properties of $\mathcal{I}_\Phi$ and $\Phi(\cdot)$ on $\mathcal{I}_\Phi$ are summarized in:

THEOREM 2.7. (a) *$\Phi(\cdot)$ is a norm on $\mathcal{I}_\Phi$ obeying*

$$\Phi(ABC) \leq \|A\|\,\|C\|\Phi(B) \tag{2.3}$$

$$\Phi(B) \geq \|B\|\Phi(1, 0, \ldots) \tag{2.4}$$

for all $B \in \mathcal{I}_\Phi$, $A, C \in \mathcal{L}(\mathcal{H})$.

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

(c) *Any norm on an operator ideal $\mathcal{I}$ obeying (2.3) agrees with a symmetric norm, Φ, on the finite rank operators. $\mathcal{I} \subset \mathcal{I}_{\tilde{\Phi}}$. If $\mathcal{I}$ is Banach in its norm, then $\mathcal{I}_{\tilde{\Phi}}^{(0)} \subset \mathcal{I}$.*

(d) (Non-commutative Fatou lemma) *If $A_m \in \mathcal{I}_\Phi$, $A_m \to A$ weakly, A is compact, and $\sup_n \Phi(A_n) < \infty$, then $A \in \mathcal{I}_\Phi$ and $\Phi(A) \leq \sup_n \Phi(A_n)$. If Φ is inequivalent to ℓ_∞, A need not be assumed to be compact a priori.*

REMARK. (a)–(c) are essentially due to Schatten [**278**] and (d) to Goh'berg-Krein [**134**].

PROOF. (a) Since Φ is a norm on sequences,

$$\Phi((\phi_n, (A+B)\psi_n) \leq \Phi((\phi_n, A\psi_n)) + \Phi((\phi_n, B\psi_n))$$

so (2.2) implies that Φ is a norm. (2.3) follows from (1.4) and Theorem 1.16(b), and (2.4) from Theorem 1.16(c).

(b) Let A_m be Cauchy in Φ-norm. By (2.4), A_m is Cauchy in operator norm so $\|A_m - A\| \to 0$ for some compact A. Moreover,

$$\begin{aligned}\Phi((\phi_n, (A_m - A)\psi_n)_{n\leq N}) &\leq \lim_{k\to\infty} \Phi((\phi_n, (A_m - A_k)\psi_n)_{n\leq N}) \\ &\leq \lim_{k\to\infty} \Phi((\phi_n, (A_m - A_k)\psi_n)) \qquad \text{(by Theorem 1.16(b))} \\ &\leq \lim_{k\to\infty} \Phi(A_m - A_k) \qquad \text{(by (2.2))}\end{aligned}$$

Thus by (2.2), $\Phi(A_m - A) \leq \lim_{k\to\infty} \Phi(A_m - A_k)$, so $A_m \to A$ in Φ-norm. This proves that $\mathfrak{I}_\Phi$ is a Banach space.

Let A be finite rank. Then, by Theorem 1.16(e), $(\phi_n, A\psi_n) \in s_\Phi^{(0)}$ for all $\phi, \psi \in \mathcal{B}$. It follows from (2.2) that if A is in the Φ-closure of the finite rank operators, then $(\phi_n, A\psi_n) \in s_\Phi^{(0)}$ for any $\phi, \psi \in \mathcal{B}$. In particular, using a canonical decomposition, $\{\mu_n(A)\} \in s_\Phi^{(0)}$. Conversely, if $\{\mu_n(A)\} \in s_\Phi^{(0)}$, the canonical decomposition will converge in Φ-norm, so A is in the norm closure of the finite rank operators.

(c) In particular, (2.3) implies that the norm is unitary invariant and thus only dependent on the singular values of A. Thus, there is a symmetric norm $\tilde{\Phi}$ on sequences with $\Phi(A) = \tilde{\Phi}(\mu_1(A), \dots, \mu_N(A), 0, \dots)$ for finite rank A's. By (2.3), $\Phi(A) \geq \Phi(\sum_{n=1}^N \mu_n(A)(\phi_n, \cdot)\psi_n) = \tilde{\Phi}(\mu_1, \dots, \mu_N, 0, \dots)$, so $\Phi(A) \geq \tilde{\Phi}(A)$, and any A in $\mathfrak{I}$ is in $\mathfrak{I}_{\tilde{\Phi}}$.

(d) For any $\phi, \psi \in \mathcal{B}$,

$$\begin{aligned}\Phi((\phi_n, A\psi_n)_{n\leq N}) &= \lim_{m\to\infty} \Phi((\phi_n, A_m\psi_n)_{n\leq N}) \\ &\leq \limsup_{m\to\infty} \Phi((\phi_n, A_m\psi_n)) \qquad \text{(by Theorem 1.16(b))} \\ &\leq \sup_m \Phi(A_m)\end{aligned}$$

The result now follows from the end of Proposition 2.6. □

REMARKS. 1. In the context of (c), it can happen that Φ and $\tilde{\Phi}$ are not equal; for example, take $\Phi(A) = \|\mu_n\|_{p,w} + \limsup_{n\to\infty} n^{1/p}\mu_n(A)$, but as in Theorem 1.16(g), if $\mathfrak{I}$ has two Banach spaces norm obeying (2.3), they must be equivalent.

