Chapter

Continuous Linear Forms in Hilbert Space

V

§ 1. *Riesz-Frechet theorem*
§ 2. *Completion*
§ 3. *Bilinear mappings*
§ 4. *Bounded bilinear mappings*
§ 5. *Sesquilinear mappings*
§ 6. *Bounded sesquilinear mappings*
§ 7. *Bounded sesquilinear forms in Hilbert space*
§ 8. *Adjoints*

§ 1. RIESZ-FRECHET THEOREM

Chapters VI onward are concerned with the study of continuous linear mappings $T: \mathfrak{H} \to \mathfrak{H}$ of a Hilbert space into itself. This study is flavored, so to speak, by the circumstance that the dual space $\mathfrak{H}'$ can in a natural way be identified with $\mathfrak{H}$ (see *Theorem 2* below); it is this "self-duality" of Hilbert space which distinguishes the theory of continuous linear mappings in Hilbert space from the theory of such mappings in Banach space.

If $\mathfrak{P}$ is any pre-Hilbert space, each vector $y \in \mathfrak{P}$ determines a continuous linear form y' on $\mathfrak{P}$, namely $y'(x) = (x|y)$. This gives rise to a mapping $y \to y'$, of $\mathfrak{P}$ into its dual space $\mathfrak{P}'$, which is *conjugate-linear* and *isometric* (see *Theorem IV.10.2*). If this mapping is surjective, $\mathfrak{P}$ is necessarily complete; we are concerned here with the converse, which is the heart of the whole chapter (see *Theorem 7.2*):

Theorem 1. (Riesz-Frechet theorem) *If f is a continuous linear form on a Hilbert space $\mathfrak{H}$, there exists one and only one vector $y \in \mathfrak{H}$ such that $f(x) = (x|y)$ for all $x \in \mathfrak{H}$.*

Proof.

That is, given $f \in \mathfrak{H}'$, the problem is to find a vector $y \in \mathfrak{H}$ such that $y' = f$.

Let $\mathfrak{N}$ be the null space of f; it is a closed linear subspace of $\mathfrak{H}$, by *Theorem IV.7.2*. If $\mathfrak{N} = \mathfrak{H}$, then $f = 0$, and one takes $y = \theta$. Assuming $f \neq 0$, $\mathfrak{N} \neq \mathfrak{H}$; since $\mathfrak{N}^{\perp\perp} = \mathfrak{N}$ by *Theorem III.6.2*, it follows that $\mathfrak{N}^{\perp} \neq \{\theta\}$. Let $z \in \mathfrak{N}^{\perp}$, $z \neq \theta$. Since $\mathfrak{N}$ and $\mathfrak{N}^{\perp}$ have only the vector θ in common, necessarily $f(z) \neq 0$; replacing z by $[f(z)]^{-1}z$, we may suppose $f(z) = 1$.

Given any $x \in \mathfrak{H}$; one has $x - f(x)z \in \mathfrak{N}$, since $f[x - f(x)z] = f(x) - f(x)f(z) = f(x) - f(x) = 0$. Since $z \perp \mathfrak{N}$, $0 = (x - f(x)z \mid z) = (x \mid z) - f(x)(z \mid z)$. Thus, $f(x) = (z \mid z)^{-1}(x \mid z) = (x \mid y)$, where $y = (z \mid z)^{-1}z$ is independent of x. That is, $f = y'$.

If also $f = w'$, $(x \mid y) = (x \mid w)$ for all x; $y = w$ results from *Theorem II.2.1*. ∎

Combining *Theorem 1* and *Theorem IV.10.2*,

Theorem 2. *Let $\mathfrak{P}$ be a pre-Hilbert space, $y \to y'$ the mapping of $\mathfrak{P}$ into $\mathfrak{P}'$ described in Theorem IV.10.2. This mapping is surjective if and only if $\mathfrak{P}$ is a Hilbert space.*

Exercises

1. Let f be a linear form on the vector space $\mathfrak{V}$, $f \neq 0$, and $\mathfrak{N}$ the null space of f. Show that there exists a vector y with the following property: every $x \in \mathfrak{V}$ has a unique representation $x = \lambda y + z$, where $z \in \mathfrak{N}$ and λ is a suitable scalar.

2. Let $\mathfrak{V}, f, \mathfrak{N}$ be as in *Exercise 1*. Suppose $\mathfrak{M}$ is a linear subspace of $\mathfrak{V}$ such that (1) $\mathfrak{M} \neq \{\theta\}$, and (2) $\mathfrak{M}$ and $\mathfrak{N}$ have only the vector θ in common. Prove that $\mathfrak{M}$ is one-dimensional, and every $x \in \mathfrak{V}$ has a unique representation $x = y + z$ with $y \in \mathfrak{M}$ and $z \in \mathfrak{N}$.

3. If $\mathfrak{P}$ is a pre-Hilbert space such that $\mathfrak{N}^{\perp\perp} = \mathfrak{N}$ for every closed linear subspace $\mathfrak{N}$ of $\mathfrak{P}$, then $\mathfrak{P}$ is a Hilbert space.

4. Let f be a linear form on the Hilbert space $\mathfrak{H}$, $\mathfrak{N}$ the null space of f. If f is not continuous, show that $\mathfrak{N}$ is a dense linear subspace of $\mathfrak{H}$.

5. If f is a linear form on a Hilbert space $\mathfrak{H}$, and $\mathfrak{N}$ is the null space of f, then f is continuous if and only if $\mathfrak{N}$ is a closed linear subspace.

6. If $\mathfrak{N}$ is a complete linear subspace of a pre-Hilbert space $\mathcal{P}$, then $\mathcal{P} = \mathfrak{N} \oplus \mathfrak{N}^{\perp}$ (this is *Theorem III.6.1*). Another (admittedly circular) proof of this can be based on *Theorem 1*.

7. Terminology as in *Exercise IV.10.9*. Prove:

(i) In a pre-Hilbert space $\mathcal{P}$, if $x_n \to x$ weakly, then $(x_n|y) \to (x|y)$ for each $y \in \mathcal{P}$.

(ii) In a Hilbert space $\mathcal{H}$, $x_n \to x$ weakly, if and only if $(x_n|y) \to (x|y)$ for each $y \in \mathcal{H}$.

(iii) If x_n is an orthonormal sequence in a Hilbert space, then $x_n \to \theta$ weakly, but one does not have $x_n \to \theta$.

8. Let $\mathcal{E}$ be a normed space, $\mathcal{H}$ a Hilbert space, and $T_n, T \in \mathfrak{L}_c(\mathcal{E},\mathcal{H})$. Then, $T_n \to T$ weakly, in the sense of *Exercise IV.10.11*, if and only if $(T_n x|y) \to (Tx|y)$ for each pair $x \in \mathcal{E}$, $y \in \mathcal{H}$.

9. If $\mathcal{H}$ is a Hilbert space, the mapping $x \to x''$ of $\mathcal{H}$ into $\mathcal{H}''$ is surjective, hence $\mathcal{H}$ is reflexive.

§ 2. COMPLETION

This section is presented as an application of the Riesz-Frechet theorem. It may be considered "optional" and highly omissible.

One knows that the pre-Hilbert space of finitely non-zero sequences (*Example II.1.3*) can be enlarged to a Hilbert space, namely l^2 (*Example II.5.1*). In what sense does this hold for an arbitrary pre-Hilbert space $\mathcal{P}$? In any case, there is a mapping $x \to x''$, of $\mathcal{P}$ into its bidual $\mathcal{P}''$, which preserves sums, scalar multiples, and norms (see *Theorem IV.10.3*). Let us denote this mapping by U; thus, $U: \mathcal{P} \to \mathcal{P}''$, $Ux = x''$. The linear subspace $U(\mathcal{P})$ of $\mathcal{P}''$ is a normed space which is in a sense "equal" to the normed space $\mathcal{P}$; more precisely, U defines an isometric vector space isomorphism of $\mathcal{P}$ with $U(\mathcal{P})$. Since $\mathcal{P}''$ is a Banach space, the closure $\overline{U(\mathcal{P})}$ is also a Banach space (see *Example IV.7.8* and *Exercise IV.7.1*). Thus, after "identifying" $\mathcal{P}$ with $U(\mathcal{P})$, we may think of the normed space $\mathcal{P}$ as having been enlarged to the Banach space $\overline{U(\mathcal{P})}$. There remains the problem of providing every pair of vectors in $\overline{U(\mathcal{P})}$ with a scalar product satisfying the axioms in *Definition II.1.1*, in such a way that the norm deduced from the scalar product is the same as the given Banach space norm on $\overline{U(\mathcal{P})}$. This is done in the following theorem, in which, incidentally, it is shown that $\overline{U(\mathcal{P})} = \mathcal{P}''$.

Theorem 1. (Completion) *If $\mathcal{P}$ is any pre-Hilbert space, there exists a Hilbert space $\mathcal{H}$, and a mapping $U: \mathcal{P} \to \mathcal{H}$, such that:*

(1) *U is linear.*

(2) $(Ux \mid Uy) = (x \mid y)$ *for all* $x,y \in \mathcal{P}$.

(3) *$U(\mathcal{P})$ is a dense linear subspace of $\mathcal{H}$.*

In fact, the norm in the bidual $\mathcal{P}''$ can be deduced from a suitable scalar product, and one can take $\mathcal{H} = \mathcal{P}''$, $Ux = x''$.

Proof.

One knows that $\mathcal{P}''$ is a Banach space (*Theorem IV.10.1*), and that the mapping $U: \mathcal{P} \to \mathcal{P}''$ defined by $Ux = x''$ is linear and isometric (*Theorem IV.10.3*). Moreover, $x''(y') = y'(x) = (x \mid y)$ for all $x,y \in \mathcal{P}$, where $y \to y'$ is the isometric conjugate linear mapping of $\mathcal{P}$ into $\mathcal{P}'$, described in *Theorem IV.10.2*.

Let $\mathcal{H}$ be the closure of $U(\mathcal{P})$ in $\mathcal{P}''$; that is, $\mathcal{H} = \overline{U(\mathcal{P})}$. $\mathcal{H}$ is a closed linear subspace of $\mathcal{P}''$ (see the proof of *Theorem III.3.2*), and is a Banach space (*Example IV.7.8*). Later in the proof, it will be seen that $\mathcal{H} = \mathcal{P}''$.

We now have a linear isometric mapping $U: \mathcal{P} \to \mathcal{H}$, where $\mathcal{H}$ is a Banach space and $U(\mathcal{P})$ is dense in $\mathcal{H}$.

A scalar product is introduced in $\mathcal{H}$ as follows. Given $u,v \in \mathcal{H}$, the problem is to define $(u \mid v)$. Choose sequences $x_n \in \mathcal{P}$, $y_n \in \mathcal{P}$, such that $Ux_n \to u$ and $Uy_n \to v$. Since $\| x_m - x_n \| = \| U(x_m - x_n) \| = \| Ux_m - Ux_n \| \to 0$, x_n is a Cauchy sequence in $\mathcal{P}$; similarly for the sequence y_n. It follows from *Theorem II.5.1* that $(x_n \mid y_n)$ is a convergent sequence of scalars. We wish to define

$$(*) \qquad (u \mid v) = \lim (x_n \mid y_n),$$

hence must check that this limit is independent of the particular sequences approximating u and v. Suppose also $\bar{x}_n \in \mathcal{P}$, $\bar{y}_n \in \mathcal{P}$, with $U\bar{x}_n \to u$ and $U\bar{y}_n \to v$. Then, $\| \bar{x}_n - x_n \| = \| U\bar{x}_n - Ux_n \| \leq \| U\bar{x}_n - u \| + \| u - Ux_n \| \to 0$, and similarly $\| \bar{y}_n - y_n \| \to 0$. The desired equality $\lim (x_n \mid y_n) = \lim (\bar{x}_n \mid \bar{y}_n)$ results from the identity $(\bar{x}_n \mid \bar{y}_n) - (x_n \mid y_n) = (\bar{x}_n - x_n \mid \bar{y}_n - y_n) + (x_n \mid \bar{y}_n - y_n) + (\bar{x}_n - x_n \mid y_n)$ (see the proof of *Theorem II.5.1*). Thus, $(u \mid v)$ is unambiguously defined by the formula (*). Since $Ux_n \to u$, one has $\| Ux_n \| \to \| u \|$, hence $(u \mid u) = \lim (x_n \mid x_n) = \lim \| x_n \|^2 = \lim \| Ux_n \|^2 = \| u \|^2$; from this it is clear that

(i) $(u|u) \geq 0$, $(u|u) = 0$ if and only if $u = \theta$.

Moreover,

(ii) $(v|u)^* = (u|v)$

(iii) $(u + v|w) = (u|w) + (v|w)$

(iv) $(\lambda u|v) = \lambda(u|v)$;

the proofs are left to the reader.

Thus, the scalar products $(u|v)$ satisfy the axioms for a pre-Hilbert space, and the relation $(u|u) = \| u \|^2$ shows that the given notion of $\| u \|$ (calculated as in *Theorem IV.10.1*) coincides with the norm $(u|u)^{\frac{1}{2}}$ deduced from the scalar product (via *Definition II.3.1*); these two (equal) norms will not be distinguished notationally. Since $\mathfrak{H}$ was already a Banach space, the scalar product $(u|v)$ makes $\mathfrak{H}$ into a Hilbert space.

One has $(Ux|Uy) = (x|y)$ for all $x,y \in \mathcal{P}$; for, with the notation of formula (*), one can take $x_n = x$ and $y_n = y$ for all n. Thus, the Hilbert space $\mathfrak{H}$ and the mapping $U: \mathcal{P} \to \mathfrak{H}$ fulfill the desired conditions (1),(2),(3).

It remains only to show that $\mathfrak{H} = \mathcal{P}''$. The proof will be broken up into a series of remarks:

1. *If $u \in \mathfrak{H}$, then $(u|x'') = u(x')$ for all $x \in \mathcal{P}$.*

For, let $x_n \in \mathcal{P}$ with $x_n'' \to u$. Since $x_n''(x') = (x_n|x)$, one has $(u|x'') = \lim (x_n''|x'') = \lim (x_n|x) = \lim x_n''(x') = u(x')$ (for the last equality, see *Exercise IV.8.2*).

2. *If $f \in \mathcal{P}'$, there is a unique vector $u \in \mathfrak{H}$ such that $f(x) = (x''|u)$ for all $x \in \mathcal{P}$.*

For, let g be the linear form on $U(\mathcal{P})$ defined by $g(x'') = f(x)$. Since $|g(x'')| = |f(x)| \leq \| f \| \, \| x \| = \| f \| \, \| x'' \|$, g is continuous. Since $U(\mathcal{P})$ is dense in $\mathfrak{H}$, by *Theorem IV.7.6* there is a continuous linear form h on $\mathfrak{H}$ such that $h(x'') = g(x'')$ for all $x \in \mathcal{P}$. By *Theorem 1.1*, there exists a vector $u \in \mathfrak{H}$ such that $h(v) = (v|u)$ for all $v \in \mathfrak{H}$. In particular, $(x''|u) = h(x'') = g(x'') = f(x)$ for all $x \in \mathcal{P}$.

Suppose also $w \in \mathfrak{H}$ satisfies $f(x) = (x''|w)$ for all $x \in \mathcal{P}$. Then, $(x''|u - w) = f(x) - f(x) = 0$ for all $x \in \mathcal{P}$. Choose a sequence $x_n \in \mathcal{P}$ such that $x_n'' \to u - w$; then, $(u - w|u - w) = \lim (x_n''|u - w) = \lim 0 = 0$, hence $u - w = \theta$.

3. *$\mathfrak{N} = \{y' : y \in \mathcal{P}\}$ is a dense linear subspace of $\mathcal{P}'$.*

$\mathfrak{N}$ is the range of the conjugate-linear mapping $y \to y'$ of $\mathcal{P}$ into $\mathcal{P}'$; it is clearly a linear subspace of $\mathcal{P}'$. Suppose $f \in \mathcal{P}'$, and $\epsilon > 0$ is given; the problem is to find a vector $y \in \mathcal{P}$ such that $\| f - y' \| \leq \epsilon$ (see *Exercise III.2.1*). By Remark 2, there is a vector $u \in \mathcal{K}$ such that $f(x) = (x''|u)$ for all $x \in \mathcal{P}$. Choose any $y \in \mathcal{P}$ such that $\| u - y'' \| \leq \epsilon$; this is possible because $U(\mathcal{P})$ is dense in $\mathcal{K}$. For any $x \in \mathcal{P}$, $|(f - y')(x)| = |f(x) - y'(x)| = |(x''|u) - (x|y)| = |(x''|u) - (x''|y'')| = |(x''|u - y'')| \leq \| x'' \| \| u - y'' \| = \| x \| \| u - y'' \| \leq \epsilon \| x \|$, hence $\| f - y' \| \leq \epsilon$ by *Theorem IV.7.4*. Finally,

4. $\mathcal{K} = \mathcal{P}''$.

