Concrete Operators, **9** (2022) 53–69

TRACE-CLASS AND NUCLEAR OPERATORS

CARLOS S. KUBRUSLY

Abstract. This paper explores the long journey from projective tensor products of a pair of Banach spaces, passing through the definition of nuclear operators still on the realm of projective tensor products, to the of notion of trace-class operators on a Hilbert space, and shows how and why these concepts (nuclear and trace-class operators, that is) agree in the end.

1. Introduction

This is an expository paper on trace-class and nuclear operators. Its purpose is to demonstrate that these classes of operators coincide on a Hilbert space. It will focus mainly on three points: (i) where these notions came from, (ii) how they are intertwined, and (iii) when they coincide. Being a plain expository paper, this has no intention to survey the subject, neither to offer an extensive bibliography on it.

To begin with, let us borrow the definitions of nuclear and trace-class operators from Sections 4 and 5 (where these definitions will be properly posed).

- Nuclear Operators. An operator T on a Banach space $\mathcal{X}$ is *nuclear* if there are $\mathcal{X}^*$-valued and $\mathcal{X}$-valued sequences $\{f_k\}$ and $\{y_k\}$ such that $\sum_k \|f_k\|\,\|y_k\| < \infty$ and $Tx = \sum_k f_k(x)\,y_k$ for every $x \in \mathcal{X}$.
- Trace-Class Operators. An operator T on a Hilbert space $\mathcal{X}$ is *trace-class* if $\sum_\gamma \big\langle |T|e_\gamma ; e_\gamma \big\rangle < \infty$ for an arbitrary orthonormal basis $\{e_\gamma\}$ for $\mathcal{X}$ and the series value does not depend on $\{e_\gamma\}$.

Perhaps the first lines in Robert Bartle's review of [9] might be a suitable start: "Grothendieck [8] showed that a Banach space $\mathcal{X}$ has the approximation property if and only if, for every nuclear operator $T\colon \mathcal{X} \to \mathcal{X}$ (i.e., operator having the form $T = \sum_k \langle f_k\,;\cdot\rangle y_k$) with $f_k \in \mathcal{X}^*$, $y_k \in \mathcal{X}$, and $\sum_k \|f_k\|\,\|y_n\| < \infty$), the number $\operatorname{tr}(T) = \sum_k \langle f_k, y_k\rangle$ is well-defined (i.e., is independent of the choice of $\{f_k\}$ and $\{y_k\}$ in the representation $T = \sum_k \langle f_k\,;\cdot\rangle y_k$) and can be used to define a trace function."

Bartle's concise description nicely summarizes the apparently long way to be covered from Grothendieck's projective tensor products (where nuclear operators originate from), to trace-class operators, and finally concluding that these classes coincide. The familiar notion of trace as the sum of eigenvalues is a fundamental result known as Lidskiĭ Theorem [17], [7, Theorem 8.4] which still remains an active research topic (e.g., [23, 11, 6, 3, 25, 20]).

These concepts (nuclear and trace-class) are linked together since their early days. Schatten in his celebrated 1950 monograph [27] (which actually is an offspring of his 1942 thesis) describes nuclear Hilbert-space operators as being precisely the trace-class: "The trace-class may be also interpreted as $\mathcal{X}\,\widehat{\otimes}_{\wedge}\,\mathcal{X}^*$" [27, Theorem 5.12] — the completion of the tensor product of a Hilbert with its dual, with respect to the greatest reasonable crossnorm (i.e., the projective norm).

Date: December 16, 2021; revised: April 11, 2022.
2010 *Mathematics Subject Classification.* 47B10, 47A80, 46M05.
Keywords. Nuclear operators, projective tensor product, trace-class operators.

The paper is organized as follows. Section 2 summarizes some common notation and terminology. Section 3 poses the necessary results on crossnormed tensor products of a pair of Banach spaces, since this is the proper setup where nuclear transformations come from. Nuclear transformations are defined in Section 4 (Theorem 4.1) as the range of a linear contraction of the projective tensor product $\mathcal{X}^*\,\widehat{\otimes}_{\wedge}\mathcal{Y}$ into the Banach space $\mathcal{B}[\mathcal{X},\mathcal{Y}]$ of all bounded linear transformations from $\mathcal{X}$ to $\mathcal{Y}$, which yields the characterization of nuclear operators mentioned above. Section 5 deals exclusively with trace-class operators on a Hilbert space, and gives a thorough view of basic properties of these operators. Section 6 shows in Theorem 6.1 that nuclear and trace-class operators in fact reduce to the same thing.

All terms and notation above will be defined here in due course.

2. Notation and Terminology

Throughout the paper all linear spaces are over the same field $\mathbb{F}$ (and the field $\mathbb{F}$ in this context means either $\mathbb{R}$ or $\mathbb{C}$). If $\mathcal{X},\mathcal{Y},\mathcal{Z}$ are linear spaces, then let $\mathcal{L}[\mathcal{X},\mathcal{Z}]$ and $b[\mathcal{X}\times\mathcal{Y},\mathcal{Z}]$ denote the linear spaces of all linear transformation of $\mathcal{X}$ into $\mathcal{Z}$, and of all bilinear maps of the Cartesian product $\mathcal{X}\times\mathcal{Y}$ into $\mathcal{Z}$. If $\mathcal{X},\mathcal{Y},\mathcal{Z}$ are normed spaces, then let $\mathcal{B}[\mathcal{X},\mathcal{Z}]$ denote the normed space of all bounded (i.e., continuous) linear transformations of $\mathcal{X}$ into $\mathcal{Z}$ equipped with its standard induced uniform norm, and let $\boldsymbol{b}[\mathcal{X}\times\mathcal{Y},\mathcal{Z}]$ denote the normed spaces of all bounded (i.e., continuous) bilinear maps of $\mathcal{X}\times\mathcal{Y}$ into $\mathcal{Z}$ equipped with its usual norm, which are both Banach spaces whenever $\mathcal{Z}$ is (see e.g., [16]).

A subspace of a normed space is a closed linear manifold of it. If $\mathcal{X}$ is a normed space, then $\mathcal{X}^*=\mathcal{B}[\mathcal{X},\mathbb{F}]$ stand for its dual; if M is a subset of an inner product space, then $M^\perp$ stands for the orthogonal complement of M. Range and kernel of a bounded linear transformation $T\in\mathcal{B}[\mathcal{X},\mathcal{Y}]$ between normed spaces $\mathcal{X}$ and $\mathcal{Y}$ will be denoted by $\mathcal{R}(T)$ — a linear manifold of $\mathcal{Y}$ — and $\mathcal{N}(T)$ — a subspace of $\mathcal{X}$ — respectively. If two normed spaces $\mathcal{X}$ and $\mathcal{Y}$ are isometrically isomorphic, and if $y\in\mathcal{Y}$ is the isometrically isomorphic image of $x\in\mathcal{X}$, then write $\mathcal{X}\cong\mathcal{Y}$ and $x\cong y$.

Let $\mathcal{X}$ and $\mathcal{Y}$ be arbitrary normed spaces. By an operator we mean a bounded linear transformation of a normed space into itself. Set $\mathcal{B}[\mathcal{X}]=\mathcal{B}[\mathcal{X},\mathcal{X}]$; the normed algebra of all operators on $\mathcal{X}$. Let $\mathcal{B}_0[\mathcal{X},\mathcal{Y}]$ and $\mathcal{B}_\infty[\mathcal{X},\mathcal{Y}]$ stand for the normed spaces of all bounded finite-rank (i.e., $\dim\mathcal{R}(T)<\infty$) and of all compact linear transformations $T\colon\mathcal{X}\to\mathcal{Y}$, respectively. Similarly, set $\mathcal{B}_0[\mathcal{X}]=\mathcal{B}_0[\mathcal{X},\mathcal{X}]$ and $\mathcal{B}_\infty[\mathcal{X}]=\mathcal{B}_\infty[\mathcal{X},\mathcal{X}]$; the ideals of the algebra $\mathcal{B}[\mathcal{X}]$ consisting of all bounded finite-rank and of all compact operators, respectively, so that $\mathcal{B}_0[\mathcal{X}]\subseteq\mathcal{B}_\infty[\mathcal{X}]\subseteq\mathcal{B}[\mathcal{X}]$. Moreover, let $\mathcal{B}_N[\mathcal{X},\mathcal{Y}]$ (accordingly, $\mathcal{B}_N[\mathcal{X}]=B_N[\mathcal{X},\mathcal{X}]$) stand for the normed spaces of all nuclear transformations. It is worth noticing that there are different, also common, notations such as $\mathcal{K}$ for compact, $\mathcal{F}$ for finite-rank, and $\mathcal{N}$ for nuclear operators — at the end of Section 5 it will become clear the reason for our choice in favor of the above sub-indexed-$\mathcal{B}$ notation (see also, e.g., [29, Sections 6.1 and 7.1]). A Banach space $\mathcal{Y}$ has the *approximation property* if $\mathcal{B}_0[\mathcal{X},\mathcal{Y}]$ is dense in $\mathcal{B}_\infty[\mathcal{X},\mathcal{Y}]$ for every normed space $\mathcal{X}$ — every Banach space with a Schauder basis has the approximation property, in particular, since the range of a compact linear transformation is separable, every Hilbert space has the approximation property.

3. Preliminaries on Crossnormed Tensor Product Spaces

The algebraic *tensor product* of linear spaces $\mathcal{X}$ and $\mathcal{Y}$ is a linear space $\mathcal{X}\otimes\mathcal{Y}$ for which there is a bilinear map $\theta\colon \mathcal{X}{\times}\mathcal{Y}\to\mathcal{X}\otimes\mathcal{Y}$ (called the *natural bilinear map* associated with $\mathcal{X}\otimes\mathcal{Y}$) whose range spans $\mathcal{X}\otimes\mathcal{Y}$ with the following additional property: for every bilinear map $\phi\colon \mathcal{X}{\times}\mathcal{Y}\to\mathcal{Z}$ into any linear space $\mathcal{Z}$ there exists a (unique) linear transformation $\Phi\colon \mathcal{X}\otimes\mathcal{Y}\to\mathcal{Z}$ for which the diagram

$$\begin{array}{ccc} \mathcal{X}{\times}\mathcal{Y} & \xrightarrow{\ \phi\ } & \mathcal{Z} \\ & {}_{\theta}\searrow & \uparrow{\scriptstyle\Phi} \\ & & \mathcal{X}\otimes\mathcal{Y} \end{array}$$

commutes. Set $x\otimes y=\theta(x,y)$ for each $(x,y)\in\mathcal{X}{\times}\mathcal{Y}$. These are the *single tensors*. An arbitrary element F in the linear space $\mathcal{X}\otimes\mathcal{Y}$ is a finite sum $\sum_i x_i\otimes y_i$ of single tensors, and the representation of F as a finite sum of single tensors $F=\sum_i x_i\otimes y_i$ is not unique. (For an exposition on algebraic tensor products see, e.g., [15].)

If $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces and $\mathcal{X}^*$ and $\mathcal{Y}^*$ are their duals, then let $x\otimes y$ and $f\otimes g$ be single tensors in the tensor product spaces $\mathcal{X}\otimes\mathcal{Y}$ and $\mathcal{X}^*\otimes\mathcal{Y}^*$. A norm $\|\cdot\|_\alpha$ on $\mathcal{X}\otimes\mathcal{Y}$ is a *reasonable crossnorm* if, for every $x\in\mathcal{X}$, $y\in\mathcal{Y}$, $f\in\mathcal{X}^*$, $g\in\mathcal{Y}^*$,

(a) $\|x\otimes y\|_\alpha\le\|x\|\,\|y\|$,

(b) $f\otimes g$ lies in $(\mathcal{X}\otimes\mathcal{Y})^*$, and $\|f\otimes g\|_{*\alpha}\le\|f\|\,\|g\|$ (where $\|\cdot\|_{*\alpha}$ is the norm on the dual $(\mathcal{X}\otimes\mathcal{Y})^*$ when $(\mathcal{X}\otimes\mathcal{Y})$ is equipped with the norm $\|\cdot\|_\alpha$),

so that $\mathcal{X}^*\otimes\mathcal{Y}^*\subseteq(\mathcal{X}\otimes\mathcal{Y})^*$. It can be verified that (i) $\|x\otimes y\|_\alpha=\|x\|\,\|y\|$ whenever $\|\cdot\|_\alpha$ is a reasonable crossnorm, and (ii) when restricted to $\mathcal{X}^*\otimes\mathcal{Y}^*$ the norm $\|\cdot\|_{*\alpha}$ on $(\mathcal{X}\otimes\mathcal{Y})^*$ is a reasonable crossnorm (with respect to $(\mathcal{X}^*\otimes\mathcal{Y}^*)^*$). Two special norms on $\mathcal{X}\otimes\mathcal{Y}$ are the so-called *injective* $\|\cdot\|_\vee$ and *projective* $\|\cdot\|_\wedge$ norms,

$$\|F\|_\vee=\sup_{\|f\|\le1,\,\|g\|\le1,\;f\in\mathcal{X}^*,\,g\in\mathcal{Y}^*}\Big|\sum\nolimits_i f(x_i)\,g(y_i)\Big|,$$

$$\|F\|_\wedge=\inf_{\{x_i\}_i,\,\{y_i\}_i,\;F=\sum_i x_i\otimes y_i}\sum\nolimits_i\|x_i\|\,\|y_i\|,$$

for every $F=\sum_i x_i\otimes y_i\in\mathcal{X}\otimes\mathcal{Y}$, where the infimum is taken over all (finite) representations of $F\in\mathcal{X}\otimes\mathcal{Y}$. It can also be shown that (iii) these are indeed norms on $\mathcal{X}\otimes\mathcal{Y}$, that (iv) both $\|\cdot\|_\vee$ and $\|\cdot\|_\wedge$ are reasonable crossnorms and, moreover, that (v) a norm $\|\cdot\|_\alpha$ on $\mathcal{X}\otimes\mathcal{Y}$ is a reasonable crossnorm if and only if

$$\|F\|_\vee\le\|F\|_\alpha\le\|F\|_\wedge\quad\text{for every}\quad F\in\mathcal{X}\otimes\mathcal{Y}.$$

Let $\mathcal{X}\otimes_\alpha\mathcal{Y}=(\mathcal{X}\otimes\mathcal{Y},\|\cdot\|_\alpha)$ stand for the tensor product space of a pair of Banach spaces equipped with a reasonable crossnorm $\|\cdot\|_\alpha$, which is not necessarily complete. Their completion (see, e.g., [13, Section 4.7]) is denoted by $\mathcal{X}\widehat{\otimes}_\alpha\mathcal{Y}$ (same notation $\|\cdot\|_\alpha$ for the extended norm on $\mathcal{X}\widehat{\otimes}_\alpha\mathcal{Y}$). In particular, $\mathcal{X}\widehat{\otimes}_\vee\mathcal{Y}$ and $\mathcal{X}\widehat{\otimes}_\wedge\mathcal{Y}$ are referred to as the *injective and projective tensor products*. For the theory of the Banach space $\mathcal{X}\widehat{\otimes}_\alpha\mathcal{Y}$ (in particular, $\mathcal{X}\widehat{\otimes}_\vee\mathcal{Y}$ and $\mathcal{X}\widehat{\otimes}_\wedge\mathcal{Y}$) see, e.g., [10], [2], [26], [4].

