CHAPTER 11

Classes of Compact Operators

1. Canonical Representation and Singular Numbers of Compact Operators

In the study of compact self-adjoint operators a basic role is played by their expansions into the orthogonal series of finite rank operators (9.2.1). This expansion is just the spectral theorem for such operators. For an arbitrary operator of class $\mathbf{S}_\infty$ the spectral analysis, clearly, cannot be exhausting. For example, the simplest integral Volterra operator $(Tu)(t) = \int_0^t u(\tau)\, d\tau$ on $L_2(0, 1)$ has no eigen-values at all (it satisfies $\sigma(T) = \sigma_c(T) = \{0\}$). Nevertheless, the operators of class $\mathbf{S}_\infty$ admit another (non-spectral) expansion into a series of rank one operators which in many questions substitutes the spectral theorem. This expansion is called *the canonical representation* (or *Schmidt expansion*) of a compact operator.

1. We proceed from the polar decomposition (8.1.5) which holds for any closed densely defined operator. For continuous operators this decomposition can be obtained easier and we do this here.

Let $T \in \mathbf{B}(H)$, $M = (T^*T)^{1/2} =: |T|$. Then $\|Tx\| = \|Mx\|$ and $N(T) = N(M)$. It follows that $\overline{R(M)} = H \ominus N(M) = H \ominus N(T) = \overline{R(T^*)}$. Let $g = Mx$, $h = Tx$. Consider the linear operator W: $g \to h$. This operator is well-defined since $g = 0$ implies $x \in N(T)$ and $h = 0$. It follows from $\|Tx\| = \|Mx\|$ that W extends by continuity to an isometry from $\overline{R(M)} = \overline{R(T^*)}$ onto $\overline{R(T)}$. Defining W to be zero on $N(T)$, *we obtain a partial isometry with initial space* $\overline{R(T^*)}$ *and range* $\overline{R(T)}$. Then $Tx = h = Wy = WMx$ and so T admits *the polar decomposition*

$$T = WM, \qquad M = (T^*T)^{1/2}. \tag{1}$$

2. Let now $T \in \mathbf{S}_\infty$ (and so $M \in \mathbf{S}_\infty$). Let us enumerate the positive eigen-values $\lambda_k = \lambda_k(M)$ of M counted with multiplicities in a non-increasing sequence. We write down the spectral expansion of M in the form $M = \Sigma_k \lambda_k(\cdot, \varphi_k)\varphi_k$ where $\{\varphi_k\}$ is the orthonormal sequence of eigen-vectors of M which is complete in $\overline{R(T^*)}$. The numbers

$$s_k(T) := [\lambda_k(T^*T)]^{1/2} = \lambda_k(|T|) \tag{2}$$

are called *the singular numbers* (*s-numbers*) *of T*. The definition given means that

$s_k(T) > 0$. If rank $T = r < \infty$, it is sometimes convenient to define $s_k(T)$, with $k > r$, to be zero. In such cases we shall do this without special mention. The system $\psi_k = W\varphi_k$ is obviously orthonormal and complete in $\overline{R(T)}$. Substituting the spectral expansion of M in (1), we obtain a representation of T as a sum (finite or infinite) of rank one operators

$$T = \sum_k s_k(\cdot, \varphi_k)\psi_k \quad (s_k = s_k(T)). \tag{3}$$

It follows directly from (3) that

$$T^* = \sum_k s_k(T)(\cdot, \psi_k)\varphi_k. \tag{4}$$

This, in turn, implies

$$T^*T = \sum_k s_k^2(T)(\cdot, \varphi_k)\varphi_k, \qquad TT^* = \sum_k s_k^2(T)(\cdot, \psi_k)\psi_k, \tag{5}$$

$$T\varphi_l = s_l(T)\psi_l, \qquad T^*\psi_l = s_l(T)\varphi_l, \tag{6}$$

$$T^*T\varphi_k = s_k^2(T)\varphi_k, \qquad TT^*\psi_k = s_k^2(T)\psi_k. \tag{7}$$

Representation (3) is called *the canonical representation* (*Schmidt expansion*) of T. It follows from (5)–(7) that this representation is essentially unique. The freedom is only in the choice of eigen-vectors in (7) but they must satisfy the 'accordance conditions' (6). Thus, we have proved the following result.

THEOREM 1. *Each operator $T \in \mathbf{S}_\infty$ admits a representation* (3), *where $\{\varphi_k\}$, $\{\psi_k\}$ are orthonormal systems of the same cardinality, $\{s_k\}$ is a positive non-increasing sequence tending to zero. Representation* (3) *is unique in the sense that $\{s_k\}$, $\{\varphi_k\}$ and $\{\psi_k\}$ must satisfy* (6) *and* (7). *The numbers $\{s_k\}$ are defined by* (2). *The systems $\{\varphi_k\}$ and $\{\psi_k\}$ are complete in $\overline{R(T^*)}$ and $\overline{R(T)}$. The Schmidt expansion of T^* is given by* (4).

REMARK. Incidentally, we have also found that the non-zero eigen-values of T^*T and TT^* coincide (together with multiplicities): $0 \neq \lambda_k(T^*T) = \lambda_k(TT^*)$. The eigen-vectors are related to each other by (6). These facts follow also from Theorem 8.1.4. Note that the dimensions of $N(T^*T) = N(T)$ and $N(TT^*) = N(T^*)$ are not necessarily equal.

The following result naturally supplements Theorem 1.

THEOREM 2. *Let $\{\varphi_k\}$, $\{\psi_k\}$ be orthonormal systems in H of the same cardinality, $\{s_k\}$ be a non-increasing sequence, $s_k > 0$, $s_k \to 0$ as $k \to \infty$. Then* (3) *defines an operator $T \in \mathbf{S}_\infty$ whose Schmidt expansion is given by* (3).

Proof. Denote by T_n the n-th partial sum of (3). Then

$$\|(T_{n+q} - T_n)x\|^2 = \sum_{n+1}^{n+q} s_k^2 |(x, \varphi_k)|^2 \leq s_{n+1}^2 \sum_k |(x, \varphi_k)|^2 \leq s_{n+1}^2 \|x\|^2. \tag{8}$$

Thus, the sequence $T_n \in \mathbf{K}$ is a Cauchy sequence in norm and so (3) converges to an operator $T \in \mathbf{S}_\infty$. Since (3) implies (4)–(7), it follows that (3) is the Schmidt expansion of T. □

REMARK. Inequality (8) with $q \to \infty$ yields an estimate for $\| T - T_n \|$. Since (8) turns into equality with $x = \varphi_{n+1}$, we have

$$\| T - T_n \| = s_{n+1}(T). \tag{9}$$

3. Let us mention some elementary properties of s-numbers. Clearly $s_k(T) = s_k(|T|)$. It follows from (3) and (4) that

$$s_k(T) = s_k(T^*). \tag{10}$$

For $n = 0$ (9) means that $s_1(T) = \| T \|$. Definition (2) together with the minimax principle (Theorem 9.2.4) for the spectrum of T^*T yields

$$s_{n+1}(T) = \min_L \max_{x \in L\setminus\{0\}} (\| Tx \| / \| x \|) \quad (\operatorname{def} L \leqslant n). \tag{11}$$

If $R \in \mathbf{B}(H)$ then $\| RTx \| \leqslant \| R \| \cdot \| Tx \|$. This together with (11) implies

$$s_k(RT) \leqslant \| R \| s_k(T), \qquad s_k(TR) \leqslant \| R \| s_k(T), \tag{12}$$

the second inequality being a consequence of the first in view of (10). It follows easily from (12) that for unitary operators U_1, U_2

$$s_k(U_1 T U_2) = s_k(T).$$

Let us mention one more extremal characteristic of s-numbers:

$$s_{n+1}(T) = \min_K \| T - K \| \quad (\operatorname{rank} K \leqslant n). \tag{13}$$

Indeed, in this case def $N(K)$ = rank K^* = rank $K \leqslant n$, and so (11) implies

$$s_{n+1}(T) \leqslant \max_{x \in N(K)} (\| Tx \| / \| x \|) = \max_{x \in N(K)} (\|(T - K)x \| / \| x \|) \leqslant \| T - K \|.$$

The equality in (13) is attained at $K = T_n$ (see (9)).

Let $T_1, T_2 \in \mathbf{S}_\infty$, then in view of (13)

$$s_{m+n-1}(T_1 + T_2) \leqslant s_m(T_1) + s_n(T_2). \tag{14}$$

Indeed, if $T = T_1 + T_2$ and rank $K_1 \leqslant m - 1$, rank $K_2 \leqslant n - 1$ then

$$s_{m+n-1}(T) \leqslant \| T - (K_1 + K_2) \| \leqslant \| T_1 - K_1 \| + \| T_2 - K_2 \|.$$

Besides, K_1 and K_2 can be chosen so that the right-hand side turns into $s_m(T_1) + s_n(T_2)$. Putting $n = 1$ in (14), we see that $s_m(T) \leqslant \| T_2 \| + s_m(T_1) = \| T - T_1 \| + s_m(T_1)$. Interchanging the roles of T and T_1, we obtain an inequality similar to (9.2.19):

$$| s_m(T) - s_m(T_1) | \leqslant \| T - T_1 \| \quad (T, T_1 \in \mathbf{S}_\infty). \tag{15}$$

Therefore the s-numbers depend on T continuously.

Sometimes inequality (14) is more convenient to express in terms of *the distribution function* of s-numbers. For $T \in \mathbf{S}_\infty$ set

$$\pi_T(\sigma) = \sum_{s_k(T) > \sigma} 1, \quad \sigma > 0. \tag{16}$$

Then an inequality similar to (9.2.20) holds. Namely, let T_1, $T_2 \in \mathbf{S}_\infty$, $T = T_1 + T_2$. Then

$$\pi_{T_1+T_2}(\sigma) \leqslant \pi_{T_1}(\sigma_1) + \pi_{T_2}(\sigma_2), \quad \sigma = \sigma_1 + \sigma_2, \sigma_1 > 0, \sigma_2 > 0. \tag{17}$$

Indeed, put $m = \pi_{T_1}(\sigma_1)$, $n = \pi_{T_2}(\sigma_2)$. It follows that

$$s_{m+1}(T_1) \leqslant \sigma_1, \qquad s_{n+1}(T_2) \leqslant \sigma_2$$

and in view of (14) $s_{m+n+1}(T) \leqslant \sigma_1 + \sigma_2 = \sigma$. Thus, $\pi_T(\sigma) \leqslant m + n$ which coincides with (17).

Inequalities for s-numbers of products are often useful as well. Let T_1, $T_2 \in \mathbf{S}_\infty$. Then

$$s_{m+n-1}(T_1T_2) \leqslant s_m(T_1)s_n(T_2), \quad m, n = 1, 2, \ldots \tag{18}$$

Indeed, let $T = T_1T_2$ and K_1, K_2 be chosen as in the proof of (14). Since $T - K = (T_1 - K_1)(T_2 - K_2)$, where $K = K_1(T - K_2) + T_1K_2$, we have

$$\| T - K \| \leqslant \| T_1 - K_1 \| \cdot \| T_2 - K_2 \| = s_m(T_1)s_n(T_2).$$

It remains to notice that rank $K \leqslant$ rank K_1 + rank $K_2 \leqslant m + n - 2$ and to use (13). It follows from (18) that

$$\pi_{T_1T_2}(\sigma) \leqslant \pi_{T_1}(\sigma_1) + \pi_{T_2}(\sigma_2), \quad \sigma = \sigma_1\sigma_2, \sigma_1 > 0, \sigma_2 > 0. \tag{19}$$

The derivation of (19) from (18) is essentially the same as that of (17) from (14).

