JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematics and Mathematical Physics
Vol. 68B, No. 4, October–December 1964

The Greatest Crossnorm

Robert Schatten [1]

(September 8, 1964)

Let $\mathfrak{H}$ be a complex Hilbert space. If T is a completely continuous operator on $\mathfrak{H}$ then $(T^*T)^{1/2}$ is also completely continuous and nonnegative. If $\lambda_1, \lambda_2, \ldots$ represent all the nonzero eigenvalues of $(T^*T)^{1/2}$—each eigenvalue repeated in the sequence the number of times equal to its multiplicity—we may form the sum $\Sigma_i \lambda_i$ which we denote by $\tau(T)$. By definition, the *trace-class* (τc) consists of all those operators T for which $\tau(T)$ is finite. (τc) forms a linear space and $\tau(T)$ defines there a norm. The resulting normed linear space turns out to be complete, and the operators of finite rank form a dense set in (τc).

It is of significance to observe that for operators T of finite rank, $\tau(T)$ may be also expressed via concepts meaningful in a perfectly general Banach space. This observation permits then to carry over to perfectly general Banach spaces the concept of a trace-class of operators: One considers the linear space of all the operators T of finite rank on the given Banach space. There one defines $\tau(T)$ via the concepts meaningful in general Banach spaces. The customary metric completion of the so resulting normed linear space furnishes then the desired trace-class of operators.

1. Introduction

Let $\mathfrak{H}$ be a complex Hilbert space. If T is a completely continuous operator on $\mathfrak{H}$, then $(T^*T)^{1/2}$ is not only completely continuous but also nonnegative (hence Hermitean). If $\lambda_1, \lambda_2, \ldots$ represent all the nonzero eigenvalues of $(T^*T)^{1/2}$—each eigenvalue repeated in the sequence the number of times equal to its multiplicity—we may form the sum $\Sigma_i\lambda_i$ which—to indicate its dependence on the operator T—will be also denoted by $\tau(T)$. By definition, the *trace-class* (τc) consists of all those operators T for which $\tau(T)$ is finite. It is not a simple argument to prove that (τc) forms a linear space and that $\tau(T)$ defines there a norm. Incidentally, the resulting normed linear space turns out to be complete, that is, forms a Banach space. It is also true that the operators of finite rank form a dense set in (τc).

We remark that for an operator T, the above definition of $\tau(T)$ involves notions which are meaningful only in linear spaces with an inner product. It is of interest—and in fact of significance—to observe that for operators T of finite rank, $\tau(T)$ may be also expressed via concepts meaningful in a perfectly general Banach space. This means, for operators T on $\mathfrak{H}$ of finite rank, we have two versions for $\tau(T)$. While one immediately carries over to arbitrary Banach spaces, the other does not yield to a straightforward generalization. This observation permits one to carry over to perfectly general Banach spaces the concept of a trace-class of operators. To define the last, one simply proceeds as follows: One considers the linear space of all the operators T of finite rank on the given Banach space. There one defines $\tau(T)$ via the concepts meaningful in general Banach spaces. The customary metric completion of the resulting normed linear space furnishes then the desired trace-class of operators.

It remains thus to sketch how, for operators T of finite rank, $\tau(T)$ may be expressed via concepts meaningful in any Banach space. The argument follows: If $f_1, \ldots, f_n$ and $g_1, \ldots, g_n$ are elements in $\mathfrak{H}$, then

$$Tg = \sum_{i=1}^{n} (g, g_i) f_i$$

represents an operator T of finite rank which we shall also denote symbolically by $\sum_{i=1}^{n} f_i \otimes \overline{g_i}$. The converse is also true, that is, every operator T of finite rank admits many such representations of the form $\sum_{j=1}^{m} \varphi_j \otimes \overline{\chi_j}$; the number of terms m will vary of course with the representation of T. It can be shown that

$$\tau(T) = \inf \sum_{j=1}^{m} \|\varphi_j\|\,\|\chi_j\|$$

where the above infimum extends over the set of all sums corresponding to all possible representations of the operator T of finite rank.

The details of all that was said above form the main goal of this exposition.

[1] Department of Mathematics, Hunter College of the City University of New York. This exposition covers the topics discussed in a Seminar held at the Computation Laboratory of the National Bureau of Standards, Washington, D.C. during July, 1964.

2. The Greatest Crossnorm

The last formula expresses $\tau(T)$ in a language meaningful for any Banach space. This suggests *the desirability of investigating the above infimum for general normed linear spaces* (not necessarily inner product spaces.) The details follow:

Let $\mathscr{F}$ and $\mathscr{G}$ stand for any two normed linear spaces and $\mathscr{F}^*$, $\mathscr{G}^*$ represent their conjugate spaces, that is, the corresponding spaces of additive and bounded, that is, continuous, linear functionals on $\mathscr{F}$ and $\mathscr{G}$ respectively.

For a fixed pair of elements $G_0 \epsilon \mathscr{G}^*$ and $f_0 \epsilon \mathscr{F}$, the expression

$$Tg = G_0(g)f_0$$

defines an operator of rank 1 from $\mathscr{G}$ into $\mathscr{F}$. More generally, if $G_1, \ldots, G_n$ are in $\mathscr{G}^*$ and $f_1, \ldots, f_n$ are in $\mathscr{F}$, then

$$Tg = \sum_{i=1}^{n} G_i(g)f_i$$

represents an operator from $\mathscr{G}$ into $\mathscr{F}$ of rank at most n. The last operator we shall also denote by $\sum_{i=1}^{n} f_i \otimes G_i$; one calls then $\sum_{i=1}^{n} f_i \otimes G_i$ a representation of the operator T. It is not difficult to see that conversely, every operator T from $\mathscr{G}$ into $\mathscr{F}$ of finite rank has many such representations. With each representation $\sum_{i=1}^{n} f_i \otimes G_i$ of a given operator T of finite rank, one associates the number $\sum_{i=1}^{n} \|f_i\| \, \|G_i\|$ and then defines

$$\gamma(T) = \inf \sum_{i=1}^{n} \|f_i\| \, \|G_i\|$$

where the infimum extends over the set of numbers corresponding to all possible representations of T. We prove below that the so defined $\gamma(T)$ is a norm on the linear space of all operators from $\mathscr{G}$ into $\mathscr{F}$ of finite rank. The metric completion of the above normed linear space is then defined as the trace-class of operators from $\mathscr{G}$ into $\mathscr{F}$.