2. The triangle inequality for Φ and $\mathfrak{I}_\Phi$ has two other rather different proofs which are instructive. One [**278**] first establishes the formula (Φ' = conjugate norm to Φ),

$$\Phi(A) = \sup\{|\mathrm{Tr}(AB)| \mid \Phi'(B) = 1\}$$

from which $\Phi(A + C) \leq \Phi(A) + \Phi(C)$ is easy. The other [**134**] first establishes that $\|A + B\|_1 \leq \|A\|_1 + \|B\|_1$ for finite rank operators. Given this and $A + B = \sum_n \mu_n(A+B)(\phi_n, \cdot)\psi_n$, let P_N be the projection onto the space spanned by $\psi_1, \dots, \psi_N$. Then $\|P_N(A+B)\|_1 = \sum_{n=1}^N \mu_n(A+B)$ while $\mu_n(P_N A) \leq \mu_n(A)$ for $n = 1, \dots, N$ and $\mu_n(P_N A) = 0$ for $n > N$. Thus

$$\sum_{n=1}^N \mu_n(A+B) \leq \sum_{n=1}^N [\mu_n(A) + \mu_n(B)]$$

from which $\Phi(A+B) \leq \Phi(A) + \Phi(B)$ by Theorem 1.16(b).

Corresponding to the spaces ℓ_p and $\ell_{p,w}$, we introduce the norm ideals $\mathcal{I}_p$ and $\mathcal{I}_{p,w}$. Before beginning their explicit study, we discuss two additional general properties of the $\mathcal{I}_\Phi$: Hölder's inequality and interpolation. In Chapter 3, we will discuss the duality theory for the $\mathcal{I}_\Phi$.

Hölder's inequality for the operator norms $\|\cdot\|_p$ has been around for many years (see, e.g., [**134**]) — its simplest proof is probably by complex interpolation (see [**251**, appendix to Section IX.4]). Using other interpolation methods, Simon [**296**] proved a Hölder inequality for the $\|\cdot\|_{p,w}$ operator norms. By mimicking the Goh'berg-Krein proof [**134**] of the $\|\cdot\|_p$ Hölder inequality, we are able to obtain a general result which includes the $\|\cdot\|_{p,w}$ result:

Theorem 2.8 (Abstract Hölder Inequality). *Let Φ_1, Φ_2, Φ_3 be three symmetric sequence norms with the property that if $a \in s_2$, $b \in s_3$, then ab, given by $(ab)_n = a_n b_n$, is in s_1 and*

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

Then, if $A \in \mathcal{I}_{\Phi_2}$, $B \in \mathcal{I}_{\Phi_3}$, it follows that $AB \in \mathcal{I}_{\Phi_1}$ and

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

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

$$p^{-1} = q^{-1} + r^{-1} \tag{2.5a}$$

and $A \in \mathcal{I}_q$, $B \in \mathcal{I}_r$, then $AB \in \mathcal{I}_p$ and

$$\|AB\|_p \leq \|A\|_q \|B\|_r \tag{2.5b}$$

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

$$\|AB\|_{p,w} \leq \frac{p}{(p-1)} \|A\|_{q,w} \|B_{p,w} \tag{2.5c}$$

Proof. By Horn's inequality, (1.12), and Corollary 1.10,

$$\sum_{n=1}^N \mu_n(AB) \leq \sum_{n=1}^N \mu_n(A)\mu_n(B)$$

Then, using Theorem 1.9,

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

(2.5b) then follows from the usual Hölder inequality and (2.5c) from the obvious inequalities

$$\left[\frac{(p-1)}{p}\right] \|ab\|_{p,w} \leq \|a\|^*_{p,w} \|b\|^*_{r,w} \leq \|a\|_{q,w} \|b\|_{r,w}$$ □

Conjecture. Let $\|A\|^*_{p,w} \equiv \|\mu_n(A)\|^*_{p,w} = \sup_n [n^{1/p}\mu_n(A)]$. Then

$$\|AB\|^*_{p,w} \leq \|A\|^*_{q,w} \|B\|^*_{r,w}$$

The history of interpolation ideas in the context of operator ideals is somewhat murky. Kunze [**180**] introduced the complex method, including Riesz-Thorin and Stein theorems for $\mathcal{I}_p$ spaces, in the context of general non-commutative integration. The abstract complex interpolation theory of Calderón [**65**] and Lions [**205, 206**] is based on "wandering vector" theorems so that, given Kunze's work, such theorems were implicit in Calderón-Lions. Apparently unaware of this earlier work, Goh'berg and Krein [**134**] proved such theorems for $\mathcal{I}_p$ and a class of spaces including $\mathcal{I}_{p,w}$. Marcinkiewicz-type theorems were first discussed by Bennett [**29**]. The two abstract theorems below are essentially due to Simon [**296**].