Given $F \in \mathcal{P}''$, let us show that $F \in \mathfrak{N}$. For all $y \in \mathcal{P}$, one has $|F(y')| \leq \| F \| \| y' \| = \| F \| \| y \| = \| F \| \| y'' \|$. Defining $G(y'') = [F(y')]^*$, G is a continuous linear form on $U(\mathcal{P})$. Repeating the argument in Remark 2, there is a vector $u \in \mathcal{K}$ such that $G(y'') = (y''|u)$ for all $y \in \mathcal{P}$, hence $(u|y'') = [G(y'')]^* = F(y')$. Quoting Remark 1, $u(y') = F(y')$ for all $y \in \mathcal{P}$. That is, $u(f) = F(f)$ for all $f \in \mathfrak{N}$, where $\mathfrak{N}$ is the dense linear subspace of $\mathcal{P}'$ described in Remark 3; $u = F$ results from *Theorem IV.7.6* (see also *Exercise IV.6.1*). Thus, $F = u \in \mathcal{K}$. ∎

Exercises

1. If x_n is an orthonormal sequence in a pre-Hilbert space $\mathcal{P}$, then $x_n \to \theta$ weakly (see also *Exercise 1.7*).

2. If $\mathcal{K}$ is a Hilbert space, the mapping $x \to x''$ of $\mathcal{K}$ into $\mathcal{K}''$ is surjective, hence $\mathcal{K}$ is reflexive.

3. If $\mathcal{E}$ is a normed space, call a linear subspace $\mathfrak{N}$ *quadratic* in case there exists a scalar product $(x|y)$ defined for $x,y \in \mathfrak{N}$, such that (1) $\mathfrak{N}$ is a pre-Hilbert space, and (2) $(x|x)^{1/2}$ coincides with the given norm $\| x \|$ for $x \in \mathfrak{N}$. Then:

*(i) A linear subspace $\mathfrak{N}$ of $\mathcal{E}$ is quadratic if and only if $\| x + y \|^2 + \| x - y \|^2 = 2 \| x \|^2 + 2 \| y \|^2$ for all $x,y \in \mathfrak{N}$.

(ii) If $\mathfrak{N}$ is a quadratic linear subspace of $\mathcal{E}$, its closure $\overline{\mathfrak{N}}$ is also quadratic.

4. Notation as in the proof of *Theorem 1;* in particular $\mathfrak{N}$ is the dense linear subspace of $\mathcal{P}'$ described in Remark 3 of the proof. Prove:

(i) $\mathfrak{N}$ is quadratic in the sense of *Exercise 3*.

(ii) The Banach space $\mathcal{P}'$ is a Hilbert space.

(iii) Hence, in turn, $\mathcal{P}''$ is a Hilbert space. Let $\mathcal{K} = \mathcal{P}'$ be the Hilbert

space described in (ii). There is a mapping $f \to f'$ of $\mathfrak{X}$ *onto* $\mathfrak{X}' = \mathcal{P}''$ by *Theorem 1.2*. Show: if $x \in \mathcal{P}$, then $(x')' = x''$.

5. Let $\mathcal{E}$ be a normed space. A *completion* for $\mathcal{E}$ is a pair $(\mathcal{B},U)$, where
(1) $\mathcal{B}$ is a Banach space,
(2) $U: \mathcal{E} \to \mathcal{B}$ is a linear mapping,
(3) U is isometric,
(4) $U(\mathcal{E})$ is dense in $\mathcal{B}$.
Then:
*(i) It can be shown that every normed space $\mathcal{E}$ has a completion. Specifically, the mapping $Ux = x''$ of $\mathcal{E}$ into $\mathcal{E}''$ can be shown to be isometric, and one can take for $\mathcal{B}$ the closure in $\mathcal{E}''$ of the linear subspace $U(\mathcal{E})$.
(ii) Suppose $(\mathcal{B},U)$ is a completion for $\mathcal{E}$. If $\mathcal{F}$ is a Banach space, and $S: \mathcal{E} \to \mathcal{F}$ is a continuous linear mapping, there is one and only one continuous linear mapping $T: \mathcal{B} \to \mathcal{F}$ such that $TU = S$. Moreover, $\| T \| = \| S \|$.
(iii) If $(\mathcal{B}_1,U_1)$ and $(\mathcal{B}_2,U_2)$ are two completions for $\mathcal{E}$, there exists a unique mapping $T: \mathcal{B}_1 \to \mathcal{B}_2$ such that: (1) T is a vector space isomorphism, (2) T is isometric, and (3) $U_2 = TU_1$. In this sense, a completion for $\mathcal{E}$ is unique.

6. If $\mathcal{P}$ is a pre-Hilbert space, a *completion* for $\mathcal{P}$ is a pair $(\mathfrak{K},U)$, where
(1) $\mathfrak{K}$ is a Hilbert space,
(2) $U: \mathcal{P} \to \mathfrak{K}$ is a linear mapping,
(3) $(Ux|Uy) = (x|y)$ for all $x,y \in \mathcal{P}$,
(4) $U(\mathcal{P})$ is dense in $\mathfrak{K}$.
Then:
(i) According to *Theorem 1*, every pre-Hilbert space has a completion.
(ii) Discuss the analogs of parts (ii) and (iii) of *Exercise 5*.
(iii) Regarding $\mathcal{P}$ as a normed space, any completion for $\mathcal{P}$, in the sense of *Exercise 5*, leads to a completion of the pre-Hilbert space $\mathcal{P}$.

*7. Let $\mathfrak{X}$ be a metric space. A *completion* for $\mathfrak{X}$ is a pair $(\mathfrak{Y},U)$, where
(1) $\mathfrak{Y}$ is a complete metric space,
(2) $U: \mathfrak{X} \to \mathfrak{Y}$,
(3) $d(Ux,Uy) = d(x,y)$ for all $x,y \in \mathfrak{X}$ (i.e. U is isometric),
(4) $U(\mathfrak{X})$ is dense in $\mathfrak{Y}$.
It can be shown that every metric space $\mathfrak{X}$ has a completion $(\mathfrak{Y},U)$.

8. Let $\mathcal{P} = \mathfrak{M} \oplus \mathfrak{N}$ in the sense of *Definition III.2.5.* Viewing $\mathcal{P}$ as a linear subspace of its completion $\mathcal{H}$, one has $\mathcal{H} = \overline{\mathfrak{M}} \oplus \overline{\mathfrak{N}}$.

§ 3. BILINEAR MAPPINGS

If $\mathfrak{X}$ and $\mathfrak{Y}$ are non-empty sets, the *Cartesian product* of $\mathfrak{X}$ and $\mathfrak{Y}$ is the set of all symbols (x,y), called *ordered pairs*, where $x \in \mathfrak{X}$ and $y \in \mathfrak{Y}$. If (x_1,y_1) and (x_2,y_2) are ordered pairs, the relation $(x_1,y_1) = (x_2,y_2)$ means that both $x_1 = x_2$ and $y_1 = y_2$. The Cartesian product of $\mathfrak{X}$ and $\mathfrak{Y}$ is denoted $\mathfrak{X} \times \mathfrak{Y}$. If $u = (x,y) \in \mathfrak{X} \times \mathfrak{Y}$, x is the *first coordinate* of u, y is the *second coordinate* of u. Observe that $\mathfrak{Y} \times \mathfrak{X}$ is conceptually distinct from $\mathfrak{X} \times \mathfrak{Y}$.

Examples

1. If $\mathcal{R}$ is the set of all real numbers, $\mathcal{R} \times \mathcal{R}$ is the "Cartesian plane." Any mapping $T\colon \mathcal{R} \times \mathcal{R} \to \mathcal{R}$ is called a "real-valued function of two independent real variables"; for example, $T(x,y) = x^2 - y^2$.

2. If $\mathcal{P}$ is any pre-Hilbert space, the scalar product in $\mathcal{P}$ can be thought of as a mapping $T\colon \mathcal{P} \times \mathcal{P} \to \mathcal{C}$, namely $T(x,y) = (x|y)$.

Definition 1. *If* $\mathfrak{U},\mathfrak{V},\mathfrak{W}$ *are vector spaces, a mapping* $\varphi\colon \mathfrak{U} \times \mathfrak{V} \to \mathfrak{W}$ *is said to be* **bilinear** *in case the relations*

$$\begin{aligned} \varphi(x_1 + x_2,y) &= \varphi(x_1,y) + \varphi(x_2,y) \\ \varphi(\lambda x,y) &= \lambda\varphi(x,y) \\ \varphi(x,y_1 + y_2) &= \varphi(x,y_1) + \varphi(x,y_2) \\ \varphi(x,\lambda y) &= \lambda\varphi(x,y) \end{aligned}$$

hold identically. If moreover $\mathfrak{W} = \mathcal{C}$, φ *is called a* **bilinear form** *on* $\mathfrak{U} \times \mathfrak{V}$.

Thus, φ is a bilinear mapping of $\mathfrak{U} \times \mathfrak{V}$ into $\mathfrak{W}$, if and only if: (1) for each fixed $y \in \mathfrak{V}$, $x \to \varphi(x,y)$ is a linear mapping of $\mathfrak{U}$ into $\mathfrak{W}$, and (2) for each fixed $x \in \mathfrak{U}$, $y \to \varphi(x,y)$ is a linear mapping of $\mathfrak{V}$ into $\mathfrak{W}$. Briefly, φ is linear in each coordinate. Clearly $\varphi(x,\theta) = \varphi(\theta,y) = \theta$.

Examples

3. Let $\mathfrak{U}$ be the vector space of n-ples $x = (\lambda_1,\cdots,\lambda_n)$, $y = (\mu_1,\cdots,\mu_n)$, $\cdots$, and define $\varphi(x,y) = \sum_1^n \lambda_k\mu_k$. Then, φ is a bilinear form on $\mathfrak{U} \times \mathfrak{U}$.

4. Let $\mathfrak{U}$ and $\mathfrak{W}$ be vector spaces, $T: \mathfrak{U} \to \mathfrak{W}$ a linear mapping. Let $\mathfrak{V} = \mathfrak{L}(\mathfrak{W},\mathfrak{C})$ be the vector space of all linear forms on $\mathfrak{W}$, and define $\varphi: \mathfrak{U} \times \mathfrak{V} \to \mathfrak{C}$ by the formula $\varphi(x,f) = f(Tx)$. Then, φ is a bilinear form on $\mathfrak{U} \times \mathfrak{V}$.

5. Let $\mathfrak{U},\mathfrak{V},\mathfrak{W}$ be vector spaces, f a linear form on $\mathfrak{U}$, and $T: \mathfrak{V} \to \mathfrak{W}$ a linear mapping. Define $\varphi: \mathfrak{U} \times \mathfrak{V} \to \mathfrak{W}$ by the formula $\varphi(x,y) = f(x)(Ty)$. Then, φ is a bilinear mapping of $\mathfrak{U} \times \mathfrak{V}$ into $\mathfrak{W}$.

6. If $\mathfrak{A}$ is an algebra, the mapping $\varphi: \mathfrak{A} \times \mathfrak{A} \to \mathfrak{A}$ defined by $\varphi(a,b) = ab$ is bilinear. In particular, if $\mathfrak{V}$ is any vector space, the mapping $(S,T) \to ST$ of $\mathfrak{L}(\mathfrak{V}) \times \mathfrak{L}(\mathfrak{V})$ into $\mathfrak{L}(\mathfrak{V})$ is bilinear (see *Theorem IV.5.1*).

Exercises

1. If $\mathfrak{X}$ and $\mathfrak{Y}$ are non-empty sets, there exists a natural bijective mapping $T: \mathfrak{X} \times \mathfrak{Y} \to \mathfrak{Y} \times \mathfrak{X}$.

2. If $\mathfrak{X},\mathfrak{Y},\mathfrak{Z}$ are non-empty sets, there exists a natural bijective mapping $T: \mathfrak{X} \times (\mathfrak{Y} \times \mathfrak{Z}) \to (\mathfrak{X} \times \mathfrak{Y}) \times \mathfrak{Z}$.

3. Any mapping $T: \mathfrak{X} \to \mathfrak{Y}$ determines a subset of $\mathfrak{X} \times \mathfrak{Y}$, namely the set of ordered pairs $\mathcal{G}_T = \{(x,Tx): x \in \mathfrak{X}\}$, called the *graph* of T. Prove:

(i) If $S: \mathfrak{X} \to \mathfrak{Y}$ and $T: \mathfrak{X} \to \mathfrak{Y}$, then $S = T$ if and only if $\mathcal{G}_S = \mathcal{G}_T$.

(ii) Given $\mathcal{G} \subset \mathfrak{X} \times \mathfrak{Y}$. There exists a mapping $T: \mathfrak{X} \to \mathfrak{Y}$ whose graph is $\mathcal{G}$, if and only if: (1) $\mathcal{G} \neq \varnothing$, and (2) given $(x,y_1) \in \mathcal{G}$ and $(x,y_2) \in \mathcal{G}$, necessarily $y_1 = y_2$.

4. If $\mathfrak{U},\mathfrak{V},\mathfrak{W}$ are vector spaces, denote by $\mathfrak{B}(\mathfrak{U},\mathfrak{V};\mathfrak{W})$ the set of all bilinear mappings $\varphi: \mathfrak{U} \times \mathfrak{V} \to \mathfrak{W}$. For two such mappings φ,ψ, write $\varphi = \psi$ in case $\varphi(x,y) = \psi(x,y)$ for all $x \in \mathfrak{U}$ and $y \in \mathfrak{V}$. Then, $\mathfrak{B}(\mathfrak{U},\mathfrak{V};\mathfrak{W})$ is a vector space, relative to the operations

$$(\varphi + \psi)(x,y) = \varphi(x,y) + \psi(x,y)$$

$$(\lambda\varphi)(x,y) = \lambda\varphi(x,y).$$

5. Notation as in *Exercise 4*. Suppose $\varphi \in \mathfrak{B}(\mathfrak{U},\mathfrak{V};\mathfrak{W})$. Each $x \in \mathfrak{U}$ determines a linear mapping $\varphi_x: \mathfrak{V} \to \mathfrak{W}$, defined by $\varphi_x(y) = \varphi(x,y)$. Each $y \in \mathfrak{V}$ determines a linear mapping $\varphi^y: \mathfrak{U} \to \mathfrak{W}$, defined by $\varphi^y(x) = \varphi(x,y)$. Prove:

(i) For each fixed $\varphi \in \mathfrak{B}(\mathfrak{U},\mathfrak{V};\mathfrak{W})$, $x \to \varphi_x$ is a linear mapping of $\mathfrak{U}$ into $\mathfrak{L}(\mathfrak{V},\mathfrak{W})$.

(ii) For each fixed $\varphi \in \mathfrak{B}(\mathfrak{U},\mathfrak{V};\mathfrak{W})$, $y \to \varphi^y$ is a linear mapping of $\mathfrak{V}$ into $\mathfrak{L}(\mathfrak{U},\mathfrak{W})$.

(iii) For each fixed $x \in \mathfrak{U}$, $\varphi \to \varphi_x$ is a linear mapping of $\mathfrak{B}(\mathfrak{U},\mathfrak{V};\mathfrak{W})$ into $\mathfrak{L}(\mathfrak{V},\mathfrak{W})$.

(iv) For each fixed $y \in \mathfrak{V}$, $\varphi \to \varphi^y$ is a linear mapping of $\mathfrak{B}(\mathfrak{U},\mathfrak{V};\mathfrak{W})$ into $\mathfrak{L}(\mathfrak{U},\mathfrak{W})$.

(v) For each fixed pair $(x,y) \in \mathfrak{U} \times \mathfrak{V}$, $\varphi \to \varphi(x,y)$ is a linear mapping of $\mathfrak{B}(\mathfrak{U},\mathfrak{V};\mathfrak{W})$ into $\mathfrak{W}$.