As was already pointed out in the Introduction, one of our objectives is to show, that for the special case when both $\mathscr{F}$ and $\mathscr{G}$ represent a Hilbert space $\mathfrak{H}$, then $\gamma(T) = \tau(T)$ for every operator T on $\mathfrak{H}$ of finite rank.

Similarly, if $f_1, \ldots, f_n$ are in $\mathscr{F}$ and $g_1, \ldots, g_n$ are in $\mathscr{G}$, the expression $\sum_{i=1}^{n} f_i \otimes g_i$ represents an operator T of finite rank from $\mathscr{G}^*$ into $\mathscr{F}$ whose defining equation is given by $T(G) = \sum_{i=1}^{n} G(g_i)f_i$. Moreover, every operator T of finite rank from $\mathscr{G}^*$ into $\mathscr{F}$ has many such representations.

At this stage one should ask the following question: Assume that $f_1, \ldots, f_n$ are in $\mathscr{F}$ and $g_1, \ldots, g_n$ are in $\mathscr{G}$. Assume also that $f'_1, \ldots, f'_m$ are in $\mathscr{F}$ and $g'_1, \ldots, g'_m$ are in $\mathscr{G}$. When do the expressions $\sum_{i=1}^{n} f_i \otimes g_i$ and $\sum_{j=1}^{m} f'_j \otimes g'_j$ represent the same operator of finite rank from $\mathscr{G}^*$ into $\mathscr{F}$. To answer this question observe (writing $f_1 \otimes g_1 + \ldots + f_n \otimes g_n$ instead of $\sum_{i=1}^{n} f_i \otimes g_i$) that

$$f_1 \otimes g_1 + f_2 \otimes g_2 + \ldots + f_n \otimes g_n = f_{1'} \otimes g_{1'} + f_{2'} \otimes g_{2'} + \ldots + f_{n'} \otimes g_{n'} \quad (1)$$

where $1', 2', \ldots, n'$ is any permutation of the $1, 2, \ldots, n$.

$$f_1 \otimes (g_1 + \tilde{g}_1) + f_2 \otimes g_2 + \ldots + f_n \otimes g_n = f_1 \otimes g_1 + f_1 \otimes \tilde{g}_1 + f_2 \otimes g_2 + \ldots + f_n \otimes g_n. \quad (2a)$$

$$(f_1 + \tilde{f}_1) \otimes g_1 + f_2 \otimes g_2 + \ldots + f_n \otimes g_n = f_1 \otimes g_1 + \tilde{f}_1 \otimes g_1 + f_2 \otimes g_2 + \ldots + f_n \otimes g_n. \quad (2b)$$

$$(a_1 f_1) \otimes g_1 + (a_2 f_2) \otimes g_2 + \ldots + (a_n f_n) \otimes g_n = f_1 \otimes (a_1 g_1) + f_2 \otimes (a_2 g_2) + \ldots + f_n \otimes (a_n g_n) \quad (3)$$

where $a_1, a_2, \ldots, a_n$ are arbitrary scalars.

It is not difficult to see that two expressions $\sum_{i=1}^{n} f_i \otimes g_i$ and $\sum_{j=1}^{m} f'_j \otimes g'_j$ define the same operator T of finite rank if and only if one can be derived from the other by a finite number of successive applications of the above relations 1, 2a, 2b, 3.

The problem is to furnish some "enlightening" information concerning $\gamma(T) = \inf \sum_{i=1}^{n} \|f_i\| \, \|g_i\|$ where the infimum is extended over the set of sums corresponding to all possible representations $\sum_{i=1}^{n} f_i \otimes g_i$ of the given operator T of finite rank. In other words, we are interested in some characterizations of the above infimum. Perhaps in particular Banach spaces, it is possible to express $\gamma(T)$ *directly in terms of* T just as it was done in the case of completely continuous operators T on a Hilbert space. In the last, $\gamma(T) = \tau(T) = \sum_i \lambda_i$, where the λ_i represent all the nonzero eigenvalues of $(T^*T)^{1/2}$; each eigenvalue appearing in the last sum the number of times equal to its multiplicity.

In this connection one is tempted to add an additional problem: Suppose that $\mathscr{F}_1$ and $\mathscr{G}_1$ are subspaces of the normed linear spaces $\mathscr{F}$ and $\mathscr{G}$. If $f_1, \ldots, f_n$ are in $\mathscr{F}_1$ and $g_1, \ldots, g_n$ are in $\mathscr{G}_1$, then $\sum_{i=1}^{n} f_i \otimes g_i$ may be considered as an operator T_1 from $\mathscr{G}_1^*$ into $\mathscr{F}_1$, as well as an operator T from $\mathscr{G}^*$ into $\mathscr{F}$. Since all the representations of T_1 are among the representations of

T, it is clear that $\gamma(T) \leq \gamma(T_1)$. When are they equal? The preceding discussion and the equality $\gamma(T) = \tau(T)$ which we promise to prove later, imply that this is always the case when both $\mathscr{F}$ and $\mathscr{G}$ are Hilbert spaces.

EXAMPLE Let $\mathscr{L}$ stand for the linear space of all continuous functions $f(s)$ on $0 \leq s \leq 1$. There,

$$\|f\| = \max_{0 \leq s \leq 1} |f(s)|$$

defines a norm. Moreover, the resulting normed linear space is complete.

If $f(s)$ and $g(s)$ are both in $\mathscr{L}$, then the "product" $f(s)g(t)$ is defined and continuous on the square $0 \leq s,\ t \leq 1$. A function $K(s, t)$ defined on $0 \leq s, t \leq 1$ will be termed "degenerate" if it can be represented as a finite sum of products, that is, if it admits a representation of the form

$$K(s, t) = \sum_{i=1}^{n} f_i(s) g_i(t)$$

with $f_1(s), \ldots, f_n(s)$ and $g_1(s), \ldots, g_n(s)$ in $\mathscr{L}$.