The next two fundamental results on the projective tensor product will be need along the next two sections. The first one is referred to as Grothendieck Theorem (see, e.g., [4, Proposition 1.1.4] and [26, Proposition 2.8]). In fact, most results mentioned in this section (and beyond) are Grothendieck's, and there are different theorems named after Grothendieck (the one in Theorem 3.1 below is not the classical Grothendieck Theorem as in, for instance, [4, Chapter 4] and [22]). The second result in Theorem 3.2 below is called the Universal Mapping Principle (see, e,g., [4, Theorem 1.1.8] and [26, Theorem 2.9]). Proofs are included for sake of completeness.

Theorem 3.1. (Grothendieck). *If $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces, then for every $\digamma \in \mathcal{X}\widehat{\otimes}_{\wedge}\mathcal{Y}$ there exist $\mathcal{X}$-valued and $\mathcal{Y}$-valued sequences $\{x_k\}$ and $\{y_k\}$, respectively, for which the real sequence $\{\|x_k\|\,\|y_k\|\}$ is summable and*

$$\digamma = \sum\nolimits_k x_k \otimes y_k$$

(i.e., $\digamma \in \mathcal{X}\widehat{\otimes}_{\wedge}\mathcal{Y}$ is representable in the form of a countable sum $\digamma = \sum_k x_k \otimes y_k$ in the sense that it is either a finite or countably infinite representation). Moreover, the projective norm $\|\cdot\|_{\wedge}$ on $\mathcal{X}\widehat{\otimes}_{\wedge}\mathcal{Y}$ is given by

$$\|\digamma\|_{\wedge} = \inf \sum\nolimits_k \|x_k\|\,\|y_k\|,$$

where the infimum is taken over all representations $\sum_k x_k \otimes y_k$ of $\digamma \in \mathcal{X}\widehat{\otimes}_{\wedge}\mathcal{Y}$.

Proof. Let $\mathcal{X}\widehat{\otimes}_{\wedge}\mathcal{Y}$ be the completion of $\mathcal{X}\otimes_{\wedge}\mathcal{Y}$. Identify $\mathcal{X}\otimes_{\wedge}\mathcal{Y}$ with isometrically isomorphic images of it, and so regard $\mathcal{X}\otimes_{\wedge}\mathcal{Y}$ as being dense in $\mathcal{X}\widehat{\otimes}_{\wedge}\mathcal{Y}$. Thus take $\digamma \in \mathcal{X}\widehat{\otimes}_{\wedge}\mathcal{Y}\backslash\mathcal{X}\otimes_{\wedge}\mathcal{Y}$ (otherwise the resulting finite sum is trivially obtained) so that $\digamma$ is arbitrarily close to elements in $\mathcal{X}\otimes_{\wedge}\mathcal{Y}$. Take an arbitrary $\varepsilon > 0$. For each positive integer k take $\digamma_k = \sum_{i=1}^{n_k} x_i \otimes y_i$ in $\mathcal{X}\otimes_{\wedge}\mathcal{Y}$ such that

$$\|\digamma - \digamma_k\|_{\wedge} < \tfrac{\varepsilon}{2}\tfrac{1}{2^k}.$$

In particular, $\big|\|\digamma_1\|_{\wedge} - \|\digamma\|_{\wedge}\big| \le \|\digamma_1 - \digamma\|_{\wedge} \le \frac{\varepsilon}{2}\frac{1}{2}$ and hence $\|\digamma_1\|_{\wedge} \le \|\digamma\|_{\wedge} + \frac{\varepsilon}{2}\frac{1}{2}$. Take a representation $\sum_{i=1}^{n_1} x_i \otimes y_i$ for $\digamma_1$ for which $\sum_{i=1}^{n_1}\|x_i\|\,\|y_i\|$ is close enough to $\|\digamma_1\|_{\wedge} = \inf_{\{x_i\}_{i=1}^{n_1},\{y_i\}_{i=1}^{n_1}} \sum_{i=1}^{n_1}\|x_i\|\,\|y_i\|$, say

$$\sum\nolimits_{i=1}^{n_1}\|x_i\|\,\|y_i\| \le \|\digamma_1\|_{\wedge} + \tfrac{3\varepsilon}{4} \le \|\digamma\|_{\wedge} + \varepsilon.$$

Computing the norm of $\digamma_{k+1} - \digamma_k \in \mathcal{X}\otimes_{\wedge}\mathcal{Y}$ we get

$$\|\digamma_{k+1} - \digamma_k\|_{\wedge} \le \|\digamma_{k+1} - \digamma\|_{\wedge} + \|\digamma - \digamma_k\|_{\wedge} < \tfrac{\varepsilon}{2}\tfrac{1}{2^{k+1}} + \tfrac{\varepsilon}{2}\tfrac{1}{2^k} = \tfrac{\varepsilon}{2}\tfrac{1}{2^k}\tfrac{3}{2} < \varepsilon\tfrac{1}{2^k},$$

so we can take a representation $\sum_{i=n_k+1}^{n_{k+1}} x_i \otimes y_i$ for $\digamma_{k+1} - \digamma_k \in \mathcal{X}\otimes_{\vee}\mathcal{Y}$ for which $\sum_{i=n_k+1}^{n_{k+1}}\|x_i\|\,\|y_i\|$ is close enough to $\|\digamma_{k+1} - \digamma_k\|_{\wedge}$, say

$$\sum\nolimits_{i=n_k+1}^{n_{k+1}}\|x_i\|\,\|y_i\| \le \|\digamma_{k+1} - \digamma_k\|_{\wedge} + \tfrac{\varepsilon}{4}\tfrac{1}{2^k} < \tfrac{\varepsilon}{2}\tfrac{1}{2^k}\tfrac{3}{2} + \tfrac{\varepsilon}{2}\tfrac{1}{2^k}\tfrac{1}{2} = \varepsilon\tfrac{1}{2^k}.$$

Since $\|\cdot\|_{\wedge}$ is a reasonable crossnorm we get for each $k > 1$

$$\begin{aligned}\sum\nolimits_{i=1}^{n_{k+1}}\|x_i \otimes y_i\|_{\wedge} &= \sum\nolimits_{i=1}^{n_1}\|x_i\|\,\|y_i\| + \sum\nolimits_{j=1}^{k}\sum\nolimits_{i=n_j+1}^{n_{j+1}}\|x_i\|\,\|y_i\|\\ &< \|\digamma\|_{\wedge} + \varepsilon + \varepsilon\sum\nolimits_{j=1}^{k}\tfrac{1}{2^j} < \|\digamma\|_{\wedge} + 2\varepsilon.\end{aligned}$$

Thus the sequence $\{x_i \otimes y_i\}$ is absolutely summable with $\sum_i\|x_i \otimes y_i\|_{\wedge} < \|\digamma\|_{\wedge} + 2\varepsilon$. Therefore, since $\mathcal{X}\widehat{\otimes}_{\wedge}\mathcal{Y}$ is a Banach space, the sequence is summable. This means the sequence of partial sums $\{\sum_{i=1}^{n_k} x_i \otimes y_i\}$ converges in $\mathcal{X}\widehat{\otimes}_{\wedge}\mathcal{Y}$. Hence the limit $\digamma \in \mathcal{X}\widehat{\otimes}_{\wedge}\mathcal{Y}$ of $\digamma_k = \sum_{i=1}^{n_k} x_i \otimes y_i \in \mathcal{X}\otimes_{\wedge}\mathcal{Y}$ is represented as

$$\digamma = \sum\nolimits_i x_i \otimes y_i,$$

and

$$\|F\|_\wedge \le \sum_i \|x_i \otimes y_i\|_\wedge = \sum_i \|x_i\|\,\|y_i\| \le \|F\|_\wedge + 2\varepsilon,$$

so that the projective norm $\|F\|_\wedge$ of F is the infimum of $\sum_i \|x_i\|\,\|y_i\|$ over all representations of F of the form $\sum_i x_i \otimes y_i$. □

Theorem 3.2. Universal Mapping Principle. *If* $(\mathcal{X},\mathcal{Y},\mathcal{Z})$ *is an arbitrary triple of Banach spaces, then*

$$\boldsymbol{b}[\mathcal{X}\times\mathcal{Y},\mathcal{Z}] \cong \mathcal{B}[\mathcal{X}\,\widehat{\otimes}_\wedge\,\mathcal{Y},\mathcal{Z}]$$

(*i.e., the Banach spaces* $\boldsymbol{b}[\mathcal{X}\times\mathcal{Y},\mathcal{Z}]$ *and* $\mathcal{B}[\mathcal{X}\,\widehat{\otimes}_\wedge\,\mathcal{Y},\mathcal{Z}]$ *are isometrically isomorphic*).

Proof. Take the natural bilinear map $\theta \in b[\mathcal{X}\times\mathcal{Y},\mathcal{X}\otimes\mathcal{Y}]$ associated with the tensor product space $\mathcal{X}\otimes\mathcal{Y}$. It is readily verified that the composition with θ, namely

$$C_\theta(\Phi) = \Phi\circ\theta \in b[\mathcal{X}\times\mathcal{Y},\mathcal{Z}] \quad \text{for every} \quad \Phi\in\mathcal{L}[\mathcal{X}\otimes\mathcal{Y},\mathcal{Z}],$$

defines a linear-space isomorphism $C_\theta\colon \mathcal{L}[\mathcal{X}\otimes\mathcal{Y},\mathcal{Z}] \to b[\mathcal{X}\times\mathcal{Y},\mathcal{Z}]$. Equip the linear space $\mathcal{X}\otimes\mathcal{Y}$ with the projective norm to get the normed space $\mathcal{X}\otimes_\wedge\mathcal{Y}$. Since $\mathcal{Z}$ is a Banach space, $\mathcal{B}[\mathcal{X}\otimes_\wedge\mathcal{Y},\mathcal{Z}]$ is a Banach space which is a linear manifold of the linear space $\mathcal{L}[\mathcal{X}\otimes_\wedge\mathcal{Y},\mathcal{Z}]$. Let $\mathcal{J}$ be the restriction to $\mathcal{B}[\mathcal{X}\otimes_\wedge\mathcal{Y},\mathcal{Z}]$ of the linear-space isomorphism C_θ on $\mathcal{L}[\mathcal{X}\otimes_\wedge\mathcal{Y},\mathcal{Z}]$, which remains linear and injective,

$$\mathcal{J} = C_\theta|_{\mathcal{B}[\mathcal{X}\otimes_\wedge\mathcal{Y},\mathcal{Z}]}\colon \mathcal{B}[\mathcal{X}\otimes_\wedge\mathcal{Y},\mathcal{Z}] \to \mathcal{R}(\mathcal{J}) \subseteq b[\mathcal{X}\times\mathcal{Y},\mathcal{Z}].$$

Next we show that (a) $\mathcal{R}(\mathcal{J}) = \boldsymbol{b}[\mathcal{X}\times\mathcal{Y},\mathcal{Z}]$ and (b) $\mathcal{J}$ is an isometry onto $\mathcal{R}(\mathcal{J})$. (Let $\|\cdot\|_{\mathcal{B}}$ and $\|\cdot\|_{\boldsymbol{b}}$ stand for the norms in $\mathcal{B}[\mathcal{X}\otimes_\wedge\mathcal{Y},\mathcal{Z}]$ and $\boldsymbol{b}[\mathcal{X}\times\mathcal{Y},\mathcal{Z}]$.)

(a_1) If $\Phi\in\mathcal{B}[\mathcal{X}\otimes_\wedge\mathcal{Y},\mathcal{Z}]$, then $\mathcal{J}(\Phi) = C_\theta(\Phi) = \Phi\circ\theta \in b[\mathcal{X}\times\mathcal{Y},\mathcal{Z}]$ is such that

$$\phi(x,y) = \mathcal{J}(\Phi)(x,y) = (\Phi\circ\theta)(x,y) = \Phi(x\otimes y) \quad \text{for every} \quad (x,y)\in\mathcal{X}\times\mathcal{Y},$$

and hence

$$\|\mathcal{J}(\Phi)(x,y)\|_{\mathcal{Z}} = \|\Phi(x\otimes y)\|_{\mathcal{Z}} \le \|\Phi\|_{\mathcal{B}}\|x\otimes y\|_\wedge = \|\Phi\|_{\mathcal{B}}\|x\|_{\mathcal{X}}\|y\|_{\mathcal{Y}}. \qquad (*)$$

Then the bilinear map $\mathcal{J}(\Phi)$ is bounded. Thus $\mathcal{J}(\Phi)\in\boldsymbol{b}[\mathcal{X}\times\mathcal{Y},\mathcal{Z}]$ for every Φ in $\mathcal{B}[\mathcal{X}\otimes_\wedge\mathcal{Y},\mathcal{Z}]$, and so $\mathcal{R}(\mathcal{J})\subseteq\boldsymbol{b}[\mathcal{X}\times\mathcal{Y},\mathcal{Z}]$.

(a_2) Conversely, if $\phi\in\boldsymbol{b}[\mathcal{X}\times\mathcal{Y},\mathcal{Z}]$, then there is a unique $\Phi\in\mathcal{B}[\mathcal{X}\otimes_\wedge\mathcal{Y},\mathcal{Z}]$ such that $C_\theta(\Phi) = \Phi\circ\theta = \phi$. So

$$\|\Phi(x\otimes y)\|_{\mathcal{Z}} = \|\Phi\circ\theta(x,y)\|_{\mathcal{Z}} = \|\phi(x,y)\|_{\mathcal{Z}} \le \|\phi\|_{\boldsymbol{b}}\|x\|_{\mathcal{X}}\|y\|_{\mathcal{Y}}$$

for every $(x,y)\in\mathcal{X}\times\mathcal{Y}$. Hence for an arbitrary $F = \sum_i x_i\otimes y_i \in \mathcal{X}\otimes_\wedge\mathcal{Y}$,

$$\begin{aligned}\|\Phi(F)\|_{\mathcal{Z}} &= \Big\|\Phi\sum_i x_i\otimes y_i\Big\|_{\mathcal{Z}} = \Big\|\sum_i \Phi(x_i\otimes y_i)\Big\|_{\mathcal{Z}}\\ &\le \sum_i \|\phi\|_{\boldsymbol{b}}\|x_i\|_{\mathcal{X}}\|y_i\|_{\mathcal{Y}} = \|\phi\|_{\boldsymbol{b}}\sum_i \|x_i\|_{\mathcal{X}}\|y_i\|_{\mathcal{Y}}.\end{aligned}$$

Since this holds for every (finite) representation of $F = \sum_i x_i\otimes y_i$, and since $\|F\|_\wedge = \inf\sum_i \|x_i\|_{\mathcal{X}}\|y_i\|_{\mathcal{Y}}$ over all representations, then for every $F\in\mathcal{X}\otimes_\wedge\mathcal{Y}$

$$\|\Phi(F)\|_{\mathcal{Z}} \le \|\phi\|_{\boldsymbol{b}}\|F\|_\wedge, \qquad (**)$$

and the linear Φ is bounded: $\Phi\in\mathcal{B}[\mathcal{X}\otimes_\wedge\mathcal{Y},\mathcal{Z}]$. So $\phi = C_\theta(\Phi) = \mathcal{J}(\Phi) \subseteq \mathcal{R}(\mathcal{J})$. Thus, as this holds for every $\phi\in\boldsymbol{b}[\mathcal{X}\times\mathcal{X},\mathcal{Z}]$, we get $\boldsymbol{b}[\mathcal{X}\times\mathcal{X},\mathcal{Z}]\subseteq\mathcal{R}(\mathcal{J})$.