It can easily be deduced from the spectral resolution (9.2.4) of a normal operator that the sequence of s-numbers of such an operator is the non-increasing rearrangement of the moduli of the eigen-values (counted with multiplicities). In this case φ_k and ψ_k coincide up to a unimodular factor.

2. Nuclear Operators. Trace of an Operator

1. The class $\mathbf{S}_\infty$ contains the important (in many respects) subclass $\mathbf{S}_1$ of *nuclear operators*. An operator $T \in \mathbf{S}_\infty$ is called nuclear if

$$\| T \|_1 := \sum_k s_k(T) < \infty. \tag{1}$$

Clearly, $\| T \| = s_1(T) \leqslant \| T \|_1$, the equality being attained only on the operators of rank 1. Next, $\| T \|_1 = \| |T| \|_1$, T is nuclear if and only if so is T^* and $\| T \|_1 = \| T^* \|_1$ is also in view of (1.10). If R_1, $R_2 \in \mathbf{B}(H)$ and $T \in \mathbf{S}_1$ then in accordance with

(1.12) we have $R_1TR_2 \in \mathbf{S}_1$ and

$$\| R_1TR_2 \|_1 \leq \| R_1 \| \cdot \| R_2 \| \cdot \| T \|_1. \tag{2}$$

Our next goal is to obtain a nuclearity criterion which does not involve s-numbers. First of all, we prove the following result.

LEMMA 1. *Let $T \in \mathbf{B}(H)$. Suppose that for some orthonormal basis $\{g_l\}$ we have*

$$\sum_l \| Tg_l \|^2 < \infty. \tag{3}$$

Then $T \in \mathbf{S}_\infty$.

Proof. It suffices to show that $x_n \xrightarrow{w} 0$ implies $T^*x_n \to 0$. (Then in view of Theorem 2.6.1 $T^* \in \mathbf{S}_\infty$, and so $T \in \mathbf{S}_\infty$.) Note that $\| x_n \| \leq c$. Expanding T^*x_n in the system $\{ g_l \}$, we obtain

$$\| T^*x_n \|^2 = \sum_l | (T^*x_n, g_l) |^2 = \sum_l | (x_n, Tg_l) |^2.$$

The terms of the last series tend to zero as $n \to \infty$ and are majorized by the terms of the converging series $\Sigma_l c^2 \| Tg_l \|^2$. □

THEOREM 2. *Let $T \in \mathbf{B}(H)$, $T > 0$. If for some orthonormal basis $\{g_l\}$ the series $\Sigma_l (Tg_l, g_l)$ converges then $T \in \mathbf{S}_1$. In this case for any orthonormal basis $\{h_l\}$*

$$\sum_l (Th_l, h_l) = \sum_k s_k (T) = \| T \|_1. \tag{4}$$

Proof. By the hypotheses $\Sigma_l \| T^{1/2}g_l \|^2 < \infty$ and by Lemma 1 $T^{1/2} \in \mathbf{S}_\infty$ and $T \in \mathbf{S}_\infty$. Let us now use (1.3) and the fact that $\varphi_k = \psi_k$ for $T > 0$. Then

$$\sum_l (Th_l, h_l) = \sum_l \sum_k s_k | (h_l, \varphi_k) |^2 = \sum_k s_k \| \varphi_k \|^2 = \sum_k s_k. \qquad \square$$

The case of an arbitrary operator of class $\mathbf{S}_1$ is more difficult. The question is settled by the two following theorems.

THEOREM 3. *Let $T \in \mathbf{S}_1$ and $\{g_l\}$, $\{h_l\}$ be arbitrary orthonormal sequences in H. Then*

$$\sum_l | (Tg_l, h_l) | \leq \| T \|_1, \tag{5}$$

the equality in (5) *being attained at $g_l = \varphi_l$, $h_l = \psi_l$.*

Proof. It follows from (1.3) that

$$\begin{aligned} \sum_l | (Tg_l, h_l) | &\leq \sum_{k,l} s_k | (g_l, \varphi_k) (\psi_k, h_l) | \\ &\leq \sum_k s_k [\sum_l | (g_l, \varphi_k) |^2]^{1/2} \cdot [\sum_l | (\psi_k, h_l) |^2]^{1/2} \\ &= \sum_k s_k \| \varphi_k \| \cdot \| \psi_k \| = \sum_k s_k = \| T \|_1. \end{aligned}$$

For $g_l = \varphi_l$, $h_l = \psi_l$ in view of (1.6) we have equality in (5). □

THEOREM 4. *If* $T \in \mathbf{B}(H)$ *and* $\Sigma_l (Tg_l, h_l)$ *converges for any orthonormal systems* $\{ g_l \}$, $\{ h_l \}$ *then* $T \in \mathbf{S}_1$.

Proof. We use the polar decomposition (1.1). Let $\{ g_l \}$ be an orthonormal basis in $\overline{R(M)}$ and $h_l = Wg_l$. Since W is isometric on $\overline{R(M)}$, $\{ h_l \}$ is also orthonormal. Now, the series

$$\sum_l (Mg_l, g_l) = \sum_l (WMg_l, Wg_l) = \sum_l (Tg_l, h_l) < \infty \tag{6}$$

converges by the assumptions. The system $\{g_l\}$ can be completed to a basis in H. This does not change the sum in (6) since $H \ominus \overline{R(M)} = N(M)$. It follows from Theorem 2 that $M \in \mathbf{S}_1$ and so $T \in \mathbf{S}_1$. □

The following sufficient condition for nuclearity is often convenient in practice.

THEOREM 5. *If* $T \in \mathbf{B}(H)$ *and for some orthonormal basis* $\{g_l\}$ *the series* $\Sigma_l \| Tg_l \|$ *converges then* $T \in \mathbf{S}_1$.

Proof. The equality $\| Tx \| = \| Mx \|$ implies that we can restrict ourselves to the case $T > 0$. But then

$$\sum_l (Tg_l, g_l) \leqslant \sum_l \| Tg_l \| < \infty$$

and the result follows from Theorem 2. □

2. The following Theorem will 'justify' notation (1).

THEOREM 6. *The functional* (1) *is a norm on* $\mathbf{S}_1$. *The space* $\mathbf{S}_1$ *endowed with this norm is a Banach space.*

Proof. Let $T_1, T_2 \in \mathbf{S}_1$ and $T = T_1 + T_2$. In view of (5) we have for any orthonormal systems $\{g_l\}$, $\{h_l\}$

$$\sum_l | (Tg_l, h_l) | \leqslant \sum_l | (T_1 g_l, h_l) | + \sum_l | (T_2 g_l, h_l) | \leqslant \| T_1 \|_1 + \| T_2 \|_1.$$

It follows from Theorem 4 that $T/\in \mathbf{S}_1$ while Theorem 3 implies the triangle inequality

$$\| T_1 + T_2 \|_1 \leqslant \| T_1 \|_1 + \| T_2 \|_1. \tag{7}$$

Therefore $\mathbf{S}_1$ is a linear subset of $\mathbf{B}(H)$ and the functional (1) is a norm on $\mathbf{S}_1$. It remains to show that $\mathbf{S}_1$ is complete. Let $\{T_l\}$ be a Cauchy sequence in $\mathbf{S}_1$. Then it is norm convergent and its limit T belongs to $\mathbf{S}_\infty$. Given $\varepsilon > 0$ we have

$$\sum_k s_k(T_l - T_m) \leqslant \varepsilon \quad (l, m \geqslant n_\varepsilon). \tag{8}$$

In view of (1.15) the limit passage in (8) as $l \to \infty$ yields

$$\sum_k s_k(T - T_m) \leqslant \varepsilon \quad (m \geqslant n_\varepsilon). \tag{9}$$

Thus, $T \in \mathbf{S}_1$ since $T_m \in \mathbf{S}_1$ and $\| T - T_m \|_1 \to 0$ as $m \to \infty$. □

Note that the set **K** of finite rank operators is dense in $\mathbf{S}_1$. Indeed, if T_n is the n-th partial sum of the series (1.3) then $s_k(T_n) = s_k(T)$ for $k \leqslant n$ and $s_k(T_n) = 0$ for $k > n$. Therefore $\| T - T_n \|_1 = \Sigma_{k > n} s_k(T) \to 0$ as $n \to \infty$.

It follows easily from the fact that **K** is dense in $\mathbf{S}_1$ that $\mathbf{S}_1$ *is separable*.

3. The notion of the matrix trace of an operator on a finite-dimensional space can naturally be generalized to the class $\mathbf{S}_1$. This generalization is based on the following result.

THEOREM 7. *If $T \in \mathbf{S}_1$ then the series $\Sigma_l (Tg_l, g_l)$ converges absolutely for any orthonormal basis $\{g_l\}$ and its sum does not depend on the choice of the basis $\{g_l\}$.*

Proof. Let us apply expansion (1.3). Then

$$\sum_l (Tg_l, g_l) = \sum_l \sum_k s_k(g_l, \varphi_k)(\psi_k, g_l) = \sum_k s_k \sum_l (\psi_k, g_l)(g_l, \varphi_k)$$
$$= \sum_k s_k (\psi_k, \varphi_k) = \sum_k (T\varphi_k, \varphi_k).$$

The inversion of the order of summation is justified by the absolute convergence of the double series which was actually ascertained in the proof of Theorem 3. It remains to note that the right-hand side does not involve $\{g_l\}$. □

Now we can introduce the following functional on $\mathbf{S}_1$ by

$$\operatorname{Tr} T := \sum_l (Tg_l, g_l), \tag{10}$$

which is called *the trace of* T. The trace of T is the sum of the diagonal entries of the matrix of T in the basis $\{g_l\}$. Theorem 7 means that this sum does not depend on the choice of matrix representation.

The functional (10) is obviously linear. It follows from (5) that $|\operatorname{Tr} T| \leqslant \| T \|_1$. Clearly, $\operatorname{Tr} T = \| T \|_1$ for $T > 0$. Hence $\operatorname{Tr} T$ is a continuous functional on $\mathbf{S}_1$ with norm 1. Next,

$$\operatorname{Tr} T^* = \sum_l (T^* g_l, g_l) = \sum_l \overline{(Tg_l, g_l)} = \overline{\operatorname{Tr} T}.$$

The trace of a product of two operators does not depend on the order of the factors.

THEOREM 8. *Let $T \in \mathbf{S}_\infty$, $R \in \mathbf{B}(H)$, $TR \in \mathbf{S}_1$, $RT \in \mathbf{S}_1$. Then*

$$\operatorname{Tr} TR = \operatorname{Tr} RT. \tag{11}$$

Proof. As a basis we choose the system $\{\psi_l\}$ from (1.3) supplemented with an orthonormal basis of $N(T^*)$. Then

$$\operatorname{Tr} TR = \sum_l (TR\psi_l, \psi_l) = \sum_l (R\psi_l, T^*\psi_l) = \sum_l s_l(T)(R\psi_l, \varphi_l). \tag{12}$$

Here we have used (1.6). Similarly, taking $\{\varphi_l\}$, we obtain

$$\operatorname{Tr} RT = \sum_l (RT\varphi_l, \varphi_l) = \sum_l s_l(T)(R\psi_l, \varphi_l). \qquad \square$$

Theorem 8 implies that the trace is a *similarity invariant*. Moreover the following obviously holds.

COROLLARY 9. *Let $T \in \mathbf{S}_1$ and $Z_1, Z_2 \in \mathbf{B}(H)$, $Z_2Z_1 = I$. Then*

$$\operatorname{Tr} Z_1TZ_2 = \operatorname{Tr} T. \tag{13}$$

Among the assumptions of Theorem 8 the condition $T \in \mathbf{S}_\infty$ does not look essential, though it is essential for the proof given above. In fact, Theorem 8 can be given in a more precise form.