It is not difficult to see that not every continuous function $K(s, t)$ is degenerate. It is also clear that every degenerate $K(s, t)$ admits an infinite number of representations as a finite sum of products; the number of terms in each sum will vary with the particular representation.

For each representation $\sum_{i=1}^{n} f_i(s) g_i(t)$ of a degenerate $H(s,t)$ form the number

$$\sum_{i=1}^{n} \|f_i\| \|g_i\| = \sum_{i=1}^{n} \max_{0 \leq s \leq 1} |f_i(s)| \cdot \max_{0 \leq s \leq 1} |g_i(s)|$$

and then consider the infimum over the set of numbers so obtained corresponding to all possible representations of the degenerate $H(s, t)$.

We are interested in some characterizations of the above infimum. In particular, we wonder whether the above infimum can be expressed *directly* in terms of a given (degenerate) $H(s, t)$.

THEOREM 1. Let $\mathscr{F}$ and $\mathscr{G}$ represent two normed linear spaces and $\mathfrak{R}$ stand for the linear space of all operators from $\mathscr{G}^*$ into $\mathscr{F}$ of finite rank. Then, the above defined function $\gamma(T)$ represents a norm on $\mathscr{R}$. Also, $\gamma(T)$ has the "cross property," that is, $\gamma(T)$ coincides with the bound $\|T\|$ for all operators T of rank ≤ 1. The last condition characterizes $\gamma(T)$ completely in the following sense: $\gamma(T)$ is the greatest norm on $\mathscr{R}$ having the cross property.

PROOF: Let $T \epsilon \mathscr{R}$ and $G \epsilon \mathscr{G}^*$. Then for any representation $\sum_{i=1}^{n} f_i \otimes g_i$ of T we have

$$\|T(G)\| = \|\sum_{i=1}^{n} G(g_i) f_i\| \leq \|G\| \sum_{i=1}^{n} \|f_i\| \|g_i\|.$$

Thus,

$$\|T(G)\| \leq \|G\| \gamma(T).$$

The last inequality implies

$$\|T\| \leq \gamma(T)$$

for all operators T in $\mathscr{R}$.

(i). If $T=0$, then obviously $\gamma(T)=0$. If $T \neq 0$, then $0 < \|T\| \leq \gamma(T)$, and thus $\gamma(T) > 0$.

(ii). It is also clear that for any scalar a we have $\gamma(aT) = |a| \gamma(T)$.

(iii). To prove that $\gamma(T_1 + T_2) \leq \gamma(T_1) + \gamma(T_2)$ for any two operators T_1 and T_2 in $\mathscr{R}$, we argue as follows: Let $\epsilon > 0$ be given. Choose a representation $\sum_{i=1}^{n} f_i \otimes g_i$ of T_1, such that

$$\sum_{i=1}^{n} \|f_i\| \|g_i\| \leq \gamma(T_1) + \frac{\epsilon}{2}.$$

Similarly, we can find a representation $\sum_{j=1}^{m} f'_j \otimes g'_j$ of T_2 such that

$$\sum_{j=1}^{m} \|f'_j\| \|g'_j\| \leq \gamma(T_2) + \frac{\epsilon}{2}.$$

But, then $\sum_{i=1}^{n} f_i \otimes g_i + \sum_{j=1}^{m} f'_j \otimes g'_j$ is a representation for $T_1 + T_2$, and therefore

$$\gamma(T_1 + T_2) \leq \sum_{i=1}^{n} \|f_i\| \|g_i\| + \sum_{j=1}^{m} \|f'_j\| \|g'_j\| \leq \gamma(T_1) + \gamma(T_2) + \epsilon.$$

(iv). Assume now that T is of rank ≤ 1. Then T admits a representation in the form $T = f \otimes g$, and thus

$$\gamma(T) \leq \|f\| \|g\| = \|T\|.$$

We already know that $\|T\| \leq \gamma(T)$ holds. Thus $\gamma(T) = \|T\|$.

Finally, suppose that for a norm α we have $\alpha(f \otimes g) = \|f\| \|g\|$. Let $T \epsilon \mathscr{R}$. For every representation $\sum_{i=1}^{n} f_i \otimes g_i$ of T we have

$$\alpha(T) = \alpha\left(\sum_{i=1}^{n} f_i \otimes g_i\right) \leq \sum_{i=1}^{n} \alpha(f_i \otimes g_i) = \sum_{i=1}^{n} \|f_i\| \|g_i\|.$$

Thus, $\alpha(T) \leq \gamma(T)$.

3. Some Preliminaries on Hilbert Space

The setting for our discussion is a fixed complex linear space with a complex inner product (f, g). An inner product generates a norm $\|f\|=(f, f)^{1/2}$. Of course, every norm $\|f\|$ generates a metric; the distance of two vectors f and g being given by $\|f-g\|$. We shall assume thus that our space is a Hilbert space $\mathfrak{H}$, that is, a linear space with an inner product for which the resulting metric space is complete.

In an inner product space, the following identities may be readily verified to hold for any pair of vectors f and g:

$$\|f+g\|^2+\|f-g\|^2=2\|f\|^2+2\|g\|^2,$$

$$4(f, g)=\|f+g\|^2-\|f-g\|^2+i\|f+ig\|^2-i\|f-ig\|^2.$$

The first is known as the parallelogram law. The second is known as the polarization identity.

The parallelogram law suggests that the norm on a Hilbert space is of a special kind. Obviously, not every norm on a linear space satisfies this law, and thus could not be obtained from an inner product. An early result of P. Jordan and J. von Neumann states that a Banach space (normed linear space for which the resulting metric space is complete) is a Hilbert space, that is, has a norm derivable from an inner product, if and only if the norm satisfies the parallelogram law.

A *basis* $\{\varphi_j\}$ in $\mathfrak{H}$ is by definition a maximal orthonormal family of vectors; the index set J of subscripts j is not necessarily countable. It is a consequence of Zorn's lemma (which is equivalent to Zermelo's axiom of choice) that in every Hilbert space there are bases. The cardinal number of elements in any two bases of the given space is always the same and defines the dimension of the space. Any orthonormal family of vectors in $\mathfrak{H}$ can be made part of a basis.