(a) By (a_1) and (a_2) we get $\mathcal{R}(\mathcal{J}) = \boldsymbol{b}[\mathcal{X}\times\mathcal{Y},\mathcal{Z}]$.

(b) Take an arbitrary $\Phi \in \mathcal{B}[\mathcal{X} \otimes_{\wedge} \mathcal{Y}, \mathcal{Z}]$. As we saw in $(*)$,

$$\|\mathcal{J}(\Phi)(x,y)\|_{\mathcal{Z}} \leq \|\Phi\|_{\mathcal{B}} \|x\|_{\mathcal{X}} \|y\|_{\mathcal{Y}}$$

for every $(x,y) \in \mathcal{X} \times \mathcal{Y}$. Then the bilinear map $\mathcal{J}(\Phi)$ is bounded with norm

$$\|\mathcal{J}(\Phi)\|_{\boldsymbol{b}} \leq \|\Phi\|_{\mathcal{B}}.$$

Conversely, the unique $\phi = C_\theta(\Phi) = \mathcal{J}(\Phi)$ is such that $\phi \in \boldsymbol{b}[\mathcal{X} \times \mathcal{Y}, \mathcal{Z}]$ because $\mathcal{R}(\mathcal{J}) \subseteq \boldsymbol{b}[\mathcal{X} \times \mathcal{Y}, \mathcal{Z}]$ by (a_1) and so, as we saw in $(**)$,

$$\|\Phi(F)\|_{\mathcal{Z}} \leq \|\phi\|_{\boldsymbol{b}} \|F\|_{\wedge}$$

for every $F \in \mathcal{X} \otimes_{\wedge} \mathcal{Y}$. Therefore,

$$\|\Phi\|_{\mathcal{B}} \leq \|\phi\|_{\boldsymbol{b}} = \|\mathcal{J}(\Phi)\|_{\boldsymbol{b}}.$$

Thus $\|\mathcal{J}(\Phi)\|_{\boldsymbol{b}} = \|\Phi\|_{\mathcal{B}}$ for every $\Phi \in \mathcal{B}[\mathcal{X} \otimes_{\wedge} \mathcal{Y}, \mathcal{Z}]$. That is, $\mathcal{J}$ is an isometry.

(c) Then the linear transformation $\mathcal{J} : \mathcal{B}[\mathcal{X} \otimes_{\wedge} \mathcal{Y}] \to \boldsymbol{b}[\mathcal{X} \times \mathcal{Y}, \mathcal{Z}]$ is a surjective isometry by (a) and (b), which means $\mathcal{J}$ is an isometric isomorphism. Hence $\mathcal{B}[\mathcal{X} \otimes_{\wedge} \mathcal{Y}, \mathcal{Z}]$ and $\boldsymbol{b}[\mathcal{X} \times \mathcal{Y}, \mathcal{Z}]$ are isometrically isomorphic Banach spaces,

$$\mathcal{B}[\mathcal{X} \otimes_{\wedge} \mathcal{Y}, \mathcal{Z}] \cong \boldsymbol{b}[\mathcal{X} \times \mathcal{Y}, \mathcal{Z}].$$

Consider the completion $\mathcal{X} \widehat{\otimes}_{\wedge} \mathcal{Y}$ of $\mathcal{X} \otimes_{\wedge} \mathcal{Y}$. Since $\mathcal{Z}$ is complete, it follows that

$$\mathcal{B}[\mathcal{X} \widehat{\otimes}_{\wedge} \mathcal{Y}, \mathcal{Z}] \cong \mathcal{B}[\mathcal{X} \otimes_{\wedge} \mathcal{Y}, \mathcal{Z}].$$

So the stated result follows by transitivity. □

4. Nuclear Operators

Let $\mathcal{X}, \mathcal{Y}, \mathcal{V}$ be Banach spaces. As a starting point, consider the following expression involving the projective tensor product.

$$\mathcal{X}^* \widehat{\otimes}_{\wedge} \mathcal{V}^* \subseteq (\mathcal{X} \widehat{\otimes}_{\wedge} \mathcal{V})^* \cong \boldsymbol{b}[\mathcal{X} \times \mathcal{V}, \mathbb{F}] \cong \mathcal{B}[\mathcal{X}, \mathcal{V}^*].$$

The above inclusion comes from the definition of a reasonable crossnorm, the first isometric isomorphism is a particular case of the *universal mapping principle* for the projective norm as in Theorem 3.2, and the second one is the classical identification of bounded bilinear forms with bounded linear transformations (see, e.g., [2, Section 1.4, p.6]). For the particular case where $\mathcal{Y}$ is isometrically isomorphic to the dual of some Banach space $\mathcal{V}$ (for instance, if is $\mathcal{Y}$ is reflexive), then we get $\mathcal{X}^* \widehat{\otimes}_{\wedge} \mathcal{Y} \subseteq \mathcal{B}[\mathcal{X}, \mathcal{Y}]$, where the inclusion means algebraic embedding. On the other hand, this not only holds in general (no restriction to Banach spaces being the dual of some Banach space) but is strengthened to an isometric embedding for the injective norm instead: *the injective tensor product $\mathcal{X}^* \widehat{\otimes}_{\vee} \mathcal{Y}$ is isometrically embedded in the Banach space $\mathcal{B}[\mathcal{X}, \mathcal{Y}]$, and so it is viewed as a subspace of $\mathcal{B}[\mathcal{X}, \mathcal{Y}]$*:

$$\mathcal{X}^* \widehat{\otimes}_{\vee} \mathcal{Y} \subseteq \mathcal{B}[\mathcal{X}, \mathcal{Y}]$$

(see, e.g., [4, Proposition 1.1.5]). This fails in general for the projective norm, and the theorem below shows how far one can get along this line in the general case.

Theorem 4.1. *There is a natural transformation $K : \mathcal{X}^* \widehat{\otimes}_{\wedge} \mathcal{Y} \to \mathcal{B}[\mathcal{X}, \mathcal{Y}]$ such that*

(a) *K is a linear contraction with $\|K\| = 1$,*

(b) *the range* $\mathcal{R}(K)$ *of* K *is characterized as follows*: $T \in \mathcal{B}[\mathcal{X},\mathcal{Y}]$ *lies in* $\mathcal{R}(K)$ *if and only if there are* $\mathcal{X}^*$*-valued and* $\mathcal{Y}$*-valued sequences* $\{f_k\}$ *and* $\{y_k\}$ *for which the real sequence* $\{\|f_k\|\,\|y_k\|\}$ *is summable and*

$$Tx = \sum\nolimits_k f_k(x)\,y_k \quad \textit{for every} \quad x \in \mathcal{X},$$

and

(c) *for each* $T \in \mathcal{R}(K)$,

$$\|T\|_N = \inf_{\{f_k\},\{y_k\},\,T=\Sigma_k f_k(\cdot)\,y_k} \sum\nolimits_k \|f_k\|\,\|y_k\|$$

defines a norm on the linear space $\mathcal{R}(K)$ *for which* $\|T\| \le \|T\|_N$.

Proof. Let $\mathcal{X}^*$ be the dual of $\mathcal{X}$. Take an arbitrary $\digamma = \sum_k f_k \otimes y_k \in \mathcal{X}^* \widehat{\otimes}_\wedge \mathcal{Y}$ so that $\|\digamma\|_\wedge = \inf \sum_k \|f_k\|\,\|y_k\|$ by Theorem 3.1. Associated with $\digamma$ consider the natural transformation $\Psi_\digamma : \mathcal{X} \to \mathcal{Y}$ given by

$$\Psi_\digamma x = \sum\nolimits_k f_k(x)\,y_k \quad \text{for every} \quad x \in \mathcal{X},$$

which is linear and does not depend on the representation $\sum_k f_k \otimes y_k$ of $\digamma$. Also, $\Psi_\digamma$ is bounded. In fact $\|\Psi_\digamma x\| \le \sum_k \|f_k\|\,\|y_k\|\,\|x\|$. So $\|\Psi_\digamma x\| \le \|\digamma\|_\wedge \|x\|$, for every $x \in \mathcal{X}$. Thus $\Psi_\digamma$ lies in $\mathcal{B}[\mathcal{X},\mathcal{Y}]$ for each $\digamma \in \mathcal{X}^* \widehat{\otimes}_\wedge \mathcal{Y}$. This defines a transformation

$$K : \mathcal{X}^* \widehat{\otimes}_\wedge \mathcal{Y} \to \mathcal{B}[\mathcal{X},\mathcal{Y}] \quad \text{such that} \quad K(\digamma) = \Psi_\digamma \quad \text{for every} \quad \digamma \in \mathcal{X}^* \widehat{\otimes}_\wedge \mathcal{Y}.$$

(a) K is clearly linear. It is contraction as well. In fact, as $\|\Psi_\digamma x\| \le \|\digamma\|_\wedge \|x\|$ for every $x \in \mathcal{X}$, $\|K(\digamma)\| = \|\Psi_\digamma\| = \sup_{\|x\|=1} \|\Psi_\digamma x\| \le \|\digamma\|_\wedge$ for every $\digamma \in \mathcal{X}^* \widehat{\otimes}_\wedge \mathcal{Y}$. So $K \in \mathcal{B}\big[\mathcal{X}^* \widehat{\otimes}_\wedge \mathcal{Y}, \mathcal{B}[\mathcal{X},\mathcal{Y}]\big]$ with $\|K\| = \sup_{\|\digamma\|_\wedge=1} \|K(\digamma)\| \le 1$. Reversely, for $f \otimes y$ in $\mathcal{X}^* \widehat{\otimes}_\wedge \mathcal{Y}$ with $\|f\| = \|y\| = 1$ we get $\|\Psi_{f\otimes y} x\| = |f(x)|\,\|y\| = |f(x)|$, and hence $\|K(f \otimes y)\| = \|\Psi_{f\otimes y}\| = \sup_{\|x\|=1} |f(x)| = \|f\| = 1 = \|f\|\,\|y\| = \|f \otimes y\|_\wedge$. Then with $\digamma = f \otimes y$ in $\mathcal{X}^* \otimes_\wedge \mathcal{Y}$ we get $\|\digamma\|_\wedge = 1$ and $\|K(\digamma)\| = \|\digamma\|_\wedge$. Thus $\|K\| \ge 1$.

(b) By definition, $K(f \otimes y) = T \in \mathcal{B}[\mathcal{X},\mathcal{Y}]$ if an only if $Tx = \Psi_{f\otimes y}(x) = f(x)y$ for every $x \in \mathcal{X}$. Then for an arbitrary $\digamma = \sum_k f_k \otimes y_k \in \mathcal{X}^* \widehat{\otimes}_\wedge \mathcal{Y}$,

$$K(\digamma) = T \in \mathcal{B}[\mathcal{X},\mathcal{Y}] \quad \text{if an only if} \quad Tx = \Psi_\digamma(x) = \sum\nolimits_k f_k(x)\,y_k$$

for every $x \in \mathcal{X}$, since $K(\digamma) = \sum_k K(f_k \otimes y_k)$ because K is linear and bounded. This characterizes the range $\mathcal{R}(K)$ of K.

(c) As for $\|\cdot\|_N$ being a norm on $\mathcal{R}(K)$, we verify the triangle inequality only. Take $T, T' \in \mathcal{R}(K)$ so that $T + T' = \sum_k f_k(\cdot)y_k + \sum_k f'_k(\cdot)y'_k = \sum_k f''_k(\cdot)y''_k$. Thus $\|T + T'\|_N = \inf_{f_k,y_k,f'_k,y'_k} (\sum_k \|f_k\|\,\|y_k\| + \|f'_k\|\,\|y'_k\|) = \inf_{f''_k,y''_k} \sum_k \|f''_k\|\,\|y''_k\| \le \inf_{f_k,y_k} \sum_k \|f_k\|\,\|y_k\| + \inf_{f'_k,y'_k} \sum_k \|f'_k\|\,\|y'_k\| = \|T\|_N + \|T'\|_N$. Now for the norm inequality: if $T \in \mathcal{R}(K)$, then there are $\{f_k\}$ and $\{y_k\}$ with $\{\|f_k\|\,\|y_k\|\}$ summable such that $T(x) = \sum_k f_k(x)y_k$, and so $\|Tx\| = \|\sum_k f_k(x)y_k\|$, for every $x \in \mathcal{X}$. Thus

$$\|T\| = \sup_{\|x\|=1} \|Tx\| \le \sup_{\|x\|=1} \sum\nolimits_k \|f_k\|\,\|y_k\|\,\|x\| = \sum\nolimits_k \|f_k\|\,\|g_k\|$$

so that $\|T\| \le \inf \sum_k \|f_k\|\,\|g_k\| = \|T\|_N$. □

A transformation $T \in \mathcal{B}[\mathcal{X},\mathcal{Y}]$ is *nuclear* if it lies in the range $\mathcal{R}(K)$ of such a natural contraction with unit norm $K \in \mathcal{B}[\mathcal{X}^* \widehat{\otimes}_\wedge \mathcal{Y},\, \mathcal{B}[\mathcal{X},\mathcal{Y}]]$ defined in Theorem 4.1, and the range $\mathcal{R}(K)$ is called *the linear space of nuclear transformations*. Thus set

$$\mathcal{B}_N[\mathcal{X},\mathcal{Y}]=\mathcal{R}(K)\subseteq\mathcal{B}[\mathcal{X},\mathcal{Y}].\quad\text{In particular,}\quad\mathcal{B}_N[\mathcal{X}]=\mathcal{B}_N[\mathcal{X},\mathcal{X}]\subseteq\mathcal{B}[\mathcal{X}].$$

Theorem 4.1 prompts the following redefinition:

> A linear transformation $T\in\mathcal{B}[\mathcal{X},\mathcal{Y}]$ between Banach spaces $\mathcal{X}$ and $\mathcal{Y}$ is *nuclear* (i.e., $T\in\mathcal{B}_N[\mathcal{X},\mathcal{Y}]$) if there are $\mathcal{X}^*$-valued and $\mathcal{Y}$-valued sequences $\{f_k\}$ and $\{y_k\}$ such that $\sum_k\|f_k\|\,\|y_k\|<\infty$ and $Tx=\sum_k f_k(x)\,y_k$ for every $x\in\mathcal{X}$.

The expression $Tx=\sum_k f_k(x)\,y_k$ is a *nuclear representation* of T and

$$\|T\|_N=\inf\sum\nolimits_k\|f_k\|\,\|y_k\|\quad\text{for every}\quad T\in\mathcal{B}_N[\mathcal{X},\mathcal{Y}]$$

(where the infimum is taken over all nuclear representations of T) defines a norm on the linear space $\mathcal{B}_N[\mathcal{X},\mathcal{Y}]$, the *nuclear norm*, such that $\|T\|\le\|T\|_N$.

The linear contraction K of norm one in Theorem 4.1 is not necessarily injective (thus it may not be an isometry) and $\mathcal{R}(K)$ is identified with the quotient space of $\mathcal{X}^*\widehat{\otimes}_\wedge\mathcal{Y}$ modulo $\mathcal{N}(K)$, that is, $\mathcal{B}_N[\mathcal{X},\mathcal{Y}]\cong(\mathcal{X}^*\widehat{\otimes}_\wedge\mathcal{Y})/\mathcal{N}(K)$ (see, e.g., [26, p.41]), and therefore $\mathcal{B}_N[\mathcal{X},\mathcal{Y}]$ is a Banach space. However, if one of $\mathcal{X}^*$ or $\mathcal{Y}$ is a Banach space with the approximation property, then $\mathcal{N}(K)=\{0\}$ (see, e.g.,[26, Proposition 4.6]), and so in this case $\mathcal{B}_N[\mathcal{X},\mathcal{Y}]\cong\mathcal{X}^*\widehat{\otimes}_\wedge\mathcal{Y}$. In particular,

$$\mathcal{X}\;\text{ is a Hilbert space }\;\Longrightarrow\;\mathcal{B}_N[\mathcal{X}]\cong\mathcal{X}^*\widehat{\otimes}_\wedge\mathcal{X}\cong\mathcal{X}\widehat{\otimes}_\wedge\mathcal{X}^*.$$

Moreover, $\mathcal{B}_N[\mathcal{X}]$ is a two-sided ideal of the Banach algebra $\mathcal{B}[\mathcal{X}]$ by the next result.