Let Φ_0 and Φ_1 be two symmetric forms on finite sequences. Let Φ_t be the complex interpolated norms; we assume the reader is familiar with this notion (see, e.g., [**251**, appendix to Section IX.4]). It is not hard to see that each Φ_t is also a symmetric norm. We use $\mathcal{I}_{(t)}$ to stand for $\mathcal{I}_{\Phi_t}$.

THEOREM 2.9. *Let $\mathcal{I}_{(t)}$ be given as above. Let $z \mapsto A(z)$ be a map from the strip $0 \le \operatorname{Re} z \le 1$ to the bounded operators so that for any $\phi, \psi \in \mathcal{H}$, $(\phi, A(z)\psi)$ is continuous in the strip, analytic in its interior, and bounded. If $A(iy)$ (resp. $A(1+iy)$) lie in $\mathcal{I}_{(0)}$ (resp. $\mathcal{I}_{(1)}$) and*

$$M_t = \sup_y \Phi_t(A(t+iy)) < \infty$$

for $t = 0, 1$, then $A(z) \in \mathcal{I}_{(\operatorname{Re} z)}$ for all z in the strip and M_t is convex.

PROOF. As usual, by replacing $A(z)$ by $e^{a+bz}A(z)$, we can suppose that $M_0 = M_1 = 1$. Let $\phi, \psi \in \mathcal{B}$. Then $(\phi, A(z)\psi)_{n\le N}$ is an analytic f-valued function, so by definition of Φ_t,

$$\begin{aligned}\Phi_t((\phi, A(t+iy)\psi)_{n\le N}) &\le \sup_{\substack{y \text{ real} \\ x=0 \text{ or } 1}} (\Phi_x((\phi, A(x+iy)\psi)_{n\le N}) \\ &\le \sup_{\substack{y \text{ real} \\ x=0 \text{ or } 1}} (\Phi_x(A(x+iy)) \le 1\end{aligned}$$

by Proposition 2.6. Taking $N \to \infty$ and using Proposition 2.6 again, the theorem results if Φ_t is inequivalent to $\|\cdot\|_\infty$. The result for the case $\Phi_0 = \Phi_1 = \|\cdot\|_\infty$ requires an additional argument, which can be found in [**254**, p. 115]. □

THEOREM 2.10. *Let Φ_1, Φ_2, Φ_3 be symmetric norms and X_1, X_2, X_3 Banach spaces with $X_1 \cap X_2$ dense in X_3.*
(a) *If every linear contraction from X_1 to s_{Φ_1} and X_2 to s_{Φ_2} is automatically a contraction from X_3 to s_{Φ_3}, then the same is true if s_Φ is replaced by $\mathcal{I}_\Phi$.*
(b) *If every linear contraction from s_{Φ_1} to X_1 and x_{Φ_2} to X_2 is automatically a contraction from s_{Φ_3} to X_3, then the same is true if s_Φ is replaced by $\mathcal{I}_\Phi$.*

PROOF. (a) Let T be a contraction from X_i to $\mathcal{I}_{\Phi_i}$ $(i = 1, 2)$ and let $\phi, \psi \in \mathcal{B}$. Define $T_{\phi,\psi}(x) = \{\phi_n, T(x)\psi_n)\}$. Then, by Proposition 2.6, $T_{\phi,\psi}$ is a contraction from X_i to s_{Φ_1} $(i = 1, 2)$ and so, by hypothesis, from X_3 to s_{Φ_3}. Using Proposition 2.6 again, T is a contraction from X_3 to $\mathcal{I}_{\Phi_3}$.

(b) Fix $A = \sum \mu_n(A)(\phi_n, \cdot\,)\psi_n$ in $\mathcal{I}_\infty$. Given a contraction from $\mathcal{I}_{\Phi_i}$ $(i = 1, 2)$ to X_i, let

$$S_{\phi,\psi}(\{a_n\}) = T\left(\sum a_n(\phi_n, \cdot\,)\psi_n\right)$$

Then $S_{\phi,\psi}$ is a contraction from s_{Φ_i} to X_i $(i = 1, 2)$ and so from s_{Φ_3} to X_3. In particular,

$$\|T(A)\|_{X_3} = \|S_{\phi,\psi}(\mu_n(A))\|_{X_3} \le \Phi_3(\mu_n(A)) = \Phi_3(A) \qquad \square$$

REMARKS. 1. This theorem immediately implies Riesz-Thorin and Marcinkiewicz theorems for $\mathcal{I}_p$ and $\mathcal{I}_{p,w}$. While a Stein interpolation theorem is not explicitly included, it easily follows by the method of proof of Theorems 2.9 and 2.10.

2. Theorems 2.9, 2.10, and the proof of the triangle inequality in Theorem 2.7 show the power of Proposition 2.6. For $\{\mu_n(A)\}$ is a complicated non-linear function of A while $(\phi_n, A\psi_n)$ is linear.