6. Notation as in *Exercise 5*. Suppose $\varphi \in \mathfrak{B}(\mathfrak{U},\mathfrak{V};\mathfrak{W})$. The mapping $x \to \varphi_x$ is a linear mapping of $\mathfrak{U}$ into $\mathfrak{L}(\mathfrak{V},\mathfrak{W})$; denote it by $\hat{\varphi}$, that is, $\hat{\varphi}(x) = \varphi_x$. Thus, $\hat{\varphi} \in \mathfrak{L}[\mathfrak{U},\mathfrak{L}(\mathfrak{V},\mathfrak{W})]$. Prove: $\varphi \to \hat{\varphi}$ is a vector space isomorphism of $\mathfrak{B}(\mathfrak{U},\mathfrak{V};\mathfrak{W})$ with $\mathfrak{L}[\mathfrak{U},\mathfrak{L}(\mathfrak{V},\mathfrak{W})]$.

7. Notation as in *Exercise 5*. Using the techniques of *Exercise 6*, obtain a vector space isomorphism $\varphi \to \tilde{\varphi}$ of $\mathfrak{B}(\mathfrak{U},\mathfrak{V};\mathfrak{W})$ with $\mathfrak{L}[\mathfrak{V},\mathfrak{L}(\mathfrak{U},\mathfrak{W})]$.

8. Let $\mathfrak{U},\mathfrak{W}$ be vector spaces, $\mathfrak{V} = \mathfrak{L}(\mathfrak{U},\mathfrak{C})$ the vector space of linear forms f on $\mathfrak{U}$. Prove:

(i) For each fixed $z \in \mathfrak{W}$, $(x,f) \to f(x)z$ is a bilinear mapping of $\mathfrak{U} \times \mathfrak{V}$ into $\mathfrak{W}$.

(ii) For each fixed $f \in \mathfrak{V}$, $(x,z) \to f(x)z$ is a bilinear mapping of $\mathfrak{U} \times \mathfrak{W}$ into $\mathfrak{W}$.

(iii) For each fixed $x \in \mathfrak{U}$, $(f,z) \to f(x)z$ is a bilinear mapping of $\mathfrak{V} \times \mathfrak{W}$ into $\mathfrak{W}$.

9. Given a fixed bilinear mapping φ of $\mathfrak{U} \times \mathfrak{V}$ into $\mathfrak{W}$. For subsets $\mathfrak{S} \subset \mathfrak{U}$ and $\mathfrak{T} \subset \mathfrak{V}$, define

$$\mathfrak{S}^\perp = \{y \in \mathfrak{V}: \varphi(x,y) = \theta \text{ for all } x \in \mathfrak{S}\}$$

$$^\perp\mathfrak{T} = \{x \in \mathfrak{U}: \varphi(x,y) = \theta \text{ for all } y \in \mathfrak{T}\}.$$

Prove:

(i) $\mathfrak{S}^\perp$ and $^\perp\mathfrak{T}$ are linear subspaces of $\mathfrak{V}$ and $\mathfrak{U}$, respectively.

(ii) $\mathfrak{S} \subset {}^\perp(\mathfrak{S}^\perp)$, and $\mathfrak{T} \subset ({}^\perp\mathfrak{T})^\perp$.

(iii) If $\mathfrak{S}_1 \subset \mathfrak{S}_2$, then $\mathfrak{S}_1{}^\perp \supset \mathfrak{S}_2{}^\perp$; if $\mathfrak{T}_1 \subset \mathfrak{T}_2$, then ${}^\perp\mathfrak{T}_1 \supset {}^\perp\mathfrak{T}_2$.

(iv) $({}^\perp(\mathfrak{S}^\perp))^\perp = \mathfrak{S}^\perp$, and ${}^\perp(({}^\perp\mathfrak{T})^\perp) = {}^\perp\mathfrak{T}$.

10. Let $\mathfrak{V}$ and $\mathfrak{W}$ be vector spaces. Prove:

(i) $\mathfrak{V} \times \mathfrak{W}$ becomes a vector space, on defining $(x_1,y_1) + (x_2,y_2) = (x_1 + x_2, y_1 + y_2)$ and $\lambda(x,y) = (\lambda x, \lambda y)$, with zero element (θ,θ), and $-(x,y) = (-x,-y)$. $\mathfrak{V} \times \mathfrak{W}$ is called the *vector space direct product* of $\mathfrak{V}$ and $\mathfrak{W}$.

(ii) A mapping $T: \mathcal{U} \to \mathcal{W}$ is linear if and only if its graph $\mathcal{G}_T$ is a linear subspace of $\mathcal{U} \times \mathcal{W}$ (see *Exercise 3*).

(iii) $x \to (x,\theta)$ is an injective linear mapping of $\mathcal{U}$ into $\mathcal{U} \times \mathcal{W}$; denote its range by $\mathcal{U}_o$. Similarly $y \to (\theta,y)$ is an injective linear mapping of $\mathcal{W}$ into $\mathcal{U} \times \mathcal{W}$; denote its range by $\mathcal{W}_o$. Then: $\mathcal{U}_o$ and $\mathcal{W}_o$ are linear subspaces of $\mathcal{U} \times \mathcal{W}$, and every vector in $\mathcal{U} \times \mathcal{W}$ has a unique representation in the form $u_o + w_o$, with $u_o \in \mathcal{U}_o$ and $w_o \in \mathcal{W}_o$.

(iv) Suppose $\mathfrak{M}$ and $\mathfrak{N}$ are linear subspaces of a vector space $\mathfrak{U}$, such that (1) $\mathfrak{M} + \mathfrak{N} = \mathfrak{U}$, and (2) the only vector common to $\mathfrak{M}$ and $\mathfrak{N}$ is θ. Then, there is a natural vector space isomorphism of $\mathfrak{U}$ with the direct product vector space $\mathfrak{M} \times \mathfrak{N}$.

11. Let $\mathcal{P},\mathcal{Q}$ be pre-Hilbert spaces, and let $\mathcal{P} \times \mathcal{Q}$ be their vector space direct product, as defined in *Exercise 10*. Prove:

(i) $\mathcal{P} \times \mathcal{Q}$ becomes a pre-Hilbert space, on defining $((x_1,y_1) \mid (x_2,y_2)) = (x_1 \mid x_2) + (y_1 \mid y_2)$. $\mathcal{P} \times \mathcal{Q}$ is called the *pre-Hilbert space direct product* of $\mathcal{P}$ and $\mathcal{Q}$.

(ii) One has $\| (x,y) \|^2 = \| x \|^2 + \| y \|^2$.

(iii) Let $\mathcal{P}_o$ be the range of the isometric linear mapping $x \to (x,\theta)$ of $\mathcal{P}$ into $\mathcal{P} \times \mathcal{Q}$, and $\mathcal{Q}_o$ the range of $y \to (\theta,y)$. Then $\mathcal{P}_o = \mathcal{Q}_o^{\perp}$, $\mathcal{Q}_o = \mathcal{P}_o^{\perp}$, and $\mathcal{P} \times \mathcal{Q} = \mathcal{P}_o \oplus \mathcal{Q}_o$ in the sense of *Definition III.2.5*.

(iv) One has $(x_n,y_n) \to (x,y)$ if and only if both $x_n \to x$ and $y_n \to y$.

(v) $\mathcal{P} \times \mathcal{Q}$ is a Hilbert space if and only if both $\mathcal{P}$ and $\mathcal{Q}$ are Hilbert spaces.

12. Let $\mathfrak{X}$ and $\mathfrak{Y}$ be metric spaces. Prove:

(i) $\mathfrak{X} \times \mathfrak{Y}$ becomes a metric space, on defining $d((x_1,y_1),(x_2,y_2)) = d(x_1,x_2) + d(y_1,y_2)$. $\mathfrak{X} \times \mathfrak{Y}$ is called *a* metric space *direct product* of $\mathfrak{X}$ and $\mathfrak{Y}$ (there are other natural ways of defining distances in $\mathfrak{X} \times \mathfrak{Y}$).

(ii) One has $(x_n,y_n) \to (x,y)$ if and only if both $x_n \to x$ and $y_n \to y$.

(iii) If $T: \mathfrak{X} \to \mathfrak{Y}$ is a continuous mapping, its graph $\mathcal{G}_T$ is a closed subset of $\mathfrak{X} \times \mathfrak{Y}$.

(iv) For each fixed $y_o \in \mathfrak{Y}$, $x \to (x,y_o)$ is an isometric mapping of $\mathfrak{X}$ onto a closed subset $\mathfrak{X}_o$ of $\mathfrak{X} \times \mathfrak{Y}$. Similarly for the mapping $y \to (x_o,y)$, where $x_o \in \mathfrak{X}$ is fixed.

(v) $\mathfrak{X} \times \mathfrak{Y}$ is complete if and only if both $\mathfrak{X}$ and $\mathfrak{Y}$ are complete.

(vi) Given completions for $\mathfrak{X}$ and $\mathfrak{Y}$, in the sense of *Exercise 2.7*, there is a natural way of obtaining a completion for $\mathfrak{X} \times \mathfrak{Y}$.

13. Let $\mathcal{E}$ and $\mathcal{F}$ be normed spaces, and let $\mathcal{E} \times \mathcal{F}$ be their vector space direct product, as defined in *Exercise 10*. Prove:

(i) $\mathcal{E} \times \mathcal{F}$ becomes a normed space, on defining $\| (x,y) \| = \| x \| + \| y \|$.

(ii) $x \to (x,\theta)$ and $y \to (\theta,y)$ are isometric linear mappings.

(iii) If distances in $\mathcal{E},\mathcal{F}$, and $\mathcal{E} \times \mathcal{F}$ are defined as in § 7 of Chapter IV, $\mathcal{E} \times \mathcal{F}$ is the metric space direct product of $\mathcal{E}$ and $\mathcal{F}$, in the sense of *Exercise 12*.

(iv) If $T: \mathcal{E} \to \mathcal{F}$ is a continuous linear mapping, its graph $\mathcal{G}_T$ is a closed linear subspace of $\mathcal{E} \times \mathcal{F}$.

(v) $\mathcal{E} \times \mathcal{F}$ is a Banach space if and only if both $\mathcal{E}$ and $\mathcal{F}$ are Banach spaces.

(vi) If $\mathcal{E}$ and $\mathcal{F}$ happen to be pre-Hilbert spaces, $(x_n,y_n) \to (x,y)$ in the above sense if and only if $(x_n,y_n) \to (x,y)$ in the sense of *Exercise 11*.

14. If $\mathfrak{X}$ and $\mathfrak{Y}$ are metric spaces, $T: \mathfrak{X} \to \mathfrak{Y}$ is a bijective mapping, $S = T^{-1}$, and $\mathcal{G}_T$ is a closed subset of $\mathfrak{X} \times \mathfrak{Y}$ (see *Exercise 12*), then $\mathcal{G}_S$ is also closed.

15. (i) If $\mathcal{E}$ is a Banach space, $\mathcal{F}$ is a normed space, and $T: \mathcal{E} \to \mathcal{F}$ is a continuous linear mapping, then the normed space $\mathcal{G}_T$ (see *Exercise 13*) is a Banach space.

*(ii) If $\mathcal{E}$ and $\mathcal{F}$ are Banach spaces, and $T: \mathcal{E} \to \mathcal{F}$ is a continuous vector space isomorphism, it can be shown that T^{-1} is also continuous.

*(iii) If $\mathcal{E}$ and $\mathcal{F}$ are Banach spaces, and $T: \mathcal{E} \to \mathcal{F}$ is a linear mapping such that $\mathcal{G}_T$ is closed, it can be shown that T is continuous.

§ 4. BOUNDED BILINEAR MAPPINGS

Definition 1. *If $\mathcal{E},\mathcal{F},\mathcal{G}$ are normed spaces, a bilinear mapping $\varphi: \mathcal{E} \times \mathcal{F} \to \mathcal{G}$ is said to be* **bounded** *in case there exists a constant M such that $\| \varphi(x,y) \| \leq M \| x \| \| y \|$ for all $x \in \mathcal{E}$ and $y \in \mathcal{F}$. If moreover $\mathcal{G} = \mathcal{C}$, φ is called a* **bounded bilinear form** *on $\mathcal{E} \times \mathcal{F}$.*

Examples

1. Let $\mathcal{E}$ and $\mathcal{G}$ be normed spaces, $T: \mathcal{E} \to \mathcal{G}$ a continuous linear mapping, and $\mathcal{F} = \mathcal{G}'$ the dual space of $\mathcal{G}$. Define $\varphi: \mathcal{E} \times \mathcal{F} \to \mathcal{C}$ by the formula $\varphi(x,f) = f(Tx)$. Then, φ is a bilinear form on $\mathcal{E} \times \mathcal{F}$ (see also *Example 3.4*), and $|\varphi(x,f)| = |f(Tx)| \leq \| f \| \| Tx \| \leq \| T \| (\| x \| \| f \|)$ shows that φ is bounded.

2. If $\mathfrak{A}$ is any normed algebra, then the bilinear mapping $\varphi: \mathfrak{A} \times \mathfrak{A} \to \mathfrak{A}$ defined by $\varphi(a,b) = ab$ is bounded: $\| \varphi(a,b) \| \leq \| a \| \| b \|$. In particular, if $\mathcal{E}$ is a normed space, $(S,T) \to ST$ defines a bounded bilinear mapping of $\mathfrak{L}_c(\mathcal{E}) \times \mathfrak{L}_c(\mathcal{E})$ into $\mathfrak{L}_c(\mathcal{E})$ (see *Theorem IV.9.1*). ▮

Using arguments similar to those in *Theorems IV.7.3* and *IV.7.4*, the condition for boundedness of a bilinear mapping can be reformulated as follows:

Theorem 1. *If $\mathcal{E},\mathcal{F},\mathcal{G}$ are normed spaces, and $\varphi: \mathcal{E} \times \mathcal{F} \to \mathcal{G}$ is a bilinear mapping, the following conditions on φ are equivalent:*

(a) *φ is bounded.*

(b) $M_1 = \text{LUB}\,\{\| \varphi(x,y) \|: \| x \| \leq 1, \| y \| \leq 1\}$ *is finite.*

(c) $M_2 = \text{LUB}\,\{\| \varphi(x,y) \|: \| x \| < 1, \| y \| < 1\}$ *is finite.*

(d) $M_3 = \text{LUB}\,\{\| \varphi(x,y) \|: \| x \| = 1, \| y \| = 1\}$ *is finite.*

If φ is bounded, $M_1 = M_2 = M_3$; denoting the common value by $\| \varphi \|$, one has

(1) $\| \varphi(x,y) \| \leq \| \varphi \| \| x \| \| y \|$ *for all* $x \in \mathcal{E}$, $y \in \mathcal{F}$.

Moreover, $\| \varphi \|$ is minimal in this property:

(2) *If* $M \geq 0$ *and* $\| \varphi(x,y) \| \leq M \| x \| \| y \|$ *for all* $x \in \mathcal{E}$ *and* $y \in \mathcal{F}$, *necessarily* $\| \varphi \| \leq M$.

Definition 2. *Notation as in Theorem 1. The number $\| \varphi \|$ is called the* **norm** *of the bounded bilinear mapping $\varphi: \mathcal{E} \times \mathcal{F} \to \mathcal{G}$. Thus,*

$$\| \varphi \| = \text{LUB}\,\{\| \varphi(x,y) \|: \| x \| \leq 1, \| y \| \leq 1\}.$$

Clearly $\| \varphi \| = 0$ if and only if $\varphi(x,y) = \theta$ for all $x \in \mathcal{E}$, $y \in \mathcal{F}$; otherwise, $\| \varphi \| > 0$. In condition (d) of *Theorem 1*, one assumes $\mathcal{E} \neq \{\theta\}$, $\mathcal{F} \neq \{\theta\}$, so as to have a supply of unit vectors.

Exercises

1. If $\mathcal{E},\mathcal{F},\mathcal{G}$ are normed spaces, and $\varphi: \mathcal{E} \times \mathcal{F} \to \mathcal{G}$ is a bilinear mapping, the following conditions on φ are equivalent:
(a) φ is bounded.
(b) If $x_n \to x$ and $y_n \to y$, then $\varphi(x_n,y_n) \to \varphi(x,y)$.
(c) If $x_n \to \theta$ and $y_n \to \theta$, then $\varphi(x_n,y_n) \to \theta$.