Corollary 4.2. *Let $\mathcal{X},\mathcal{Y},\mathcal{V},\mathcal{W}$ be Banach spaces. If $T\in\mathcal{B}_N[\mathcal{X},\mathcal{Y}]$, $R\in\mathcal{B}[\mathcal{V},\mathcal{X}]$ and $L\in\mathcal{B}[\mathcal{Y},\mathcal{W}]$, then $LTR\in\mathcal{B}_N[\mathcal{V},\mathcal{W}]$ and $\|LTR\|_N\le\|L\|\,\|T\|_N\|R\|$.*

Proof. This is a straightforward consequence of the definition of nuclear transformation: as $\sum_k\|f_k\|\,\|y_k\|<\infty$ and $Tx=\sum_k f_k(x)\,y_k$, set $g_k=f_kR\in\mathcal{V}^*$ and $w_k=Ly_k\in\mathcal{W}$. So $\sum_k\|g_k\|\,\|y_k\|\le\sum_k\|f_k\|\,\|y_k\|\,\|R\|<\infty$ and $TRx=\sum_k g_k(x)\,y_k$. Also $\sum_k\|f_k\|\,\|w_k\|\le\sum_k\|f_k\|\,\|w_k\|\,\|L\|<\infty$ and $LTx=\sum_k f_k(x)\,w_k$. □

From now on let $\mathcal{X}$ be a Hilbert space. Denote the inner product in $\mathcal{X}$ by $\langle\cdot\,;\cdot\rangle$, and take the Banach algebra $\mathcal{B}[\mathcal{X}]$ of all operators on $\mathcal{X}$. Consider the Fourier Series and the Riesz Representation Theorems (see, e.g., [13, Theorems 5.48 and 5.62]). A functional f lies in $\mathcal{X}^*$ if and only if there is a unique z in $\mathcal{X}$, called the Riesz representation of f, such that $f(x)=\langle x\,;z\rangle$ for every $x\in\mathcal{X}$ and $\|z\|=\|f\|$. Thus on a Hilbert space the previous redefinition can be rewritten as follows:

> An operator $T\in\mathcal{B}[\mathcal{X}]$ on a Hilbert space $\mathcal{X}$ is *nuclear* (i.e., $T\in\mathcal{B}_N[\mathcal{X}]$) if there are $\mathcal{X}$-valued sequences $\{z_k\}$ and $\{y_k\}$ such that $\sum_k\|z_k\|\,\|y_k\|<\infty$ and $Tx=\sum_k\langle x\,;z_k\rangle\,y_k$ for every $x\in\mathcal{X}$.

For each $T\in\mathcal{B}[\mathcal{X}]$ let $T^*\in\mathcal{B}[\mathcal{X}]$ stand for its Hilbert-space adjoint.

Corollary 4.3. *Let $T\in\mathcal{B}[\mathcal{X}]$ be an operator on a Hilbert space $\mathcal{X}$.*

(a) *If there exist $\mathcal{X}$-valued sequences $\{z_k\}$ and $\{y_k\}$ such that*

$$Tx=\sum\nolimits_k\langle x\,;z_k\rangle y_k\quad\textit{for every}\quad x\in\mathcal{X},$$

then

$$T^*x=\sum\nolimits_k\langle x\,;y_k\rangle z_k\quad\textit{for every}\quad x\in\mathcal{X}.$$

(b) *T is nuclear if an only if its adjoint T^* is nuclear and* $\|T^*\|_N = \|T\|_N$.

Proof. Let $T \in \mathcal{B}[\mathcal{X}]$ be an operator on a Hilbert space $\mathcal{X}$.
(a) Suppose there are $\mathcal{X}$-valued sequences $\{z_k\}$ and $\{y_k\}$ such that

$$Tx = \sum\nolimits_k \langle x\,; z_k\rangle\, y_k \quad \text{for every} \quad x \in \mathcal{X}.$$

Then $\langle Tx\,; y\rangle = \big\langle \sum_k \langle x\,; z_k\rangle\, y_k; y\big\rangle = \sum_k \langle x\,; z_k\rangle\,\langle y_k; y\rangle = \sum_k \big\langle x\,; \langle y\,; y_k\rangle\, z_k\big\rangle = \big\langle x\,; \sum_k \langle y\,; y_k\rangle\, z_k\big\rangle$ for every $x, y \in \mathcal{X}$, and therefore (by uniqueness of the adjoint)

$$T^*y = \sum\nolimits_k \langle y\,; y_k\rangle\, z_k \quad \text{for every} \quad y \in \mathcal{X}.$$

(b) Immediate from item (a) and the definition of nuclear operator on $\mathcal{X}$. □

5. Trace-Class Operators

Let $\mathcal{X}$ be a Hilbert space. A summary of the elementary expressions required in this section goes as follows. For each $T \in \mathcal{B}[\mathcal{X}]$ set $|T| = (T^*T)^{\frac{1}{2}} \in \mathcal{B}[\mathcal{X}]$. Then

$$\|Tx\|^2 = \langle Tx\,; Tx\rangle = \langle T^*Tx\,; x\rangle = \big\langle |T|^2x\,; x\big\rangle = \big\langle |T|x\,; |T|x\big\rangle = \big\|\,|T|x\big\|^2$$

for every $x \in \mathcal{X}$. So $\|T\| = \big\|\,|T|\,\big\|$. Since $\|T\|^2 = \|T^*T\| = \|TT^*\| = \|T^*\|^2$, we get

$$\big\|\,|T|\,\big\|^2 = \|T\|^2 = \big\|\,|T|^2\big\| = \big\|\,|T^*|^2\big\| = \|T^*\|^2 = \big\|\,|T^*|\,\big\|^2.$$

Moreover, since $\|Q^{\frac{1}{2}}\|^2 = \|Q\| = \|Q^2\|^{\frac{1}{2}}$ for every nonnegative operator $Q \in \mathcal{B}[\mathcal{X}]$,

$$\big\|\,|T|^{\frac{1}{2}}\big\|^2 = \big\|\,|T|\,\big\| = \big\|\,|T|^2\big\|^{\frac{1}{2}}.$$

Hence for each $x \in \mathcal{X}$,

$$\big\|\,|T|x\big\|^2 = \big\|\,|T|^{\frac{1}{2}}|T|^{\frac{1}{2}}x\big\|^2 \le \big\|\,|T|^{\frac{1}{2}}\big\|^2\big\|\,|T|^{\frac{1}{2}}x\big\|^2 = \big\|\,|T|\,\big\|\,\big\|\,|T|^{\frac{1}{2}}x\big\|^2 = \|T\|\,\big\langle |T|x\,; x\big\rangle.$$

Now let $\{e_\gamma\}_{\gamma\in\Gamma}$ and $\{f_\gamma\}_{\gamma\in\Gamma}$ be arbitrary orthonormal bases for a Hilbert space $\mathcal{X}$, indexed by an arbitrary nonempty index set Γ (alternate notation: $\{e_\gamma\}_\gamma$ or $\{e_\gamma\}$). By the Parseval identity (viz., $\|x\|^2 = \sum_\gamma |\langle x\,; e_\gamma\rangle|^2$ for every $x \in \mathcal{X}$) we get

$$\sum\nolimits_{\gamma\in\Gamma}\|Te_\gamma\|^2 = \sum\nolimits_{(\gamma,\delta)\in\Gamma^2}|\langle Te_\gamma\,; f_\delta\rangle|^2 = \sum\nolimits_{(\gamma,\delta)\in\Gamma^2}|\langle T^*f_\delta\,; e_\gamma\rangle|^2 = \sum\nolimits_{\delta\in\Gamma}\|T^*f_\delta\|^2$$

whenever any of the families $\{\|Te_\gamma\|\}_\gamma$ or $\{\|T^*f_\gamma\|\}_\gamma$ is square summable (i.e., if $\sum_\gamma \|Te_\gamma\|^2 < \infty$ or $\sum_\gamma \|T^*f_\gamma\|^2 < \infty$) for some orthonormal bases for $\mathcal{X}$. Applying the above displayed identity to $|T|^{\frac{1}{2}} \in \mathcal{B}[\mathcal{X}]$ instead of $T \in \mathcal{B}[\mathcal{X}]$ we get

$$\sum\nolimits_\gamma \big\langle |T|e_\gamma\,; e_\gamma\big\rangle = \sum\nolimits_\gamma \big\langle |T|f_\gamma\,; f_\gamma\big\rangle.$$

Thus if the family of positive numbers $\big\{\big\langle |T|e_\gamma\,; e_\gamma\big\rangle\big\}_\gamma = \big\{\big\|\,|T|^{\frac{1}{2}}e_\gamma\big\|^2\big\}_\gamma$ is summable, then its sum does not depend on the orthonormal basis $\{e_\gamma\}_\gamma$ for $\mathcal{X}$.

An operator $T \in \mathcal{B}[\mathcal{X}]$ on a Hilbert space $\mathcal{X}$ is *trace-class* if $\sum_\gamma \big\langle |T|e_\gamma; e_\gamma\big\rangle < \infty$ for an arbitrary orthonormal basis $\{e_\gamma\}$ for $\mathcal{X}$.

Let $\mathcal{B}_1[\mathcal{X}]$ denote the subset of $\mathcal{B}[\mathcal{X}]$ consisting of all trace-class operators. Since $\big\langle |T|x\,; x\big\rangle = \big\|\,|T|^{\frac{1}{2}}x\big\|^2$ for every $x \in \mathcal{X}$, an operator T lies in $\mathcal{B}_1[\mathcal{X}]$ if and only if $\sum_\gamma \big\|\,|T|^{\frac{1}{2}}e_\gamma\big\|^2 < \infty$ for an arbitrary orthonormal basis $\{e_\gamma\}$. So for $T \in \mathcal{B}_1[\mathcal{X}]$ set

$$\|T\|_1 = \sum\nolimits_\gamma \big\langle |T|e_\gamma\,; e_\gamma\big\rangle = \sum\nolimits_\gamma \big\|\,|T|^{\frac{1}{2}}e_\gamma\big\|^2.$$

Clearly, $\big||T|\big| = |T|$. Then $|T|^2 = T^*T \in \mathcal{B}[\mathcal{X}]$ is trace-class (i.e., $|T|^2$ lies in $\mathcal{B}_1[\mathcal{X}]$) if and only if $\sum_\gamma \big\||T|e_\gamma\big\|^2 < \infty$; equivalently, if $\sum_\gamma \|Te_\gamma\|^2 < \infty$.

> An operator $T \in \mathcal{B}[\mathcal{X}]$ on a Hilbert space $\mathcal{X}$ is *Hilbert–Schmidt* if $|T|^2 \in \mathcal{B}_1[\mathcal{X}]$; equivalently if $\sum_\gamma \|Te_\gamma\|^2 < \infty$ for an arbitrary orthonormal basis $\{e_\gamma\}$ for $\mathcal{X}$.

Let $\mathcal{B}_2[\mathcal{X}]$ denote the subset of $\mathcal{B}[\mathcal{X}]$ consisting of all Hilbert–Schmidt operators. If T lies in $\mathcal{B}_2[\mathcal{X}]$, then set $\|T\|_2^2 = \sum_\gamma \|Te_\gamma\|^2$ so that from what we have seen above,

$$\|T\|_2 = \Big(\sum\nolimits_\gamma \|Te_\gamma\|^2\Big)^{\frac{1}{2}} = \Big(\sum\nolimits_\gamma \big\||T|e_\gamma\big\|^2\Big)^{\frac{1}{2}} = \big\||T|^2\big\|_1^{\frac{1}{2}} = \|T^*T\|_1^{\frac{1}{2}},$$

and so (recall: $\big||T|\big| = |T|$) we may infer:

$$T \in \mathcal{B}_1[\mathcal{X}] \iff |T| \in \mathcal{B}_1[\mathcal{X}] \iff |T|^{\frac{1}{2}} \in \mathcal{B}_2[\mathcal{X}] \quad \text{and} \quad \|T\|_1 = \big\||T|\big\|_1 = \big\||T|^{\frac{1}{2}}\big\|_2^2,$$

$$T \in \mathcal{B}_2[\mathcal{X}] \iff |T| \in \mathcal{B}_2[\mathcal{X}] \iff |T|^2 \in \mathcal{B}_1[\mathcal{X}] \quad \text{and} \quad \|T\|_2^2 = \big\||T|\big\|_2^2 = \big\||T|^2\big\|_1.$$

Lemma 5.1. *Let $\mathcal{X}$ be a Hilbert space. The following assertions hold true.*

(a) *The set $\mathcal{B}_2[\mathcal{X}]$ is a linear space and $\|\cdot\|_2 : \mathcal{B}_2[\mathcal{X}] \to \mathbb{R}$ is a norm on $\mathcal{B}_2[\mathcal{X}]$.*

(b) *If $S \in \mathcal{B}[\mathcal{X}]$ and $T \in \mathcal{B}_2[\mathcal{X}]$, then* $\max\{\|ST\|_2, \|TS\|_2\} \le \|S\|\,\|T\|_2$.

(c) *$\mathcal{B}_2[\mathcal{X}]$ is an ideal of the algebra $\mathcal{B}[\mathcal{X}]$.*

(d) *$\mathcal{B}_2[\mathcal{X}] \subseteq \mathcal{B}_\infty[\mathcal{X}]$ and $\|T\| \le \|T\|_2$ for every $T \in \mathcal{B}_2[\mathcal{X}]$.*

(e) *$T^* \in \mathcal{B}_2[\mathcal{X}]$ if and only if $T \in \mathcal{B}_2[\mathcal{X}]$ and $\|T\|_2 = \|T^*\|_2$.*

Proof. Suppose S and T lie in $\mathcal{B}[\mathcal{X}]$ and consider the following assertion.

Claim 1. If $S, T \in \mathcal{B}_2[\mathcal{X}]$, then $S + T \in \mathcal{B}_2[\mathcal{X}]$ and $\|S + T\|_2 \le \|S\|_2 + \|T\|_2$.