The remainder of these lectures concerns the theory and properties of $\mathcal{I}_p$ and $\mathcal{I}_{p,w}$. We next want to discuss some general criteria for an integral operator on $L^2(M, d\mu)$ to be in $\mathcal{I}_p$ (mainly for $p = 1, 2$). See also Addendum D. In Chapter 4, we return to this question and prove deep results for $(M, d\mu) = (\mathbb{R}^\nu, dx)$. (*Note:* $\mathcal{I}_1 \equiv$ "trace class"; $\mathcal{I}_2 \equiv$ "Hilbert-Schmidt operators.") The first result is classical:

THEOREM 2.11. *Let $\mathcal{H} = L^2(M, d\mu)$ for some separable measure space (i.e., one with $\mathcal{H}$ separable). If $A \in \mathcal{I}_2$, then there exists a unique function $K \in L^2(M \times M, d\mu \otimes d\mu)$ with*

$$(A\phi)(x) = \int K(x,y)\phi(y)\, d\mu(y) \tag{2.6}$$

Conversely, any $K \in L^2(M \times M)$ defines an operator A by (2.6) which is in $\mathcal{I}_2$ and

$$\|A\|_2 = \|K\|_2 \tag{2.7}$$

REMARK. (2.6) is intended in the sense that the integral converges absolutely for a.e. x.

PROOF. Let $A \in \mathcal{I}_2$ and let $A = \sum \mu_n(A)(\phi_n, \cdot\,)\psi_n$ be its canonical expansion. The functions $\overline{\phi_n(y)}\,\psi_n(x)$ are orthonormal in $L^2(M \times M)$, so there is a function K in L^2 obeying (2.7) so that

$$K = L^2 - \lim_{n\to\infty} \left(\sum_{n=1}^{N} \mu_n(A)\,\overline{\phi_n(y)}\,\psi_n(x) \right)$$

(2.6) is easily seen to hold.

Conversely, let $K \in L^2(M \times M)$. Then, by the Schwarz inequality,

$$\left| \int K(x,y)\phi(y)\, d\mu(y) \right| \leq \left(\int |K(x,y)|^2\, d\mu(y) \right)^{1/2} \|\phi\|$$

so (2.6) defines a bounded operator. Moreover, given any $\phi, \psi \in \mathcal{B}$, $\overline{\phi_n}\,(x)\psi_n(y)$ are orthonormal in L^2, so

$$\sum |(\phi_n, A\psi_n)|^2 \leq \|K\|_2^2$$

Thus, $A \in \mathcal{I}_2$ by Proposition 2.6 and $\|A\|_2 \leq \|K\|_2$. This implies that the K representing A is unique and thus the equality in (2.7). □

Unfortunately, there is no simple necessary and sufficient condition for $A \in \mathcal{I}_1$ in terms of its integral kernel. Since we will not need the following, we quote it without proof (for $X = \mathbb{R}$, it is proven in Goh'berg-Krein [**134**, Section III.10] and for general X in Reed-Simon [**253**, Section XI.4]):

Theorem 2.12. *Let μ be a Baire measure on a locally compact space X. Let K be a function on $X \times X$ which is continuous and Hermitian positive, that is, $\sum_{i=1}^{N} \bar{z}_i z_j K(x_i, x_j) \geq 0$ for any $x_1, \ldots, x_N \in X$, $z \in \mathbb{C}^N$, and any N. Then $K(x,x) \geq 0$ for all x. Suppose that, in addition,*

$$\int K(x,x)\, d\mu(x) < \infty$$

Then there exists a unique operator A in $\mathcal{I}_1$ obeying (2.6) *and*

$$\|A\|_1 = \int K(x,x)\, d\mu(x)$$

Remark. Continuity is clearly essential since (2.6) only determines K a.e., and if μ has no pure points, $\{(x,x) \mid x \in X\}$ has $\mu \otimes \mu$ measure zero. Moreover, positivity is essential. For Carleman [**67**] has constructed a continuous real periodic function, f, on $[0, 2\pi]$ whose Fourier coefficients do not lie in any ℓ_p for $p < 2$. Thus $(Ag)(\theta) = \int f(\theta - \phi) g(\phi)\, d\phi$ defines an operator with $e^{in\theta}$ as eigenfunctions. Since $\hat{f} \notin \ell_p$, $p < 2$, $A \notin \mathcal{I}_p$ for $p < 2$. See Addendum D.

The following comparison theorem is sometimes useful:

Definition. Let A and B be operators on $L^2(M, d\mu)$. We say that B pointwise dominates A if and only if

$$|(A\phi)(x)| \leq (B|\phi|)(x)$$

for all ϕ.