We mention in passing that every finite-dimensional normed linear space generates a necessarily complete metric space. Moreover, any two norms on the same finite-dimensional linear space generate equivalent topologies. More precisely: Any two normed linear spaces of the same finite dimension (and over the same scalar field) are necessarily linearly homeomorphic.

Also the following result due to Klee points out a striking difference between a finite and an infinite dimensional Hilbert space: Every infinite-dimensional Hilbert space is homeomorphic to the surface of its unit sphere, that is to the set of all f's such that $\|f\|=1$.

To include in our discussion nonseparable spaces, that is, spaces whose bases are uncountable, we adhere to the following convention: For a family $\{\alpha_j\}$ of nonnegative numbers (where j varies over some possibly uncountable set of indices), we shall write $\sum_j \alpha_j=\alpha$ if $\alpha_j\neq 0$ for at most a countable number of indices j, and the sum formed from all the nonzero α_j, converges to α. In all other cases, we shall write $\sum_j \alpha_j=+\infty$.

Bessel's inequality

$$\sum_{i=1}^{k}|(f, \chi_i)|^2\leq\|f\|^2$$

valid for any finite orthonormal family of vectors $\{\chi_1, \ldots, \chi_k\}$, implies that if $\{\varphi_j\}$ is a basis, then for any $f\epsilon\mathfrak{H}$, we have $(f, \varphi_j)\neq 0$ for at most a countable number of subscripts j.

The following conditions on an orthonormal family of vectors $\{\varphi_j\}$ are equivalent:

1. $\{\varphi_j\}$ is a basis in $\mathfrak{H}$.
2. $(f, \varphi_j)=0$ for all j, implies $f=0$.
3. For every $f\epsilon\mathfrak{H}$ we have the Fourier expansion

$$f=\sum_j (f, \varphi_j)\varphi_j.$$

4. For every pair f, g in $\mathfrak{H}$ we have Parseval's identity:

$$(f,g)=\sum_j (f,\varphi_j)(\varphi_j,g).$$

5. For every $f\epsilon\mathfrak{H}$ we have

$$\|f\|^2=\sum_j |(f, \varphi_j)|^2.$$

We shall consider exclusively linear transformations defined on all of $\mathfrak{H}$ and having their range included in $\mathfrak{H}$. For a linear transformation A,

$$\sup_{\|f\|\leq 1}\|Af\|=\sup_{\substack{\|f\|\leq 1\\ \|g\|\leq 1}}|(Af, g)|$$

defines its bound and is denoted by $\|A\|$. We say that A is bounded if $\|A\|<+\infty$. Boundedness is equivalent to continuity. Every linear transformation of finite rank (i.e., whose range is finite dimensional), and in particular every linear transformation on a finite dimensional space, is necessarily bounded. A bounded linear transformation will be termed an *operator*. The identity operator will be denoted by I. An operator A is invertible if there is an operator B such that $AB=BA=I$. The set of complex numbers λ for which $A-\lambda I$ is not invertible defines the spectrum of A; the last is always a closed subset of the disk $|z|\leq\|A\|$ and includes, of course, all the characteristic values of A, and perhaps other complex numbers.

If A is an operator, then there exists a unique operator A^* called the adjoint of A, such that

$$(Af, g)=(f,A^*g)$$

for all pairs of vectors f,g in $\mathfrak{H}$. We have $\|A\|=\|A^*\|$ and $\|A^*A\|=\|A\|^2$. The condition $A=A^*$ defines a Hermitean operator. If A is Hermitean, then its spectrum is a subset of the real axis. The operator A is positive—in symbols $A \geqslant 0$—if $(Af,f) \geqslant 0$ for all f in $\mathfrak{H}$. Since $\mathfrak{H}$ is assumed to be a complex space, every positive operator A is necessarily Hermitean. An operator A is normal if $A^*A = AA^*$; we reserve the letter U for unitary operators, i.e., such that $U^*U = UU^* = I$.

For every positive operator A there is one and only one positive operator B such that $A = B^2$; we write $B = A^{1/2}$. In particular, for every operator A, the operator $(A^*A)^{1/2}$ is well defined; we find it more convenient to write $[A]$ instead of $(A^*A)^{1/2}$. Clearly, $[A] = [A^*]$ if and only if A is normal. If A is of finite rank (that is, has a finite dimensional range) then both A^*A and AA^* and therefore also $[A]$ and $[A^*]$ are of finite rank. We have $\|[A]\| = \|A\|$.

The polar decomposition of operators will play a central part in our later discussion. To state it, define an operator W to be partially isometric if W is isometric on a (closed) subspace $\mathfrak{M}$ of $\mathfrak{H}$ and equal to 0 on its orthogonal complement; $\mathfrak{M}$ is the initial set of W, while the range of W is its final set.

The polar decomposition. Let A be an operator. There exists a partially isometric operator W whose initial set is the closure of the range of $[A]$ and whose final set is the closure of the range of A, satisfying the following relations:

(i) $A = W[A]$,
(ii) $[A] = W^*A$,
(iii) $A^* = W^*[A^*]$,
(iv) $[A^*] = W[A]W^*$.

The above decomposition is unique in the following sense: If $A = W_1B_1$ where $B_1 \geqslant 0$ and W_1 is partially isometric having as its initial set the closure of the range of B_1, then $B_1 = [A]$ and $W_1 = W$.

In the case A is of finite rank, we may assume that W is unitary (not necessarily unique, however).

Our exposition will deal almost exclusively with completely continuous operators. Recall that a sequence $\{f_n\}$ of elements in $\mathfrak{H}$ is said to be weakly convergent in f, in symbols $f_n \rightharpoonup f$, if $(f_n,g) \to (f,g)$ for all $g \epsilon \mathfrak{H}$; as usual $f_n \to f$ will symbolize strong convergence, that is, $\|f_n - f_n\| \to 0$.

An operator A is termed completely continuous if for every bounded sequence of vectors $f_1, f_2, f_3, \ldots$ the transformed sequence $Af_1, Af_2, Af_3, \ldots$ contains a subsequence convergent (in the strong sense) to some element of the space. Equivalently, an operator A is completely continuous if it transforms every weakly convergent sequence of elements into a strongly convergent sequence, that is, $f_n \rightharpoonup f$ implies $Af_n \to Af$.