THEOREM 8a. *Equality* (11) *remains valid if in the hypotheses of Theorem* 8 *$T \in \mathbf{S}_\infty$ is replaced by $T \in \mathbf{B}(H)$.*

The proof of Theorem 8a is rather simple but makes use of a comparatively difficult *theorem of Lidskii's* which claims that the matrix trace and the spectral trace of a nuclear operator coincide. This theorem will be presented in Section 7 where we give the proof of Theorem 8a (see Sub-§5) as well.

4. We presently find the general form of a linear functional on $\mathbf{S}_1$ and $\mathbf{S}_\infty$. We proceed from the following lemma.

LEMMA 10. *Let $T \in \mathbf{B}(H)$ and let*

$$\sup_{Q \in \mathbf{K}} \frac{|\operatorname{Tr} QT|}{\|Q\|} =: a < \infty. \tag{14}$$

Then $T \in \mathbf{S}_1$ and $\|T\|_1 = a$.

Proof. We use (1.1), let $\{g_l\}$ be an orthonormal basis in $\overline{R(M)}$ and let $Q = \Sigma_{l \leq r} (\cdot, Wg_l)g_l$, then $QW = \Sigma_{l \leq r} (\cdot, g_l)\, g_l$ is the projection onto $\vee_{l \leq r}\, g_l$. In accordance with (12) we have

$$\operatorname{Tr} QT = \operatorname{Tr} QWM = \sum_{l \leq r} (Mg_l, g_l) \leq a\|Q\| = a.$$

Then Theorem 2 implies that $M \in \mathbf{S}_1$ and $\|M\|_1 \leq a$. It remains for us to note that $\|T\|_1 = \|M\|_1$ and $a \leq \|T\|_1$ in view of the estimate

$$|\operatorname{Tr} QT| \leq \|QT\|_1 \leq \|Q\| \cdot \|T\|_1, \qquad \square \tag{15}$$

THEOREM 11. (*a*) *Each continuous linear functional on $\mathbf{S}_1$ can uniquely be represented by*

$$l(T) = \operatorname{Tr} QT, \tag{16}$$

where $Q \in \mathbf{B}(H)$. Besides, $\| l \| = \| Q \|$.

(b) Each continuous linear functional on $\mathbf{S}_\infty$ can uniquely be represented by (16) *with $Q \in \mathbf{S}_1$. Besides, $\| l \| = \| Q \|_1$.*

Proof. (a) For $f, g \in H$ put $T_{f,g} = (\cdot, g)f$ and $\Omega(f, g) = l(T_{f,g})$. The sesqui-linear form Ω is continuous: $| \Omega(f, g) | \leq \| l \| \cdot \| T_{f,g} \| = \| l \| \cdot \| f \| \cdot \| g\|$. Therefore $\Omega(f, g) = (Qf, g)$, where $Q \in \mathbf{B}(H)$ and $\| Q \| \leq \| l \|$. Since $\operatorname{Tr} QT_{f,g} = (Qf, g)$, it follows that (16) holds with $T = T_{f,g}$. By linearity (16) extends to the set of $T \in \mathbf{K}$ which is dense in $\mathbf{S}_1$ and then by continuity to all $T \in \mathbf{S}_1$. It follows from (15) that $\| l \| = \| Q \|$. The representation is unique since Q is uniquely determined by its sesqui-linear form.

(b) The operator Q is introduced in the same way as in (a). One should have in mind here that $\| T_{f,g} \|_1 = \|T_{f,g} \| = \| f \| \cdot \| g \|$. After obtaining (16) for $T \in \mathbf{K}$ one should use Lemma 10 (interchanging the roles of Q and T). Then $Q \in \mathbf{S}_1$ and $\| Q \|_1 \leq \| l \|$. The rest of the proof is the same as in (a). □

Clearly an analogue of Theorem 11 for antilinear functionals on $\mathbf{S}_1$ and $\mathbf{S}_\infty$ holds. The only distinction is that (16) should be replaced by

$$l(T) = \operatorname{Tr} QT^*. \tag{17}$$

It follows from (17) that the *dual space to $\mathbf{S}_\infty$ is $\mathbf{S}_1$ and the dual to $\mathbf{S}_1$ is $\mathbf{B}(H)$. Therefore the spaces $\mathbf{S}_1$, $\mathbf{S}_\infty$, $\mathbf{B}(H)$ are non-reflexive.*

3. Hilbert–Schmidt Operators

1. Another important class of compact operators is the class of *Hilbert–Schmidt operators*. This class (denoted by $\mathbf{S}_2$) consists of the operators for which the series (2.3) converges for some orthonormal basis. The situation is clarified by the following theorem.

THEOREM 1. *If the series* (2.3) *converges for some orthonormal basis then it converges for any orthonormal basis and its sum*

$$\| T \|_2^2 := \sum_l \| Tg_l \|^2 \tag{1}$$

does not depend on the choice of basis. The inclusion $\mathbf{S}_2 \subset \mathbf{S}_\infty$ holds. An operator $T \in \mathbf{S}_\infty$ belongs to $\mathbf{S}_2$ if and only if its sequence of singular numbers belongs to l^2. Besides,

$$\| T \|_2^2 = \sum_k s_k^2 (T). \tag{2}$$

Proof. For orthonormal bases $\{g_l\}$, $\{h_m\}$ we have

$$\sum_l \| Tg_l \|^2 = \sum_l \sum_m | (Tg_l, h_m) |^2 = \sum_m \sum_l | (g_l, T^*h_m) |^2 = \sum_m \| T^*h_m \|^2.$$

Equality (3) shows that the sum (1) does not depend on the choice of a basis $\{g_l\}$. The fact that $\mathbf{S}_2 \subset \mathbf{S}_\infty$ is established in Lemma 2.1. For $T \in \mathbf{S}_\infty$ we consider the series (1.3). Let us substitute $g_l = \varphi_l$ in (1) and supplement this sequence by an orthonormal basis in $N(T)$. Then (1.6) implies

$$\sum_l \| Tg_l \|^2 = \sum_l s_l^2(T) \| \psi_l \|^2 = \sum_l s_l^2 (T). \qquad \square$$

Equality (2) implies that $\mathbf{S}_1 \subset \mathbf{S}_2 \subset \mathbf{S}_\infty$ accompanied with the inequalities $\| T \| \leq \| T \|_2 \leq \| T \|_1$. It is also clear that $\| T \|_2 = \| T^* \|_2$ and that for $T \in \mathbf{S}_2$, $R_1, R_2 \in \mathbf{B}(H)$ we have

$$\| R_1 T R_2 \|_2 \leq \| R_1 \| \cdot \|R_2 \| \cdot \|T \|_2. \tag{4}$$

Now let $T = T_1 + T_2$ and $T_1, T_2 \in \mathbf{S}_2$. Let us show that $T \in \mathbf{S}_2$:

$$\sum_l \| Tg_l \|^2 \leq \sum_l (\| T_1 g_l \| + \| T_2 g_l \|)^2$$
$$\leq \left[\Big(\sum_l \| T_1 g_l \|^2\Big)^{1/2} + \Big(\sum_l \| T_2 g_l \|^2\Big)^{1/2}\right]^2 = (\| T_1 \|_2 + \| T_2 \|_2)^2.$$

Thus $T \in \mathbf{S}_2$ and

$$\| T_1 + T_2 \|_2 \leq \| T_1 \|_2 + \| T_2 \|_2. \tag{5}$$

THEOREM 2. *The space* $\mathbf{S}_2$ *is a Banach space with respect to the norm* $\| \cdot \|_2$.

Proof. As soon as (5) is proved, it is clear that $\mathbf{S}_2$ is a linear subset of $\mathbf{B}(H)$ and $\| \cdot \|_2$ is a norm on $\mathbf{S}_2$. The completeness of $\mathbf{S}_2$ can be established after the pattern of $\mathbf{S}_1$ (compare with the derivation of (2.9) and (2.8)). □

It is easy to see that $\mathbf{K}$ is dense in $\mathbf{S}_2$ and that $\mathbf{S}_2$ is separable. The arguments are the same as in the case of $\mathbf{S}_1$.

2. Let us dwell on the relations between $\mathbf{S}_1$ and $\mathbf{S}_2$ in more detail.

THEOREM 3. *Let* $T_1, T_2 \in \mathbf{S}_2$. *Then* $T = T_1 T_2 \in \mathbf{S}_1$ *and*

$$\| T \|_1 \leq \| T_1 \|_2 \|T_2 \|_2. \tag{6}$$

Proof. For arbitrary orthonormal bases $\{g_l\}$, $\{h_l\}$ we have

$$\sum_l | (Tg_l, h_l) | = \sum_l | (T_1 g_l, T_2^* h_l) | \leq \sum_l \| T_1 g_l \| \cdot \| T_2^* h_l \|$$
$$\leq \Big(\sum_l \| T_1 g_l \|^2\Big)^{1/2} \Big(\sum_l \| T_2^* h_l \|^2\Big)^{1/2} = \| T_1 \|_2 \| T_2 \|_2.$$

It follows from Theorem 2.4 that $T \in \mathbf{S}_1$ while (6) follows from Theorem 2.3. □

REMARK. If $T \in \mathbf{S}_1$ then there exist $T_1, T_2 \in \mathbf{S}_2$ such that $T = T_1 T_2$. For example, basing on (1.1), we can put $T_1 = WM^{1/2}$, $T_2 = M^{1/2}$. Here we should have in mind that $M = | T | \in \mathbf{S}_1$ for $T \in \mathbf{S}_1$ and so $M^{1/2} \in \mathbf{S}_2$.

The class $\mathbf{S}_2$ can naturally be endowed with a Hilbert structure. For $T_1, T_2 \in \mathbf{S}_2$ set

$$\langle T_1, T_2 \rangle = \operatorname{Tr} T_1 T_2^* = \operatorname{Tr} T_2^* T_1. \tag{7}$$

It follows directly from the properties of the trace that (7) is a sesqui-linear functional. Now,

$$\langle T, T \rangle = \operatorname{Tr} T^*T = \sum_k \lambda_k(T^*T) = \sum_k s_k^2(T) = \| T \|_2^2,$$

i.e. (7) is an inner product on $\mathbf{S}_2$ which agrees with $\| \cdot \|_2$. This together with Theorem 2 implies that $\mathbf{S}_2$ *is a (separable) Hilbert space with respect to the inner product* (7).

Let us mention the following simple but useful result.

THEOREM 4. *Let $\{g_k\}$, $\{h_l\}$ be orthonormal bases in H. Then the system of rank one operators $T_{kl} = (\cdot, g_k)h_l$ forms an orthonormal basis in $\mathbf{S}_2 = \mathbf{S}_2(H)$.*

Proof. Obviously, $\| T_{kl} \|_2 = 1$. For $T \in \mathbf{S}_2$ we have

$$\langle T, T_{kl} \rangle = (Tg_k, h_l). \tag{8}$$

Taking $T = T_{mn}$ in (8), we see that the system $\{T_{kl}\}$ is orthogonal in $\mathbf{S}_2$. It follows from (3) that this system is complete. □

Relation (3) can also be interpreted as follows. Let T be an operator represented by its matrix in a basis $\{g_k\}$. Then $T \in \mathbf{S}_2$ if and only if the entries of the matrix are square summable. Indeed, the entries of the matrix equal (Tg_k, g_l) and the result follows directly from (3). In Subsection 3 we obtain a 'continual' analogue of this fact which relates to integral operators.