Theorem 2.13. *If B pointwise dominates A and $B \in \mathcal{I}_{2n}$ for some integer n, then $A \in \mathcal{I}_{2n}$ and $\|A\|_{2n} \leq \|B\|_{2n}$.*

Proof. The intuition is based on the idea that if A and B have integral kernels K and L, then domination says that $|K(x,y)| \leq L(x,y)$ and, by Theorem 2.11, $\|A\|_{2n}^{2n} = \int K(x_1, x_2) K^*(x_2, x_3) \ldots K^*(x_{2n}, x_1)$ where $K^*(x,y) = \overline{K(y,x)}$. Let P be the projection onto a space spanned by the characteristic functions of a finite number of disjoint sets. Then PAP is pointwise dominated by PBP and each is given by an integral kernel so that $\|PAP\|_{2n} \leq \|PBP\|_{2n} \leq \|B\|_{2n}$. If one orders the P's so that $P > Q$ if and only if $\operatorname{Ran} P \supset \operatorname{Ran} Q$, then $PAP \to A$ weakly as $P \to \infty$. Thus, by Fatou's lemma (Theorem 2.7(d)), $A \in \mathcal{I}_{2n}$ and $\|A\|_{2n} \leq \|B\|_{2n}$. □

Conjecture. Theorem 2.13 remains true if $\mathcal{I}_{2n}$ is replaced by $\mathcal{I}_p$ for any $p \geq 2$. This would imply the truth of a conjecture in [**296**]; namely, for any f, g on $\mathbb{R}^\nu$, the operator with kernel $f(x-y)g(y)$ is in $\mathcal{I}_p$ ($p \geq 2$) if $|f|(x-y)|g|(y)$ is in $\mathcal{I}_p$. We also note another conjecture in [**296**]: If f^* is the symmetric decreasing rearrangement of f, then for $p \geq 2$, the $\mathcal{I}_p$ norm of $f^*(x-y)g^*(y)$ dominates the $\mathcal{I}_p$ norm of $f(x-y)g(y)$; for $p = 2n$, this is true by rearrangement inequalities [**56**]. The first conjecture here is now known to be false; see Addendum E.

Remarks. 1. Theorem 2.13 is not true if $\mathcal{I}_{2n}$ is replaced by $\mathcal{I}_1$. For example, if f is the function on $[0, 2\pi]$ which is 1 on $[0, \pi]$ and -1 on $[\pi, 2\pi]$, then $\phi \mapsto |f| * \phi$ is rank one and so in $\mathcal{I}_1$, while $\phi \mapsto f * \phi$ is *not* in $\mathcal{I}_1$ since the Fourier coefficients of f go like n^{-1}. Presumably, Theorem 2.13 is not true if $\mathcal{I}_{2n}$ is replaced by any $\mathcal{I}_p$ with $p < 2$. Recently, Pitt [**239**] and Dodds-Fremlin [**98**] have proven the $p = \infty$ analog of Theorem 2.13.

2. The above examples show that proving an integral operator is in $\mathcal{I}_1$ can be quite difficult — this is unfortunate for $\mathcal{I}_1$ is often the natural space, for example, for defining $\det(1+A)$ and in the scattering theory of Chapter 6. However, the counterexamples which prevent nice theorems from holding are generally rather contrived, so that I have found the following to be true: If an integral operator with kernel K occurs in some "natural" way and $\int |K(x,x)|\,dx < \infty$, then the operator can (almost always) be proven to be trace class (although sometimes only after some considerable effort).

Finally, we note the following criteria, the first of which is used as the basis of the definition of $\mathcal{I}_1$ in [**250**]:

THEOREM 2.14. *Let A be a bounded, positive operator on $\mathcal{H}$ and let $\{\phi_n\}$ be an orthonormal basis. Then*

$$\sum_{n=1}^{\infty} (\phi_n, A\phi_n)$$

is independent of the choice of basis. It is finite if and only if $A \in \mathcal{I}_1$ and, in that case, it equals $\|A\|_1$.

THEOREM 2.15. *Let B be a bounded operator on $\mathcal{H}$ and let $\{\phi_n\}$ and $\{\psi_n\}$ be orthonormal bases. Then*

$$\sum_{n=1}^{\infty} \|B\phi_n\|^2 \qquad \text{and} \qquad \sum_{n,m} |(\psi_m, B\phi_n)|^2$$

are independent of the choice of bases and equal. They are finite if and only if $B \in \mathcal{I}_2$ and, in that case, they equal $\|B\|_2^2$.

PROOFS. Choosing $B = \sqrt{A}$, Theorem 2.14 follows from Theorem 2.15. By Parseval's relation,

$$\sum_{n=1}^{\infty} \|B\phi_n\|^2 = \sum_{n,m} |(\psi_m, B\phi_n)|^2 = \sum_{n=1}^{\infty} \|B^*\psi_n\|^2$$

which proves the required equality and independence of basis. If $B \in \mathcal{I}_2$, then choosing the ϕ's to be a completion of the orthonormal set in $B = \sum_n \mu_n(B)(\phi_n, \cdot\,)\psi_n$, we see that the sum in question is finite and equal to $\sum \mu_n(B)^2 = \|B\|_2^2$. Conversely, suppose the sum is finite. Given any orthonormal sets ϕ, ψ, one can complete them to a basis whence $\sum_n |(\phi_n, B\psi_n)|^2 \le \sum_{n,m} |\phi_n, B\psi_m)|^2$ so $B \in \mathcal{I}_2$ by Proposition 2.6. □

We close this chapter with a discussion of convergence theorems for $\mathcal{I}_p$ analogous to two celebrated theorems for L^p: the dominated convergence theorem and the theorem that weak (Banach space) convergence and convergence of the norms implies norm convergence. In some ways the proofs are simpler than for L^p; this is because $\mathcal{I}_p^*$ is closer to the discrete ℓ_p than the general L^p space.