Proof of Claim 1. Let the index γ run over an arbitrary index set $\Gamma \ne \varnothing$, consider the Schwarz inequality on the Hilbert space ℓ^2 over Γ, and take $S, T \in \mathcal{B}_2[\mathcal{X}]$ so that $\sum_\gamma \|Se_\gamma\|\,\|Te_\gamma\| \le \big(\sum_\gamma \|Se_\gamma\|^2\big)^{\frac{1}{2}} \big(\sum_\gamma \|Te_\gamma\|^2\big)^{\frac{1}{2}} = \|S\|_2\|T\|_2$ if $S, T \in \mathcal{B}_2[\mathcal{X}]$. Thus

$$\begin{aligned}\|T + S\|_2^2 &= \sum\nolimits_\gamma \|Se_\gamma + Te_\gamma\|^2 \le \sum\nolimits_\gamma \big(\|Se_\gamma\| + \|Te_\gamma\|\big)^2 \\ &= \sum\nolimits_\gamma \|Se_\gamma\|^2 + \sum\nolimits_\gamma \|Te_\gamma\|^2 + 2\sum\nolimits_\gamma \|Se_\gamma\|\,\|Te_\gamma\| \\ &\le \|S\|_2^2 + \|T\|_2^2 + 2\|S\|_2\|T\|_2 = (\|S\|_2 + \|T\|_2)^2. \quad \square\end{aligned}$$

Since homogeneity, nonnegativity and positivity for $\|\cdot\|_2 : \mathcal{B}_2[\mathcal{X}] \to \mathbb{R}$ is readily verified, Claim 1 is enough to ensure that

$$\mathcal{B}_2[\mathcal{X}] \text{ is a linear space and } \|\cdot\|_2 \text{ is a norm on it.}$$

Also, if $S \in \mathcal{B}[\mathcal{X}]$ and $T \in \mathcal{B}_2[\mathcal{X}]$, then $\|ST\|_2^2 = \sum_\gamma \|STe_\gamma\|^2 \le \|S\|^2 \sum_\gamma \|Te_\gamma\|^2 = \|S\|^2\,\|T\|_2^2$. Similarly, $\|TS\|_2^2 = \sum_\gamma \|TSe_\gamma\|^2 = \sum_\gamma \|(TS)^*e_\gamma\|^2 = \sum_\gamma \|S^*T^*e_\gamma\|^2 \le \|S^*\|^2 \sum_\gamma \|T^*e_\gamma\|^2 = \|S\|^2 \sum_\gamma \|Te_\gamma\|^2 = \|S\|\,\|T\|_2^2$. Therefore,

$$\max\{\|ST\|_2, \|TS\|_2\} \le \|S\|\,\|T\|_2$$

and

$$\mathcal{B}_2[\mathcal{X}] \text{ is an ideal (i.e., a two-sided ideal) of } \mathcal{B}[\mathcal{X}].$$

Now take an arbitrary $T \in \mathcal{B}_2[\mathcal{X}]$ so that $\sum_{\gamma\in\Gamma} \|T^*e_\gamma\|^2 = \sum_{\gamma\in\Gamma} \|Te_\gamma\|^2 < \infty$, and take any integer $n \ge 1$. Thus there is a finite set $N_n \subseteq \Gamma$ such that $\sum_{k\in N} \|T^*e_k\|^2 < \frac{1}{n}$ for all finite sets $N \subseteq \Gamma \backslash N_n$ (Cauchy criterion for summable families — see, e.g.,[13, Theorem 5.27]). So $\sum_{\gamma\in\Gamma\backslash N_n}\|T^*e_\gamma\|^2<\frac{1}{n}$. Recall that $Tx=\sum_{\gamma\in\Gamma}\langle Tx\,;e_\gamma\rangle e_\gamma$ for every $x\in\mathcal{X}$ (Fourier series expansion). Set $T_nx=\sum_{k\in N_n}\langle Tx\,;e_k\rangle e_k$ for each $x\in\mathcal{X}$, which defines an operator T_n in $\mathcal{B}_0[\mathcal{X}]$ because N_n is finite, Hence $\|(T-T_n)x\|^2=\sum_{\gamma\in\Gamma\backslash N_n}|\langle Tx\,;e_\gamma\rangle|^2\le\big(\sum_{\gamma\in\Gamma\backslash N_n}\|T^*e_\gamma\|^2\big)\|x\|^2$ for every $x\in\mathcal{X}$. This implies that $\|T_n-T\|\to 0$, and so T is the uniform limit of a sequence of finite-rank operators on a Banach space, and therefore T is compact (see, e.g., [13, Corollary 4.55]).

Every Hilbert–Schmidt operator is compact.

As we saw above, $\sum_\gamma\|Te_\gamma\|^2=\sum_\gamma\|T^*e_\gamma\|^2$ and $\|Te\|\le\|T\|_2$ if $\|e\|=1$, and so

$$T^*\in\mathcal{B}_2[\mathcal{X}] \text{ if and only if } T\in\mathcal{B}_2[\mathcal{X}] \;\text{ and }\; \|T\|\le\|T\|_2=\|T^*\|_2. \qquad\square$$

The norm $\|\cdot\|_2$ on the linear space $\mathcal{B}_2[\mathcal{X}]$ of all Hilbert–Schmidt operators is referred to as the *Hilbert–Schmidt norm*.

Theorem 5.2. *Let $\mathcal{X}$ be a Hilbert space. The following assertions hold true.*

(a) *The set $\mathcal{B}_1[\mathcal{X}]$ is a linear space and $\|\cdot\|_1:\mathcal{B}_1[\mathcal{X}]\to\mathbb{R}$ is a norm on $\mathcal{B}_1[\mathcal{X}]$.*

(b) *$T\in\mathcal{B}_1[\mathcal{X}]$ if and only if $T=AB$ for some $A,B\in\mathcal{B}_2[\mathcal{X}]$.*

(c) *$\mathcal{B}_1[\mathcal{X}]\subseteq\mathcal{B}_2[\mathcal{X}]$ and $\|T\|_2^2\le\|T\|\,\|T\|_1$ so that $\|T\|_2\le\|T\|_1$ for $T\in\mathcal{B}_1[\mathcal{X}]$.*

(d) *$\mathcal{B}_1[\mathcal{X}]$ is an ideal of the algebra $\mathcal{B}[\mathcal{X}]$.*

(e) *If $S\in\mathcal{B}[\mathcal{X}]$ and $T\in\mathcal{B}_1[\mathcal{X}]$, then $\max\{\|ST\|_1,\|TS\|_1\}\le\|S\|\,\|T\|_1$.*

(f) *$T^*\in\mathcal{B}_1[\mathcal{X}]$ if and only if $T\in\mathcal{B}_1[\mathcal{X}]$ and $\|T\|_1=\|T^*\|_1$.*

(g) *$\mathcal{B}_0[\mathcal{X}]\subseteq\mathcal{B}_1[\mathcal{X}]$.*

Proof. Suppose S and T lie in $\mathcal{B}[\mathcal{X}]$ and consider the following assertion.

Claim 2. If $S,T\in\mathcal{B}_1[\mathcal{X}]$, then $S+T\in\mathcal{B}_1[\mathcal{X}]$ and $\|S+T\|_1\le\|S\|_1+\|T\|_1$.

Proof of Claim 2. Consider the polar decompositions $T+S=W|T+S|$, $T=W_1|T|$ and $S=W_2|S|$, where W,W_1,W_2 are partial isometries in $\mathcal{B}[\mathcal{X}]$, so that $|T+S|=W^*(T+S)$, $|T|=W_1^*T$, and $|S|=W_2^*T$. Hence (Schwartz inequality on ℓ^2 over Γ)

$$\begin{aligned}
\|T+S\|_1 &= \sum_\gamma\big\langle|T+S|e_\gamma;e_\gamma\big\rangle\le\sum_\gamma|\langle Te_\gamma;W^*e_\gamma\rangle|+\sum_\gamma|\langle Se_\gamma;W^*e_\gamma\rangle|\\
&=\sum_\gamma\big|\big\langle|T|^{\frac12}e_\gamma\,;|T|^{\frac12}W_1^*We_\gamma\big\rangle\big|+\sum_\gamma\big|\big\langle|S|^{\frac12}e_\gamma\,;|S|^{\frac12}W_2^*We_\gamma\big\rangle\big|\\
&\le\sum_\gamma\big\||T|^{\frac12}e_\gamma\big\|\,\big\||T|^{\frac12}W_1^*We_\gamma\big\|+\sum_\gamma\big\||S|^{\frac12}e_\gamma\big\|\,\big\||S|^{\frac12}W_2^*We_\gamma\big\|\\
&\le\Big(\sum_\gamma\big\||T|^{\frac12}e_\gamma\|^2\Big)^{\frac12}\Big(\sum_\gamma\big\||T|^{\frac12}W_1^*We_\gamma\|^2\Big)^{\frac12}\\
&\quad+\Big(\sum_\gamma\big\||S|^{\frac12}e_\gamma\|^2\Big)^{\frac12}\Big(\sum_\gamma\big\||S|^{\frac12}W_2^*We_\gamma\|^2\Big)^{\frac12}\\
&=\big\||T|^{\frac12}\big\|_2\,\big\||T|^{\frac12}W_1^*W\big\|_2+\big\||S|^{\frac12}\big\|_2\,\big\||S|^{\frac12}W_1^*W\big\|_2\\
&\le\big\||T|^{\frac12}\big\|_2^2\,\|W_1^*W\|+\big\||S|^{\frac12}\big\|_2^2\,\|W_2^*W\|\;\le\;\|T\|_1+\|S\|_1.
\end{aligned}$$

(Recall that $T\in\mathcal{B}_1[\mathcal{X}]\implies|T|^{\frac12}\in\mathcal{B}_2[\mathcal{X}]\implies|T|^{\frac12}A\in\mathcal{B}_2[\mathcal{X}]$, $\sum_\gamma\big\||T|^{\frac12}e_\gamma\big\|^2=\big\||T|^{\frac12}\big\|_2^2=\|T\|_1$, $\big(\sum_\gamma\big\||T|^{\frac12}Ae_\gamma\big\|^2\big)^{\frac12}=\big\||T|^{\frac12}A\big\|_2\le\big\||T|^{\frac12}\big\|_2\|A\|$ by Lemma 5.1(b), and $\|W\|=\|W_1\|=\|W_2\|=1$ since these are partial isometries). $\square$

Claim 2 is enough to ensure that

$$\mathcal{B}_1[\mathcal{X}] \text{ is a linear space and } \|\cdot\|_1 \text{ is a norm on it,}$$

since $\|\cdot\|_1$ is trivially homogeneous (so every multiple of an operator in $\mathcal{B}_1[\mathcal{X}]$ lies in $\mathcal{B}_1[\mathcal{X}]$), and nonnegativeness and positiveness for $\|\cdot\|_1$ are readily verified. Consider again the polar decomposition $T = W_1|T| = W_1|T|^{\frac{1}{2}}|T|^{\frac{1}{2}}$. If $T \in \mathcal{B}_1[\mathcal{X}]$, then $|T|^{\frac{1}{2}}$ lies in $\mathcal{B}_2[\mathcal{X}]$ and so $W_1|T|^{\frac{1}{2}}$ lies in $\mathcal{B}_2[\mathcal{X}]$ as well according to Lemma 5.1(c). Conversely, suppose $T = AB$ with $A, B \in \mathcal{B}_2[\mathcal{X}]$. Thus T lies in $\mathcal{B}_2[\mathcal{X}]$ by Lemma 5.1(c) again. Since $|T| = W_1^*T$ we get $|T| = W^*AB$ with $A^*W \in \mathcal{B}_2[\mathcal{X}]$ (according to Lemma 5.1(c) once again). Therefore

$$\sum\nolimits_\gamma \big\langle |T|e_\gamma\,;e_\gamma\big\rangle \le \sum\nolimits_\gamma \|Be_\gamma\|\|A^*We_\gamma\| \le \Big(\sum\nolimits_\gamma \|Be_\gamma\|^2\Big)^{\frac{1}{2}} \Big(\sum\nolimits_\gamma \|A^*We_\gamma\|^2\Big)^{\frac{1}{2}}$$

(by the Schwarz inequality on both Hilbert spaces $\mathcal{X}$ and ℓ^2 over Γ, as we did before). Hence $T \in \mathcal{B}_1[\mathcal{X}]$ with $\|T\|_1 \le \|B\|_2\|A^*W\|_2 \le \|B\|_2\|A\|_2$. Summing up:

$$T \in \mathcal{B}_1[\mathcal{X}] \iff T = AB \text{ for some } A, B \in \mathcal{B}_2[\mathcal{X}] \;\text{ and }\; \|T\|_1 \le \|A\|_2\|B\|_2.$$

Also, since the product AB lies in $\mathcal{B}_2[\mathcal{X}]$ by Lemma 5.1(c), $\mathcal{B}_1[\mathcal{X}] \subseteq \mathcal{B}_2[\mathcal{X}]$:

Every trace-class operator is Hilbert–Schmidt.

Moreover, since $\big\| |T|x\big\|^2 \le \|T\| \, \big\| |T|^{\frac{1}{2}}x\big\|^2$ for every $x \in \mathcal{X}$ and $\|T\| \le \|T\|_2$ (by Lemma 5.1(d)), we get $\|T\|_2^2 = \sum_\gamma \big\| |T|e_\gamma\big\|^2 \le \|T\| \sum_\gamma \big\| |T|^{\frac{1}{2}}e_\gamma\big\|^2 = \|T\|\,\|T\|_1$. So

$$\|T\|_2^2 \le \|T\|\|T\|_1 \implies \|T\|_2 \le \|T\|_1 \quad \text{for every} \quad T \in \mathcal{B}_1[\mathcal{X}].$$

Hence if T lies in $\mathcal{B}_1[\mathcal{X}]$ and S lies in $\mathcal{B}[\mathcal{X}]$, then $ST = (SA)B$ and $TS = A(BS)$ for some $A, B \in \mathcal{B}_2[\mathcal{X}]$. Thus $ST = CB$ and $TS = AD$ with both $C = SA$ and $D = BS$ in $\mathcal{B}_2[\mathcal{X}]$ according to Proposition A1(c). So ST and TS lie in $\mathcal{B}_1[\mathcal{X}]$. Therefore

$$\mathcal{B}_1[\mathcal{X}] \text{ is an ideal (i.e., a two-sided ideal) of } \mathcal{B}[\mathcal{X}].$$

Claim 3. Let $\{e_\gamma\}$ be any orthonormal basis for $\mathcal{X}$. If $T \in \mathcal{B}_1[\mathcal{X}]$ and $S \in \mathcal{B}[\mathcal{X}]$, then

(i) $\sum_\gamma |\langle Te_\gamma\,;e_\gamma\rangle| \le \|T\|_1$ and $\sum_\gamma \langle Te_\gamma\,;e_\gamma\rangle$ does not depend on $\{e_\gamma\}$,

(ii) $\sum_\gamma \langle TSe_\gamma\,;e_\gamma\rangle = \sum_\gamma \langle STe_\gamma\,;e_\gamma\rangle$,

(iii) $\big|\sum_\gamma \langle S|T|e_\gamma\,;e_\gamma\rangle\big| = \big|\sum_\gamma \big\langle |T|Se_\gamma\,;e_\gamma\big\rangle\big| \le \|S\|\,\|T\|_1$.