3. Let us characterize the class $\mathbf{S}_2$ of operators on the Hilbert space $H_\mu = L_2(Y, \mu)$. Put $Z = Y \times Y$, $\nu = \mu \times \mu$ and $H_\nu = L_2(Z, \nu)$. A function $C(x, y)$ $(x, y \in Y)$ is called a *Hilbert–Schmidt kernel* if it is defined ν-a.e. on Z and $\int_Z |C(x, y)|^2 \, d\mu(x) \, d\mu(y) < \infty$. In other words the Hilbert–Schmidt kernels are just the functions in H_ν. To each such kernel we assign the integral operator $\hat{C}$ on H_μ:

$$(\hat{C}u)(x) = \int_Y C(x, y)u(y) \, d\mu(y). \tag{9}$$

The integral in (9) is finite for μ-a.e. $x \in Y$ in view of the inequality

$$|(\hat{C}u)(x)|^2 \leq \int_Y |C(x, y)|^2 \, d\mu(y) \| u \|^2. \tag{10}$$

Integrating the estimate (10), we see that $\hat{C}$ is bounded on H_μ. Its sesqui-linear form can be written as follows:

$$(\hat{C}u, v) = \int_Z C(x, y)u(y)\overline{v(x)} \, d\mu(x) \, d\mu(y). \tag{11}$$

Let us show that the mapping J: $C \to \hat{C}$ is an isometry from H_ν onto $\mathbf{S}_2(H_\mu)$. Let $\{u_k\}$ be an orthonormal basis in H_μ. The functions $\{w_{kl}\}$, $w_{kl}(x, y) = \overline{u_k(y)}\, u_l(x)$ form an orthonormal basis in H_ν (see §2.3, Sub-§6), whereas the operators $\hat{w}_{kl} = (\cdot, u_k)u_l$ form an orthonormal basis in $\mathbf{S}_2(H_\mu)$ (see Theorem 4). Clearly, $Jw_{kl} = \hat{w}_{kl}$. This implies directly that J is a unitary mapping from H_ν onto $\mathbf{S}_2(H_\mu)$. Thus, we have proved the following theorem, where $\| \cdot \|_\nu$ denotes the norm in H_ν.

THEOREM 5. *Let $C \in H_\nu$. Then the corresponding integral operator (9) belongs to $\mathbf{S}_2$ and*

$$\| \hat{C} \|_2 = \| C \|_\nu = \Big(\int_Z |C(x, y)|^2 \, d\mu(x) \, d\mu(y) \Big)^{1/2}.$$

Conversely, if $T \in \mathbf{S}_2(H_\mu)$ then there exists a unique Hilbert–Schmidt kernel $C \in H_\nu$ such that $T = \hat{C}$.

Theorem 5 admits various generalizations. For example, it is possible to give a similar characterization of the class $\mathbf{S}_2$ for the space H decomposed into a direct integral. We restrict ourselves to the case of the spaces $H_{\mu, F} = L_2(Y, \mu; F)$.

Put $G = \mathbf{S}_2(F)$ and $H_{\nu, G} = L_2(Z, \nu; G)$. Operator-functions $C \in H_{\nu, G}$ are called *operator-valued Hilbert–Schmidt kernels in* $H_{\mu, F}$. The operator $\hat{C} = JC$ is defined as before by (9). But now $C(x, y)u(y)$ means the value of the operator $C(x, y) \in \mathbf{S}_2(F)$ at $u(y) \in F$. Equality (11) is replaced by

$$(\hat{C}u, v) = \int_Z (C(x, y)u(y), v(x))_F \, d\nu(x, y).$$

THEOREM 6. *Let $C \in H_{\nu, G}$. Then the integral operator $\hat{C} = JC$ defined by* (9) *belongs to $\mathbf{S}_2$ and*

$$\| \hat{C} \|_2 = \| C \|_{\nu, F} = \Big(\int_Z \| C(x, y) \|^2_{\mathbf{S}_2(F)} \, d\mu(x) \, d\mu(y) \Big)^{1/2}.$$

Conversely, if $T \in \mathbf{S}_2(H_{\mu, F})$ then there exists a unique abstract Hilbert–Schmidt kernel $C \in H_{\nu, G}$ such that $T = JC$.

Proof. Let $\{u_k\}$ and $\{w_{kl}\}$ be the same as in the proof of Theorem 5. Let $\{f_i\}$ be an orthonormal basis in F and $\{g_{ij}\}$, $g_{ij} = (\cdot, f_i)f_j$, be the corresponding basis in G. Define the operators $\hat{B}_{ijkl}$ on $H_{\mu, F}$ by

$$(\hat{B}_{ijkl}u)(x) = \int_Y (u(y), f_i)_F \, \overline{u_k(y)} \, d\mu(y) u_l(x) f_j,$$

and the operator-functions B_{ijkl} by $B_{ijkl}(x, y) = w_{kl}(x, y)g_{ij}$. Then, in view of Theorem 2.3.7, $\{B_{ijkl}\}$ is a basis in $H_{\nu, G}$, $\{\hat{B}_{ijkl}\}$ is a basis in $\mathbf{S}_2(H_{\mu, F})$ and $\hat{B}_{ijkl} = JB_{ijkl}$. The rest can be proved as in the proof of Theorem 5. □

4. $\mathbf{S}_p$ classes

1. The spaces $\mathbf{S}_1$, $\mathbf{S}_2$ can naturally be embedded in the continual 'scale' of spaces

$\mathbf{S}_p$, $0 < p < \infty$. The class $\mathbf{S}_p$ is the subclass of $\mathbf{S}_\infty$ consisting of the operators satisfying

$$\| T \|_p := \left[\sum_k s_k^p(T) \right]^{1/p} < \infty. \tag{1}$$

The space $\mathbf{S}_\infty$ is a limit point of this scale but is not included in the scale. Note that $\mathbf{S}_{p_1} \subset \mathbf{S}_{p_2} \subset \mathbf{S}_\infty$, $0 < p_1 < p_2 < \infty$ and

$$\| T \| \leq \| T \|_{p_2} \leq \| T \|_{p}. \tag{2}$$

For rank one operators T all $\| T \|_p$ are equal.

If $T \in \mathbf{S}_p$ then $T^* \in \mathbf{S}_p$ and $\| T \|_p = \| T^* \|_p$. Estimates (2.2) and (3.4) can also be generalized: if $T \in \mathbf{S}_p$; $R_1, R_2 \in \mathbf{B}(H)$ then $R_1 T R_2 \in \mathbf{S}_p$ and $\| R_1 T R_2 \|_p \leq \| R_1 \| \cdot \| R_2 \| \cdot \| T \|_p$. This follows directly from (1.10) and (1.12).

The spaces $\mathbf{S}_p$ are linear: if $T_1, T_2 \in \mathbf{S}_p$ then $T = T_1 + T_2 \in \mathbf{S}_p$. Indeed, $T \in \mathbf{S}_\infty$ and in view of (1.14)

$$s_{2n}(T) \leq s_{2n-1}(T) \leq s_n(T_1) + s_n(T_2),$$

hence $T \in \mathbf{S}_p$.

The spaces $\mathbf{S}_p$ are normed only for $p \geq 1$. As we shall see in Subsection 2, the functional (1) in this case is a norm. It follows from (2) that $\mathbf{S}_1$ is the smallest class among the normed $\mathbf{S}_p$ classes and

$$\| T \| \leq \| T \|_p \leq \| T \|_1, \quad 1 < p < \infty.$$

Some information about the non-normed case $0 < p < 1$ will be given in Section 5, Subsection 4.

The proof of the triangle inequality for $\| \cdot \|_p$, $1 < p < \infty$ is not quite elementary. It is based on the following property of s-numbers which is of independent interest.

THEOREM 1. *Let $T, Q \in \mathbf{S}_\infty$. Then*

$$\sum_{k \leq r} s_k(TQ) \leq \sum_{k \leq r} s_k(T) s_k(Q), \quad r = 1, 2, \ldots \tag{3}$$

The proof of Theorem 1 will be given in Subsection 4.

2. The analogy between the classes $\mathbf{S}_p$, $1 < p < \infty$, and the sequence spaces l^p appears, in particular, in the following 'Hölder inequality'.

THEOREM 2. *Let $T \in \mathbf{S}_p, p > 1, Q \in \mathbf{S}_q$, where $p^{-1} + q^{-1} = 1$. Then $TQ \in \mathbf{S}_1$ and*

$$\| TQ \|_1 \leq \| T \|_p \| Q \|_q \quad (p > 1, p^{-1} + q^{-1} = 1). \tag{4}$$

Proof. It suffices to apply the Hölder inequality to the right-hand side of (3) and then to pass to the limit as $r \to \infty$. □

It follows from (4) and from the inequality $|\operatorname{Tr} T| \leqslant \| T \|_1$ that

$$|\operatorname{Tr} QT| \leqslant \| T \|_p \| Q \|_q \quad (p>1, p^{-1}+q^{-1}=1). \tag{5}$$

For a given $T \in \mathbf{S}_p$ the equality in (5) is attained for an appropriate Q. Namely, write down T in the form (1.1) and put $Q^* = WM^{p-1}$. It is easy to see that $Q \in \mathbf{S}_q$ and (5) turns into equality. Therefore

$$\| T \|_p = \sup_{Q \in \mathbf{S}_q} \frac{|\operatorname{Tr} QT|}{\| Q \|_q} = \sup_{\| Q \|_q = 1} |\operatorname{Tr} QT| \quad (p>1). \tag{6}$$

Now we are in a position to prove the triangle inequality. If $T_1 \in \mathbf{S}_p$, $T_2 \in \mathbf{S}_p$, $p > 1$, then (we have already proved that $T_1 + T_2 \in \mathbf{S}_p$)

$$\| T_1 + T_2 \|_p \leqslant \| T_1 \|_p + \| T_2 \|_p. \tag{7}$$

Inequality (7) immediately follows from (6) and the obvious inequality

$$|\operatorname{Tr}(T_1 + T_2)Q| \leqslant |\operatorname{Tr} T_1 Q| + |\operatorname{Tr} T_2 Q|.$$

Actually, we have shown that the functional (1) is a norm on $\mathbf{S}_p$. The proof of the completeness of $\mathbf{S}_p$ is the same as for the case $p = 1$ (see Theorem 2.6). Therefore, the following theorem holds.

THEOREM 3. *The class* $\mathbf{S}_p$, $1 \leqslant p < \infty$, *endowed with the norm* (1) *forms a Banach space.*

As in the case $p = 1, 2$ it can be shown that $\mathbf{K}$ *is dense in* $\mathbf{S}_p$ and $\mathbf{S}_p$ *is separable.*

3. We presently describe the general form of a continuous linear functional on $\mathbf{S}_p$ with $p > 1$. We begin with an analogue of Lemma 2.10 where a relation similar to (6) is obtained without the *a priori* assumption $T \in \mathbf{S}_p$.

LEMMA 4. *Let* $T \in \mathbf{B}(H)$ *and*

$$a := \sup_{Q \in \mathbf{K}} \frac{|\operatorname{Tr} QT|}{\| Q \|_q} < \infty \quad (q>1). \tag{8}$$

Then $T \in \mathbf{S}_p$, $p^{-1} + q^{-1} = 1$ *and* $\| T \|_p = a$.