It is an immediate consequence of the polar decomposition, that one of the operators A, A^*, $[A]$, $[A^*]$ is completely continuous if and only if the same is true for the remaining three.

Definition: If φ and χ are two elements of $\mathfrak{H}$ let $\varphi \otimes \bar{\chi}$ represent the operator whose defining equation is given by

$$(\varphi \otimes \bar{\chi})f = (f, \chi)\varphi$$

for all f in $\mathfrak{H}$.

Clearly, the transformation $\varphi \otimes \bar{\chi}$ defined above, is an operator; its bound $\|\varphi \otimes \bar{\chi}\| = \|\varphi\|\,\|\chi\|$. The range of $\varphi \otimes \bar{\chi}$ is of dimension 1 or 0.

Remark. Observe the slight change of notation from the one introduced in 2. Riesz' representation theorem states that a linear functional $F_0(\varphi)$ defined for all $\varphi \epsilon \mathfrak{H}$, is bounded if and only if there is a unique $\chi_0 \epsilon \mathfrak{H}$ such that

$$F_0(\varphi) = (\varphi, \chi_0) \text{ for all } \varphi \epsilon \mathfrak{H}.$$

The correspondence $\chi_0 \leftrightarrow F_0$ is not an isomorphism but a conjugate-isomorphism between $\mathfrak{H}$ and $\mathfrak{H}^*$, because $\chi_0 \leftrightarrow F_0$ implies $\lambda\chi_0 \leftrightarrow \bar{\lambda}F_0$ for any complex λ. We shall thus write $\varphi \otimes \bar{\chi}$, to remind us that the χ plays the part of a bounded linear functional on $\mathfrak{H}$.

The following relations are immediate consequences of the definition of $\varphi \otimes \chi$:

(i) $(\varphi \otimes \bar{\chi})^* = \chi \otimes \bar{\varphi}$.
(ii) $(\lambda\varphi) \otimes \bar{\chi} = \lambda(\varphi \otimes \bar{\chi})$.
(ii′) $\varphi \otimes \overline{(\lambda\chi)} = \bar{\lambda}(\varphi \otimes \bar{\chi})$.
(iii) $(\varphi_1 + \varphi_2) \otimes \bar{\chi} = \varphi_1 \otimes \bar{\chi} + \varphi_2 \otimes \bar{\chi}$.
(iii′) $\varphi \otimes \overline{(\chi_1 + \chi_2)} = \varphi \otimes \bar{\chi}_1 + \varphi \otimes \bar{\chi}_2$.
(iv) $(\varphi_1 \otimes \bar{\chi}_1)(\varphi_2 \otimes \bar{\chi}_2) = (\varphi_2, \chi_1)\varphi_1 \otimes \bar{\chi}_2$.
(v) $A(\varphi \otimes \bar{\chi}) = (A\varphi) \otimes \bar{\chi}$.
(vi) $(\varphi \otimes \bar{\chi})A = \varphi \otimes \overline{(A^*\chi)}$.

The meaning of the symbol $\sum_{i=1}^{n} \lambda_i\varphi_i \otimes \bar{\chi}_i$ is then clear; it represents an operator of rank at most n (i.e., whose range is at most n-dimensional).

In general, analogous infinite sums have no meaning. However, the following theorem will be useful for all our purposes:

THEOREM 2. Let $\{\varphi_i\}$ and $\{\chi_i\}$ stand for any two orthonormal familes of vectors and $\{\mu_i\}$ a *bounded* family of complex numbers indexed by the same set of subscripts. Then,

$$Tf = \sum_i \mu_i(f, \chi_i)\varphi_i$$

is meaningful for every f in $\mathfrak{H}$ and represents an operator T which we shall also denote by

$$\sum_i \mu_i\varphi_i \otimes \bar{\chi}_i.$$

The bound of T is given by

$$\|T\| = \sup_i |\mu_i|.$$

PROOF: We have:

$$\sum_i |\mu_i(f, \chi_i)|^2 \leqslant$$

$$\sup_i |\mu_i|^2 \sum_i |(f, \chi_i)|^2 \leqslant \|f\|^2 \sup_i |\mu_i|^2.$$

Thus Tf is meaningful and

$$\|Tf\|^2 \leqslant \|f\|^2 \sup_i |\mu_i|^2.$$

This of course means that $\|T\| \leqslant \sup_i |\mu_i|$. On the other hand, we have $\|T\chi_i\| = |\mu_i|$. And therefore $\|T\| \geqslant \sup_i |\mu_i|$. Thus, $\|T\| = \sup_i |\mu_i|$.

The fundamental theorem of algebra implies that the characteristic equation of a complex matrix possesses at least one (in general, complex) root. It follows that an operator on a finite-dimensional (complex) space has at least one proper value. In general if $\lambda_1, \lambda_2, \ldots, \lambda_k$ are different proper values of A and $\varphi_1\ \varphi_2, \ldots, \varphi_k$ corresponding proper vectors (that is, $A\varphi_i = \lambda_i\varphi_i$) then $\varphi_1, \varphi_2, \ldots, \varphi_k$ are linearly independent. In the case an inner product space is infinite dimensional, it is always possible to construct completely continuous operators which do not have a single proper value. It may be added that thus far it is not known whether every operator T on a Hilbert space $\mathfrak{H}$ possesses a proper invariant subspace. We mean hereby, a subspace $\mathfrak{M}$ such that $0 \neq \mathfrak{M} \neq \mathfrak{H}$ and $T(\mathfrak{M}) \subset \mathfrak{M}$.

Added in proof: The September 1964 issue of the Bulletin of the American Mathematical Society, carries a research communication by Louis de Branges and James Rovnyak to the effect that every operator (linear and bounded) in a Hilbert space admits proper invariant subspaces.

The story is however quite different when one deals with Hermitean operators. The spectral theorem in its simplest version states that a *Hermitean operator* A *on a finite-dimensional unitary space admits a basis (orthonormal) in that space, made up of proper vectors of* A. We mean hereby, that there is an orthonormal basis $\varphi_1, \ldots \varphi_n$ for which $A\varphi_i = \lambda_i\varphi_i$ for $1 \leq i \leq n$, and thus

$$A = \sum_{i=1}^{n} \lambda_i \varphi_i \otimes \bar{\varphi}_i.$$

The corresponding proper values $\lambda_1, \ldots, \lambda_n$ are necessarily real.