2. There is a normed space analog for *Exercise 3.8.*

3. There is a normed space analog for *Exercise 3.9:* assuming φ bounded, $\mathcal{S}^{\perp}$ and ${}^{\perp}\mathcal{J}$ are *closed* linear subspaces.

4. If $\mathcal{E},\mathcal{F},\mathcal{G}$ are normed spaces, denote by $\mathfrak{B}_c(\mathcal{E},\mathcal{F};\mathcal{G})$ the set of all bounded bilinear mappings $\varphi\colon \mathcal{E}\times\mathcal{F}\to\mathcal{G}$. Then, $\mathfrak{B}_c(\mathcal{E},\mathcal{F};\mathcal{G})$ is a linear subspace of $\mathfrak{B}(\mathcal{E},\mathcal{F};\mathcal{G})$ (see *Exercise 3.4*), and becomes a normed space when $\|\varphi\|$ is defined as in *Definition 2.*

5. Let $\mathcal{E},\mathcal{F},\mathcal{G}$ be normed spaces, $\varphi\colon \mathcal{E}\times\mathcal{F}\to\mathcal{G}$ a bounded bilinear mapping. Show:

(i) For each fixed $x\in\mathcal{E}$, $y\to\varphi_x(y)=\varphi(x,y)$ is a continuous linear mapping of $\mathcal{F}$ into $\mathcal{G}$, and $\|\varphi_x\|\leq\|\varphi\|\,\|x\|$. Thus, $\varphi_x\in\mathfrak{L}_c(\mathcal{F},\mathcal{G})$.

(ii) For each fixed $y\in\mathcal{F}$, $x\to\varphi^y(x)=\varphi(x,y)$ is a continuous linear mapping of $\mathcal{E}$ into $\mathcal{G}$, and $\|\varphi^y\|\leq\|\varphi\|\,\|y\|$. Thus, $\varphi^y\in\mathfrak{L}_c(\mathcal{E},\mathcal{G})$.

(iii) The mapping $x\to\hat{\varphi}(x)=\varphi_x$ is a continuous linear mapping of $\mathcal{E}$ into $\mathfrak{L}_c(\mathcal{F},\mathcal{G})$, and $\|\hat{\varphi}\|\leq\|\varphi\|$. Thus, $\hat{\varphi}\in\mathfrak{L}_c[\mathcal{E},\mathfrak{L}_c(\mathcal{F},\mathcal{G})]$.

(iv) The mapping $y\to\tilde{\varphi}(y)=\varphi^y$ is a continuous linear mapping of $\mathcal{F}$ into $\mathfrak{L}_c(\mathcal{E},\mathcal{G})$, and $\|\tilde{\varphi}\|\leq\|\varphi\|$. Thus, $\tilde{\varphi}\in\mathfrak{L}_c[\mathcal{F},\mathfrak{L}_c(\mathcal{E},\mathcal{G})]$.

(v) In fact, $\|\hat{\varphi}\|=\|\varphi\|=\|\tilde{\varphi}\|$.

6. Notation as in *Exercise 5.* Prove:

(i) The mapping $\varphi\to\hat{\varphi}$ is an isometric vector space isomorphism of $\mathfrak{B}_c(\mathcal{E},\mathcal{F};\mathcal{G})$ with $\mathfrak{L}_c[\mathcal{E},\mathfrak{L}_c(\mathcal{F},\mathcal{G})]$.

(ii) The mapping $\varphi\to\tilde{\varphi}$ is an isometric vector space isomorphism of $\mathfrak{B}_c(\mathcal{E},\mathcal{F};\mathcal{G})$ with $\mathfrak{L}_c[\mathcal{F},\mathfrak{L}_c(\mathcal{E},\mathcal{G})]$.

7. Let $\mathcal{E}$ and $\mathcal{F}$ be normed spaces, $\mathfrak{B}$ a Banach space. If $\mathfrak{M}$ is a dense linear subspace of $\mathcal{E}$, $\mathfrak{N}$ is a dense linear subspace of $\mathcal{F}$, and $\varphi\colon\mathfrak{M}\times\mathfrak{N}\to\mathfrak{B}$ is a bounded bilinear mapping, there exists one and only one bounded bilinear mapping $\psi\colon\mathcal{E}\times\mathcal{F}\to\mathfrak{B}$ such that $\psi(x,y)=\varphi(x,y)$ for all $x\in\mathfrak{M}$ and $y\in\mathfrak{N}$.

*8. Let $\mathcal{E},\mathcal{F},\mathcal{G}$ be normed spaces, and $\varphi\colon\mathcal{E}\times\mathcal{F}\to\mathcal{G}$ a bilinear mapping. Suppose that for each $x\in\mathcal{E}$, $y\to\varphi(x,y)$ is continuous, and that for each $y\in\mathcal{F}$, $x\to\varphi(x,y)$ is continuous. Assuming either $\mathcal{E}$ or $\mathcal{F}$ is a Banach space, it can be shown that φ is bounded.

*9. Let $\mathcal{E}$ and $\mathcal{G}$ be normed spaces, $\mathcal{F}=\mathcal{G}'$ the normed space dual of $\mathcal{G}$. Each $T\in\mathfrak{L}_c(\mathcal{E},\mathcal{G})$ determines a bounded bilinear form $\varphi_T\colon\mathcal{E}\times\mathcal{F}\to\mathcal{C}$, namely $\varphi_T(x,f)=f(Tx)$ (see *Example 1*), and one has $\|\varphi_T\|\leq\|T\|$.

(i) It can be shown that $\|\varphi_T\|=\|T\|$.

(ii) $T \to \varphi_T$ is an isometric linear mapping of $\mathfrak{L}_c(\mathcal{E},\mathcal{G})$ into $\mathfrak{B}_c(\mathcal{E},\mathcal{F};\mathcal{C})$.

(iii) If moreover $\mathcal{G}$ is reflexive, the mapping $T \to \varphi_T$ is surjective, hence is an isometric vector space isomorphism.

10. The following is a specialization of *Exercise 9* which can be worked out using material in the text.

Let $\mathcal{H}$ be a Hilbert space. Each $T \in \mathfrak{L}_c(\mathcal{H})$ determines a bounded bilinear form $\varphi_T\colon \mathcal{H} \times \mathcal{H}' \to \mathcal{C}$, namely $\varphi_T(x,f) = f(Tx)$. In view of *Theorem 1.1*, this means $\varphi_T(x,y') = y'(Tx) = (Tx\,|\,y)$ for all $x,y \in \mathcal{H}$. The relation $\| \varphi_T \| = \| T \|$ is elementary (see *Example IV.7.5*). Moreover, $\mathcal{H}$ is reflexive, that is, the mapping $x \to x''$ of $\mathcal{H}$ into $\mathcal{H}''$ is surjective (*Exercise 1.9*). It follows that $T \to \varphi_T$ is an isometric vector space isomorphism of $\mathfrak{L}_c(\mathcal{H})$ with $\mathfrak{B}_c(\mathcal{H},\mathcal{H}';\mathcal{C})$.

11. If $\mathcal{E}$ is a normed space, and $\mathcal{H}$ is a Hilbert space, there exists an isometric vector space isomorphism $T \to \varphi_T$ of $\mathfrak{L}_c(\mathcal{E},\mathcal{H})$ with $\mathfrak{B}_c(\mathcal{E},\mathcal{H}';\mathcal{C})$.

12. If $\mathcal{E}$ and $\mathcal{F}$ are normed spaces, and $\mathcal{G}$ is a Banach space, then $\mathfrak{B}_c(\mathcal{E},\mathcal{F};\mathcal{G})$ is a Banach space. In particular, $\mathfrak{B}_c(\mathcal{E},\mathcal{F};\mathcal{C})$ is a Banach space.

§ 5. SESQUILINEAR MAPPINGS

Definition 1. *If $\mathcal{U},\mathcal{V},\mathcal{W}$ are vector spaces, a mapping $\varphi\colon \mathcal{U} \times \mathcal{V} \to \mathcal{W}$ is said to be* **sesquilinear** *in case the relations*

$$\varphi(x_1 + x_2,y) = \varphi(x_1,y) + \varphi(x_2,y)$$

$$\varphi(\lambda x,y) = \lambda\varphi(x,y)$$

$$\varphi(x,y_1 + y_2) = \varphi(x,y_1) + \varphi(x,y_2)$$

$$\varphi(x,\lambda y) = \lambda^*\varphi(x,y)$$

hold identically. If moreover $\mathcal{W} = \mathcal{C}$, φ is called a **sesquilinear form** *on $\mathcal{U} \times \mathcal{V}$.*

Thus, φ is a sesquilinear mapping of $\mathcal{U} \times \mathcal{V}$ into $\mathcal{W}$, if and only if: (1) for each fixed $y \in \mathcal{V}$, $x \to \varphi(x,y)$ is a linear mapping of $\mathcal{U}$ into $\mathcal{W}$, and (2) for each fixed $x \in \mathcal{U}$, $y \to \varphi(x,y)$ is a semilinear mapping of $\mathcal{V}$ into $\mathcal{W}$ (see *Definition IV.1.3*). Briefly, φ is linear in the first coordinate, semilinear in the second coordinate. Clearly $\varphi(\theta,y) = \varphi(x,\theta) = \theta$.

Examples

1. If $\mathcal{P}$ is any pre-Hilbert space, the mapping $\varphi: \mathcal{P} \times \mathcal{P} \to \mathcal{C}$ defined by $\varphi(x,y) = (x|y)$ is a sesquilinear form on $\mathcal{P} \times \mathcal{P}$.

2. If $\mathcal{U}$ is a vector space, $\mathcal{P}$ is a pre-Hilbert space, and $T: \mathcal{U} \to \mathcal{P}$ is a linear mapping, then $\varphi(x,y) = (Tx|y)$ defines a sesquilinear form on $\mathcal{U} \times \mathcal{P}$.

3. If $\varphi: \mathcal{U} \times \mathcal{V} \to \mathcal{C}$ is a sesquilinear form, then the mapping $\psi: \mathcal{V} \times \mathcal{U} \to \mathcal{C}$ defined by $\psi(y,x) = [\varphi(x,y)]^*$ is sesquilinear. ▮

Obviously, one could paraphrase the results of the preceding two sections for sesquilinear mappings. However, there is a simple way of deducing the theory of sesquilinear mappings from the theory of bilinear mappings, via the concept of the *complex-conjugate* of a vector space:

Definition 2. *Let $\mathcal{V}$ be a vector space, with sums denoted $x + y$, and scalar multiples denoted λx. The* **complex-conjugate** *of $\mathcal{V}$ is the vector space $\mathcal{V}^*$ obtained as follows:*

(1) *The set $\mathcal{V}^*$ is the same as the set $\mathcal{V}$.*

(2) *If $x,y \in \mathcal{V}^*$, the* **sum** *of x and y is defined to be $x + y$ (as calculated in $\mathcal{V}$).*

(3) *If λ is a scalar, and $x \in \mathcal{V}^*$, the* **multiple** *of x by λ is defined to be $\lambda^* x$ (as calculated in $\mathcal{V}$); this will be denoted $\lambda o x$.*

It is easy to check that $\mathcal{V}^*$ satisfies the axioms for a vector space listed in *Definition I.1.1*. For example: $(\lambda\mu)ox = (\lambda\mu)^*x = (\lambda^*\mu^*)x = \lambda^*(\mu^* x) = \lambda^*(\mu o x) = \lambda o(\mu o x)$. It is clear that $\mathcal{V}^*$ is the only vector space structure which can be defined on the set $\mathcal{V}$ in such a way that the identity mapping $I: \mathcal{V} \to \mathcal{V}^*$ is *conjugate*-linear.

Examples

4. If $\mathcal{P}$ is any pre-Hilbert space, the mapping $y \to y'$ of $\mathcal{P}$ into $(\mathcal{P}')^*$ is *linear;* for, $(y + z)' = y' + z'$, and $(\lambda y)' = \lambda^* y' = \lambda o y'$ (see *Theorem IV.10.2*).

5. If $\mathcal{V}$ and $\mathcal{W}$ are vector spaces, a mapping $T: \mathcal{V} \to \mathcal{W}$ is conjugate-linear if and only if $T : \mathcal{V}^* \to \mathcal{W}$ is linear (if and only if $T: \mathcal{V} \to \mathcal{W}^*$ is linear).

6. If $\mathcal{E}$ is a normed space, then $\mathcal{E}^*$ is also a normed space with $\| x \|$ defined the same as in $\mathcal{E}$. The point is that $\| \lambda o x \| = \| \lambda^* x \| = |\lambda^*| \cdot \| x \| = |\lambda| \cdot \| x \|$. The identity mapping $I: \mathcal{E} \to \mathcal{E}^*$ is then isometric and conjugate-linear; $\mathcal{E}$ is a Banach space if and only if $\mathcal{E}^*$ is a Banach space.

7. Combining *Examples 4* and *6*, $y \to y'$ is an isometric linear mapping of $\mathcal{P}$ into the normed space $(\mathcal{P}')^*$.

8. If $\mathcal{P}$ is a pre-Hilbert space with scalar products denoted $(x|y)$, the vector space $\mathcal{P}^*$ becomes a pre-Hilbert space on defining scalar products $[x|y]$ by the formula $[x|y] = (x|y)^* = (y|x)$. For instance, $[\lambda o x|y] = [\lambda^* x|y] = (y|\lambda^* x) = \lambda(y|x) = \lambda[x|y]$. Thus, $\mathcal{P}^*$ acquires a norm from its scalar product, as well as the norm inherited from $\mathcal{P}$ via *Example 6;* these norms coincide: $[x|x]^{1/2} = (x|x)^{1/2}$. $\mathcal{P}$ is a Hilbert space if and only if $\mathcal{P}^*$ is a Hilbert space. ∎

The reason for introducing the space $\mathcal{U}^*$ is the following theorem, whose proof is immediate from *Example 5:*

Theorem 1. *Let* $\mathcal{U},\mathcal{V},\mathcal{W}$ *be vector spaces, and* $\varphi: \mathcal{U} \times \mathcal{V} \to \mathcal{W}$. *Then,* φ *is a sesquilinear mapping of* $\mathcal{U} \times \mathcal{V}$ *into* $\mathcal{W}$ *if and only if it is a bilinear mapping of* $\mathcal{U} \times \mathcal{V}^*$ *into* $\mathcal{W}$.

The "polarization identity" (see *Theorem II.3.3*) has a generalization for sesquilinear mappings:

Theorem 2. *If* $\mathcal{V}$ *and* $\mathcal{W}$ *are vector spaces, and* $\varphi: \mathcal{V} \times \mathcal{V} \to \mathcal{W}$ *is sesquilinear, then*

$$\varphi(x,y) = \tfrac{1}{4}\{\varphi(x+y,x+y) - \varphi(x-y,x-y) + i\varphi(x+iy,x+iy) - i\varphi(x-iy,x-iy)\}$$

for all $x,y \in \mathcal{V}$.

Proof.

One has $\varphi(x+y,x+y) = \varphi(x,x) + \varphi(y,y) + \varphi(x,y) + \varphi(y,x)$. The proof proceeds as in *Theorem II.3.3.* ∎

An immediate consequence of *Theorem 2* is the following:

Theorem 3. *If* $\varphi: \mathcal{V} \times \mathcal{V} \to \mathcal{W}$ *and* $\psi: \mathcal{V} \times \mathcal{V} \to \mathcal{W}$ *are sesquilinear mappings such that* $\varphi(x,x) = \psi(x,x)$ *for all* $x \in \mathcal{V}$, *then* $\varphi = \psi$, *that is,* $\varphi(x,y) = \psi(x,y)$ *for all* $x,y \in \mathcal{V}$.

Definition 3. *Let* $\mathcal{V}$ *be a vector space,* $\varphi: \mathcal{V} \times \mathcal{V} \to \mathcal{C}$ *a sesquilinear form.* φ *is said to be* **Hermitian** *in case* $\varphi(y,x) = [\varphi(x,y)]^*$ *for all* $x,y \in \mathcal{V}$.

Theorem 4. *A sesquilinear form* $\varphi: \mathcal{V} \times \mathcal{V} \to \mathcal{C}$ *is Hermitian if and only if* $\varphi(x,x)$ *is* **real,** *for all* $x \in \mathcal{V}$.