Proof. Let us represent T in the form (1.1). Let $\{g_k\}$ be an orthonormal basis in $\overline{R(M)}$ and $Q_r = \Sigma_{k \leqslant r} \sigma_k(\cdot, Wg_k)g_k$, $\sigma_k > 0$. Then $\operatorname{Tr} Q_r T = \operatorname{Tr} Q_r WM = \Sigma_{k \leqslant r} \sigma_k (Mg_k, g_k)$ and in view of (8)

$$\sum_{k \leqslant r} \sigma_k (Mg_k, g_k) \leqslant a \| Q_r \|_q = a\Big[\sum_{k \leqslant r} \sigma_k^q\Big]^{1/q}. \tag{9}$$

Let us first show that $T \in \mathbf{S}_\infty$. Otherwise there would exist an interval $[\alpha, \beta]$, $\alpha > 0$, such that $\dim E_M[\alpha, \beta]H = \infty$. If $g_k \in E_M[\alpha, \beta]H$, then $\Sigma_{k \leqslant r}(Mg_k, g_k) \geqslant \alpha r$ which contradicts (9): putting $\sigma_k = 1$, we would have $\alpha r \leqslant ar^{1/q}$ for big r. Thus,

$T \in \mathbf{S}_\infty$. Set $g_k = \varphi_k$ in (9), where φ_k's are the eigen-vectors of M. Then (9) yields

$$\sum_{k \leq r} \sigma_k s_k(T) \leq a\Big[\sum_{k \leq r} \sigma_k^q\Big]^{1/q}.$$

For $\sigma_k = s_k^{p-1}$ and $r \to \infty$ this means that $T \in \mathbf{S}_p$ and $\| T \|_p \leq a$. The inequality $a \leq \| T \|_p$ follows from (5). □

THEOREM 5. *Each continuous linear functional on* $\mathbf{S}_p$, $1 < p < \infty$ *can uniquely be represented in the form* (2.16), *where* $Q \in \mathbf{S}_q$, $p^{-1} + q^{-1} = 1$ *and* $\| l \| = \| Q \|_q$.

Proof. The proof is very similar to that of Theorem 2.11(b). Here Lemma 4 plays the role of Lemma 2.10. □

Antilinear functionals on $\mathbf{S}_p$ can also be represented in the form (2.17). Thus for $p > 1$, $p^{-1} + q^{-1} = 1$, *the spaces* $\mathbf{S}_p$ *and* $\mathbf{S}_q$ *are dual to each other*. Consequently, it follows from (2.17) that *the spaces* $\mathbf{S}_p$ *are reflexive for* $p > 1$.

4. For the proof of Theorem 1 we need the following lemma.

LEMMA 6. *Let* $V = \{v_{kl}\}$, $1 \leq k, l \leq r$ *be a matrix such that*

$$\sum_{k \leq r} | v_{kl} | \leq 1 \quad (\forall l); \qquad \sum_{l \leq r} | v_{kl} | \leq 1 \quad (\forall k). \tag{10}$$

Let $x = \{x_k\}$, $y = \{y_k\}$, $1 \leq k \leq r$, *and let* $x_1 \geq \cdots \geq x_r \geq 0$, $y_1 \geq \cdots \geq y_r \geq 0$. *Then*

$$|(Vx, y)| \leq (x, y). \tag{11}$$

Proof. Set $f_k = (1, \ldots, 1, 0, \ldots, 0)$ (1 repeated k times). It follows from (10) that

$$|(Vf_k, f_l)| = \Big| \sum_{i \leq k,\, j \leq l} v_{ij} \Big| \leq \min(k, l) = (f_k, f_l) \quad (1 \leq k, l \leq r). \tag{12}$$

The assumptions imposed on x, y mean that $x = \Sigma_{k \leq r} \alpha_k f_k$, $y = \Sigma_{k \leq r} \beta_k f_k$ for some non-negative α_k, β_k. Taking into account (12), we see that

$$|(Vx, y)| = \Big| \sum_{k, l \leq r} \alpha_k \beta_l (Vf_k, f_l) \Big| \leq \sum_{k, l \leq r} \alpha_k \beta_l (f_k, f_l) = (x, y). \qquad \square$$

Proof of Theorem 1. Let WM be the polar decomposition of TQ and P be the projection onto the linear span of the first r eigen-vectors of M. Then $M = W^*TQ$ and

$$\sum_{k \leq r} s_k(TQ) = \sum_{k \leq r} s_k(M) = \operatorname{Tr} MP = \operatorname{Tr} PW^*TQP = \operatorname{Tr} \tilde{T}\tilde{Q}, \tag{13}$$

where $\tilde{T} = PW^*T$, $\tilde{Q} = QP$ and rank $\tilde{T} \leq r$, rank $\tilde{Q} \leq r$. Let

$$\tilde{T} = \sum_{k \leq r} \tilde{s}_k(\cdot, \varphi_k)\psi_k, \qquad \tilde{Q} = \sum_{k \leq r} \tilde{\sigma}_k(\cdot, \omega_k)\theta_k,$$

be their canonical representations. In accordance with (2.12) we obtain from (13)

$$\sum_{n\leq r} s_n(TQ) = \sum_{k\leq r} \tilde{s}_k(\tilde{Q}\psi_k, \varphi_k) = \sum_{k,l\leq r} \tilde{s}_k\tilde{\sigma}_l(\psi_k, \omega_l)(\theta_l, \varphi_k) = (Vx, y),$$

where $V = \{v_{kl}\}$; $v_{kl} \doteq (\psi_k, \omega_l)(\theta_l, \varphi_k)$; $x = \{\tilde{s}_k\}$, $y = \{\tilde{\sigma}_k\}$; $1 \leq k, l \leq r$. The matrix V satisfies (10): for every $l \leq r$

$$\sum_{k\leq r} |v_{kl}| \leq \Big(\sum_k |(\psi_k, \omega_l)|^2\Big)^{1/2}\Big(\sum_k |(\theta_l, \varphi_k)|^2\Big)^{1/2} \leq \|\omega_l\|\,\|\theta_l\| = 1 \tag{14}$$

and similarly for the sums over l. Now (11) implies that

$$\sum_{k\leq r} s_k(TQ) = (Vx, y) \leq (x, y) = \sum_{k\leq r} \tilde{s}_k\tilde{\sigma}_k,$$

and it remains to note that in view of (1.12) $\tilde{s}_k \leq s_k(T)$, $\tilde{\sigma}_k \leq s_k(Q)$, $1 \leq k \leq r$. □

5. The following simple assertion is useful for estimating s-numbers (see §8).

THEOREM 7. *Let* $T \in \mathbf{S}_p$, $0 < p < \infty$, rank $T' \leq k$. *Then*

$$\| T - T' \|_p \geq \Big(\sum_{k+1}^{\infty} s_m^p(T)\Big)^{1/p}. \tag{15}$$

The equality is attained if T' *is the k-th partial sum of the Schmidt expansion* (1.3) *of* T.

Proof. Applying inequality (1.14) with $n = k + 1$, we obtain

$$s_{m+k}(T) \leq s_m(T - T'), \quad m = 1, 2, \ldots$$

This implies (15) straightforwardly. The second assertion is obvious. □

For $p \to \infty$ the limit case of Theorem 7 is equality (1.13).

5. Additional Information on Singular Numbers of Compact Operators

In Sections 1–4 we have already met some important inequalities for s-numbers. The simplest examples are inequalities (1.14) and (1.18). More difficult examples are inequality (4.3) and the triangle inequality (4.7) for the $\mathbf{S}_p$ classes, $p \geq 1$. These examples, however, do not exhaust properties of s-numbers which are used for the investigation of compact operators. Here we present additional material on this subject.

1. We start with the important relations (discovered by H. Weyl) between the eigen-values and the singular values of compact operators.

Let $\{\lambda_k\} = \{\lambda_k(T)\}$ be the sequence of eigen-values of an operator T, enumerated so that the moduli are non-increasing, and counted with their algebraic multiplicities (see §3.9). If the sequence $\{\lambda_k\}$ is finite, we continue it by zeros (cf. similar agreements in the Remark to Theorem 10.2.5 and in §1, Sub-§2). Notice that in accordance with this agreement $\lambda_n(T) = 0$ does not necessarily mean that $0 \in \sigma_p(T)$.

We need the following technical result.

LEMMA 1. *Let* $T \in \mathbf{S}_\infty$, $\lambda_r(T) \neq 0$. *Then there exists a T-invariant subspace G such that* $\dim G = r$ *and the part* T_G *of T on G satisfies*

$$\lambda_k(T_G) = \lambda_k(T), \quad k = 1, \ldots, r.$$

Proof. Let $\lambda \in \sigma_p'(T)$, $H_\lambda^{(k)}$ be the subspaces (3.5.9) and $\tilde{H}$ be such a subspace that $H_\lambda^{(k)} \subset \tilde{H} \subset H_\lambda^{(k+1)}$ for some $k \geqslant 0$. Then $(T - \lambda I)\tilde{H} \subset H_\lambda^{(k)} \subset \tilde{H}$ and so $\tilde{H}$ is T-invariant. If $\tilde{T}$ is the part of T on $\tilde{H}$, then $(\tilde{T} - \lambda I)^{k+1} = 0$ and (see Corollary 3.7.12) the spectrum of $\tilde{T}$ consists of the unique eigen-value λ of algebraic multiplicity $\dim \tilde{H}$. By an appropriate choice of $\tilde{H}$ we can make this dimension equal an arbitrary number j between 1 and $\varkappa_T(\lambda)$.

Consider now a subspace of the form

$$G = \sum_{j<r;\, \lambda_j \neq \lambda_r} M_T(\lambda_j) + \tilde{H},$$

where $M_T(\lambda_j)$ are the root spaces and the subspace $\tilde{H} \subset M_T(\lambda_r)$ is chosen as above. By Theorem 3.5.2 the root spaces are linearly independent. Therefore, $\tilde{H}$ can be chosen so that $\dim G = r$. Then the spectrum of T_G consists of $\lambda_k(T)$, $k = 1, \ldots, r$. □

THEOREM 2. *If* $T \in \mathbf{S}_\infty$ *then*

$$\prod_1^r |\lambda_k(T)| \leqslant \prod_1^r s_k(T), \quad r = 1, 2, \ldots. \tag{1}$$

Proof. Inequality (1) is meaningful only if $\lambda_r \neq 0$. Let G be the subspace from Lemma 1, $\dim G = r$, $P = P_G$, T_G be the part of T on G. Obviously $TP = T_G \oplus 0$, hence $s_k(T_G) = s_k(TP) \leqslant s_k(T)$, $k = 1, \ldots, r$. Let t_r be the matrix of T_G in some orthonormal basis in G. Then

$$\prod_1^r |\lambda_k(T)|^2 = \prod_1^r |\lambda_k(T_G)|^2 = |\det t_r|^2 = \det(t_r^* t_r)$$
$$= \prod_1^r \lambda_k(T_G^* T_G) = \prod_1^r s_k^2(T_G) \leqslant \prod_1^r s_k^2(T). \qquad □$$

If $r = 1$, inequality (1) turns into the elementary estimate $|\lambda_1| \leqslant \|T\|$. For $k > 1$ the 'individual' estimates $|\lambda_k| \leqslant s_k$ are false in general. For example, for an operator on $\mathbf{C}^r$ the inequalities $|\lambda_k| \leqslant s_k$, $\forall k \leqslant r$, and the identity $\Pi_1^r |\lambda_k| = \Pi_1^r s_k$ would imply that $|\lambda_k| = s_k$, $\forall k$. It can be shown (see [6], §2.3) that this is equivalent to the fact that the operator is normal.

As an application of inequality (1) we present the following result.

THEOREM 3. *Let $T \in \mathbf{S}_\infty$, $s_n(T) \leqslant C \cdot n^{-\alpha}$, $\alpha > 0$. Then*

(a) $|\lambda_n(T)| \leqslant C\,\mathrm{e}^\alpha n^{-\alpha}$;

(b) $\limsup_{n\to\infty} n^\alpha |\lambda_n(T)| \leqslant \mathrm{e}^\alpha \limsup_{n\to\infty} n^\alpha s_n(T)$. (2)

In particular, if $s_n(T) = o(n^{-\alpha})$ then $\lambda_n(T) = o(n^{-\alpha})$.