Proof of Claim 3. Let $T = W_1|T|$ be the polar decomposition of $T \in \mathcal{B}_1[\mathcal{X}]$. Since $|T|^{\frac{1}{2}} \in \mathcal{B}_2[\mathcal{X}]$ with $\big\| |T|^{\frac{1}{2}}\big\|_2 = \|T\|_1^{\frac{1}{2}}$ and $\|W_1^*\| = 1$ we get by Lemma 5.1(b),

$$
\begin{aligned}
(\mathrm{i}_1) \qquad \sum\nolimits_\gamma |\langle Te_\gamma\,;e_\gamma\rangle| &= \sum\nolimits_\gamma \big|\big\langle |T|^{\frac{1}{2}}e_\gamma\,;|T|^{\frac{1}{2}}W_1^*e_\gamma\big\rangle\big| \le \sum\nolimits_\gamma \big\| |T|^{\frac{1}{2}}e_\gamma\big\| \, \big\| |T|^{\frac{1}{2}}W_1^*e_\gamma\big\| \\
&\le \Big(\sum\nolimits_\gamma \big\| |T|^{\frac{1}{2}}e_\gamma\big\|^2\Big)^{\frac{1}{2}} \Big(\sum\nolimits_\gamma \big\| |T|^{\frac{1}{2}}W_1^*e_\gamma\big\|^2\Big)^{\frac{1}{2}} \\
&\le \|T\|_1^{\frac{1}{2}} \, \big\| |T|^{\frac{1}{2}}W_1^*\big\|_2 \le \|T\|_1^{\frac{1}{2}} \, \big\| |T|^{\frac{1}{2}}\big\|_2 \, \|W_1^*\| \le \|T\|_1
\end{aligned}
$$

(by the Schwarz inequality on both Hilbert spaces $\mathcal{X}$ and ℓ^2 over Γ again). Thus $\{\langle Te_\gamma\,;e_\gamma\rangle\}$ is a summable family, and by taking an arbitrary orthonormal basis $\{f_\gamma\}$ for the Hilbert space $\mathcal{X}$, and applying the Fourier Series Theorem, we get

$$(\mathrm{i}_2) \qquad \sum\nolimits_\gamma \langle Te_\gamma\,;e_\gamma\rangle = \sum\nolimits_{\gamma,\delta} \langle Te_\gamma\,;f_\delta\rangle\,\langle f_\delta\,;e_\gamma\rangle, = \sum\nolimits_\delta \langle f_\delta\,;T^*f_\delta\rangle = \sum\nolimits_\delta \langle Tf_\delta\,;f_\delta\rangle.$$

By (i_1) and (i_2) we get (i). As ST and TS lie in $\mathcal{B}_1[\mathcal{X}]$ by (d), it follows by (i) that the sums in (ii) exist and do not depend on the orthonormal basis. So let $\{f_\gamma\}$ be any orthonormal basis for $\mathcal{X}$ and consider again the Fourier Series Theorem. Thus

(ii) $$\begin{aligned}\sum\nolimits_\gamma\langle TSe_\gamma\,;e_\gamma\rangle &= \sum\nolimits_\gamma\langle Se_\gamma\,;T^*e_\gamma\rangle = \sum\nolimits_{\gamma,\delta}\langle Se_\gamma\,;f_\delta\rangle\langle Tf_\delta\,;e_\gamma\rangle\\ &= \sum\nolimits_{\gamma,\delta}\langle Te_\gamma\,;f_\delta\rangle\langle Sf_\delta\,;e_\gamma\rangle = \sum\nolimits_\gamma\langle Te_\gamma\,;S^*e_\gamma\rangle = \sum\nolimits_\gamma\langle STe_\gamma\,;e_\gamma\rangle.\end{aligned}$$

Recall: $T\in\mathcal{B}_1[\mathcal{X}] \iff |T|\in\mathcal{B}_1[\mathcal{X}] \iff |T|^{\frac{1}{2}}\in\mathcal{B}_2[\mathcal{X}]$ with $\|T\|_1 = \big\|\,|T|\,\big\|_1 = \big\|\,|T|^{\frac{1}{2}}\big\|_2^2$ (so that $S|T|$ and $|T|S$ lie in $\mathcal{B}_1[\mathcal{X}]$ by (d)), and if $A\in\mathcal{B}_2[\mathcal{X}]$ and $B\in\mathcal{B}[\mathcal{X}]$, then $\|AB\|_2\le\|A\|_2\|B\|$ by Lemma 5.1(b). Therefore, before applying (ii), we get

(iii) $$\begin{aligned}\Big|\sum\nolimits_\gamma\langle S|T|e_\gamma\,;e_\gamma\rangle\Big| &= \Big|\sum\nolimits_\gamma\big\langle |T|^{\frac{1}{2}}e_\gamma\,;|T|^{\frac{1}{2}}S^*e_\gamma\big\rangle\Big|\\ &\le \Big(\sum\nolimits_\gamma\big\|\,|T|^{\frac{1}{2}}e_\gamma\|^2\Big)^{\frac{1}{2}}\Big(\sum\nolimits_\gamma\big\|\,|T|^{\frac{1}{2}}S^*e_\gamma\|^2\Big)^{\frac{1}{2}}\\ &= \big\|\,|T|^{\frac{1}{2}}\big\|_2\,\big\|\,|T|^{\frac{1}{2}}S^*\big\|_2 \le \big\|\,|T|^{\frac{1}{2}}\big\|_2\,\big\|\,|T|^{\frac{1}{2}}\big\|_2\,\|S^*\|\\ &= \|T\|_1\,\|S\|. \quad\square\end{aligned}$$

Now apply Claim 3(iii) to support the following argument. Consider again the polar decompositions $T=W_1|T|$, $ST=W_L|ST|$, and $TS=W_R|TS|$, where W_1, W_L, W_R are partial isometries in $\mathcal{B}[\mathcal{X}]$ (with norm one as well as their adjoint), and so $|ST|=W_L^*ST=W_L^*SW_1|T|$ and $|TS|=W_R^*TS=W_R^*W_1|T|S$. Since T lies in $\mathcal{B}_1[\mathcal{X}]$, Claim 3(iii) ensures $\|ST\|_1=\sum_\gamma\big\langle|ST|e_\gamma\,;e_\gamma\big\rangle=\sum_\gamma\langle W_L^*SW_1|T|e_\gamma\,;e_\gamma\rangle\le\|W_L^*SW_1\|\,\|T\|_1\le\|W_L^*\|\,\|S\|\,\|W_1\|\,\|T\|_1=\|S\|\,\|T\|_1$. Since $W_R^*W_1|T|\in\mathcal{B}_1[\mathcal{X}]$, Claim 3(iii) also ensures $\|TS\|_1=\sum_\gamma\big\langle|TS|e_\gamma\,;e_\gamma\big\rangle=\sum_\gamma\langle W_R^*W_1|T|Se_\gamma\,;e_\gamma\rangle=\sum_\gamma\langle SW_R^*W_1|T|e_\gamma\,;e_\gamma\rangle\le\|SW_R^*W_1\|\,\|T\|_1\le\|S\|\,\|W_R^*\|\,\|W_1\,\|T\|_1=\|S\|\,\|T\|_1$. Thus

$$\max\{\|ST\|_1,\|TS\|_1\}\le\|S\|\,\|T\|_1.$$

If $T\in\mathcal{B}_1[\mathcal{X}]$, then $T=AB$ with $A,B\in\mathcal{B}_2[\mathcal{X}]$ by (b) (and so $A^*,B^*\in\mathcal{B}_2[\mathcal{X}]$ by Lemma 5.1(e)). Then, using (b) again, $T^*=B^*A^*\in\mathcal{B}_1[\mathcal{X}]$. Dually, if $T^*\in\mathcal{B}_1[\mathcal{X}]$, then $T=T^{**}\in\mathcal{B}_1]\mathcal{X}]$. Now by taking the polar decompositions $T=W_1|T|$ and $T^*=W_1'|T^*|$ we get $|T^*|=W_1'^*T^*=W_1'^*|T|W_1^*$. Therefore $\|T^*\|_1=\big\|\,|T^*|\,\big\|_1=\|W_1'^*|T|W_1^*\|_1\le\|W_1'^*\|\,\big\|\,|T|\,\big\|_1\,\|W_1^*\|_1=\big\|\,|T|\,\big\|_1=\|T\|_1$ by (e) (proved above). Dually, $\|T\|_1=\|T^{**}\|_1\le\|T^*\|_1$. Thus

$$T\in\mathcal{B}_1[\mathcal{X}]\iff T^*\in\mathcal{B}_1[\mathcal{X}]\quad\text{and}\quad\|T^*\|_1=\|T\|_1.$$

Finally, recall that $\mathcal{R}(T)^-=\mathcal{N}(T^*)^\perp$. If $\dim\mathcal{R}(T)$ is finite, then so is $\dim\mathcal{N}(T^*)^\perp$. Let $\{e_\delta\}$ be an orthonormal basis for $\mathcal{N}(T^*)$ and let $\{e_k\}$ be a finite orthonormal basis for $\mathcal{N}(T^*)^\perp$. As $\mathcal{X}=\mathcal{N}(T^*)+\mathcal{N}(T^*)^\perp$, then $\{e_\gamma\}=\{e_\delta\}\cup\{e_k\}$ is an orthonormal basis for $\mathcal{X}$. Now either $T^*e_\gamma=0$ or $T^*e_\gamma=T^*e_k$. Therefore $\sum_\gamma\big\langle|T^*|e_\gamma;e_\gamma\big\rangle=\sum_k\big\langle|T^*|e_k;e_k\big\rangle<\infty$. Thus $T^*\in\mathcal{B}_1[\mathcal{X}]$. So $T\in\mathcal{B}_1[\mathcal{X}]$ by (f). Hence $\mathcal{B}_0[\mathcal{X}]\subseteq\mathcal{B}_1[\mathcal{X}]$:

Every finite-rank operator is trace-class. $\square$

To proceed we need the following auxiliary result which will support Remark 5.4 and Theorem 6.1. It is a standard application of the Spectral Theorem for compact operators (for similar versions see, e.g., [19, Theorem 6.14.1], [29, Theorem 7.6]).

Proposition 5.3. *If T is compact, then there exist an orthonormal basis $\{e_\gamma\}$ for $\mathcal{X}$ and a family of nonnegative numbers $\{\mu_\gamma\}$ such that*

$$|T|x = \sum\nolimits_\gamma \mu_\gamma \langle x\,;e_\gamma\rangle e_\gamma \quad \textit{for every} \quad x \in \mathcal{X}.$$

Proof. The operator $|T| \in \mathcal{B}[\mathcal{X}]$ is nonnegative (so normal) and compact. (As the class of compact operators from $\mathcal{B}[\mathcal{X}]$ is an ideal of $\mathcal{B}[\mathcal{X}]$, the nonnegative square root $|T|$ of the nonnegative compact $|T|^2$ is compact since $|T|^2 = T^*T$ is compact — see e.g., [13, Problem 5.62].) Since $\big\||T|x\big\| = \|Tx\|$ for every $x \in \mathcal{X}$ we get $\mathcal{N}\big(|T|\big) = \mathcal{N}(T)$. Then by the Spectral Theorem there is a countable orthonormal basis $\{e_k\}$ for the separable Hilbert space $\mathcal{H} = \mathcal{N}(T)^\perp$ consisting of eigenvectors of $|T|$ associated with positive eigenvalues $\{\mu_k\}$ of $|T|$ such that $|T|u = \sum_k \mu_k\langle u\,;e_k\rangle e_k$ for every $u \in \mathcal{H}$ (see, e.g., [14, Corollary 3.4]). Since $\mathcal{X} = \mathcal{H} \oplus \mathcal{N}$ with $\mathcal{N} = \mathcal{N}(T)$, there is an orthonormal basis $\{e_\gamma\} = \{e_k\} \cup \{e_\delta\}$ for $\mathcal{X}$. Here $\{e_\delta\}$ is an orthonormal basis for the (not necessarily separable) Hilbert space $\mathcal{N}$, where $|T|v = 0$ for $v \in \mathcal{N}$ so that the above expansion on $\mathcal{H}$ describes $|T|x$ for all $x = u \oplus v$ in the orthogonal direct sum $\mathcal{X} = \mathcal{H} \oplus \mathcal{N}$ (with $|T|e_\delta = 0$, $Te_k = \mu_k e_k$, and $\mu_\gamma = 0$ if $\gamma \neq k$). □

Remark 5.4. Take $T \in \mathcal{B}[\mathcal{X}]$ so that $\|Tx\| = \big\||T|x\big\|$ for every $x \in \mathcal{X}$. If $\{e_\gamma\}$ is any orthonormal basis for $\mathcal{X}$, then $\sum_\gamma \big\langle |T|e_\gamma\,;e_\gamma\big\rangle \le \sum_\gamma \big\||T|e_\gamma\big\| = \sum_\gamma \|Te_\gamma\|$ and so

$$\sum\nolimits_\gamma \|Te_\gamma\| < \infty \implies \sum\nolimits_\gamma \big\langle |T|e_\gamma\,;e_\gamma\big\rangle < \infty \iff T \in \mathcal{B}_1[\mathcal{X}].$$

Conversely, Proposition 5.3 ensures the existence of an orthonormal basis $\{e_\gamma\}$ for $\mathcal{X}$ such that $|T|e_\gamma = \mu_\gamma e_\gamma$. So $\sum_\gamma \|Te_\gamma\| = \sum_\gamma \big\||T|e_\gamma\big\| = \sum_\gamma \mu_\gamma = \sum_\gamma \big\langle |T|e_\gamma\,;e_\gamma\big\rangle$. Since $\mathcal{B}_1[\mathcal{X}] \subseteq \mathcal{B}_\infty[\mathcal{X}]$ according to Theorem 5.2(c) and Lemma 5.1(d), we get

$$T \in \mathcal{B}_1[\mathcal{X}] \iff \sum\nolimits_\gamma \big\langle |T|e_\gamma\,;e_\gamma\big\rangle < \infty \implies \sum\nolimits_\gamma \|Te_\gamma\| < \infty.$$

Therefore

(a) $\quad T \in \mathcal{B}_1[\mathcal{X}] \iff \sum_\gamma \|Te_\gamma\| < \infty$ *for some orthonormal basis* $\{e_\gamma\}$.

Again, suppose $\sum_\gamma \big\langle |T|e_\gamma; e_\gamma\big\rangle < \infty$, which means T lies in $\mathcal{B}_1[\mathcal{X}]$. By Theorem 5.2(b) $T = AB$ with $A, B \in \mathcal{B}_2[\mathcal{X}]$ and so $\sum_\gamma \|Ae_\gamma\|^2 < \infty$ and $\sum_\gamma \|Be_\gamma\|^2 < \infty$ for an arbitrary orthonormal basis $\{e_\gamma\}$ for $\mathcal{X}$. Since $2|\langle Te_\gamma, e_\gamma\rangle| = 2|\langle ABe_\gamma; e_\gamma\rangle| = 2|\langle Be_\gamma; A^*e_\gamma\rangle| \le 2\|Be_\gamma\|\,\|A^*e_\gamma\| \le \|Be_\gamma\|^2 + \|A^*e_\gamma\|^2$, we get $2\sum_\gamma |\langle Te_\gamma, e_\gamma\rangle| \le \sum_\gamma \|Be_\gamma\|^2 + \sum_\gamma \|A^*e_\gamma\|^2 = \sum_\gamma \|Be_\gamma\|^2 + \sum_\gamma \|Ae_\gamma\|^2 < \infty$. Hence

(b) $\quad T \in \mathcal{B}_1[\mathcal{X}] \implies \sum_\gamma |\langle Te_k; e_\gamma\rangle| < \infty$ *for every orthonormal basis* $\{e_\gamma\}$.

(Actually, Claim 3(i) in the proof of Theorem 5.2 has shown by a different proof that $T \in \mathcal{B}_1[\mathcal{X}]$ implies $\sum_\gamma |\langle Te_k; e_\gamma\rangle| < \|T\|_1$.) However, the converse of (b) fails:

(c) $\quad \sum_\gamma |\langle Te_k; e_\gamma\rangle| < \infty$ *for every orthonormal basis* $\{e_\gamma\} \;\not\Longrightarrow\; T \in \mathcal{B}_1[\mathcal{X}]$.