The non-commutative Fatou lemma (Theorem 2.7(d)) suggests that the analog of pointwise convergence should be weak operator convergence. The analog is seen to be good by considering ℓ_∞ embedded in $\mathcal{L}(\ell_2)$ as the family of diagonal matrices. Then $\mathcal{I}_p \cap \ell_\infty = \ell_p$ and a sequence $A^{(n)}$ with $\sup \|A^{(n)}\| < \infty$ converges to A in the weak operator topology if and only if $a_m^{(n)} \to a_m$ for each m.

The following is clearly a direct analog of the dominated convergence theorem:

THEOREM 2.16. *Let A_m, A, B be bounded operators on a Hilbert space, $\mathcal{H}$, with $B^* = B \geq 0$. Suppose that $|A_m| \leq B$ and $|A_m^*| \leq B$ for all m, $|A| \leq B$, $|A^*| \leq B$, and that $A_m \to A$ weakly. If $p < \infty$ and $B \in \mathcal{I}_p$, then $\|A - A_m\|_p \to 0$.*

PROOF. Fix ε. We can find a finite rank projection P so that for $Q = 1 - P$, we have that

$$\|QBQ\|_p < \varepsilon$$

for example, by choosing P to be the projection onto the span of the first few eigenvectors of B in a canonical expansion. Thus, for m large, $\|Q|A_m|Q\|_p \leq \varepsilon$ so $\||A_m|^{1/2}Q\|_{2p} \leq \varepsilon^{1/2}$ and similarly, $\||A_m^*|^{1/2}Q\|_{2p} \leq \varepsilon^{1/2}$. Using Hölder's inequality, $\|A_mQ\|_p \leq \||A_m|^{1/2}\|_{2p}\||A_m|^{1/2}Q\|_{2p} \leq \|B\|_p\varepsilon^{1/2}$ and similarly, $\|QA_mP\|_p \leq \|A_m^*Q\|_p \leq \|B\|_p\varepsilon^{1/2}$. Thus

$$\|A - A_m\|_p \leq 4\varepsilon^{1/2}\|B\|_p + \|P(A - A_m)P\|_p$$

Since P is finite rank, $\|P(A - A_m)P\|_p \to 0$ as $m \to \infty$ by the weak convergence. Thus

$$\lim \|A - A_m\|_p \leq 4\varepsilon^{1/2}\|B\|_p$$

Since ε is arbitrary, $\|A - A_m\|_p \to 0$. □

REMARK. If ϕ_n is an orthonormal set and $A_m = (\phi_1, \cdot)\phi_m$, then $A_m \to 0$ weakly and $|A_m| \leq (\phi_1, \cdot)\phi_1 \in \mathcal{I}_1$ but $\|A_m\| \nrightarrow 0$. This shows that the condition $|A_m^*| \leq B$ cannot be dropped.

In applications, this theorem is not as useful as its commutative analog for the following reason: Typically in the commutative case, one has $\|f_n - f_{n+1}\| \leq 2^{-n}$ and takes $g = |f_1| + \sum_{n=1}^{\infty}|f_{n+1} - f_n|$ whence $|f_m| \leq g$. Because the triangle inequality fails for $|\cdot|$ on operators, this trick will not work for $\mathcal{I}_p$ spaces. For this reason, the following device of Simon [**293**] is often useful.

DEFINITION. We write $A \underset{\mu}{<} B$ if $\mu_n(A) \leq \mu_n(B)$ for all n. We write $A \xrightarrow{\mu} B$ if $\mu_n(A) \to \mu_n(B)$ for all n.

Notice that $A \xrightarrow{\mu} 0$ if and only if $\|A\| \to 0$ since $\mu_n(A) \leq \mu_1(A) = \|A\|$. Moreover, by Theorem 1.20 (or a more elementary argument [**293**]), $|\mu_n(A) - \mu_n(B)| \leq \|A - B\|$ so that $A \xrightarrow{\mu} B$ if $\|A - B\| \to 0$.