Proof.

If φ is Hermitian, $\varphi(x,x) = [\varphi(x,x)]^*$, hence $\varphi(x,x)$ is real.

Suppose, conversely, that $\varphi(x,x)$ is real for all x. Define $\psi(x,y) = [\varphi(y,x)]^*$; as noted in *Example 3*, ψ is a sesquilinear form on $\mathcal{V} \times \mathcal{V}$. By assumption, $\psi(x,x) = \varphi(x,x)$ for all x, hence $\psi = \varphi$ by *Theorem 3*. In other words, φ is Hermitian. ∎

Definition 4. *Let* $\mathcal{V}$ *be a vector space,* $\varphi: \mathcal{V} \times \mathcal{V} \to \mathcal{C}$ *a sesquilinear form.* φ *is said to be* **positive** *in case* $\varphi(x,x) \geq 0$ *for all* $x \in \mathcal{V}$.

In particular, a positive sesquilinear form is Hermitian, by *Theorem 4*. There is a "Cauchy-Schwarz inequality" for positive sesquilinear forms:

Theorem 5. *If* φ *is a positive sesquilinear form on* $\mathcal{V} \times \mathcal{V}$, *then* $|\varphi(x,y)|^2 \leq \varphi(x,x)\varphi(y,y)$ *for all* $x,y \in \mathcal{V}$.

Proof.

By analogy with scalar products, write $\| z \| = \sqrt{\varphi(z,z)}$, only for convenience in the present proof.

If $\varphi(x,x) > 0$ or $\varphi(y,y) > 0$, the proof proceeds as in *Theorem II.3.4*.

Suppose $\varphi(x,x) = \varphi(y,y) = 0$. For all scalar λ, one has $0 \leq \varphi(x + \lambda y, x + \lambda y) = \varphi(x,x) + \lambda^*\varphi(x,y) + \lambda\varphi(y,x) + |\lambda|^2 \varphi(y,y) = \lambda^*\varphi(x,y) + \lambda\varphi(y,x)$. In particular, for $\lambda = -\varphi(x,y)$, $0 \leq -2|\varphi(x,y)|^2$. Thus, $\varphi(x,y) = 0$, and the desired inequality is obvious. ∎

Exercises

1. Let $\mathcal{P},\mathcal{Q}$ be pre-Hilbert spaces, and suppose the pre-Hilbert spaces $\mathcal{P}^*$ and $\mathcal{Q}^*$ are constructed as in *Example 8*. The following conditions are equivalent:

(a) $J: \mathcal{P} \to \mathcal{Q}$ is conjugate-linear, and $(Jx|Jy) = (y|x)$ for all $x,y \in \mathcal{P}$.

(b) $J: \mathcal{P} \to \mathcal{Q}^*$ is linear, and $[Jx|Jy] = (x|y)$ for all $x,y \in \mathcal{P}$.

(c) $J: \mathcal{P}^* \to \mathcal{Q}$ is linear, and $(Jx|Jy) = [x|y]$ for all $x,y \in \mathcal{P}^*$.

If $\mathcal{P} = \mathcal{Q}$, a surjective mapping J satisfying condition (a) is called a *conjugation* of the pre-Hilbert space $\mathcal{P}$. For example, $J(\lambda_k) = (\lambda_k^*)$ is a conjugation in the Hilbert space l^2.

2. Let $\mathfrak{H}$ be a Hilbert space. Show:

(i) The mapping $U: \mathfrak{H} \to (\mathfrak{H}')^*$ defined by $Ux = x'$ is a vector space isomorphism.

(ii) The Banach space $\mathfrak{H}'$ becomes a Hilbert space, on defining $(x'|y') = (y|x)$.

(iii) If $\mathfrak{H}'$ is the Hilbert space described in (ii), $(\mathfrak{H}')^*$ becomes a Hilbert space on setting $[x'|y'] = (y'|x')$. The mapping $J: \mathfrak{H} \to \mathfrak{H}'$, defined by $Jx = x'$, satisfies condition (a) of *Exercise 1*.

(iv) With the Hilbert space structure of $(\mathfrak{H}')^*$ defined as in (iii), the mapping U of part (i) is a Hilbert space isomorphism.

3. If $\mathcal{U}$ is a vector space, the vector space $(\mathcal{U}^*)^*$ is entitled to be called *equal* to the vector space $\mathcal{U}$. Analogously for normed spaces and pre-Hilbert spaces (see *Examples 6* and *8*).

4. If $\varphi: \mathcal{U} \times \mathcal{U} \to \mathcal{W}$ is bilinear or sesquilinear, the "parallelogram law" holds:

$$\varphi(x + y, x + y) + \varphi(x - y, x - y) = 2\varphi(x,x) + 2\varphi(y,y).$$

5. If $\varphi: \mathcal{U} \times \mathcal{U} \to \mathcal{W}$ is bilinear, then

$$\varphi(x + y, x + y) - \varphi(x - y, x - y) + i\varphi(x + iy, x + iy) - i\varphi(x - iy, x - iy) = \theta.$$

6. If $\mathcal{U}, \mathcal{V}, \mathcal{W}$ are vector spaces, the set $\mathcal{S}(\mathcal{U}, \mathcal{V}; \mathcal{W})$ of all sesquilinear mappings $\varphi: \mathcal{U} \times \mathcal{V} \to \mathcal{W}$ is a vector space relative to the operations

$$(\varphi + \psi)(x,y) = \varphi(x,y) + \psi(x,y)$$

$$(\lambda\varphi)(x,y) = \lambda\varphi(x,y).$$

The identity mapping $\varphi \to \varphi$ is a vector space isomorphism of $\mathcal{S}(\mathcal{U}, \mathcal{V}; \mathcal{W})$ with $\mathcal{B}(\mathcal{U}, \mathcal{V}^*; \mathcal{W})$. [See *Exercise 3.4*.]

7. Let $\mathcal{V}$ be the vector space of 2-ples $x = (\lambda_1, \lambda_2)$, $y = (\mu_1, \mu_2), \cdots$. Define $\varphi(x,y) = \lambda_1\mu_2$, $\psi(x,y) = \lambda_2\mu_1$. Then, φ and ψ are bilinear forms, $\varphi(x,x) = \psi(x,x)$ for all x, but $\varphi \neq \psi$.

8. Let $\mathcal{U}, \mathcal{V}$ be vector spaces. If $\varphi: \mathcal{U} \times \mathcal{V} \to \mathcal{C}$ is a sesquilinear form, define $\varphi^\#: \mathcal{V} \times \mathcal{U} \to \mathcal{C}$ by the formula $\varphi^\#(y,x) = [\varphi(x,y)]^*$.

Then:

(i) $\varphi^{\#}$ is a sesquilinear form on $\mathfrak{V} \times \mathfrak{U}$.

(ii) $\varphi \to \varphi^{\#}$ is a bijective conjugate-linear mapping of $\mathfrak{S}(\mathfrak{U},\mathfrak{V}; \mathfrak{C})$ onto $\mathfrak{S}(\mathfrak{V},\mathfrak{U}; \mathfrak{C})$.

(iii) If $\mathfrak{V} = \mathfrak{U}$, $\varphi^{\#} = \varphi$ if and only if φ is Hermitian.

9. Let $\mathcal{P}$ and $\mathcal{Q}$ be pre-Hilbert spaces, $J: \mathcal{P} \to \mathcal{Q}$ a mapping such that $(Jx|Jy) = (y|x)$ for all $x,y \in \mathcal{P}$. Prove:

(i) If $J(\mathcal{P})$ is a linear subspace of $\mathcal{Q}$, then J is conjugate-linear.

(ii) In particular, if $\mathcal{P} = \mathcal{Q}$ and J is surjective, then J is a conjugation of $\mathcal{P}$ in the sense of *Exercise 1*.

(iii) If $\mathcal{P} = \mathfrak{K}$ is a Hilbert space, and $J(\mathfrak{K})$ is a total subset of $\mathcal{Q}$, then J is surjective.

10. Let φ be a positive sesquilinear form on $\mathfrak{V} \times \mathfrak{V}$, and let $\mathfrak{N} = \{x \in \mathfrak{V}: \varphi(x,x) = 0\}$. Then, $\mathfrak{N} = \{x \in \mathfrak{V}: \varphi(x,y) = 0 \text{ for all } y \in \mathfrak{V}\}$, hence $\mathfrak{N}$ is a linear subspace of $\mathfrak{V}$.

11. If φ is a positive sesquilinear form, under what conditions on x and y does $|\varphi(x,y)|^2 = \varphi(x,x)\varphi(y,y)$ hold?

§ 6. BOUNDED SESQUILINEAR MAPPINGS

Definition 1. *If $\mathcal{E},\mathcal{F},\mathcal{G}$ are normed spaces, a sesquilinear mapping $\varphi: \mathcal{E} \times \mathcal{F} \to \mathcal{G}$ is said to be* **bounded** *in case there exists a constant M such that $\| \varphi(x,y) \| \leq M \| x \| \| y \|$ for all $x \in \mathcal{E}$, $y \in \mathcal{F}$. If moreover $\mathcal{G} = \mathfrak{C}$, φ is called a* **bounded sesquilinear form.**

In other words, φ is a bounded sesquilinear mapping of $\mathcal{E} \times \mathcal{F}$ into $\mathcal{G}$, if and only if it is a bounded bilinear mapping of $\mathcal{E} \times \mathcal{F}^*$ into $\mathcal{G}$ (see *Definition 4.1*, *Example 5.6*, and *Theorem 5.1*). The number

$$\text{LUB } \{\| \varphi(x,y) \|: \| x \| \leq 1, \| y \| \leq 1\}$$

is denoted $\| \varphi \|$, and is called the *norm* of φ. By *Theorem 4.1*,

Theorem 1. *If $\varphi: \mathcal{E} \times \mathcal{F} \to \mathcal{G}$ is a bounded sesquilinear mapping,*

(1) $$\| \varphi \| = \text{LUB } \{\| \varphi(x,y) \|: \| x \| < 1, \| y \| < 1\}$$
$$= \text{LUB } \{\| \varphi(x,y) \|: \| x \| = 1, \| y \| = 1\};$$

(2) $\| \varphi(x,y) \| \leq \| \varphi \| \| x \| \| y \|$ *for all* $x \in \mathcal{E}$, $y \in \mathcal{F}$;

(3) *if* $M \geq 0$ *and* $\| \varphi(x,y) \| \leq M \| x \| \| y \|$ *for all* $x \in \mathcal{E}$ *and* $y \in \mathcal{F}$, *necessarily* $\| \varphi \| \leq M$.

Examples

1. If $\mathcal{P}$ is a pre-Hilbert space, the sesquilinear form $\varphi: \mathcal{P} \times \mathcal{P} \to \mathcal{C}$ defined by $\varphi(x,y) = (x \mid y)$ is bounded, by the Cauchy-Schwarz inequality. Assuming $\mathcal{P} \neq \{\theta\}$, $\| \varphi \| = 1$.

2. If $\mathcal{E}$ is a normed space, $\mathcal{P}$ is a pre-Hilbert space, and $T: \mathcal{E} \to \mathcal{P}$ is a continuous linear mapping, then $\varphi(x,y) = (Tx \mid y)$ is a bounded sesquilinear form, with $\| \varphi \| = \| T \|$ (see *Example IV.7.5* or *Exercise IV.7.5*).

3. If $\mathcal{E},\mathcal{F}$ are normed spaces, and $\varphi: \mathcal{E} \times \mathcal{F} \to \mathcal{C}$ is a bounded sesquilinear form, the sesquilinear form $\psi: \mathcal{F} \times \mathcal{E} \to \mathcal{C}$ defined by $\psi(y,x) = [\varphi(x,y)]^*$ is also bounded, and $\| \psi \| = \| \varphi \|$. ∎

The following result plays an important role in the exposition of spectral theory given in Chapter VIII:

Theorem 2. *If $\mathcal{P}$ is a pre-Hilbert space, the norm of a* **Hermitian** *bounded sesquilinear form $\varphi: \mathcal{P} \times \mathcal{P} \to \mathcal{C}$ is given by the formula*

$$\| \varphi \| = \text{LUB} \{ |\varphi(x,x)| : \| x \| \leq 1 \}.$$

Proof.

By assumption, $\varphi(x,x)$ is real, for all x (see *Theorem 5.4*).

If $\| x \| \leq 1$, $|\varphi(x,x)| \leq \| \varphi \|$ by the definition of $\| \varphi \|$; hence the indicated LUB is a finite real number M, and $M \leq \| \varphi \|$. Clearly $|\varphi(x,x)| \leq M \| x \| \| x \|$ for all $x \in \mathcal{P}$.

Fix $x,y \in \mathcal{P}$, with $\| x \| \leq 1$, $\| y \| \leq 1$. It will suffice to show that $|\varphi(x,y)| \leq M$.

Case 1. Suppose $\varphi(x,y)$ is a real number. By *Theorem 5.2*, $\varphi(x,y) = \frac{1}{4}\{\varphi(x+y,x+y) - \varphi(x-y,x-y) + i\varphi(x+iy,x+iy) - i\varphi(x-iy,x-iy)\}$. Since $\varphi(z,z)$ is real for all z, and since $\varphi(x,y)$ is by assumption real, necessarily $\varphi(x,y) = \frac{1}{4}\{\varphi(x+y,x+y) - \varphi(x-y,x-y)\}$. Quoting the parallelogram law at the appropriate step, $|\varphi(x,y)| \leq \frac{1}{4}\{|\varphi(x+y,x+y)| + |\varphi(x-y,x-y)|\} \leq \frac{1}{4}\{M \| x+y \|^2 + M \| x-y \|^2\} = \frac{M}{4}\{2 \| x \|^2 + 2 \| y \|^2\} \leq \frac{M}{4}\{2+2\} = M$.

Case 2. In general, one can write $|\varphi(x,y)| = \lambda\varphi(x,y)$, where λ is a suitable complex number of absolute value 1. Then, $\varphi(\lambda x,y) = \lambda\varphi(x,y) = |\varphi(x,y)|$ is real; since $\| \lambda x \| = |\lambda| \cdot \| x \| = \| x \| \leq 1$, $|\varphi(\lambda x,y)| \leq M$ by case 1, thus $|\varphi(x,y)| \leq M$. ∎

Exercises

1. If $\mathcal{E},\mathcal{F},\mathcal{G}$ are normed spaces, the set $\mathcal{S}_c(\mathcal{E},\mathcal{F};\mathcal{G})$ of all bounded sesquilinear mappings $\varphi\colon \mathcal{E}\times\mathcal{F}\to\mathcal{G}$ is a linear subspace of $\mathcal{S}(\mathcal{E},\mathcal{F};\mathcal{G})$, and becomes a normed space on defining $\|\varphi\|$ as above. The identity mapping $\varphi\to\varphi$ is an isometric vector space isomorphism of $\mathcal{S}_c(\mathcal{E},\mathcal{F};\mathcal{G})$ with $\mathcal{B}_c(\mathcal{E},\mathcal{F}^*;\mathcal{G})$. (See *Exercises 5.6* and *4.4*)

2. Let $\mathcal{E}$ be a normed space, $\mathcal{P}$ a pre-Hilbert space. Each continuous linear mapping $T\colon \mathcal{E}\to\mathcal{P}$ determines a bounded sesquilinear form $\varphi_T\colon \mathcal{E}\times\mathcal{P}\to\mathcal{C}$, namely $\varphi_T(x,y)=(Tx|y)$, and one has $\|\varphi_T\|=\|T\|$ (see *Example 2*). Then, $T\to\varphi_T$ is an isometric linear mapping of $\mathcal{L}_c(\mathcal{E},\mathcal{P})$ into $\mathcal{S}_c(\mathcal{E},\mathcal{P};\mathcal{C})$ (see *Exercise 1*).

3. If $\mathcal{P}$ is a pre-Hilbert space, $\mathcal{G}$ is a normed space, and $\varphi\colon\mathcal{P}\times\mathcal{P}\to\mathcal{G}$ is a bounded bilinear mapping such that $\varphi(y,x)=\varphi(x,y)$, then $\|\varphi\|=\text{LUB}\{\|\varphi(x,x)\|\colon\|x\|\leq 1\}$.