Conversely, every sum of the above form with $\varphi_1, \ldots, \varphi_n$ orthonormal and $\lambda_1, \ldots, \lambda_n$ real, represents a Hermitean operator.

The infinite-dimensional extension of the above result which follows is also well known:

THEOREM 3. Every Hermitean completely continuous operator A on a Hilbert space admits in that space an orthonormal basis of characteristic vectors. The corresponding nonzero (necessarily real) proper values are each of finite multiplicity and may be arranged either in a finite or denumerably infinite sequence $\lambda_1, \lambda_2, \ldots$ (each nonzero characteristic value being repeated in the sequence the number of times equal to its miltiplicity) such that $\lambda_i \to 0$. If $\varphi_1, \varphi_2, \ldots$ is a corresponding orthonormal sequence of proper vectors (that is, $A\varphi_i = \lambda_i\varphi_i$ for $i = 1, 2, \ldots$), then

$$A = \sum_i \lambda_i \varphi_i \otimes \bar{\varphi}_i$$

Conversely, every sum of the above form, that is, with $\{\varphi_i\}$ orthonormal, λ_i real and $\lambda_i \to 0$, represents a completely continuous Hermitean operator.

The above yields a representation characterizing the completely continuous Hermitean operators. To obtain an analogous representation valid for all completely continuous operators, one makes use of the polar decomposition for operators.

THEOREM 4. An operator A is completely continuous if and only if it admits a *polar representation.*

$$A = \sum_i \lambda_i \varphi_i \otimes \bar{\chi}_i$$

where both $\{\varphi_i\}$ and $\{\chi_i\}$ are orthonormal sequences and the λ_i's are positive. The sum has either a finite or denumerably infinite number of terms. In the last case, we have also $\lambda_i \to 0$. The above representation is unique in the sense that the λ_i's are necessarily all the positive proper values (each represented in the sequence $\{\lambda_i\}$ the number of times equal to its multiplicity) of $[A]$.

PROOF: Since A is completely continuous, the same is true for $[A]$. Thus,

$$[A] = \sum_i \lambda_i \chi_i \otimes \bar{\chi}_i.$$

Now, if $A = W[A]$ is its polar decomposition, then W is isometric on the closed linear manifold determined by $\{\chi_i\}$. Thus, $\{W\chi_i\}$ is also an orthonormal family. Put $W\chi_i = \varphi_i$. Then

$$A = W[A] = W\left(\sum_i \lambda_i \chi_i \otimes \bar{\chi}_i\right)$$

$$= \sum_i \lambda_i (W\chi_i) \otimes \bar{\chi}_i = \sum_i \lambda_i \varphi_i \otimes \chi_i.$$

4. The Schmidt-Class and the Trace-Class of Operators

LEMMA. Let A be a given operator and $\{\varphi_i\}$, $\{\chi_j\}$ be any two bases. Then

$$\sum_i \|A\varphi_i\|^2$$

$$\sum_i \sum_j |(A\varphi_i, \chi_j)|^2$$

$$\sum_j \|A^*\chi_j\|^2$$

represent the same (finite or infinite) value; we denote the last by $|A|^2$.

PROOF. For a fixed i, the "Pythagorean theorem" implies

$$\|A\varphi_i\|^2 = \sum_j |(A\varphi_i, \chi_j)|^2$$

and thus,

$$\sum_i \|A\varphi_i\|^2 = \sum_i \sum_j |(A\varphi_i, \chi_j)|^2.$$

Replacing in the last, A by A^*, $\{\varphi_i\}$ by $\{\chi_j\}$, and $\{\chi_j\}$

by $\{\varphi_i\}$ one gets

$$\sum_j \|A^*\chi_j\|^2 = \sum_i \sum_j |(A^*\chi_j, \varphi_i)|^2$$

$$= \sum_i \sum_j |(\chi_j, A\varphi_i)|^2 = \sum_i \sum_j |(A\varphi_i, \chi_j)|^2.$$

Thus, the values of the above three sums are the same. Observe finally that, $A^{**} = A$. This concludes the proof.

Definition. For a given operator A, let $|A|^2$ stand for the common value of the three "sums" determined above.

LEMMA. For any operator A, we have $\|A\| \leq |A|$.

PROOF. It is sufficient to prove that $\|A\varphi\| \leq |A|$ for every vector φ such that $\|\varphi\| = 1$. This is easy: Choose a basis $\{\varphi_i\}$ with φ as one of its elements. Then

$$\|A\varphi\|^2 \leq \sum_i \|A\varphi_i\|^2 = |A|^2.$$

LEMMA. For any two operators A and B we have

$$|A+B| \leq |A| + |B|.$$

PROOF. It is sufficient to consider the case in which both $|A| < +\infty$ and $|B| < +\infty$. Choose a fixed basis $\{\varphi_i\}$. Then

$$|A+B| = \left(\sum_i \|(A+B)\varphi_i\|^2\right)^{1/2}$$

$$\leq \left(\sum_i (\|A\varphi_i\| + \|B\varphi_i\|)^2\right)^{1/2}$$

$$\leq \left(\sum_i \|A\varphi_i\|^2\right)^{1/2} + \left(\sum_i \|B\varphi_i\|^2\right)^{1/2} = |A| + |B|.$$

THEOREM 5. Let $\mathfrak{B}$ stand for the set of all operators A for which $|A| < +\infty$. With the obvious definitions of addition and scalar multiplication, $\mathfrak{B}$ is a complex linear space. There $|A|$ represents a norm. The norm also satisfies the parallelogram law,

$$|A+B|^2 + |A-B|^2 = 2|A|^2 + 2|B|^2.$$

PROOF. That $|A|$ is a norm on $\mathfrak{B}$ follows from the preceding propositions. The last equality is also true, since if $\{\varphi_j\}$ is a fixed basis, then for every j

$$\|A\varphi_j + B\varphi_j\|^2 + \|A\varphi_j - B\varphi_j\|^2 = 2\|A\varphi_j\|^2 + 2\|B\varphi_j\|^2.$$

THEOREM 6. Let A and B be two operators in $\mathfrak{B}$ and $\{\varphi_j\}$ a given basis. Then

$$(A, B) = \sum_j (A\varphi_j, B\varphi_j)$$

is a well-defined complex number, independent of the chosen basis. Also, (A, B) defines an inner product on $\mathfrak{B}$ and thus

$$|A| = (A, A)^{1/2}$$

is the norm which it generates.