Indeed, take a unilateral shift $S \in \mathcal{B}[\mathcal{X}]$ of multiplicity one on an infinite-dimensional separable Hilbert space $\mathcal{X}$. Then S shifts a countable orthonormal basis for $\mathcal{X}$. Say $Se_k = e_{k+1}$ for each integer $k > 0$ for some orthonormal basis $\{e_k\}$ for $\mathcal{X}$. Observe that $\langle Sf_k; f_k\rangle = 0$ for every orthonormal basis $\{f_k\}$ for $\mathcal{X}$. In fact, take any orthonormal basis $\{f_k\}$ for $\mathcal{X}$ and consider the Fourier expansion of f_k in terms of

$\{e_k\}$, viz., $f_k = \sum_j \langle f_k; e_j\rangle e_j$, and so $Sf_k = \sum_j \langle f_k; e_j\rangle Se_j$. Then

$$\begin{aligned}\langle Sf_k; f_k\rangle &= \Big\langle \sum\nolimits_j \langle f_k; e_j\rangle e_{j+1}; \sum\nolimits_i \langle f_k; e_i\rangle e_i \Big\rangle = \sum\nolimits_{i,j} \langle f_k; e_j\rangle \overline{\langle f_k; e_i\rangle} \langle e_{j+1}; e_i\rangle \\ &= \sum\nolimits_j \langle f_k; e_j\rangle \overline{\langle f_k; e_{j+1}\rangle} = \sum\nolimits_j \langle e_{j+1}; f_k\rangle \overline{\langle e_j; f_k\rangle} = \langle e_{j+1}; e_j\rangle = 0,\end{aligned}$$

by taking the Fourier expansion of each e_k in terms of $\{f_k\}$. Thus $\sum_k |\langle Sf_k; f_k\rangle| = 0$ for every orthonormal basis $\{f_k\}$. But S is an isometry, so $S^*S = I$ and hence $|S| = I$, the identity on $\mathcal{X}$. Thus $S \notin \mathcal{B}_1[\mathcal{X}]$ (it is not even compact).

Theorem 5.5. *Let $\mathcal{X}$ be a Hilbert space. The following assertions hold true.*

(a) $(\mathcal{B}_1[\mathcal{X}], \|\cdot\|_1)$ *is a Banach space.*

(b) $\mathcal{B}_0[\mathcal{X}]$ *is dense in* $(\mathcal{B}_1[\mathcal{X}], \|\cdot\|_1)$.

Proof. (a) Essentially the same argument that proves completeness of $(\ell_1, \|\cdot\|_1)$. Let $\{T_n\}$ be an arbitrary $\mathcal{B}_1[\mathcal{X}]$-valued Cauchy sequence in $(\mathcal{B}_1[\mathcal{X}], \|\cdot\|_1)$. Then it is a Cauchy sequence in the Banach space $(\mathcal{B}[\mathcal{X}], \|\cdot\|)$ (since $\|\cdot\| \le \|\cdot\|_1$), and so

$$\|T_n - T\| \to 0 \quad \text{for some} \quad T \in \mathcal{B}[\mathcal{X}].$$

Recall that the product of a pair of uniformly convergent sequences of operators converges uniformly to the product of the limits, and also that uniform convergence is preserved both under the adjoint and under the square root operations (see, e.g., [13, Problems 4.46, 5.26, 5.63]). Thus $\|T_n - T\| = \big\| |T_n - T|^{\frac{1}{2}} \big\|^2 \to 0$ implies

$$\big\| |T_n - T|^{\frac{1}{2}} \big\| \to 0 \quad \text{and} \quad \big\| |T_n|^{\frac{1}{2}} - |T|^{\frac{1}{2}} \big\| \to 0.$$

Let $\{e_\gamma\}_{\gamma\in\Gamma}$ be any orthonormal basis for $\mathcal{X}$. Thus $\big\| |T_n|^{\frac{1}{2}} e_\gamma \big\|^2 \to \big\| |T|^{\frac{1}{2}} e_\gamma \big\|^2$ and so

$$\|T\|_1 = \sum\nolimits_{\gamma\in\Gamma} \big\| |T|^{\frac{1}{2}} e_\gamma \big\|^2 \le \sup_n \sum\nolimits_{\gamma\in\Gamma} \big\| |T_n|^{\frac{1}{2}} e_\gamma \big\|^2 = \sup_n \|T_n\|_1 < \infty,$$

since $\{T_n\}$ is bounded in $(\mathcal{B}_1[\mathcal{X}], \|\cdot\|_1)$ because it is Cauchy in $(\mathcal{B}_1[\mathcal{X}], \|\cdot\|_1)$. Then $T \in \mathcal{B}_1[\mathcal{X}]$. (Recall that $\|T_n - T\| \to 0$ and $T \in \mathcal{B}_1[\mathcal{X}]$ does not imply $\|T_n - T\|_1 \to 0$ — see [12] for all possible implications along this line.) Now take an arbitrary $\varepsilon > 0$. Again, since $\{T_n\}$ is Cauchy in $(\mathcal{B}_1[\mathcal{X}], \|\cdot\|_1)$, there exists a finite positive integer n_ε such that for every $m, n \ge n_\varepsilon$,

$$\sum\nolimits_{\gamma\in J} \big\| |T_n - T_m|^{\frac{1}{2}} e_\gamma \big\|^2 \le \sum\nolimits_{\gamma\in\Gamma} \big\| |T_n - T_m|^{\frac{1}{2}} e_\gamma \big\|^2 = \|T_n - T_m\|_1 < \varepsilon$$

for every finite set $J \subseteq \Gamma$. Since $\lim_m \big\| |T_n - T_m|^{\frac{1}{2}} e_\gamma \big\|^2 = \big\| |T_n - T|^{\frac{1}{2}} e_\gamma \big\|^2$ for each n and e_γ, it follows that $\sum_{\gamma\in J} \big\| |T_n - T|^{\frac{1}{2}} e_\gamma \big\|^2 < \varepsilon$ for every finite set $J \subseteq \Gamma$ and so

$$\|T_n - T\|_1 = \sum\nolimits_{\gamma\in\Gamma} \big\| |T_n - T|^{\frac{1}{2}} e_\gamma \big\|^2 = \sup_J \sum\nolimits_{\gamma\in J} \big\| |T_n - T|^{\frac{1}{2}} e_\gamma \big\|^2 \le \varepsilon$$

whenever $n \ge n_\varepsilon$, where the supremum is taken over all finite sets $J \subseteq \Gamma$. This means $\|T_n - T\|_1 \to 0$. So every Cauchy sequence in $(\mathcal{B}_1[\mathcal{X}], \|\cdot\|_1)$ converges in it.

(b) Recall: (i) $\mathcal{B}_0[\mathcal{X}] \subseteq \mathcal{B}_1[\mathcal{X}] \subseteq \mathcal{B}_\infty[\mathcal{X}]$, and (ii) Hilbert spaces have the approximation property which means $\mathcal{B}_0[\mathcal{X}]$ is dense in $(\mathcal{B}_\infty[\mathcal{X}], \|\cdot\|)$, and so $\mathcal{B}_0[\mathcal{X}]$ is dense in $(\mathcal{B}_1[\mathcal{X}], \|\cdot\|)$. To verify that $\mathcal{B}_0[\mathcal{X}]$ is dense in $(\mathcal{B}_1[\mathcal{X}], \|\cdot\|_1)$ proceed as follows. Let $\{e_\gamma\}_{\gamma\in\Gamma}$ be any orthonormal basis for $\mathcal{X}$ and take an arbitrary $T \in \mathcal{B}_1[\mathcal{X}]$. So

$$\|T\|_1 = \sum\nolimits_{\gamma\in\Gamma} \big\langle |T| e_\gamma \,; e_\gamma \big\rangle = \sup_J \sum\nolimits_{\gamma\in J} \big\langle |T| e_\gamma \,; e_\gamma \big\rangle < \infty,$$

where the supremum is taken over all finite sets $J \subseteq \Gamma$. Take an arbitrary $\varepsilon > 0$. The above expression (asserting that the family $\{\langle Te_\gamma, e_\gamma\rangle\}_\gamma$ is summable) ensures the existence of a finite set $J_\varepsilon \subseteq \Gamma$ for which $\sum_{\gamma\in\Gamma\setminus J_\varepsilon} \big\langle |T|e_\gamma\,;e_\gamma\big\rangle < \varepsilon$. Then set $\mathcal{X}_\varepsilon = \operatorname{span}\{e_\gamma \in \mathcal{X} \colon \gamma \in J_\varepsilon\}$, a finite-dimensional subspace of $\mathcal{X}$, and take $T_\varepsilon = T|_{\mathcal{X}_\varepsilon}$ in $\mathcal{B}_0[\mathcal{X}]$. (Indeed, $\mathcal{R}(T_\varepsilon) = T(\mathcal{X}_\varepsilon)$ is finite-dimensional since $\mathcal{X}_\varepsilon$ is.) Thus

$$\|T - T_\varepsilon\|_1 = \sum\nolimits_{\gamma\in\Gamma}\big\langle |T - T_\varepsilon|e_\gamma\,; e_\gamma\big\rangle = \sum\nolimits_{\gamma\in\Gamma\setminus J_\varepsilon}\big\langle |T|e_\gamma\,; e_\gamma\big\rangle < \varepsilon,$$

and therefore $\mathcal{B}_0[\mathcal{X}]$ is dense in $(\mathcal{B}_1[\mathcal{X}], \|\cdot\|_1)$. □

Either Claim 3 in the proof of Theorem 5.2 or Remark 5.4(b) ensure that the family $\{\langle Te_\gamma, e_\gamma\rangle\}_\gamma$ is summable (i.e., the series $\{\sum_\gamma\langle Te_\gamma, e_\gamma\rangle\}_\gamma$ converges in $(\mathbb{F}, |\cdot|)$) for every orthonormal basis $\{e_\gamma\}$ for the Hilbert space $\mathcal{X}$, and the limit does not depend on the choice of the orthonormal basis. Therefore if $T \in \mathcal{B}_1[\mathcal{X}]$ and $\{e_\gamma\}$ is any orthonormal basis for $\mathcal{X}$, then set

$$\operatorname{tr}(T) = \sum\nolimits_\gamma\langle Te_\gamma\,; e_\gamma\rangle \quad \text{and hence} \quad \|T\|_1 = \operatorname{tr}\big(|T|\big).$$

The number $\operatorname{tr}(T) \in \mathbb{F}$ is the *trace* of $T \in \mathcal{B}_1[\mathcal{X}]$ (so the terminology *trace-class*). The norm $\|\cdot\|_1 = \operatorname{tr}\big(|\cdot|\big)$ on the linear space $\mathcal{B}_1[\mathcal{X}]$ of all trace-class operators is referred to as the *trace norm*. Thus the trace-class itself can be written as

$$\mathcal{B}_1[\mathcal{X}] = \{T \in \mathcal{B}[\mathcal{X}] \colon\ \operatorname{tr}\big(|T|\big) < \infty\}.$$

By Theorem 5.2 (and Claim 3 in its proof), for every $T \in \mathcal{B}_1[\mathcal{X}]$ and $S \in \mathcal{B}[\mathcal{X}]$,

$$|\operatorname{tr}(T)| \le \|T\|_1, \qquad \operatorname{tr}(T^*) = \overline{\operatorname{tr}(T)}, \qquad \operatorname{tr}(TS) = \operatorname{tr}(ST),$$

$$\big|\operatorname{tr}\big(S|T|\big)\big| = \big|\operatorname{tr}\big(|T|S\big)\big| \le \|S\|\,\|T\|_1.$$

Actually, if $T \in \mathcal{B}_2[\mathcal{X}]$ so that $|T|^2 \in \mathcal{B}_1[\mathcal{X}]$, then

$$\operatorname{tr}\big(|T^2|\big) = \operatorname{tr}(T^*T) = \|T\|_2^2 \le \|T\|_1^2.$$

Also, since inner products are linear in the first argument,

$$\operatorname{tr}(\cdot)\colon \mathcal{B}_1[\mathcal{X}] \to \mathbb{F} \ \text{ is a bounded linear functional} \qquad (\text{i.e., } \operatorname{tr}(\cdot) \in \mathcal{B}_1[\mathcal{X}]^*).$$

Remark 5.6. As $\operatorname{tr}(\cdot)\colon \mathcal{B}_1[\mathcal{X}] \to \mathbb{F}$ is bounded and linear, and according to Theorem 5.2(b,c) and Lemma 5.1(e), consider the function $\langle\ ;\ \rangle_2 \colon \mathcal{B}_2[\mathcal{X}]\times\mathcal{B}_2[\mathcal{X}] \to \mathbb{F}$ given by

$$\langle T; S\rangle_2 = \operatorname{tr}(S^*T) \quad \text{for every} \quad S, T \in \mathcal{B}_2[\mathcal{X}].$$

Equivalently, for every $S, T \in \mathcal{B}_2[\mathcal{X}]$ and for an arbitrary orthonormal basis $\{e_\gamma\}$,

$$\langle T; S\rangle_2 = \sum\nolimits_\gamma\langle Te_\gamma\,; Se_\gamma\rangle.$$

Since $\operatorname{tr}(\cdot)$ is linear, $\operatorname{tr}(T^*) = \overline{\operatorname{tr}(T)}$, and $\langle T; T\rangle_2 = \operatorname{tr}(T^*T) = \|T\|_2^2$, the function $\langle\cdot\,;\cdot\rangle_2$ is a Hermitian symmetric sesquilinear form which induces a quadratic form. In other words, $\langle\cdot\,;\cdot\rangle_2$ is an inner product on $\mathcal{B}_2[\mathcal{X}]$ that induces the norm $\|\cdot\|_2$:

> $(\mathcal{B}_2[\mathcal{X}], \|\cdot\|_2)$ is an inner product space where the norm $\|\cdot\|_2$ is generated by the inner product defined by $\langle T; S\rangle_2 = \operatorname{tr}(S^*T)$ for every $S, T \in \mathcal{B}_2[\mathcal{X}]$.

Note that the Schwartz inequality for this inner product on $\mathcal{B}_2[\mathcal{X}]$ is a straightforward consequence of the definition of the Hilbert–Schmidt norm $\|\cdot\|_2$ (and, of course, of the Schwartz inequality on both Hilbert spaces, $\mathcal{X}$ and ℓ^2 over Γ):

$$\begin{aligned}|\langle T;S\rangle_2| &= |\mathrm{tr}(S^*T)| = \Big|\sum\nolimits_\gamma \langle Te_\gamma\,;Se_\gamma\rangle\Big| \le \sum\nolimits_\gamma \|Te_\gamma\|\,\|Se_\gamma\| \\ &\le \Big(\sum\nolimits_\gamma \|Te_\gamma\|^2\Big)^{\frac{1}{2}}\Big(\sum\nolimits_\gamma \|Se_\gamma\|^2\Big)^{\frac{1}{2}} = \|T\|_2\,\|S\|_2 \quad \text{for every} \quad S,T\in\mathcal{B}_2[\mathcal{X}].\end{aligned}$$

Similarly to Theorem 5.5, it can be verified that

(a) $(\mathcal{B}_2[\mathcal{X}],\|\cdot\|_2)$ is a Hilbert space,

(b) $\mathcal{B}_0[\mathcal{X}]$ is dense in $(\mathcal{B}_2[\mathcal{X}],\|\cdot\|_2)$.