Proof. (a) It follows from (1) and from the inequality $n! > n^n\,\mathrm{e}^{-n}$ that

$$|\lambda_n|^n \leqslant \prod_{k \leqslant n} |\lambda_k| \leqslant \prod_{k \leqslant n} s_k \leqslant C^n (n!)^{-\alpha} \leqslant C^n\,\mathrm{e}^{n\alpha} n^{-n\alpha}.$$

(b) Let $C = \sup(n^\alpha s_n)$, $M = \limsup(n^\alpha s_n)$ and let $s_n \leqslant (M + \varepsilon)n^{-\alpha}$ for $n > N = N(\varepsilon)$. Then

$$|\lambda_n|^n \leqslant C^N (M + \varepsilon)^{n-N} (n!)^{-\alpha} \leqslant C^N (M + \varepsilon)^{n-N}\,\mathrm{e}^{n\alpha} n^{-n\alpha}$$

or

$$n^\alpha |\lambda_n| \leqslant (C(M + \varepsilon)^{-1})^{N/n} (M + \varepsilon)\varepsilon^\alpha, \quad n > N(\varepsilon).$$

Taking $n \to \infty$ and then $\varepsilon \to 0$, we obtain (2). □

2. In what follows we need some preliminaries from the theory of sequences. Let $\{a_k\}_1^\nu$, $\{b_k\}_1^\nu$ be non-increasing sequences of real numbers, where ν is a positive integer or $\nu = \infty$. We say that $\{a_k\} \prec \{b_k\}$ if for any r, $0 \leqslant r - 1 < \nu$,

$$\sum_1^r a_k \leqslant \sum_1^r b_k. \tag{3}$$

The sequences ordered by '$\prec$' often occur in the theory of compact operators. For example, (4.3) means that $\{s_k(TQ)\} \prec \{s_k(T)s_k(Q)\}$ and (1) means that

$$\{\log|\lambda_k(T)|\} \prec \{\log s_k(T)\}. \tag{4}$$

Below we give other examples. The following lemma yields an extensive class of functions preserving $\prec$.

A function $\varphi(t)$, $t \in \mathbf{R}$ *is said to belong to the class* Φ if it is convex (i.e. $\varphi(\alpha t_1 + (1-\alpha)t_2) \leqslant \alpha\varphi(t_1) + (1-\alpha)\varphi(t_2)$, $\forall t_1, t_2 \in \mathbf{R}$, $\alpha \in (0, 1)$) and $\lim_{t\to-\infty} \varphi(t) = 0$. It is well known that for any convex function φ the derivative φ' exists almost everywhere and does not decrease. The functions φ in Φ can be represented by

$$\varphi(t) = \int_{-\infty}^t (t - x)\,\mathrm{d}\varphi'(x) = \int_{\mathbf{R}} (t - x)_+\,\mathrm{d}\varphi'(x). \tag{5}$$

LEMMA 4. *If $\{a_k\} \prec \{b_k\}$ and $\varphi \in \Phi$ then*

$$\{\varphi(a_k)\} \prec \{\varphi(b_k)\}. \tag{6}$$

Proof. Consider first the function $\psi_0(t) = t_+ := \max(t, 0)$. Clearly $\psi_0 \in \Phi$. Fix a number r. Let p be the number of positive terms in $\{a_k\}_1^r$, q be the number of positive terms in $\{b_k\}_1^r$, $0 \leqslant p, q \leqslant r$. Then

$$Q := \sum_1^r \psi_0(b_k) - \sum_1^r \psi_0(a_k) = \sum_1^q b_k - \sum_1^p a_k$$
$$= \sum_1^p (b_k - a_k) + \sum_{p+1}^q b_k = \sum_1^q (b_k - a_k) - \sum_{q+1}^p b_k.$$

The first representation is meaningful for $p \leqslant q$ while the second for $p \geqslant q$. Since $b_k > 0$ for $k \leqslant q$, $b_k \leqslant 0$ for $k > q$ and $\{a_k\} \prec \{b_k\}$, it follows that in both cases $Q \geqslant 0$. This proves (6) for $\psi = \psi_0$. Since the translation $a_k \mapsto a_k - x$, $b_k \mapsto b_k - x$ preserves $\prec$, it follows that (6) also holds for $\psi_x(t) = (t - x)_+$ for every $x \in \mathbf{R}$. Now, for $\varphi \in \Phi$ we obtain from (5) (recall that φ' is non-decreasing) that

$$\sum_1^r \varphi(a_k) = \int_{\mathbf{R}} \sum_1^r \psi_x(a_k)\, \mathrm{d}\varphi'(x) \leqslant \int_{\mathbf{R}} \sum_1^r \psi_x(b_k)\, \mathrm{d}\varphi'(x) = \sum_1^r \varphi(b_k). \qquad \square$$

The following result follows directly from (4) and Lemma 4.

THEOREM 5. *Let $T \in \mathbf{S}_\infty$. Then*

$$\sum_1^r |\lambda_k(T)|^p \leqslant \sum_1^r s_k^p(T), \quad p > 0, r = 1, 2, \ldots \tag{7}$$

Proof. It is sufficient to apply (8) to the sequences (4) with $\varphi(t) = e^{pt}$. $\square$

COROLLARY 6. *Let $T \in \mathbf{S}_p$, $0 < p < \infty$. Then $\{\lambda_k(T)\} \in l_p$ and*

$$\sum_k |\lambda_k(T)|^p \leqslant \sum_k s_k^p(T). \tag{8}$$

In particular,

$$\sum_k |\lambda_k(T)| \leqslant \|T\|_1, \quad T \in \mathbf{S}_1. \tag{9}$$

Inequalities (8) are close in spirit to the estimates of Theorem 3. Inequality (9) ascertains the absolute convergence of the series $\Sigma_k \lambda_k(T)$ for $T \in \mathbf{S}_1$.

3. In Section 1, Subsection 3, we discussed extremal characteristics of singular numbers. There are similar characteristics not only for individual s-numbers but also for their finite sums.

THEOREM 7. *Let $T \in \mathbf{S}_\infty$. Then*

$$\sum_1^r s_k(T) = \sup \sum_1^r |(Tg_l, h_l)|, \quad r = 1, 2, \ldots \tag{10}$$

where sup *is taken over the set of all orthonormal systems $\{g_k\}_1^r$, $\{h_k\}_1^r$.*

Proof. Put $g_k = \varphi_k$, $h_k = \psi_k$, $k = 1, \ldots, r$, where φ_k, ψ_k are the elements of the Schmidt expansion (1.3). Then the sums on both sides of (10) are equal. Since $\mathbf{K}$ is dense in $\mathbf{S}_\infty$, it suffices to verify the inequality $\geq$ for $T \in \mathbf{K}$. Let rank $T \leq d < \infty$. Clearly, we can suppose that $d \geq r$. We have

$$\sum_{k \leq r} |(Tg_k, h_k)| = \sum_{k \leq r} \Big|\sum_{l \leq d} s_l(g_k, \varphi_l)(\psi_l, h_k)\Big| \leq \sum_{k,l} v_{kl}\, s_l,$$

where $v_{kl} = |(g_k, \varphi_l)| \cdot |(\psi_l, h_k)|$, $1 \leq k \leq r$, $1 \leq l \leq d$. Putting $v_{kl} = 0$ for $r < k \leq d$, we obtain a $d \times d$-matrix V. The matrix V satisfies (4.10). This can be verified in the same way as (4.14). Applying Lemma 4.6 to $x, f_r \in \mathbf{R}^d$, where $x = (s_1, \ldots, s_d)$, $f_r = (1, \ldots, 1, 0, \ldots, 0)$ (1 is repeated r times), we obtain

$$\sum_1^r |(Tg_k, h_k)| \leq \sum_{k,l} v_{kl} s_l = (Vx, f_r) \leq (x, f_r) = \sum_1^r s_k(T). \qquad \square$$

THEOREM 8. *Let $T_1, T_2 \in \mathbf{S}_\infty$. Then for every $r = 1, 2, \ldots$ the triangle inequality*

$$\sum_1^r s_k(T_1 + T_2) \leq \sum_1^r s_k(T_1) + \sum_1^r s_k(T_2) \tag{11}$$

holds.

Proof. The sums on the right-hand side of (10) satisfy the triangle inequality. It remains to apply Theorem 7. □

It follows from Theorem 8 that for any r the functional $\Sigma_1^r s_k(T)$ is a norm on $\mathbf{S}_\infty$. This norm is equivalent to the usual operator norm since $\|T\| \leq \Sigma_1^r s_k(T) \leq r\, \|T\|$. In many questions of the theory of compact operators it is useful to consider all the norms $\Sigma_1^r s_k$ and their behaviour as $r \to \infty$ (see e.g. Theorem 6.3).

Inequalities (11) mean that $\{s_k(T_1 + T_2)\} \prec \{s_k(T_1) + s_k(T_2)\}$. This permits us, making use of Lemma 4, to obtain another proof of the triangle inequality for $\mathbf{S}_p$, $1 \leq p < \infty$. It suffices to apply Lemma 4 for $\varphi(t) = t_+^p$ and then to use the Minkowski inequality in l_p.

4. To conclude this section we present some assertions about s-numbers without proofs and discuss their corollaries. Besides Theorem 8 there is an important class of inequalities described by Rotfel'd's theorem (see [24]). Before formulating this theorem we recall that a function F: $\mathbf{R}_+ \to \mathbf{R}$ is called concave if

$$F(\alpha t_1 + (1 - \alpha)t_2) \geq \alpha F(t_1) + (1 - \alpha)F(t_2), \quad \forall t_1, t_2 \geq 0, \forall \alpha \in (0, 1).$$

THEOREM 9. *Let F: $\mathbf{R}_+ \to \mathbf{R}$ be a concave non-decreasing function such that $F(0) = F(+0) = 0$ and let $T_1, T_2 \in \mathbf{S}_\infty$, $T = T_1 + T_2$. Then for every $r = 1, 2, \ldots$*

$$\sum_{k \leq r} F(s_k(T)) \leq \sum_{k \leq r} F(s_k(T_1)) + \sum_{k \leq r} F(s_k(T_2)). \tag{12}$$

In the most important case $F(t) = t^p$, $0 < p \leq 1$ (12) yields

$$\sum_1^r s_k^p(T_1 + T_2) \leqslant \sum_1^r (s_k^p(T_1) + s_k^p(T_2)), \quad 0 < p \leqslant 1, r = 1, 2, \ldots . \tag{13}$$

For $p = 1$ this coincides with (11).

Inequalities (13) play an important role in the theory of $\mathbf{S}_p$ classes with $0 < p < 1$. As we have already mentioned, these spaces are not normed. Instead of the triangle inequality (4.7) the functional (4.1) satisfies the inequality

$$\| T_1 + T_2 \|_p^p \leqslant \| T_1 \|_p^p + \| T_2 \|_p^p$$

which follows from (13) by $r \to \infty$. Supplying $\mathbf{S}_p$, $p < 1$, with the metric

$$\varrho_p(T_1 - T_2) = \| T_1 - T_2 \|_p^p,$$

we turn $\mathbf{S}_p$ into a metric linear space. The completeness and separability of $\mathbf{S}_p$ as well as the density of $\mathbf{K}$ in $\mathbf{S}_p$ for $p < 1$ can be established in the same way as for $p \geqslant 1$.

The following important inequalities for s-numbers are called *Horn's inequalities* (they are proved, e.g., in [6], §2.4). Let us state the result.

THEOREM 10. *Let* $T_1, T_2 \in \mathbf{S}_\infty$. *Then*

$$\prod_{k \leqslant r} s_k(T_1 T_2) \leqslant \prod_{k \leqslant r} s_k(T_1) s_k(T_2), \quad r = 1, 2, \ldots . \tag{14}$$

Inequalities (14) can be rewritten in the form

$$\{\log s_k(T_1 T_2)\} \prec \{\log s_k(T_1) + \log s_k(T_2)\}.$$

This implies (cf. the derivation of (7)) the inequalities

$$\sum_1^r s_k^p(T_1 T_2) \leqslant \sum_1^r s_k^p(T_1) s_k^p(T_2), \quad r = 1, 2, \ldots, 0 < p < \infty. \tag{15}$$

The inequalities obtained contain (for $p = 1$) Theorem 4.1. Let us give one more example of the application of (14).