4. Let $\mathcal{P}$ be a pre-Hilbert space, and define $\varphi(x,y)=i(x|y)$. Then, φ is a bounded sesquilinear form, $\|\varphi\|=\text{LUB}\{|\varphi(x,x)|\colon\|x\|\leq 1\}$, but φ is not Hermitian.

5. Let $\mathcal{P}$ be a pre-Hilbert space, and suppose $U\colon\mathcal{P}\to\mathcal{P}$ is a continuous linear mapping such that $\|U\|=1$ and $Uz=z$ for some $z\neq\theta$. Define $\varphi_U(x,y)=(Ux|y)$. Then, $\|\varphi_U\|=\|U\|=1=\text{LUB}\{|\varphi_U(x,x)|\colon\|x\|\leq 1\}$, but φ_U need not be Hermitian.

§7. BOUNDED SESQUILINEAR FORMS IN HILBERT SPACE

As a special case of *Example 6.2*,

Theorem 1. *If $\mathcal{H}$ and $\mathcal{K}$ are Hilbert spaces, and $T\colon\mathcal{H}\to\mathcal{K}$ is a continuous linear mapping, the formula $\varphi_T(x,y)=(Tx|y)$ defines a bounded sesquilinear form φ_T on $\mathcal{H}\times\mathcal{K}$, such that $\|\varphi_T\|=\|T\|$.*

The point of stating *Theorem 1* is that its converse is true:

Theorem 2. *If $\mathcal{H}$ and $\mathcal{K}$ are Hilbert spaces, and $\varphi\colon\mathcal{H}\times\mathcal{K}\to\mathcal{C}$ is a bounded sesquilinear form, there exists one and only one continuous linear mapping $T\colon\mathcal{H}\to\mathcal{K}$ such that $\varphi(x,y)=(Tx|y)$ for all $x\in\mathcal{H}$, $y\in\mathcal{K}$. In the notation of Theorem 1, $\varphi=\varphi_T$, and $\|\varphi\|=\|\varphi_T\|=\|T\|$.*

Proof.

Given $x \in \mathcal{H}$, the problem is to define $Tx \in \mathcal{K}$. Now, $y \to [\varphi(x,y)]^*$ is a linear form f_x on $\mathcal{K}$, and $|f_x(y)| = |[\varphi(x,y)]^*| = |\varphi(x,y)| \leq \| \varphi \| \, \| x \| \, \| y \|$ shows that f_x is continuous, and $\| f_x \| \leq \| \varphi \| \, \| x \|$. By *Theorem 1.1*, there is a unique vector $z \in \mathcal{K}$ such that $f_x(y) = (y|z)$ for all $y \in \mathcal{K}$. Defining $Tx = z$, one has $(y|Tx) = f_x(y) = [\varphi(x,y)]^*$, thus $\varphi(x,y) = (Tx|y)$. Incidentally, $\| Tx \| = \| f_x \|$ by *Example IV.7.4*, thus $\| Tx \| \leq \| \varphi \| \, \| x \|$. It will now be shown that the mapping $T: \mathcal{H} \to \mathcal{K}$ is linear.

T is *additive:* given any $x_1, x_2 \in \mathcal{H}$, $T(x_1 + x_2) = Tx_1 + Tx_2$. For, given any $y \in \mathcal{K}$, $(T(x_1 + x_2)|y) = \varphi(x_1 + x_2, y) = \varphi(x_1,y) + \varphi(x_2,y) = (Tx_1|y) + (Tx_2|y) = (Tx_1 + Tx_2|y)$; see *Theorem II.2.1.*

T is *homogeneous:* If $x \in \mathcal{H}$, λ is scalar, and $y \in \mathcal{K}$, $(T(\lambda x)|y) = \varphi(\lambda x, y) = \lambda\varphi(x,y) = \lambda(Tx|y) = (\lambda(Tx)|y)$, hence $T(\lambda x) = \lambda(Tx)$.

T is *unique:* Suppose $S: \mathcal{H} \to \mathcal{K}$ is any mapping such that $(Sx|y) = \varphi(x,y)$ for all $x \in \mathcal{H}$, $y \in \mathcal{K}$. Then, $(Tx - Sx|y) = 0$ for all $x \in \mathcal{H}$, $y \in \mathcal{K}$; hence $Tx - Sx = \theta$ for all $x \in \mathcal{H}$, that is, $T = S$.

Since $\| Tx \| \leq \| \varphi \| \, \| x \|$, the linear mapping T is continuous. Since $\varphi = \varphi_T$, $\| \varphi \| = \| T \|$ by *Theorem 1.* ∎

Exercises

1. If $\mathcal{H}$ and $\mathcal{K}$ are Hilbert spaces, the correspondence $T \to \varphi_T$ described in *Theorem 1* is an isometric vector space isomorphism of $\mathfrak{L}_c(\mathcal{H},\mathcal{K})$ with $\mathfrak{S}_c(\mathcal{H},\mathcal{K};\mathbb{C})$. [See *Exercise 6.2.*]

2. If $\mathcal{E}$ is a normed space, and $\mathcal{K}$ is a Hilbert space, generalize *Theorem 2* and *Exercise 1* for bounded sesquilinear forms $\varphi: \mathcal{E} \times \mathcal{K} \to \mathbb{C}$.

§ 8. ADJOINTS

Throughout this section, $\mathcal{H},\mathcal{K},\mathcal{L}$ denote Hilbert spaces.

Theorem 1. *If $T: \mathcal{H} \to \mathcal{K}$ is a continuous linear mapping, there exists one and only one continuous linear mapping $T^*: \mathcal{K} \to \mathcal{H}$ such that $(Tx|y) = (x|T^*y)$ for all $x \in \mathcal{H}$, $y \in \mathcal{K}$. One has $\| T^* \| = \| T \|$.*

Proof.

Define $\psi(y,x) = (y|Tx)$. Clearly ψ is a sesquilinear form on $\mathcal{K} \times \mathcal{H}$, and $|\psi(y,x)| = |(y|Tx)| = |(y|Tx)^*| = |(Tx|y)|$ shows that ψ is

bounded and $\| \psi \| = \| T \|$ (see *Theorem 7.1*). By *Theorem 7.2*, there exists a continuous linear mapping $T^*: \mathcal{K} \to \mathcal{H}$, with $\| T^* \| = \| \psi \|$, such that $\psi(y,x) = (T^*y|x)$ for all $y \in \mathcal{K}$, $x \in \mathcal{H}$. Thus, $(y|Tx) = (T^*y|x)$, $(Tx|y) = (x|T^*y)$.

If $S: \mathcal{K} \to \mathcal{H}$ is any mapping such that $(Tx|y) = (x|Sy)$ for all $x \in \mathcal{H}$, $y \in \mathcal{K}$, then $(x|Sy - T^*y) = 0$ for all x and y, hence $Sy - T^*y = \theta$ for all y; that is, $S = T^*$. ∎

Definition 1. *Notation as in Theorem 1; T^* is called the* **adjoint** *of T.*

The rest of the section is devoted to proving properties of adjoints which will be used often in the sequel.

Theorem 2. *If $S: \mathcal{H} \to \mathcal{K}$ and $T: \mathcal{H} \to \mathcal{K}$ are continuous linear mappings, and λ is scalar,*

(1) $(S + T)^* = S^* + T^*$

(2) $(\lambda T)^* = \lambda^* T^*$

(3) $(T^*y|x) = (y|Tx)$ *for all* $y \in \mathcal{K}$, $x \in \mathcal{H}$

(4) $(T^*)^* = T$

(5) $\| T^*T \| = \| TT^* \| = \| T \|^2$

(6) $T^*T = 0$ *if and only if* $T = 0$.

Proof.

(3): $(T^*y|x) = (x|T^*y)^* = (Tx|y)^* = (y|Tx)$.

(1): $((S + T)^*y|x) = (y|(S + T)x) = (y|Sx + Tx) = (y|Sx) + (y|Tx) = (S^*y|x) + (T^*y|x) = (S^*y + T^*y|x) = ((S^* + T^*)y|x)$.

(2): $((\lambda T)^*y|x) = (y|(\lambda T)x) = (y|\lambda(Tx)) = \lambda^*(y|Tx) = \lambda^*(T^*y|x) = (\lambda^*(T^*y)|x) = ((\lambda^*T^*)y|x)$.

(4): $((T^*)^*x|y) = (x|T^*y) = (Tx|y)$.

(5): Observe that the composites $T^*T: \mathcal{H} \to \mathcal{H}$ and $TT^*: \mathcal{K} \to \mathcal{K}$ are defined. If $x \in \mathcal{H}$, $\| x \| \leq 1$, one has $\| Tx \|^2 = (Tx|Tx) = (T^*Tx|x) \leq \| T^*Tx \| \, \| x \| \leq \| T^*Tx \| \leq \| T^*T \|$; taking LUB over $\| x \| \leq 1$, $\| T \|^2 \leq \| T^*T \|$. By *Theorem IV.9.1*, $\| T^*T \| \leq \| T^* \| \, \| T \| = \| T \|^2$. Thus, $\| T^*T \|^2 = \| T \|^2$. Replacing T by T^*, $\| (T^*)^*T^* \| = \| T^* \|^2$, thus $\| TT^* \| = \| T \|^2$.

(6): follows at once from (5). ∎

Theorem 3. *If* $T: \mathfrak{H} \to \mathfrak{K}$ *and* $S: \mathfrak{K} \to \mathfrak{L}$ *are continuous linear mappings, then* $(ST)^* = T^*S^*$.

Proof.

If $z \in \mathfrak{L}$ and $x \in \mathfrak{H}$, $((ST)^*z|x) = (z|(ST)x) = (z|S(Tx)) = (S^*z|Tx) = (T^*(S^*z)|x) = ((T^*S^*)z|x)$. ∎

Theorem 4. *Let* $T: \mathfrak{H} \to \mathfrak{K}$ *be a continuous linear mapping. If* $\mathfrak{S} \subset \mathfrak{H}$, $\mathfrak{J} \subset \mathfrak{K}$, *and* $T(\mathfrak{S}) \subset \mathfrak{J}$, *then* $T^*(\mathfrak{J}^\perp) \subset \mathfrak{S}^\perp$.

Proof.

Given $z \in \mathfrak{K}$ with $z \perp \mathfrak{J}$, the problem is to show that $T^*z \perp \mathfrak{S}$. If $x \in \mathfrak{S}$, then $Tx \in \mathfrak{J}$, hence $0 = (z|Tx) = (T^*z|x)$. ∎

In particular,

Theorem 5. *Suppose* $T: \mathfrak{H} \to \mathfrak{K}$ *is a continuous linear mapping,* $\mathfrak{M}$ *is a closed linear subspace of* $\mathfrak{H}$, *and* $\mathfrak{N}$ *is a closed linear subspace of* $\mathfrak{K}$. *Then,* $T(\mathfrak{M}) \subset \mathfrak{N}$ *if and only if* $T^*(\mathfrak{N}^\perp) \subset \mathfrak{M}^\perp$.

Proof.

If $T^*(\mathfrak{N}^\perp) \subset \mathfrak{M}^\perp$, then $(T^*)^*(\mathfrak{M}^{\perp\perp}) \subset \mathfrak{N}^{\perp\perp}$ by *Theorem 4;* quote *Theorem III.6.2* and part (4) of *Theorem 2.* ∎

Theorem 6. *If* $T: \mathfrak{H} \to \mathfrak{K}$ *is a continuous linear mapping,*

(1) $\{x: Tx = \theta\} = [T^*(\mathfrak{K})]^\perp$

(2) $\{x: Tx = \theta\}^\perp = \overline{T^*(\mathfrak{K})}$

(3) $\{y: T^*y = \theta\} = [T(\mathfrak{H})]^\perp$

(4) $\{y: T^*y = \theta\}^\perp = \overline{T(\mathfrak{H})}$.

Thus, the null space of T *is the orthogonal complement of the closure of the range of* T^*.

Proof.

Let $\mathfrak{M} = \{x: Tx = \theta\}$; $\mathfrak{M}$ is a closed linear subspace of $\mathfrak{H}$. Similarly, $\mathfrak{N} = \{y: T^*y = \theta\}$ is a closed linear subspace of $\mathfrak{K}$.

Since $T(\mathfrak{M}) \subset \{\theta\}$, one has $T^*(\{\theta\}^\perp) \subset \mathfrak{M}^\perp$ by *Theorem 4;* thus,

(i) $T^*(\mathfrak{K}) \subset \mathfrak{M}^\perp$.

Let $\mathfrak{S} = \mathfrak{H}$, $\mathfrak{J} = T(\mathfrak{H})$; the relation $T(\mathfrak{H}) \subset T(\mathfrak{H})$ can be written $T(\mathfrak{S}) \subset \mathfrak{J}$, hence $T^*(\mathfrak{J}^\perp) \subset \mathfrak{S}^\perp$. Since $\mathfrak{S}^\perp = \mathfrak{H}^\perp = \{\theta\}$, one has $T^*(\mathfrak{J}^\perp) = \{\theta\}$, $\mathfrak{J}^\perp \subset \mathfrak{N}$, thus

(ii) $[T(\mathfrak{H})]^\perp \subset \mathfrak{N}$.

Applying these results to T^* in place of T, one has

(i*) $T(\mathcal{H}) \subset \mathfrak{M}^\perp$

(ii*) $[T^*(\mathcal{K})]^\perp \subset \mathfrak{M}$.

From (i), $\mathfrak{M} \subset \mathfrak{M}^{\perp\perp} \subset [T^*(\mathcal{K})]^\perp$; combining this with (ii*),

(1) $\mathfrak{M} = [T^*(\mathcal{K})]^\perp$.

[Incidentally, $\mathfrak{M} = \mathfrak{M}^{\perp\perp}$ results formally from these calculations; see also *Theorem III.6.2.*]

From (1), $\mathfrak{M}^\perp = [T^*(\mathcal{K})]^{\perp\perp}$; quoting Corollary 1 of *Theorem III.6.3*,

(2) $\mathfrak{M}^\perp = \overline{T^*(\mathcal{K})}$.

The relations (3) and (4) result on replacing T by T^* in the relations (1) and (2). ∎

Exercises

1. Let $\mathcal{P}$ and $\mathcal{Q}$ be pre-Hilbert spaces, and suppose $T: \mathcal{P} \to \mathcal{Q}$ and $S: \mathcal{Q} \to \mathcal{P}$ are mappings such that $(Tx|y) = (x|Sy)$ for all $x \in \mathcal{P}$ and $y \in \mathcal{Q}$. Then, S and T are linear.

2. If $\mathcal{H}$ is a Hilbert space, $(S,T) \to ST^*$ is a bounded sesquilinear mapping of $\mathfrak{L}_c(\mathcal{H}) \times \mathfrak{L}_c(\mathcal{H})$ into $\mathfrak{L}_c(\mathcal{H})$. So is $(S,T) \to T^*S$.

3. Let $\mathcal{H}$ be classical Hilbert space, $T: \mathcal{H} \to \mathcal{H}$ a continuous linear mapping. Given an orthonormal basis e_n for $\mathcal{H}$. Let $Te_k = \sum_j \alpha_{jk} e_j$ be the expansion of Te_k in terms of the given orthonormal basis (see *Theorem II.8.1*). The array (α_{jk}) is called the *matrix* of T relative to the given orthonormal basis (see also *Exercise IV.1.6*). Prove:

(i) $(Te_k|e_j) = \alpha_{jk}$

(ii) If T^* has matrix (β_{jk}) relative to the same orthonormal basis, then $\beta_{jk} = \alpha_{kj}{}^*$. So to speak, the matrix (β_{jk}) is the "conjugate transpose" of the matrix (α_{jk}).

4. Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces, $T: \mathcal{H} \to \mathcal{K}$ a continuous linear mapping. Prove:

(i) $T'y' = (T^*y)'$ for all $y \in \mathcal{K}$ (see *Exercise IV.10.7* for the definition of T').

(ii) If $U: \mathcal{H} \to \mathcal{H}'$ is the conjugate linear isometric mapping $Ux = x'$ described in *Theorem IV.10.2*, and similarly $V: \mathcal{K} \to \mathcal{K}'$ is the mapping $Vy = y'$, then $T^* = U^{-1}T'V$.

(iii) Identifying $\mathcal{H}$ with $(\mathcal{H}')^*$ via the Hilbert space isomorphism $x \to x'$ (see *Exercise 5.2*), and similarly identifying $\mathcal{K}$ with $(\mathcal{K}')^*$, one has $T^* = T'$.

5. Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. Form the direct product Hilbert spaces $\mathcal{H} \times \mathcal{K}$ and $\mathcal{K} \times \mathcal{H}$ (see *Exercise 3.11*). Define $W: \mathcal{H} \times \mathcal{K} \to \mathcal{K} \times \mathcal{H}$ by the formula $W(x,y) = (y,-x)$. Prove:

(i) W is a Hilbert space isomorphism.

(ii) If $T: \mathcal{H} \to \mathcal{K}$ is a continuous linear mapping, the graphs of T and T^* satisfy the following relation: $\mathcal{G}_{T^*} = [W(\mathcal{G}_T)]^\perp$ (see *Exercises 3.10* and *3.13*).

What is the formula for $W^*(y,x)$?

*6. Let $\mathcal{H}$ be a Hilbert space, $\mathcal{P}$ a pre-Hilbert space, and suppose $T: \mathcal{H} \to \mathcal{P}$ and $S: \mathcal{P} \to \mathcal{H}$ are mappings such that $(Tx|y) = (x|Sy)$ for all $x \in \mathcal{H}$ and $y \in \mathcal{P}$. One knows that S and T are linear by *Exercise 1*. It can be shown that S and T are continuous. If moreover $\mathcal{P} = \mathcal{K}$ is a Hilbert space, then $S = T^*$.

7. Let $\mathcal{P}$ be the pre-Hilbert space of finitely non-zero sequences, and let e_n be the canonical basis for $\mathcal{P}$ (see *Example I.7.2*). There is a unique linear mapping $T: \mathcal{P} \to \mathcal{P}$ such that $Te_k = ke_k$ for all k. Prove: $(Tx|y) = (x|Ty)$ for all $x,y \in \mathcal{P}$, but T is not continuous. Compare this result with *Exercise 6*.

8. This exercise is preliminary to *Exercise 9*. A doubly indexed family of real numbers $\alpha_{jk} \geq 0$ ($j = 1,2,3,\cdots$; $k = 1,2,3,\cdots$) is said to be *summable* in case there is a constant $M \geq 0$ such that

$$\sum_{j=1}^{m} \sum_{k=1}^{n} \alpha_{jk} \leq M$$

for all indices m and n (equivalently, the sum of any finite number of the α_{jk} is $\leq M$). One then writes $\sum_{j,k} \alpha_{jk}$ for the LUB of all such finite sums. If $\alpha_{jk} \geq 0$ is a doubly indexed family, the symbol $\sum_{j,k} \alpha_{jk} < \infty$ signifies that the family is summable.

Show that for a family $\alpha_{jk} \geq 0$, the following conditions are equivalent:

(a) $\sum_{j,k} \alpha_{jk} < \infty$.

(b) For each j, $\sum_k \alpha_{jk} < \infty$, and $\sum_j \left(\sum_k \alpha_{jk} \right) < \infty$.

(c) For each k, $\sum_j \alpha_{jk} < \infty$, and $\sum_k \left(\sum_j \alpha_{jk} \right) < \infty$.

If either (hence all) of these conditions is fulfilled, one has

$$\sum_j \left(\sum_k \alpha_{jk} \right) = \sum_k \left(\sum_j \alpha_{jk} \right) = \sum_{j,k} \alpha_{jk}.$$

9. This exercise is preliminary to *Exercise 11*. Denote by $\mathfrak{H}$ the set of all doubly indexed families $x = (\lambda_{jk})$ of complex numbers such that $\sum_{j,k} |\lambda_{jk}|^2 < \infty$ in the sense of *Exercise 8*. If $y = (\mu_{jk})$ is another such family, define $x = y$ in case $\lambda_{jk} = \mu_{jk}$ for all j,k. With operations suitably defined, $\mathfrak{H}$ is a Hilbert space.

10. Let $\mathfrak{H}$ and $\mathfrak{K}$ be classical Hilbert spaces, e_n an orthonormal basis for $\mathfrak{H}$, and f_n an orthonormal basis for $\mathfrak{K}$. If $T: \mathfrak{H} \to \mathfrak{K}$ is a continuous linear mapping, one has $\sum_k \| Te_k \|^2 < \infty$ if and only if $\sum_j \| T^* f_j \|^2 < \infty$, and in this case one has

$$\sum_k \| Te_k \|^2 = \sum_j \| T^* f_j \|^2 = \sum_{j,k} |(Te_k | f_j)|^2.$$

11. Let $\mathfrak{H}$ and $\mathfrak{K}$ be classical Hilbert spaces, e_n a fixed orthonormal basis for $\mathfrak{H}$, f_n a fixed orthonormal basis for $\mathfrak{K}$. A continuous linear mapping $T: \mathfrak{H} \to \mathfrak{K}$ is said to be of *Hilbert-Schmidt class* if there exists an orthonormal basis g_n of $\mathfrak{H}$ such that $\sum_k \| Tg_k \|^2 < \infty$. Briefly, T is of HS-class. Prove:

(i) If $T: \mathfrak{H} \to \mathfrak{K}$ is of HS-class, then $\sum_k \| Tg_k \|^2 < \infty$ for *every* orthonormal basis g_n of $\mathfrak{H}$, and the value of this sum is independent of the particular orthonormal basis.

(ii) If $S: \mathfrak{H} \to \mathfrak{K}$ and $T: \mathfrak{H} \to \mathfrak{K}$ are of HS-class, so are $S + T$ and λT. The continuous linear mappings $T: \mathfrak{H} \to \mathfrak{K}$ of HS-class form a linear subspace $\mathfrak{L}_{hs}(\mathfrak{H},\mathfrak{K})$ of $\mathfrak{L}_c(\mathfrak{H},\mathfrak{K})$.

(iii) A continuous linear mapping $T: \mathfrak{H} \to \mathfrak{K}$ is of HS-class if and only if $T^*: \mathfrak{K} \to \mathfrak{H}$ is of HS-class.

(iv) If $T: \mathfrak{H} \to \mathfrak{K}$ is of HS-class, and $R: \mathfrak{H} \to \mathfrak{H}$, $S: \mathfrak{K} \to \mathfrak{K}$ are continuous linear mappings, then $TR: \mathfrak{H} \to \mathfrak{K}$ and $ST: \mathfrak{H} \to \mathfrak{K}$ are of HS-class.

*(v) If $T: \mathfrak{H} \to \mathfrak{K}$ is of HS-class, define $\alpha_{jk} = (Te_k | f_j)$. The correspondence $T \to (\alpha_{jk})$ is a vector space isomorphism of $\mathfrak{L}_{hs}(\mathfrak{H},\mathfrak{K})$ with the Hilbert space described in *Exercise 9*.

*(vi) If S and T are of HS-class, define $(S|T) = \sum_k (Se_k | Te_k)$. Then, $\mathfrak{L}_{hs}(\mathfrak{H},\mathfrak{K})$ becomes a pre-Hilbert space, and the scalar product $(S|T)$ is independent of the particular orthonormal basis employed in its definition. The mapping $T \to (\alpha_{jk})$ described in (v) is an

isometric vector space isomorphism; it follows that $\mathfrak{L}_{hs}(\mathfrak{H},\mathfrak{K})$ is a Hilbert space, and the mapping $T \to (\alpha_{jk})$ is a Hilbert space isomorphism.

12. If $\mathfrak{A}$ is an algebra, a mapping $a \to a^*$ of $\mathfrak{A}$ into $\mathfrak{A}$ is called an *involution* in case: (1) $a^{**} = a$, (2) $(a + b)^* = a^* + b^*$, (3) $(ab)^* = b^*a^*$, and $(\lambda a)^* = \lambda^* a^*$, for all $a,b \in \mathfrak{A}$ and scalar λ. An algebra with involution is also called a **-algebra*. A normed algebra with involution is called a *normed *-algebra*. Prove: if $\mathfrak{A}$ is a normed *-algebra such that $\| a^*a \| = \| a \|^2$ for all $a \in \mathfrak{A}$, then $\| a^* \| = \| a \|$.

13. This exercise is preliminary to *Exercise 14*. If $\mathfrak{A}$ is an algebra, a *subalgebra* of $\mathfrak{A}$ is a linear subspace $\mathfrak{S}$ such that: if $a \in \mathfrak{S}$ and $b \in \mathfrak{S}$, then $ab \in \mathfrak{S}$. Prove:

(i) In a normed algebra, the relations $a_n \to a$ and $b_n \to b$ imply $a_n b_n \to ab$.

(ii) If $\mathfrak{S}$ is a subalgebra of the normed algebra $\mathfrak{A}$, its closure $\bar{\mathfrak{S}}$ is a closed subalgebra of $\mathfrak{A}$.

(iii) If $\mathfrak{A}$ is a Banach algebra, every closed subalgebra of $\mathfrak{A}$ is itself a Banach algebra.

14. If $\mathfrak{A}$ is a *-algebra, a **-subalgebra* of $\mathfrak{A}$ is a subalgebra $\mathfrak{S}$ such that $a \in \mathfrak{S}$ implies $a^* \in \mathfrak{S}$.

(i) Suppose $\mathfrak{A}$ is a normed *-algebra such that $\| a^* \| = \| a \|$ for all $a \in \mathfrak{A}$. Then, $a_n \to a$ implies $a_n^* \to a^*$.

(ii) If $\mathfrak{A}$ is as in part (i), and $\mathfrak{S}$ is a *-subalgebra of $\mathfrak{A}$, its closure $\bar{\mathfrak{S}}$ is also a *-subalgebra of $\mathfrak{A}$.

15. A *B^*-algebra* is a Banach *-algebra $\mathfrak{A}$ such that $\| a^*a \| = \| a \|^2$ for all $a \in \mathfrak{A}$. Prove:

(i) If $\mathfrak{S}$ is a *-subalgebra of a B^*-algebra $\mathfrak{A}$, then $\bar{\mathfrak{S}}$ is a B^*-algebra.

(ii) If $\mathfrak{H}$ is a Hilbert space, and $\mathfrak{S}$ is a closed *-subalgebra of $\mathfrak{L}_c(\mathfrak{H})$, then $\mathfrak{S}$ is a B^*-algebra.

*16. There exists a converse for part (ii) of *Exercise 15*, known as the "theorem of Gelfand-Neumark": if $\mathfrak{A}$ is any B^*-algebra, there exists a Hilbert space $\mathfrak{H}$, and a mapping $\varphi: \mathfrak{A} \to \mathfrak{L}_c(\mathfrak{H})$, such that

(1) $\varphi(a + b) = \varphi(a) + \varphi(b)$
(2) $\varphi(\lambda a) = \lambda\varphi(a)$
(3) $\varphi(ab) = \varphi(a)\varphi(b)$
(4) $\| \varphi(a) \| = \| a \|$
(5) $\varphi(a^*) = [\varphi(a)]^*$ (= the adjoint of $\varphi(a)$).

It follows that $\mathcal{S} = \varphi(\mathfrak{A})$ is a B^*-algebra of the kind described in part (ii) of *Exercise 15;* hence, $\mathfrak{A}$ is essentially "equal" to a B^*-algebra of this kind.

17. Notation as in *Exercise 11*, with $\mathfrak{K} = \mathfrak{K}$. Then, $\mathfrak{L}_{hs}(\mathfrak{K},\mathfrak{K})$ is a *-subalgebra of $\mathfrak{L}_c(\mathfrak{K})$.

18. An *involution* in a vector space $\mathcal{U}$ is a mapping $x \to x^\#$ of $\mathcal{U}$ into $\mathcal{U}$ such that: (1) $x^{\#\#} = x$, (2) $(x + y)^\# = x^\# + y^\#$, and (3) $(\lambda x)^\# = \lambda^* x^\#$. Prove:

(i) If $\mathcal{U}$ is a vector space, and φ: $\mathcal{U} \times \mathcal{U} \to \mathcal{C}$ is a sesquilinear form, define $\varphi^\#(x,y) = [\varphi(y,x)]^*$. Then, $\varphi \to \varphi^\#$ is an involution in the vector space $\mathcal{S}(\mathcal{U},\mathcal{U};\mathcal{C})$ (see *Exercise 5.8*).

(ii) Let $\mathfrak{K}$ be a Hilbert space. If φ_T is defined as in *Theorem 7.1*, then $(\varphi_T)^\# = \varphi_{T^*}$.

Chapter

Operators in Hilbert Space VI

§ 1. *Manifesto*
§ 2. *Preliminaries*
§ 3. *An example*
§ 4. *Isometric operators*
§ 5. *Unitary operators*
§ 6. *Self-adjoint operators*
§ 7. *Projection operators*
§ 8. *Normal operators*
§ 9. *Invariant and reducing subspaces*

§ 1. MANIFESTO

For the rest of the book, we settle down to the study of various types of continuous linear mappings of a Hilbert space into itself. Generalities are henceforth banished to the exercises, and the reader is invited to assess each definition and proof for its possible level of generality. For example, some statements about Hilbert space will make sense for normed spaces, or vector spaces, and so on; observations of this sort will be found in the exercises. Henceforth:

1. $\mathfrak{H}$ denotes a Hilbert space $\neq \{\theta\}$.

2. Any continuous linear mapping T: $\mathfrak{H} \to \mathfrak{H}$ will be called an **operator.** In particular, I is the *identity operator*, 0 is the *zero operator*, λI is a *scalar operator*.

3. T^* denotes the adjoint of the operator T. Thus,
   (i) T^* is an operator
   (ii) $(Tx|y) = (x|T^*y)$ for all $x,y \in \mathfrak{H}$
   (iii) $T^{**} = T$
   (iv) $(S + T)^* = S^* + T^*$

(v) $(\lambda T)^* = \lambda^* T^*$
(vi) $(ST)^* = T^* S^*$
(vii) $\| T^* \| = \| T \|$
(viii) $\| T^* T \| = \| T \|^2$
(ix) $T^* T = 0$ if and only if $T = 0$.

These facts about adjoints are assumed to be known to the reader (see § 8 of Chapter V); they will be used in the sequel without explicit reference.

The special types of operators T which will be discussed in this chapter are defined as follows:

isometric operator: $T^* T = I$
unitary operator: $T^* T = TT^* = I$
self-adjoint operator: $T^* = T$
projection operator: $TT = T$ and $T^* = T$
normal operator: $T^* T = TT^*$.

Other types of operators (positive, invertible, etc....) are discussed in the exercises.

§ 2. PRELIMINARIES

The theorems of this section will be used often, and without explicit reference, in the sequel.

Theorem 1. *The null space $\mathfrak{N}$ of an operator T is a closed linear subspace of $\mathfrak{H}$.*

Proof.
See *Theorem IV.7.2.* ∎

Theorem 2. *If $\mathcal{S}$ is a total subset of $\mathfrak{H}$, and S,T are operators such that $Sx = Tx$ for all $x \in \mathcal{S}$, then $S = T$.*

Proof.
Let $\mathfrak{N}$ be the null space of $S - T$. By assumption, $\mathcal{S} \subset \mathfrak{N}$. Since $\mathcal{S}$ is total, so is $\mathfrak{N}$, hence $\mathfrak{N}^\perp = \{\theta\}$ (see *Example III.2.1*). By *Theorem III.6.2*, $\mathfrak{N} = \mathfrak{N}^{\perp\perp} = \{\theta\}^\perp = \mathfrak{H}$; that is, $S - T = 0$. ∎

Theorem 3. *If S and T are operators such that $(Sx \mid x) = (Tx \mid x)$ for all $x \in \mathfrak{H}$, then $S = T$.*

Proof.

Consider the sesquilinear forms φ and ψ defined by the formulas $\varphi(x,y) = (Sx|y)$, $\psi(x,y) = (Tx|y)$ (see *Example V.5.2*). By assumption $\varphi(x,x) = \psi(x,x)$ for all x, hence $\varphi = \psi$ by *Theorem V.5.3*. Thus, $(Sx|y) = (Tx|y)$ for all x and y; $Sx = Tx$ results from *Theorem II.2.1*. ∎

Exercises

1. If $\mathfrak{X}$ is a set, $\mathcal{P}$ is a pre-Hilbert space, and S,T are mappings of $\mathfrak{X}$ into $\mathcal{P}$ such that $(Sx|y) = (Tx|y)$ for all $x \in \mathfrak{X}$ and $y \in \mathcal{P}$, then $S = T$.