PROOF. For each j we have

$$4(A\varphi_j, B\varphi_j) = \|A\varphi_j + B\varphi_j\|^2$$

$$-\|A\varphi_j - B\varphi_j\|^2 + i\|A\varphi_j + iB\varphi_j\|^2 - i\|A\varphi_j - iB\varphi_j\|^2$$

and thus,

$$4(A, B) = |A+B|^2 - |A-B|^2 + i|A+iB|^2 - i|A-iB|^2.$$

Remark. Incidentally, the inner product space $\mathfrak{B}$ is also complete (hence a Banach space); it contains the operators of finite rank as a dense subset. Considering however, $\mathfrak{B}$ only as a *linear* set of operators on $\mathfrak{H}$, and defining the bound $\|A\|$ of an operator as a new norm on $\mathfrak{B}$, then the resulting normed linear space is *not* complete, (provided $\mathfrak{H}$ is infinite-dimensional).

Remark. Incidentally, $|A|$ also satisfies the following conditions:

(i) $|A| = |A^*|$.

(ii) For any operator B, we have

$$|AB| \leq |A|\ \|B\| \leq |A|\ |B|.$$

As a consequence of (ii), $\mathfrak{B}$ is also an algebra. In fact, $\mathfrak{B}$ is an ideal in the algebra of all operators.

Definition: An operator A in $\mathfrak{B}$ is commonly referred to as one which belongs to the *E. Schmidt-class* and $|A|$ is said to define its Hilbert-Schmidt norm.

Remark. That the Hilbert-Schmidt norm of an operator is always not smaller than its bound was proven above. Also, it can be readily verified that the Hilbert-Schmidt norm is a crossnorm; that is, $|A| = \|A\|$ wherever A is an operator of rank ≤ 1.

At this point, the following comment is in order: Let L_2 stand for the Hilbert space of all complex-valued Lebesgue measurable functions $f(x)$ defined on the interval $0 \leq x \leq 1$ for which $|f(x)|^2$ is integrable; two functions being considered identical if and only if they differ on a set of measure zero. There, the linear operations are the usual ones in function spaces; the inner product is represented by

$$(f, g) = \int f(x)\overline{g(x)}dx.$$

Similarly, let $\mathscr{L}_2$ represent the Hilbert space of complex-valued measurable functions $K(x,y)$ defined on $0 \leq x,\ y \leq 1$ for which $|K(x,y)|^2$ is integrable; the inner product being given by

$$(H, K) = \iint H(x, y)\,\overline{K(x, y)}\,dx\,dy.$$

One observes that if $K_1(x,y)$ and $K_2(x,y)$ are both in $\mathscr{L}_2$, then the function

$$H(x, y) = \int K_1(x, z)K_2(z, y)\,dz$$

is also in $\mathscr{L}_2$ and

$$\iint |H(x, y)|^2 dx\,dy$$

$$\leq \iint |K_1(x, y)|^2 dx\,dy \cdot \iint |K_2(x, y)|^2 dx\,dy.$$

744-268 O-65—4

Thus, if $H(x,y)$ is defined as the "product" of $K_1(x,y)$ and $K_2(x,y)$, the space $\mathscr{L}_2$ turns out to be an algebra.

Let $K(x,y)$ be a fixed element in $\mathscr{L}_2$. For $f(x)$ in L_2,

$$\int K(x, y) f(y)\, dy$$

is then defined for almost all x in $0 \leq x \leq 1$ and represents a function $g(x)$, again in L_2. It turns out that the equation

$$g(x) = \int K(x, y)\, f(y)\, dy$$

defines an operator K on L_2 which belongs to the Schmidt-class (of operators on L_2) and

$$|K| = \left(\int\int |K(x, y)|^2 dx\, dy\right)^{1/2}.$$

Moreover, every operator on L_2 in the Schmidt-class is obtained in such a manner. This one-to-one correspondence between $\mathscr{L}_2$ and the Schmidt-class of operators on L_2 preserves addition, scalar-multiplication, products, and the norm. This means we have the following:

THEOREM 7. The Schmidt-class of operators on L_2, and the Hilbert space $\mathscr{L}_2$, are congruent not only as Banach spaces but also as Banach algebras.

THEOREM 8. Every operator A in the Schmidt-class is necessarily completely continuous.

PROOF. Let $f_n \rightharpoonup f$. We prove the following: Given an $\epsilon > 0$, there is a natural number n_0 such that for all $n > n_0$ one has $\|Af_n - Af\| < \epsilon$.

Clearly for some constant c we have $\|f_n\| \leq c$ for all n and thus also $\|f\| \leq c$. Choose a basis $\{\varphi_j\}$. Since A is in the Schmidt-class, we can find a finite set J of indices j such that

$$\Sigma_{j \notin J} \|A\varphi_j\|^2 < \frac{\epsilon^2}{16\,c^2}.$$

Clearly,

$$f_n - f = \Sigma_j (f_n - f, \varphi_j)\varphi_j$$

and therefore,

$$Af_n - Af = \Sigma_j (f_n - f, \varphi_j) A\varphi_j.$$

Thus, for every natural n,

$$\|Af_n - Af\|^2 = \|\Sigma_{j\in J}(f_n - f, \varphi_j)A\varphi_j + \Sigma_{j\notin J}(f_n - f, \varphi_j)A\varphi_j\|^2$$

$$\leq 2\|\Sigma_{j\in J}(f_n - f, \varphi_j)A\varphi_j\|^2 + 2\|\Sigma_{j\notin J}(f_n - f, \varphi_j)A\varphi_j\|^2.$$

The second term on the right of the last inequality is

$$\leq 2\|f_n - f\|^2 \Sigma_{j\notin J}\|A\varphi_j\|^2 \leq 2(2c)^2\frac{\epsilon^2}{16c^2} = \frac{\epsilon^2}{2}.$$

On the other hand $f_n \rightharpoonup f$ implies $\lim_n (f_n - f, \varphi_j) = 0$ for all j. Thus, the sum

$$2\|\Sigma_{j\in J}(f_n - f, \varphi_j)A\varphi_j\|^2$$

having a finite number of terms can be made $< \frac{\epsilon^2}{2}$ for large enough n, say for all $n > n_0$. Therefore for all $n > n_0$, we have

$$\|Af_n - Af\|^2 < \epsilon^2.$$

The theorem which follows characterizes the completely continuous operators in the Schmidt-class.