COROLLARY 11. *Let* $T_i \in \mathbf{S}_{p_i}$, $0 < p_i \leqslant \infty$, $i = 1, 2$. *Then* $T_1 T_2 \in \mathbf{S}_p$, $p^{-1} = p_1^{-1} + p_2^{-1}$ *and*

$$\| T_1 T_2 \|_p \leqslant \| T_1 \|_p \| T_2 \|_p. \tag{16}$$

Proof. For $p_1 = \infty$ inequality (16) follows from (1.12) where one can suppose $T_1 \in \mathbf{B}(H)$ instead of $T_1 \in \mathbf{S}_\infty$. The same is true for $p_2 = \infty$. If $p_1, p_2 < \infty$ then we can apply the Hölder inequalities with $p_1 p^{-1}$, $p_2 p^{-1}$ to (15) and pass to the limit as $r \to \infty$. □

Inequality (16) contains inequality (4.4) as a special case.

6. Σ_p classes

1. Here we consider the classes of operators defined by

$$\Sigma_p := \{T \in \mathbf{S}_\infty : s_n(T) = O(n^{-1/p})\}, \quad 0 < p < \infty,$$

and the corresponding subclasses

$$\Sigma_p^{(0)} := \{T \in \mathbf{S}_\infty : s_n(T) = o(n^{-1/p})\}, \quad 0 < p < \infty.$$

An equivalent description in terms of the distribution function of s-numbers Eqn (1.16) is also useful:

$$\Sigma_p = \{T \in \mathbf{S}_\infty : \pi_T(s) = O(s^{-p}); s \to +0\};$$

$$\Sigma_p^{(0)} = \{T \in \mathbf{S}_\infty : \pi_T(s) = o(s^{-p}); s \to +0\}.$$

The functional

$$|T|_p := \sup_n \{n^{1/p} s_n(T)\} = \left(\sup_{s>0} \{s^p \pi_T(s)\}\right)^{1/p} \tag{1}$$

is finite on Σ_p. If $0 < p_1 < p_2 < \infty$ then $\Sigma_{p_1} \subset \Sigma_{p_2}^{(0)}$ and

$$\|T\| \leqslant |T|_{p_2} \leqslant |T|_{p_1}. \tag{2}$$

For rank one operators all the norms in (2) coincide.

The classes Σ_p are close to $\mathbf{S}_p$ and sometimes are called 'weak $\mathbf{S}_p$ spaces'. This closeness is expressed by

$$\Sigma_{p_1} \subset \mathbf{S}_p \subset \Sigma_p^{(0)}; \qquad |T|_p \leqslant \|T\|_p \leqslant c(p_1, p)\, |T|_{p_1}, \quad p_1 < p.$$

Indeed, if $T \in \mathbf{S}_p$ then $ns_n^p(T) \leqslant \Sigma_1^n s_k^p(T) \leqslant \|T\|_p^p$. Since the s-numbers form a monotone sequence, the inclusion $T \in \Sigma_p^{(0)}$ follows from the estimate $ns_{2n+1}^p(T) \leqslant ns_{2n}^p(T) \leqslant \Sigma_{n+1}^{2n} s_k^p(T)$. If $T \in \Sigma_{p_1}$, then $\Sigma_n s_n^p(T) \leqslant |T|_{p_1}^p \cdot \Sigma_n n^{-p/p_1}$, $p_1 < p$.

The classes Σ_p, $\Sigma_p^{(0)}$ are convenient when we need exact information on the rate of decrease of s-numbers (which in view of Theorem 5.3 yields information on the rate of decrease of eigen-values). In particular, these classes naturally appear when we study the operators with a regular power-like asymptotic of the spectrum. In this connection Theorem 4 proved below (as well as Theorems 6–8 following from it) plays a special role.

2. THEOREM 1. *The classes* Σ_p, $\Sigma_p^{(0)}$ *are linear spaces. The functional* (1) *satisfies the conditions*

$$|\lambda T|_p = |\lambda| \cdot |T|_p, \quad \lambda \in \mathbf{C}; \tag{3}$$

$$|T_1 + T_2|_p^{p/p+1} \leqslant |T_1|_p^{p/p+1} + |T_2|_p^{p/p+1}. \tag{4}$$

Proof. We apply inequality (1.17) with $\sigma_1 = \theta s$, $\sigma_2 = (1-\theta)s$, where

$$\pi_T(s) \leqslant \pi_{T_1}(\theta s) + \pi_{T_2}((1-\theta)s), \quad T = T_1 + T_2. \tag{5}$$

Multiplying by s^p and taking the suprema, we obtain

$$|T|_p^p \leqslant \theta^{-p} |T_1|_p^p + (1-\theta)^{-p} |T_2|_p^p.$$

Minimizing the right-hand side over θ, we arrive at (4). Equality (3) is obvious. The fact that Σ_p is linear follows directly from (3) and (4). It follows from (5) that if T_1, $T_2 \in \Sigma_p^{(0)}$ then $T_1 + T_2 \in \Sigma_p^{(0)}$. □

Theorem 1 shows that Σ_p, $\Sigma_p^{(0)}$ are metric spaces with respect to the metric $|T_1 - T_2|_p^{p/p+1}$. The functional (1) is a *quasi-norm*: it is homogeneous (equality (3)) and satisfies a 'weakened' triangle inequality: $|T_1 + T_2|_p \leqslant 2^{1/p} (|T_1|_p + |T_2|_p)$ which follows from (4).

THEOREM 2. *The spaces* Σ_p, $0 < p < \infty$ *are complete and non-separable. The set* $\Sigma_p^{(0)}$ *is closed in* Σ_p, *coincides with the closure of* $\mathbf{K}$ *and is separable.*

Proof. Let $\{T_k\}$ be a Cauchy sequence in Σ_p. This means that for any $\varepsilon > 0$ there exists an integer $k(\varepsilon)$ such that

$$s_n(T_k - T_l) \leqslant \varepsilon n^{-1/p}, \quad k, l > k(\varepsilon), \forall n. \tag{6}$$

In view of (2) the limit $T = u\text{-}\lim T_k \in \mathbf{S}_\infty$ exists. Making use of (1.15) we obtain from (6) with $l \to \infty$ that

$$s_n(T_k - T) \leqslant \varepsilon n^{-1/p}, \quad k > k(\varepsilon); \forall n.$$

Therefore $T_k - T \in \Sigma_p$ (and so $T \in \Sigma_p$) and $|T_k - T|_p \to 0$. This proves the completeness of Σ_p.

Let us fix an orthonormal system $\{w_j\}$ and consider the operators

$$T_k = \sum_{2^{k-1} \leqslant j \leqslant 2^k - 1} j^{-\alpha}(\cdot, w_j)w_j, \quad \alpha = p^{-1}, k = 1, 2, \ldots$$

They satisfy

$$|T_k|_p = \sup_{2^{k-1} \leqslant j \leqslant 2^k - 1} \frac{(j - 2^{k-1} + 1)^\alpha}{j^\alpha} = \frac{2^{(k-1)\alpha}}{(2^k - 1)^\alpha} > 2^{-\alpha}, \forall k.$$

If $\mu = \{\mu_k\}_1^\infty$ is an arbitrary sequence of ± 1 and $T(\mu) = \Sigma_k \mu_k T_k$ then $s_n(T(\mu)) = n^{-\alpha}$ so that $T(\mu) \in \Sigma_p$. Let $\mu \neq \mu'$ and let k be the least number for which $\mu_k = \mu_k'$. Then for $1 \leqslant j \leqslant 2^{k-1}$ we have $s_j(T(\mu) - T(\mu')) = 2s_j(T_k)$. Hence $|T(\mu) - T(\mu')|_p \geqslant 2|T_k|_p > 2^{1-\alpha}$. The fact that Σ_p is non-separable follows from the fact that the set of such sequences is non-countable.

The proof of the assertions about Σ_p^0 is conveniently given after Theorem 6. □

For $p > 1$ the classes Σ_p can be normed. Put

$$\langle T \rangle_p = \sup_r r^{p^{-1}-1} \sum_1^r s_n(T), \quad p > 1. \tag{7}$$

THEOREM 3. *The functional* (7) *is a norm on* Σ_p, $1 < p < \infty$, *and*

$$|T|_p \leqslant \langle T \rangle_p \leqslant p(p-1)^{-1} |T|_p, \quad T \in \Sigma_p.$$

Proof. The triangle inequality for $\langle\cdot\rangle_p$ follows from (5.11), whereas the remaining properties of the norm are obvious. The estimate $rs_r(T) \leqslant \Sigma_1^r s_n(T)$ implies $|T|_p \leqslant \langle T\rangle_p$. On the other hand $\Sigma_1^r s_n(T) \leqslant c\int_0^r t^{-1/p}\,dt = cp(p-1)^{-1}r^{1-1/p}$ if $s_n(T) \leqslant cn^{-1/p}$. □

3. The functionals

$$\Delta_p(T) := \limsup_{s\to 0} s^p\pi_T(s), \qquad \delta_p(T) := \liminf_{s\to 0} s^p\pi_T(s) \tag{8}$$

used for the description of the asymptotic properties of the s-numbers, are finite on Σ_p. In particular, the equality $\Delta_p(T) = \delta_p(T) = c^p$, $c > 0$ is equivalent to the existence of a 'regular' power-like asymptotics $s_n(T) \sim cn^{-1/p}$.

For operators $T = T^* \in \Sigma_p$ we introduce similar functionals which describe the behaviour of the eigen-values: let $\pi_T^\pm(\lambda)$ be the distribution function of the spectrum (9.2.11) and

$$\Delta_p^\pm(T) := \limsup_{\lambda\to +0} \lambda^p\pi_T^\pm(\lambda); \qquad \delta_p^\pm(T) = \liminf_{\lambda\to +0} \lambda^p\pi_p^\pm(\lambda). \tag{9}$$

Let us mention the obvious inequalities

$$\Delta_p(T) \leqslant |T|_p^p, \quad T \in \Sigma_p; \tag{10}$$

$$\delta_p^\pm(T) \leqslant \Delta_p^\pm(T) \leqslant \Delta_p(T), \quad T = T^* \in \Sigma_p. \tag{11}$$

To simplify the notation below we sometimes use the common notation $D_p(T)$ for each of the functionals $\Delta_p(T)$, $\Delta_p^\pm(T)$. Note that the case $D_p(T) = \Delta_p^\pm(T)$ automatically means that $T = T^*$. Similarly we use the notation d_p for δ_p, $\delta_p^\pm$.