THEOREM 2.17. *If $A_m \underset{\mu}{<} B$ with $B \in \mathcal{I}_p$ and if $\|A_m - A\| \to 0$, then $\|A_m - A\|_p \to 0$.*

PROOF. Since $\mu_n(A_m - A) \to 0$ by the above remark, it suffices to show that $\mu_n(A_m - A) \leq c_n$ for some $c \in \ell_p$, since we can then appeal to the dominated convergence theorem for ℓ_p. As in the proof of Theorem 2.5, Fan's inequalities, (1.5) imply that

$$\mu_{2n+1}(A_m - A) \leq \mu_{n+1}(A) + \mu_{n+1}(A_m) \leq 2\mu_{n+1}(B)$$
$$\mu_{2n}(A_m - A) \leq 2\mu_n(B)$$

so we can take $c = (2\mu_1(B), 2\mu_1(B), 2\mu_2(B), \dots) \in \ell_p$. □

THEOREM 2.18 (Simon [**293**]). *Let F be a continuous function on $[0,\infty)$ and let $\widetilde{F}(x) = \sup_{0<y<x}|F(y)|$. Suppose that $A_m \underset{\mu}{<} B$, $A_m \xrightarrow{\mu} A$, and that $\sum_n \widetilde{F}(\mu_n(B)) < \infty$. Then*

$$\sum_n F(\mu_n(A_m)) \to \sum_n F(\mu_n(A))$$

as $m \to \infty$.

PROOF. By continuity of F and $A_m \xrightarrow{\mu} A$, we have that $\lim_{m\to\infty} F(\mu_n(A_m)) = F(\mu_n(A))$ for each n. But, by monotonicity of $\widetilde{F}$,

$$|F(\mu_n(A_m))| \leq \widetilde{F}(\mu_n(A_m)) \leq \widetilde{F}(\mu_n(B))$$

so the result follows from the dominated convergence theorem for ℓ_1. □

The point of Theorem 2.18 is the following: Very often, one can prove inequalities for finite matrices; Theorem 2.18 provides the key for extensions of the theorem to infinite dimensions. Typically, $A_n = P_n A P_n$ with P_n suitable finite rank projections.

The following result of Grümm [**145**] is often useful (it should be mentioned that, as we shall see, for $p \neq 1$, this result follows easily from the earlier theorems of McCarthy [**219**]).

THEOREM 2.19 (Grümm's Convergence Theorem). *Fix $p < \infty$. Suppose that $A_n \to A$ and $A_n^* \to A^*$ in the strong operator topology and that $\|A_n\|_p \to \|A\|_p$. Then $\|A_n - A\|_p \to 0$.*

PROOF. Since $\sup_n \|A_n\| < \infty$, we have that $A_n^* A_n \to A^* A$ strongly. Since $\sqrt{\cdot}$ is strongly continuous (we can uniformly approximate $f(x) = \sqrt{x}$ on $[0, \sup_n \|A_n\|]$ by polynomials), we see that $|A_n| \to |A|$ strongly. Similarly, $|A_n^*| \to |A^*|$ strongly. Thus, this theorem follows from the next. □

THEOREM 2.20. *Fix $p < \infty$. Suppose that $A_n \to A$, $|A_n| \to |A|$, and $|A_n^*| \to |A^*|$ all weakly, and that $\|A_n\|_p \to \|A\|_p$. Then $\|A_n - A\|_p \to 0$.*

PROOF. Without loss, we can suppose that $\|A\|_p = \|A_n\|_p = 1$. Fix $\varepsilon > 0$. Find a finite-dimensional projection P so that $\|P|A|P\|_p \geq 1-\varepsilon$ and $\|P|A^*|P\|_p \geq 1-\varepsilon$; for example, P can be the projection onto the span of the first few eigenvectors for $|A|$ and for $|A^*|$. Now, by the weak convergence, find N so that for $n \geq N$, $\|P|A_n|P\| \geq 1 - 2\varepsilon$ and $\|P|A_n^*|P\| \geq 1 - 2\varepsilon$. By (1.20),

$$\|Q|A_n|Q\| \leq (1 - (1-2\varepsilon)^p)^{1/p}$$

and similarly for $\|Q|A_n^*|Q\|$. From this point, one need only follow the proof of Theorem 2.16. □

For $p > 1$, one can strengthen these results:

THEOREM 2.21. *Fix $1 < p < \infty$. Suppose that $A_n \to A$ weakly and that $\|A_n\|_p \to \|A\|_p$. Then $\|A - A_n\|_p \to 0$.*

FIRST PROOF. $(1 < p < \infty)$ By the weak convergence, $\mathrm{Tr}(A_n B) \to \mathrm{Tr}(AB)$ for B finite rank. Since $\|A_n\|_p$ is uniformly bounded, we can extend this to all $B \in \mathcal{I}_q$ with $q = p/(p-1)$. Thus $A_n \to A$ in the weak (Banach space) $\mathcal{I}_p$ topology. $\|A - A_p\| \to 0$ then follows from Theorems 1.23 and 1.24. □