THEOREM 9. A completely continuous operator $A = \Sigma_i \lambda_i \varphi_i \otimes \bar{\chi}_i$ belongs to the Schmidt-class if and only if $\Sigma_i \lambda_i^2 < +\infty$, that is, the series formed from the non-zero proper values of A^*A converges. In the last case we also have

$$|A| = (\Sigma_i \lambda_i^2)^{1/2}.$$

PROOF. We extend $\{\chi_i\}$ to a basis by adding $\{\omega_j\}$. Of course

$$\|A\chi_i\| = \|\lambda_i\varphi_i\| = \lambda_i \text{ and } \|A\omega_j\| = 0.$$

Hence,

$$|A|^2 = \Sigma_i \|A\chi_i\|^2 + \Sigma_j \|A\omega_j\|^2 = \Sigma_i \lambda_i^2.$$

COROLLARY. The operators of finite rank form a dense set in the Schmidt-class.

PROOF. Let $A = \Sigma_i \lambda_i \varphi_i \otimes \bar{\chi}_i$ with $\Sigma_i \lambda_i^2 < +\infty$. Let $A_n = \sum_{i=1}^{n} \lambda_i \varphi_i \otimes \bar{\chi}_i$.

Then,

$$\lim_n |A - A_n| = \lim_n (\Sigma_{i>n}\lambda_i^2)^{1/2} = 0.$$

LEMMA. Let A be a given operator and $\{\varphi_i\}$ a basis. Then

$$\Sigma_i([A]\varphi_i, \varphi_i)$$

is independent of the chosen basis.

PROOF. Since $[A] \geqslant 0$, there is a unique operator $B \geqslant 0$ such that $[A] = B^2$. Now,

$$([A]\varphi_i, \varphi_i) = (B^2\varphi_i, \varphi_i) = (B\varphi_i, B\varphi_i) = \|B\varphi_i\|^2$$

and $\Sigma_i \|B\varphi_i\|^2$ is of course independent of the chosen basis $\{\varphi_i\}$.

Definition. The operators A for which the sum in the preceding Lemma is finite, form the trace-class (τc).

THEOREM 10. Let (τc) denote the class of all operators A for which

$$\tau(A) = \Sigma_j([A]\varphi_j, \varphi_j) < +\infty$$

for a fixed basis $\{\varphi_j\}$. With the obvious definition of addition and scalar multiplication, (τc) is a linear space. The last will be normed if the above sum represents the norm of an operator A. Moreover, the resulting normed linear space is complete, hence a Banach space: it contains the operators of finite rank as a dense subset. The operators in the trace-class necessarily belong to the Schmidt-class, and thus are completely continuous. Moreover (τc) is a (two-sided) ideal in the algebra of all operators and a Banach algebra under its own norm.[2]

THEOREM 11. Every operator in the trace-class is necessarily in the Schmidt-class. Every operator in the Schmidt-class is completely continuous.

Let A be a completely continuous operator and $\Sigma_i \lambda_i \varphi_i \otimes \overline{\chi}_i$ its polar form. Then A is in the Schmidt-class if and only if $\Sigma_i \lambda_i^2 < +\infty$; we have $|A| = (\Sigma_i \lambda_i^2)^{1/2}$ The operator A is in the trace-class if and only if $\Sigma_i \lambda_i < +\infty$; we have $\tau(A) = \Sigma_i \lambda_i$.

THEOREM 12. For an operator T of finite rank

$$\tau(T) = \gamma(T) = \inf \sum_{j=1}^{m} \|f_j\| \, \|g_j\|$$

where the infimum is taken over the set of all numbers corresponding to all possible representations $\sum_{j=1}^{m} f_j \otimes \overline{g}_j$ of T.

PROOF. Since $\tau(T)$ is a crossnorm, for any representation $\sum_{j=1}^{m} f_j \otimes \overline{g}_j$ of T we have $\tau(T) \leqslant \sum_{j=1}^{m} \tau(f_j \otimes \overline{g}_j) = \sum_{j=1}^{m} \|f_j\| \, \|g_j\|$. Thus, $\tau(T) \leqslant \gamma(T)$. However, if $\sum_{i=1}^{n} \lambda_i \varphi_i \otimes \overline{\chi}_i$ is a polar representation of T, then

$$T^*T = \left(\sum_{i=1}^{n} \lambda_i \chi_i \otimes \overline{\varphi}_i \right) \left(\sum_{i=1}^{n} \lambda \varphi_i \otimes \overline{\chi}_i \right)$$

$$= \sum_{i=1}^{n} \lambda_i^2 \chi_i \otimes \overline{\chi}_i$$

and $[T] = \sum_{i=1}^{n} \lambda_i \chi_i \otimes \overline{\chi}_i$. Now, $[T]\varphi = 0$ whenever φ is orthogonal to $\chi_1, \ldots, \chi_n$. Thus,

$$\tau(T) = \sum_{i=1}^{n} ([T]\chi_i, \chi_i)$$

$$= \sum_{i=1}^{n} (\lambda_i \chi_i, \chi_i) = \sum_{i=1}^{n} \lambda_i$$

$$= \sum_{i=1}^{n} \|\lambda_i \varphi_i\| \, \|\chi_i\| \geqslant \gamma(T).$$

This concludes the proof.

(Paper 68B4–130)

[2] For the complete details of the proof of Theorem 10, the reader is advised to consult this author's monograph "Norm Ideals of Compietely Continuous Operators" Ergebnisse der Mathematik (Springer-Verlag, Berlin, 1960).