THEOREM 4. *If* $T_1, T_2 \in \Sigma_p$ *then*

$$(D_p(T_1 + T_2))^{1/(p+1)} \leqslant (D_p(T_1))^{1/(p+1)} + (D_p(T_2))^{1/(p+1)}; \tag{12}$$

$$(d_p(T_1 + T_2))^{1/(p+1)} \leqslant (D_p(T_1))^{1/(p+1)} + (d_p(T_2))^{1/(p+1)}. \tag{13}$$

Proof. To be definite we consider the functionals $\Delta_p^\pm$, $\delta_p^\pm$. Analogously to (5) we find from (9.2.20)

$$\pi_{T_1+T_2}^\pm(\lambda) \leqslant \pi_{T_1}^\pm(\theta\lambda) + \pi_{T_2}^\pm((1-\theta)\lambda), \quad 0 < \theta < 1.$$

Multiplying by λ^p and passing to lim sup (lim inf), we obtain

$$\Delta_p^\pm(T_1 + T_2) \leqslant \theta^{-p}\Delta_p^\pm(T_1) + (1-\theta)^{-p}\Delta_p^\pm(T_2);$$

$$\delta_p^\pm(T_1 + T_2) \leqslant \theta^{-p}\Delta_p^\pm(T_1) + (1-\theta)^{-p}\delta_p^\pm(T_2).$$

The minimization of the right-hand side over θ leads to (12) and (13). □

COROLLARY 5. *If* $T_1, T_2 \in \Sigma_p$ *then*

$$|(D_p(T_1))^{1/(p+1)} - (D_p(T_2))^{1/(p+1)}| \leqslant (\Delta_p(T_1 - T_2))^{1/(p+1)}, \tag{14}$$

$$|(d_p(T_1))^{1/(p+1)} - (d_p(T_2))^{1/(p+1)}| \leqslant (\Delta_p(T_1 - T_2))^{1/(p+1)}. \tag{15}$$

Proof. Let us prove, for example, (14) for Δ_p^+. In accordance with (12)

$$-(\Delta_p^+(T_2 - T_1))^{1/p+1} \leq (\Delta_p^+(T_1))^{1/p+1} - (\Delta_p^+(T_2))^{1/p+1}$$
$$\leq (\Delta_p^+(T_1 - T_2))^{1/p+1}.$$

It remains to apply the equality $\Delta_p^+(T_2 - T_1) = \Delta_p^-(T_1 - T_2)$ and inequality (11). □

Inequalities (14) and (15) are sources of many useful facts.

THEOREM 6. *The functionals* (8) *are continuous on* Σ_p. *The functionals* (9) *are continuous on the set of self-adjoint operators in* Σ_p.

Proof. It is sufficient to compare (14) and (15) with (10). □

Now we complete the proof of Theorem 2. Obviously, $\Sigma_p^{(0)} = \{T \in \Sigma_p: \Delta_p(T) = 0\}$ and hence $\Sigma_p^{(0)}$ is closed by Theorem 6. Next, $\Delta_p(T) = 0$ for $T \in \mathbf{K}$ and so $\Delta_p(T) = 0$ for $T \in \bar{\mathbf{K}}_p$, $\bar{\mathbf{K}}_p$ being the closure of $\mathbf{K}$ in Σ_p. Conversely, let $T \in \Sigma_p^{(0)}$ and let $s^p\pi_T(s) < \varepsilon^p$ for $s < s_0$, $s_0 > 0$. Denote by T_ε the operator obtained from the Schmidt expansion (1.3) of T by replacing all s-numbers $s_n(T) \leq s_0$ by zeros. Then $T_\varepsilon \in \mathbf{K}$ and $\pi_{T-T_\varepsilon}(s) = (\pi_T(s) - \pi_T(s_0))_+$. Therefore $| T - T_\varepsilon |_p < \varepsilon$ and so $T \in \bar{\mathbf{K}}_p$. We have proved that $\Sigma_p^{(0)} = \bar{\mathbf{K}}_p$. This implies automatically that $\Sigma_p^{(0)}$ is separable. □

The following assertion also follows directly from Theorem 6.

THEOREM 7. *Each of the following sets is closed in* Σ_p:

$$\{T \in \Sigma_p: \Delta_p(T) = \delta_p(T)\};$$

$$\{T = T^* \in \Sigma_p: \Delta_p^+(T) = \delta_p^+(T)\}; \qquad \{T = T^* \in \Sigma_p: \Delta_p^-(T) = \delta_p^-(T)\}.$$

Inequalities (14) and (15) contain a result due to H. Weyl (for eigen-values) and Ky Fan (for s-numbers) which is often used in practice.

THEOREM 8. *If* $T_1, T_2 \in \Sigma_p$ *and* $T_2 - T_1 \in \Sigma_p^{(0)}$ *then*

$$D_p(T_2) = D_p(T_1); \qquad d_p(T_2) = d_p(T_1). \tag{16}$$

The following theorem describes multiplicative properties of the Σ_p spaces (cf. (5.1.6)).

THEOREM 9. *Let* $T_1 \in \Sigma_{p_1}$, $T_2 \in \Sigma_{p_2}$ *and* $T = T_1T_2$. *Then* $T \in \Sigma_p$, $p^{-1} = p_1^{-1} + p_2^{-1}$ *and*

$$| T |_p \leq c \, | T_1 |_{p_1} \cdot | T_2 |_{p_2}, \tag{17}$$

$$\Delta_p(T) \leq c^p \, \Delta_{p_1}(T_1) \, \Delta_{p_2}(T_2), \tag{18}$$

where $c = c(p_1, p_2) = (p_1/p)^{1/p_1}(p_2/p)^{1/p_2}$.

In particular, if either $T_1 \in \Sigma_{p_1}^{(0)}$ or $T_2 \in \Sigma_{p_2}^{(0)}$ then $T \in \Sigma_p^{(0)}$.

Proof. In accordance with (1.19)

$$\pi_T(\sigma) \leq \pi_{T_1}(\alpha^{-1}\sigma^\theta) + \pi_{T_2}(\alpha\sigma^{1-\theta}), \quad 0 < \theta < 1, \ \alpha > 0.$$

Put here $\theta = pp_1^{-1}$. Multiplying by σ^p and taking suprema, we obtain $| T |_p^p \leq \alpha^{p_1}$

$|T_1|_{p_1}^{p_1} + \alpha^{-p_2}|T_2|_{p_2}^{p_2}$. Minimizing over α, we arrive at (17). If we take the lim sup instead of the suprema we obtain (18). □

4. Theorem 8 shows that, in fact, the functionals (8) and (9) characterize the properties of the coset $T + \Sigma_p^{(0)}$ rather than the properties of the single operator T. Let

$$\sigma_p := \Sigma_p/\Sigma_p^{(0)}, \quad 0 < p < \infty.$$

We consider these spaces in more detail.

The space σ_p is a linear space endowed with the quotient-metric $[\cdot]_{\sigma_p}^{p/p+1}$, where

$$[t]_{\sigma_p} := \inf\{|T|_p : T \in t\}, \quad t \in \sigma_p. \tag{19}$$

Since by Theorem 2 Σ_p is a complete space and Σ_p^{0} is closed in Σ_p, the space σ_p is complete. The *Calkin algebra* $\sigma_\infty = \mathbf{B}(H)/\mathbf{S}_\infty$ can be considered as the 'limit case' of the spaces σ_p as $p \to \infty$.

THEOREM 10. *The following equality holds*

$$[t]_p^p = \Delta_p(T), \quad \forall T \in t, t \in \sigma_p. \tag{20}$$

Proof. Equality (16) for $D_p = \Delta_p$ shows that the value of Δ_p does not depend on the choice of $T \in t$. Therefore the inequality $\geq$ in (20) follows directly from (19) and (10).

Let now $T \in t$ and $\gamma > \Delta_p(T)$. Then there exists $\bar{s} > 0$ such that $\pi_T(s) \leq \gamma \cdot s^{-p}$ for $s \leq \bar{s}$. In the Schmidt expansion of T (1.3) we replace all s_k, satisfying $s_k > \bar{s}$, by $\bar{s}$. Let $\tilde{T}$ be the operator obtained. Then $\tilde{T} \in t$ and $\pi_{\tilde{T}}(s) = 0$ for $s \geq \bar{s}$, $\pi_{\tilde{T}}(s) = \pi_T(s)$ for $s < \bar{s}$. Therefore $|\tilde{T}|_p^p < \gamma$, hence $[t]_p^p \leq \gamma$. Since γ is an arbitrary number greater than $\Delta_p(T)$, we have $[t]_p^p \leq \Delta_p(T)$. □

In view of Theorems 8 and 9 for $t_1 \in \sigma_{p_1}$, $t_2 \in \sigma_{p_2}$ the product $t_1t_2 \in \sigma_p$, $p^{-1} = p_1^{-1} + p_2^{-1}$ is correctly defined as the coset which contains one of the operators $T = T_1T_2$, $T_1 \in t_1$, $T_2 \in t_2$ (and then all such operators). Inequality (18) implies the following estimate (with the same c).

$$[t_1t_2]_p \leq c[t_1]_{p_1} \cdot [t_2]_{p_2}.$$

Since $s_k(T^*) = s_k(T)$, the involution $t^* := \{T^* : T \in t\}$ is defined on σ_p. If $t^* = t$ then it is natural to call t self-adjoint. An element $t \in \sigma_p$ is self-adjoint if and only if it contains a self-adjoint operator. Indeed, if $t = t^*$ then $T - T^* \in \Sigma_p^{(0)}$, $\forall T \in t$, and then $(T + T^*)/2 \in t$ is self-adjoint. The converse is evident.

For $t \in \sigma_p$ put

$$\Delta_p(t) = \Delta_p(T), \qquad \delta_p(t) = \delta_p(T), \quad \forall T \in t.$$

This definition is correct in view of Theorem 8. For $t^* = t$ the functionals $\Delta_p^{\pm}(t)$, $\delta_p^{\pm}(t)$ are defined similarly. Inequalities (14) and (15) directly imply the following result.

THEOREM 11. *The functionals Δ_p, δ_p are continuous on σ_p. The functionals $\Delta_p^{\pm}$, $\delta_p^{\pm}$ are continuous on the set of self-adjoint elements in σ_p. Each of the sets*

$\{t \in \sigma_p\colon \Delta_p(t) = \delta_p(t)\}$,

$\{t = t^* \in \sigma_p\colon \Delta_p^+(t) = \delta_p^+(t)\}$,

$\{t = t^* \in \sigma_p\colon \Delta_p^-(t) = \delta_p^-(t)\}$

is closed in σ_p.

This assertion is useful when one investigates the spectral asymptotics making use of the tools of perturbation theory.

7. Lidskii's Theorem

1. It follows from (5.9) that for an operator $T \in \mathbf{S}_1$ the series

$$\Lambda(T) := \sum_k \lambda_k(T) \tag{1}$$

converges absolutely. Recall that the eigen-values $\lambda_k(T)$ are counted with their multiplicities. The value $\Lambda(T)$ is called *the spectral trace* of T. The following deep and somewhat unexpected result was proved for the first time by V. B. Lidskii.

THEOREM 1. *The following equality holds*

$$\Lambda(T) = \operatorname{Tr} T, \quad \forall T \in \mathbf{S}_1. \tag{2}$$

In the finite dimensional case (2) is a simple consequence of the Vieta theorem for the characteristic polynomials. The extension of (2) to the case $\dim H = \infty$ is far from trivial since it is not *a priori* clear whether or not the functional $\Lambda(T)$ is linear and continuous on $\mathbf{S}_1$. The difficulty of the problem is demonstrated, in particular, by the existence of nuclear Volterra operators, which have no eigen-values at all, and so according to (2) their trace must vanish.

The first proofs of Theorem 1, based on the theory of entire functions, were comparatively complicated (see [6], §3.8). Only recently simpler approaches appeared. The proof presented here is based essentially on some ideas by Leiterer and Pietsch [25].

2. LEMMA 2. *Equality* (2) *holds for* $T \in \mathbf{K}$.

Proof. Each operator $T \in \mathbf{K}$ has the form $T = T_1 \oplus 0$, where T_1 is an operator on the finite dimensional space $H_1 = R(T) + R(T^*)$ (see §3.9, Sub-§2). Equality (2) holds for T_1 and so for T. □

To extend the result to the case of an arbitrary $T \in \mathbf{S}_1$ we obtain an estimate of the remainder of the series (1) in terms of

$$Q_m(T) := \sum_{n \geqslant m} \Big(n^{-1} \sum_{k=1}^{n} \sqrt{s_k(T)}\Big)^2, \quad m = 1, 2, \ldots \tag{3}$$