SECOND PROOF. $(2 \le p < \infty)$ Let η_n be a basis of eigenvectors for $|A|$. Then for m_0 fixed,

$$\begin{aligned}
\|A\|_p^p &= \sum_{m=1}^{\infty} \|A\eta_m\|^p \\
&\le \sum_{\substack{m=1 \\ m\neq m_0}}^{\infty} \liminf \|A_n\eta_n\|^p + \limsup \|A_n\eta_{m_0}\|^p \quad \text{(since } A_n\eta_m \to A\eta_m \text{ weakly)} \\
&\le \limsup \sum_{m=1}^{\infty} \|A_n\eta_m\|^p \quad \text{(by Fatou's lemma)} \\
&\le \limsup \|A_n\|_p^p \quad \text{(by (1.17) since } p \ge 2) \\
&= \|A\|_p^p \quad \text{(by hypothesis)}
\end{aligned}$$

Thus, $\|A\eta_{m_0}\| = \limsup \|A_n\eta_{m_0}\}$, so $A_n\eta_{m_0} \to A\eta_{m_0}$. Letting m_0 vary, we see that $A_n \to A$ strongly. Similarly, $A_n^* \to A^*$. Now use Grümm's convergence theorem. □

Open question: Does Theorem 2.21 hold for $p = 1$? This question has been settled; it does hold for $p = 1$. See Addendum H.

REMARKS. 1. All the above theorems on $\|A_n\|_p \to \|A\|_p$ fail for $p = \infty$. For example, let ϕ_n be an orthonormal set and let $A_n = (\phi_1, \cdot\,)\phi_1 + (\phi_n, \cdot\,)\phi_n$.

2. Clearly, by Grümm's theorem, $\|A_n - A\| \to 0$ and $\|A_n\|_p \to \|A\|_p$ imply $\|A_n - A\|_p \to 0$.

EXAMPLE 1. ($|\cdot|$ is continuous on each $\mathcal{I}_p$) Let $\|A_n - A\|_p \to 0$. Then $A_n^* A_n \to A^*A$ in operator norm so $|A_n| \to |A|$ in operator norm (by the argument in the proof of Theorem 2.19). Since $\||A_n|\|_p = \|A_n\|_p \to \|A\|_p = \||A_n|\|_p$, Grümm's theorem implies that $|A_n| \to |A|$ in $\|\cdot\|_p$.

EXAMPLE 2. (If $\alpha > 0$ and $p \ge \max(1, \alpha)$, then $A \to |A|^\alpha$ is continuous from $\mathcal{I}_p$ to $\mathcal{I}_{p/\alpha}$. This is a result of Brydges et al. [**64**].) As in Example 1, if $A_n \to A$ in $\mathcal{I}_p$, then $|A_n|^\alpha \to |A|^\alpha$ in operator norm. Since $\||A_n|^\alpha\|_{p/\alpha} = \|A_n\|_p^\alpha \to \|A\|_p^\alpha = \||A|^\alpha\|_{p/\alpha}$. Grümm's theorem completes the proof. Here is another proof, using instead Theorem 2.17: Without loss, we can suppose that $\|A_n - A_{n+1}\|_p \le 2^{-n}$. As noted just before Theorem 2.17, $\mu_m(A_n) \to \mu_m(A)$ so $|A_n|^\alpha \xrightarrow{\mu} |A|^\alpha$. It thus suffices to find b_m in $\ell_{p/\alpha}$ with $\mu_m(|A_n|^\alpha) \le b_m$. Equivalently, we need only find c_m in ℓ_p with $\mu_m(A_n) \le c_m$. By Theorem 1.20, $\|\mu.(A_n) - \mu.(A_{n+1})\|_{\ell_p} \le \|A_n - A_{n+1}\|_p \le 2^{-n}$, so we can take

$$c_m = \mu_m(A_1) + \sum_{n=1}^{\infty} |\mu_m(A_{n+1}) - \mu_m(A_n)|$$

EXAMPLE 3. (Grümm [**145**]) To illustrate how ad hoc methods can sometimes be useful, let us show that if $A_n \to A$ strongly, $\sup_n \|A_n\| < \infty$, and $\|B_n - B\|_p \to 0$, then $\|A_nB_n - AB\|_p \to 0$. For writing

$$\|A_nB_n - AB\|_p \le \|A_n\| \, \|B_n - B\|_p + \|(A_n - A)B\|_p$$

we see that it suffices to show that $\|(A_n - A)B\|_p \to 0$. Given ε, we can find P finite rank so that $\|(1 - P)B\|_p < \varepsilon$. By the strong convergence, $\|(A_n - A)P\| \to 0$

so

$$\limsup \|(A_n - A)B\| \leq \varepsilon(\|A\| + \limsup \|A_n\|)$$

Since ε is arbitrary, the result is proven. Note that the result is false if strong convergence is replaced by weak convergence (take $B_n = (\phi_1, \cdot)\phi_1$, $A_n = (\phi_1, \cdot)\phi_n$) or if $A_n B_n \to AB$ is replaced by $B_n A_n \to BA$ (take $B_n = (\phi_1, \cdot)\phi_1$ and $A_n = (\phi_n, \cdot)\phi_1$).