In fact, a proof of (a) follows the same argument as in the proof of Theorem 5.5(a) with $\mathcal{B}_1[\mathcal{X}]$ and$\|\cdot\|_1$ replaced by $\mathcal{B}_2[\mathcal{X}]$ and $\|\cdot\|_2$. This shows that $(\mathcal{B}_2[\mathcal{X}],\|\cdot\|_2)$ is complete, thus a Banach space, and so a Hilbert space since the norm $\|\cdot\|_2$ is induced by an inner product: $\|T\|_2 = \langle T;T\rangle_2{}^{\frac{1}{2}} = \mathrm{tr}\big(|T|^2\big)^{\frac{1}{2}}$ for every $T\in\mathcal{B}_2[\mathcal{X}]$. In the same way, a proof of (b) follows exactly as the proof of Theorem 5.5(b).

Trace and Hilbert–Schmidt classes are naturally extended in a similar fashion to classes of operators $\mathcal{B}_p[\mathcal{X}]$ for every $p>0$ so that $\mathcal{B}_0[\mathcal{X}]\subset\mathcal{B}_p[\mathcal{X}]\subset\mathcal{B}_q[\mathcal{X}]\subset\mathcal{B}_\infty[\mathcal{X}]$ for every p,q such that $0<p<q<\infty$ (see, e.g., [5, 7, 28, 29]). This, however, goes beyond the scope of the present paper. For further readings on trace-class see also, for instance, [24, 1]. In particular (cf. [24, Theorem VI.26] and [1, Theorem 19.2]),

$$(\mathcal{B}_1[\mathcal{X}],\|\cdot\|_1)\ \textit{is a the dual of}\ (\mathcal{B}_\infty[\mathcal{X}],\|\cdot\|) \qquad \big(\textit{i.e.,}\ \mathcal{B}_\infty[\mathcal{X}]^*\cong\mathcal{B}_1[\mathcal{X}]\big),$$

$$(\mathcal{B}[\mathcal{X}],\|\cdot\|)\ \textit{is a the dual of}\ \big(\mathcal{B}_1[\mathcal{X}],\|\cdot\|_1\big) \qquad \big(\textit{i.e.,}\ \mathcal{B}_1[\mathcal{X}]^*\cong\mathcal{B}[\mathcal{X}]\big),$$

which to some extent mirror the well-known classical duals $c_0^*=\ell_1$ and $\ell_1^*=\ell_\infty$ (see, e.g., [18, Examples 1.10.3.4] among many others).

6. Conclusion

Quite often the term nuclear operator is tacitly attributed to trace class operators without further explanation such as, for instance, "an operator will be called *nuclear* if it belongs to $\mathcal{B}_1[\mathcal{X}]$" [7, Section III.8]. This prompts our final result.

Theorem 6.1. *If $\mathcal{X}$ is a Hilbert space*, then

$$\mathcal{B}_1[\mathcal{X}]=\mathcal{B}_N[\mathcal{X}] \quad \textit{and} \quad \|\cdot\|_1=\|\cdot\|_N.$$

Proof. The proof is based on the he polar decomposition for Hilbert-space operators as follows. First, Proposition 5.3 is applied to show that $\mathcal{B}_1[\mathcal{X}]\subseteq\mathcal{B}_N[\mathcal{X}]$ in part (a). Then the injectiveness of T when it acts in $\mathcal{N}(A)^\perp$ is explored to obtain the reverse inclusion $\mathcal{B}_N[\mathcal{X}]\subseteq\mathcal{B}_1[\mathcal{X}]$ in part (b), which also yields the norm identity $\|\cdot\|_1=\|\cdot\|_N$. Thus, to begin with, take an operator $T\in\mathcal{B}[\mathcal{X}]$ on a Hilbert space $\mathcal{X}$.

(a) If $T\in\mathcal{B}_1[\mathcal{X}]$, then it is compact. Thus consider the setup in the proof of Proposition 5.3, where $\{e_\gamma\}$ is an orthonormal basis for the Hilbert space $\mathcal{X}$ and $\{e_k\}$ is a countable subset of it which is an orthonormal basis for the separable Hilbert space $\mathcal{H}=\mathcal{N}(A)^\perp\subseteq\mathcal{X}$. Take the direct sums $T=T|_\mathcal{H}\oplus O$ and $|T|=|T|\big|_\mathcal{H}\oplus O$ on $\mathcal{X}=\mathcal{H}\oplus\mathcal{N}$, where $T|_\mathcal{H}\in\mathcal{B}[\mathcal{H}]$ is injective. Thus $T|_\mathcal{H}$ has a polar decomposition $T|_\mathcal{H}=V|T|\big|_\mathcal{H}$ where $V\in\mathcal{B}[\mathcal{H}]$ is an isometry (see, e.g,. [13, Corollary 5.90]) so that $T = V|T|\big|_{\mathcal{H}} \oplus O \in \mathcal{B}[\mathcal{X}] = \mathcal{B}[\mathcal{H} \oplus \mathcal{N}]$. For each e_k in $\mathcal{H}$ set $f_k = Ve_k$ in $\mathcal{H}$ so that $\{f_k\}$ is an $\mathcal{H}$-valued orthonormal sequence. Then by Proposition 5.3

$$|T|x = \sum\nolimits_\gamma \mu_\gamma \langle x\,; e_\gamma\rangle e_\gamma \quad \text{and so} \quad Tx = \sum\nolimits_\gamma \mu_\gamma \langle x\,; e_\gamma\rangle f_\gamma \quad \text{for every} \quad x \in \mathcal{X},$$

with $f_\gamma = 0 \in \mathcal{N}$ if $\gamma \neq k$. Moreover, recalling that $\mu_\gamma = 0$ if $\gamma \neq k$,

$$\sum\nolimits_\gamma \mu_\gamma = \sum\nolimits_\gamma \sum\nolimits_\delta \mu_\gamma \langle e_\gamma\,; e_\delta\rangle e_\delta = \sum\nolimits_\gamma \big\langle |T|e_\gamma; e_\gamma\big\rangle.$$

Therefore if $T \in \mathcal{B}_1[\mathcal{X}]$, then

$$\sum\nolimits_\gamma \mu_\gamma = \|T\|_1 < \infty \quad \text{and} \quad Tx = \sum\nolimits_k \mu_k \langle x\,; e_k\rangle f_k \quad \text{for every} \quad x \in \mathcal{X},$$

where $\{e_k\}$ and $\{f_k\}$ are $\mathcal{H}$-valued unit sequences (thus $\mathcal{X}$-valued unit sequences).

(b) Conversely, since $\mathcal{X} = \mathcal{N}(T)^\perp \oplus \mathcal{N}(T)$, we may regard only the action of $T = T|_{\mathcal{H}} \oplus O$ and $|T| = |T|\big|_{\mathcal{H}} \oplus O$ on $\mathcal{H} = \mathcal{N}(T)^\perp$. So for notational simplicity write T and $|T|$ for the injective operators $T|_{\mathcal{H}}$ and $|T|\big|_{\mathcal{H}}$. Suppose

$$Tx = \sum\nolimits_k \alpha_k \langle x\,; z_k\rangle y_k \quad \text{for every} \quad x \in \mathcal{H},$$

for some $\mathcal{H}$-valued sequences $\{z_k\}$ and $\{y_k\}$ and for some scalar sequence $\{\alpha_k\}$ with $\|z_k\| = \|y_k\| = 1$ and $\sum_k |\alpha_k| < \infty$. Thus (by polar decomposition)

$$|T|x = V^*Tx = \sum\nolimits_k \alpha_k \langle x\,; z_k\rangle w_k \quad \text{for every} \quad x \in \mathcal{H} = \mathcal{N}(T)^\perp \subseteq \mathcal{X},$$

where $w_k = V^*y_k$ so that $\|w_k\| \leq 1$. For every orthonormal basis $\{e_j\}$ for $\mathcal{H}$,

$$\begin{aligned}\sum\nolimits_j \big\langle |T|e_j\; e_j\big\rangle &= \sum\nolimits_j \Big\langle \sum\nolimits_k \alpha_k \langle e_j; z_k\rangle w_k; e_j\Big\rangle = \sum\nolimits_j \sum\nolimits_k \alpha_k \langle e_j; z_k\rangle \langle w_k; e_j\rangle \\ &\leq \sum\nolimits_k |\alpha_k| \Big|\sum\nolimits_j \langle e_j; z_k\rangle \langle w_k; e_j\rangle\Big| = \sum\nolimits_k |\alpha_k|\, |\langle w_k; z_k\rangle| \leq \sum\nolimits_k |\alpha_k| < \infty.\end{aligned}$$

So T lies in $\mathcal{B}_1[\mathcal{H}]$. Moreover, since $Tx = \sum_k \alpha_k \langle x\,; z_k\rangle y_k$ for every $x \in \mathcal{H}$,

$$\sum\nolimits_k \mu_k = \|T\|_1 = \sum\nolimits_k \big\langle |T|e_k; e_k\big\rangle \leq \sum\nolimits_k \big\|\,|T|e_k\big\| = \sum\nolimits_k \|Te_k\| \leq \sum\nolimits_k |\alpha_k|,$$

for any orthonormal basis $\{e_k\}$ for $\mathcal{H}$. So $\|T\|_1 = \min \sum_k |\alpha_k|$, the minimum taken over all scalar summable sequences $\{\alpha_k\}$ as in the above representation of T.

From (a) and (b) we get the following statement:

> An operator T lies in $\mathcal{B}_1[\mathcal{X}]$ if and only if there are $\mathcal{X}$-valued unit sequences $\{z_k\}$ and $\{y_k\}$ (i.e., $\|z_k\| = \|y_k\| = 1$) and a scalar summable sequence $\{\alpha_k\}$ (i.e., $\sum_k |\alpha_k| < \infty$) such that $Tx = \sum_k \alpha_k \langle x\,; z_k\rangle y_k$ for every x in $\mathcal{X}$. Moreover, $\|T\|_1 = \inf \sum_k |\alpha_k|$, where the infimum is taken over all scalar summable sequences for which the above expression for Tx holds.

Such an expression for Tx is precisely a nuclear representation of T as in Section 4. Then $\mathcal{B}_1[\mathcal{X}] = \mathcal{B}_N[\mathcal{X}]$. Also, with the infimum taken over all nuclear representations of $T \in \mathcal{B}_N[\mathcal{X}]$, we get (for arbitrary unit sequences $\{z_k\}$ and $\{y_k\}$ and summable scalar sequence $\{\alpha_k\}$) $\|T\|_N = \inf \sum_k |\alpha_k| \|z_k\|\,\|y_k\| = \inf \sum_k |\alpha_k| = \|T\|_1$. □

Thus the notions of nuclear and trace-class coincide on Hilbert spaces. For their relationship beyond Hilbert spaces, the same first lines of Bartle's review in Mathematical Reviews we borrowed to open the paper can be used to close it, viz., "Grothendieck showed that a Banach space $\mathcal{X}$ has the approximation property if and only if for every nuclear operator $T = \sum_k f_k(\cdot)y_k$ the number $\mathrm{tr}(T) = \sum_k f_k(y_k)$

is well-defined" (see, e.g., [4, Theorems 1.3.6, 1.3.11, 1.4.18 and Proposition 1.4.19]). This can be regarded as a starting point for characterizing the trace property in Banach spaces. For further readings along this line see, for instance, [21, 11, 6].

Acknowledgment

The author thanks the referees for their suggestions to improve the paper.

References

1. J.B. Conway, *A Course in Operator Theory*, Graduate Studies in Mathematics Vol. 21, Amer. Math. Soc., Providence, 2000.
2. A. Defant and K. Floret, *Tensor Norms and Operator Ideals*, North-Holland, Amsterdam, 1993.
3. R. Didier, *On the trace of operators (from Grothendieck to Lidskiĭ)*, EMS Newsletter, September (2017), 26–33.
4. J. Diestel, J.H. Fourie, and J. Swart, *The Metric Theory of Tensor Products – Grothendieck's Résumé Revisited*, American Mathematical Society, Providence, 2008.
5. N. Dunford and J.T. Schwartz, *Linear Operators – Part II: Spectral Theory – Self Adjoint Operators in Hilbert Space*, Interscience, New York, 1963; reprinted: Wiley, New York, 1988.
6. T. Figiel and W. B. Johnson, *The Lidskii trace property and the nest approximation property in Banach spaces*, J. Funct. Anal. **271** (2016), 566–576.
7. I.C. Gohberg and M.G. Kreĭn, *Introduction to the Theory of Nonselfadjoint Operators*, Translations of Mathematical Monographs Vol. 18, Amer. Math. Soc., Providence, 1969.
8. A. Grothendieck, *Produits tensoriels topologiques et espaces nucléaires*, Mem. Amer. Math. Soc, **16**, (1955) 1–196, 1–140.
9. C.W. Ha, *On the trace of a class of nuclear operators*, Bull. Inst. Math. Acad. Sinica **3** (1975), 131–137.
10. H. Jarchow, *Locally Convex Spaces*, B.G. Teubner, Stuttgart, 1981.
11. W. B. Johnson and A. Szankowski, *The trace formula in Banach spaces*, Israel J. Math. **203** (2014), 389–404
12. C.S. Kubrusly, *On convergence of nuclear and correlation operators in Hilbert space*, Mat. Apl. Comput. **5** (1986), 265–282.
13. C.S. Kubrusly, *The Elements of Operator Theory*, 2nd edn. Birkhäuser-Springer, New York, 2011.
14. C.S. Kubrusly, *Spectral Theory of Bounded Linear Operators*, Birkhäuser-Springer-Switzerland, Cham, 2020.
15. C.S. Kubrusly, *Algebraic tensor products revisited: axiomatic approach*, Bull. Malays. Math. Sci. Soc. **44** (2021), 2335–2355.
16. C.S. Kubrusly, *On extensions of bilinear maps*, Math. Slovaca, to appear (2022), available at https://arxiv.org/pdf/2110.03102
17. V.B. Lidskiĭ, *Non-selfadjoint operators with a trace*, Dokl. Akad. Nauk SSSR **125** (1959), 485–587.
18. R.E. Megginson, *An Introduction to Banach Space Theory*, Springer, New York, 1998.
19. A.W. Naylor and G.R. Sell, *Linear Operator Theory in Engineering and Science*, Holt, Rinehart & Winston, New York, 1971; reprinted: Springer, New York, 1982.
20. A. Pietsch, *Traces of Hilbert space operators and their recent history*, Quest. Math. **43** (2020), 623–649.
21. G. Pisier, *Weak Hilbert spaces*, Proc. London Math. Soc. (3) **56** (1988), 547–579.
22. G. Pisier, *Grothendieck's theorem, past and present*, Bull. Amer. Math. Soc. **49** (2012), 237–323.
23. S.C. Power, *Another proof of Lidskiĭ's theorem on the trace*, Bull. London Math. Soc. **15** (1983), 146–148.
24. M. Reed and B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis*, 2nd edn. Academic Press, New York, 1980.
25. O.I. Reinov, *Some remarks on the spectra of nuclear operators*, Open Math. **16** (2018), 453–460.
26. R.A. Ryan, *Introduction to Tensor Products of Banach Spaces*, Springer, London, 2002.

27. R. Schatten, *A Theory of Cross-Spaces*, Princeton University Press, Princeton, 1950.
28. R. Schatten, *Norm Ideals of Completely Continuous Operators*, Springer, New York, 1970.
29. J. Weidmann, *Linear Operators in Hilbert Spaces*, Springer, New York, 1980.

Catholic University of Rio de Janeiro, Brazil
E-mail address: carlos@ele.puc-rio.br