2. If T is an operator such that Tx is orthogonal to x, for every vector x, then $T = 0$.

3. If $S^*S + T^*T = 0$, then $S = T = 0$.

4. Let S and T be operators, $\mathcal{S}$ a total subset of $\mathfrak{H}$. Show:
(i) If $(Sx|y) = (Tx|y)$ for all $x,y \in \mathcal{S}$, then $S = T$.
(ii) It can happen that $(Sx|x) = (Tx|x)$ for all $x \in \mathcal{S}$, and yet $S \neq T$.

5. *Theorem 3* holds for (not necessarily continuous) linear mappings in pre-Hilbert space.

6. In a normed space $\mathcal{E}$, call a subset $\mathcal{S}$ *total* in case: the only continuous linear form f on $\mathcal{E}$ such that $f(x) = 0$ for all $x \in \mathcal{S}$, is $f = 0$. Then:
*(i) It can be shown that if $\mathcal{S}$ is a total subset of $\mathcal{E}$, then $\mathcal{E}$ is the smallest closed linear subspace of $\mathcal{E}$ which contains $\mathcal{S}$. *Theorem 2* then generalizes to normed spaces.
(ii) In Hilbert space, a subset $\mathcal{S}$ is total in the above sense, if and only if it is total in the sense of *Definition II.8.1*.

7. Let $\mathcal{S}$ be a subset of a pre-Hilbert space $\mathcal{P}$.
(i) If $\mathcal{S}$ is total in the sense of *Exercise 6*, then it is total in the sense of *Definition II.8.1*.
(ii) The converse of (i) fails.

§ 3. AN EXAMPLE

The theorem of this section is a useful source of examples of operators. Suppose $\mathfrak{H}$ is classical Hilbert space, and x_n is an orthonormal

basis for $\mathfrak{H}$. Thus, every $x \in \mathfrak{H}$ has a unique representation $x = \sum_1^\infty \lambda_k x_k$, where $\sum_1^\infty |\lambda_k|^2 < \infty$, namely $\lambda_k = (x|x_k)$ (see *Theorem II.8.1*).

Theorem 1. *Let μ_n be a bounded sequence of complex numbers, $M =$ LUB $\{|\mu_k|: k = 1,2,3,\cdots\}$. There exists one and only one operator T such that*

(1) $Tx_k = \mu_k x_k$ *for all k.*

Moreover,

(2) $T\left(\sum_1^\infty \lambda_k x_k\right) = \sum_1^\infty \lambda_k \mu_k x_k$

(3) $\| T \| = M$

(4) $T^* x_k = \mu_k^* x_k$ *for all k*

(5) $T^*\left(\sum_1^\infty \lambda_k x_k\right) = \sum_1^\infty \lambda_k \mu_k^* x_k$

(6) $T^*T = TT^*$.

Proof.

(1),(2),(3): Given $x \in \mathfrak{H}$, say $x = \sum_1^\infty \lambda_k x_k$; the problem is to define Tx. Since $\sum_1^\infty |\lambda_k \mu_k|^2 \leq M^2 \sum_1^\infty |\lambda_k|^2 = M^2 \| x \|^2$ (see *Theorem II.7.2*), one can define $Tx = \sum_1^\infty \lambda_k \mu_k x_k$ (see *Theorem II.7.1*). The linearity of T results from the Lemma to *Theorem III.3.2*. Since $\| Tx \|^2 \leq M^2 \| x \|^2$, T is continuous, and $\| T \| \leq M$.

Clearly $Tx_k = \mu_k x_k$; since $\| x_k \| = 1$, $\| T \| \geq \| Tx_k \| = \| \mu_k x_k \| = |\mu_k|$ for all k, hence $\| T \| \geq M$.

(4),(5): Suppose $x = \sum_1^\infty \lambda_k x_k$ and $T^* x = \sum_1^\infty \nu_k x_k$; the problem is to show $\nu_k = \lambda_k \mu_k^*$. For all k, $\nu_k = (T^* x|x_k) = (x|Tx_k) = (x|\mu_k x_k) = \mu_k^*(x|x_k) = \mu_k^* \lambda_k$.

(6): Clearly $T^*Tx_k = |\mu_k|^2 x_k = TT^* x_k$; since the x_k are total, $T^*T = TT^*$.

Uniqueness: If S is any operator such that $Sx_k = \mu_k x_k$ for all k, then $Sx_k = Tx_k$ for all k; since the x_k are total, $S = T$. ∎

Exercises

1. Notation as in *Theorem 1*. For each $n = 1,2,3,\cdots$, let T_n be the unique operator such that $T_n x_k = \mu_k x_k$ for $k \leq n$, and $T_n x_k = \theta$ for $k > n$ (the existence of T_n is a special case of *Theorem 1*). Then, $T_n \to T$ strongly, in the sense of *Exercise IV.8.2*.

2. If $\mathcal{E}$ and $\mathcal{F}$ are normed spaces, $S: \mathcal{E} \to \mathcal{F}$ is a continuous linear mapping, $x \in \mathcal{E}$, and $x = \sum_1^\infty x_k$ in the sense of *Definition II.7.1* (generalized to normed spaces), then $Sx = \sum_1^\infty Sx_k$.

§ 4. ISOMETRIC OPERATORS

According to *Definition IV.7.3*, an operator T is *isometric* in case $\| Tx - Ty \| = \| x - y \|$ for all $x,y \in \mathfrak{H}$; by *Theorem IV.7.5*, this is equivalent to the condition $\| Tx \| = \| x \|$ for all x. In Hilbert space,

Theorem 1. *The following conditions on an operator T are equivalent:*

(a) *T is isometric*

(b) $T^*T = I$

(c) $(Tx|Ty) = (x|y)$ *for all* $x,y \in \mathfrak{H}$.

Proof.

(a) *implies* (b): By *Theorem IV.7.5*, $\| Tx \| = \| x \|$ for all x, hence $(T^*Tx|x) = (Tx|Tx) = \| Tx \|^2 = \| x \|^2 = (x|x) = (Ix|x)$.

(b) *implies* (c): $(Tx|Ty) = (T^*Tx|y) = (Ix|y) = (x|y)$.

(c) *implies* (a): $\| Tx \|^2 = (Tx|Tx) = (x|x) = \| x \|^2$. ∎

If T is isometric, it is clear from the relation $\| Tx - Ty \| = \| x - y \|$ that T is injective.

Theorem 2. *The range $T(\mathfrak{H})$ of an isometric operator T is a* **closed** *linear subspace of $\mathfrak{H}$.*

Proof.

$T(\mathfrak{H})$ is a linear subspace, by *Theorem IV.1.4*. Suppose y is adherent to $T(\mathfrak{H})$; the problem is to show that $y \in T(\mathfrak{H})$. Choose any sequence $y_n \in T(\mathfrak{H})$ such that $y_n \to y$. Say $y_n = Tx_n$. Since $\| x_m - x_n \| = \| Tx_m - Tx_n \| = \| y_m - y_n \| \to 0$, x_n is a Cauchy sequence. Since

$\mathfrak{H}$ is complete, $x_n \to x$ for suitable x. By the continuity of T, $Tx_n \to Tx$, that is, $Tx = \lim y_n = y$. ∎

Example 1. Let $\mathfrak{H}$ be classical Hilbert space, and $x_1, x_2, x_3, \cdots$ an orthonormal basis for $\mathfrak{H}$. There is a unique operator T such that $Tx_k = x_{k+1}$ for all k. Specifically, every vector x has a unique representation $x = \sum_1^\infty \lambda_k x_k$, and one defines $Tx = \sum_1^\infty \lambda_k x_{k+1}$. Clearly T is linear, and $\| Tx \|^2 = \sum_1^\infty |\lambda_k|^2 = \| x \|^2$ shows that T is isometric. T will be referred to as a *one-sided shift operator.*

The effect of T^* is as follows: $T^*x_1 = \theta$, and $T^*x_k = x_{k-1}$ for $k = 2,3,4,\cdots$. For, $(T^*x_1 | x_j) = (x_1 | Tx_j) = (x_1 | x_{j+1}) = 0 = (\theta | x_j)$ for $j = 1,2,3,\cdots$. If $k > 1$, $(T^*x_k | x_j) = (x_k | x_{j+1}) = 1$ when $j = k - 1$, and $= 0$ for all other j; thus, $(T^*x_k | x_j) = (x_{k-1} | x_j)$ for $j = 1,2,3,\cdots$.

Exercises

1. If S and T are isometric operators, so is ST.

2. Operators S and T are said to be *metrically equivalent* in case $\| Sx \| = \| Tx \|$ for all $x \in \mathfrak{H}$. Then, S and T are metrically equivalent if and only if $S^*S = T^*T$.

3. In *Example 1*, if $\mathfrak{H} = l^2$, and $x_n = e_n$ is the canonical orthonormal basis, describe the effect of T and T^* on the vector $x = (\lambda_1, \lambda_2, \lambda_3, \cdots)$.

4. Let $Tx_k = x_{k+1}$ be the one-sided shift operator (*Example 1*). The x_k are total, and for each k, Tx_k is orthogonal to x_k. Compare with *Exercises 2.2, 2.4.*

5. Conditions (a) and (c) of *Theorem 1* are equivalent, for a linear mapping in pre-Hilbert space.

6. *Theorem 2* holds for an isometric linear mapping of a Banach space.

7. If $\mathfrak{M}$ and $\mathfrak{N}$ are closed linear subspaces of $\mathfrak{H}$, T is an isometric operator, and $T(\mathfrak{M}) = \mathfrak{N}$, then $T(\mathfrak{M}^\perp) \subset \mathfrak{N}^\perp$.

8. Suppose S and T are metrically equivalent operators, in the sense of *Exercise 2.* Prove:

(i) There exists an isometric vector space isomorphism $V: T(\mathfrak{H}) \to S(\mathfrak{H})$ such that $VTx = Sx$ for all $x \in \mathfrak{H}$.

(ii) If $T(\mathfrak{H})$ is a closed linear subspace, then $S(\mathfrak{H})$ is also closed.

9. Let $S: \mathfrak{H} \to \mathfrak{H}$ be a mapping such that $(Sx|Sy) = (x|y)$ for all $x,y \in \mathfrak{H}$. Suppose, moreover, that $S(\mathfrak{H})$ is a linear subspace of $\mathfrak{H}$. Then S is an isometric operator.

10. Let $T: \mathfrak{H} \to \mathfrak{H}$ be a mapping such that $(Tx|Ty) = (y|x)$ for all $x,y \in \mathfrak{H}$. Suppose, moreover, that $T(\mathfrak{H})$ is a linear subspace of $\mathfrak{H}$. Then TT is an isometric operator.

11. Let $\mathfrak{H}$ be classical Hilbert space, x_n an orthonormal basis for $\mathfrak{H}$. An operator T is isometric if and only if Tx_n is an orthonormal sequence.

§ 5. UNITARY OPERATORS

Definition 1. *An operator T is said to be* **unitary** *in case $T^*T = TT^* = I$.*

Theorem 1. *The following conditions on an operator T are equivalent:*

(a) *T is unitary*

(b) *T^* is unitary*

(c) *T and T^* are isometric*

(d) *T is isometric and T^* is injective*

(e) *T is isometric and surjective*

(f) *T is bijective, and $T^{-1} = T^*$.*

Proof.

The equivalence of (a),(b),(c) is clear from the definitions, and the relation $T^{**} = T$.

(c) *implies* (d): Isometric operators are injective.

(d) *implies* (e): By *Theorem 4.2*, $T(\mathfrak{H})$ is a closed linear subspace of $\mathfrak{H}$. Also, T^* is injective; quoting *Theorem V.8.6*, $\mathfrak{H} = \{\theta\}^\perp = \{y: T^*y = \theta\}^\perp = \overline{T(\mathfrak{H})} = T(\mathfrak{H})$.

(e) *implies* (f): By assumption, T is bijective. Let $S = T^{-1}$; according to *Definition IV.2.2*, $ST = TS = I$. Then, $T^* = T^*I = T^*(TS) = (T^*T)S = IS = S$.

(f) *implies* (a): Clear from the definitions. ∎

Examples

1. Notation as in *Theorem 3.1*. Thus, $\mathcal{H}$ is classical Hilbert space, e_n is an orthonormal basis for $\mathcal{H}$, μ_n is a bounded sequence of complex numbers, and T is the unique operator such that $Tx_k = \mu_k x_k$ for all k. One has $T^*Tx_k = |\mu_k|^2 x_k = TT^*x_k$. Clearly, T is unitary if and only if $|\mu_k|^2 x_k = Ix_k$ for all k, in other words, $|\mu_k| = 1$ for all k.

2. Let $\mathcal{H}$, x_n be as in *Example 1*. Define $y_k = x_{2k+1}$ for $k = 0,1,2,3,\cdots$, and $y_{-k} = x_{2k}$ for $k = 1,2,3,\cdots$. Thus, y_n is an orthonormal basis indexed by the set of *all* integers. Every vector x has a unique representation $x = \sum_{-\infty}^{\infty} \lambda_k y_k$ (the sum can be interpreted as the unique vector y such that $(y|y_k) = \lambda_k$ for all k). There is a unique operator U such that $Uy_k = y_{k+1}$ for all k; specifically, one defines $U\left(\sum_{-\infty}^{\infty} \lambda_k y_k\right) = \sum_{-\infty}^{\infty} \lambda_k y_{k+1}$. It follows that $U^*y_k = y_{k-1}$ (see the calculation made in *Example 4.1*). Since $U^*Uy_k = UU^*y_k = y_k = Iy_k$, U is unitary; it will be referred to as a *two-sided shift operator*.

Exercises

1. If $\mathcal{H}$ is finite-dimensional, every isometric operator is unitary.

2. An operator T is unitary if and only if T^* is isometric and T is injective.

3. If S and T are unitary operators, so is ST.

4. If T is a unitary operator, $\mathfrak{M}$ and $\mathfrak{N}$ are closed linear subspaces, and $T(\mathfrak{M}) = \mathfrak{N}$, then $T(\mathfrak{M}^\perp) = \mathfrak{N}^\perp$.

5. An operator T is unitary if and only if $T: \mathcal{H} \to \mathcal{H}$ is a Hilbert space isomorphism. More generally, if $\mathcal{H}$ and $\mathcal{K}$ are Hilbert spaces, a mapping $T: \mathcal{H} \to \mathcal{K}$ is a Hilbert space isomorphism if and only if: T is a continuous linear mapping such that $T^*T = I$ (= identity operator in $\mathcal{H}$) and $TT^* = I$ (= identity operator in $\mathcal{K}$).

6. If $T: \mathcal{H} \to \mathcal{H}$ is a surjective mapping such that $(Tx|Ty) = (x|y)$ for all $x,y \in \mathcal{H}$, then T is a unitary operator.

7. If J and K are conjugations of $\mathcal{H}$, in the sense of *Exercise V.5.1*, then JK is a unitary operator; in particular, JJ is unitary. If J is a conjugation, and U,V are unitary operators, then UJV is a conjugation.

8. Let $\mathfrak{H}$ be classical Hilbert space, x_n an orthonormal basis for $\mathfrak{H}$, and $1',2',3',\cdots$ any rearrangement of the indices $1,2,3,\cdots$. The unique operator T such that $Tx_k = x_{k'}$ for all k, is unitary.

9. (i) If $\mathcal{S}$ is a total subset of $\mathfrak{H}$, and T is an isometric operator, then T is unitary if and only if $T(\mathcal{S})$ is a total subset of $\mathfrak{H}$.

(ii) If $\mathfrak{H}$ is classical Hilbert space, and x_n is an orthonormal basis for $\mathfrak{H}$, an operator T is unitary if and only if Tx_n is an orthonormal basis for $\mathfrak{H}$.

10. An operator S is said to be *unitarily equivalent* to the operator T in case there exists a unitary operator U such that $T = U^*SU$. Notation: $S \cong T$. Prove:

(i) $T \cong T$, for every operator T.

(ii) If $S \cong T$, then $T \cong S$.

(iii) If $R \cong S$ and $S \cong T$, then $R \cong T$. Moreover,

(iv) if $S \cong T$, then $S^* \cong T^*$.

§ 6. SELF-ADJOINT OPERATORS

Definition 1. *An operator T is said to be* **self-adjoint** *(or "Hermitian") in case $T^* = T$.*

Theorem 1. *The following conditions on an operator T are equivalent:*

(a) *T is self-adjoint*

(b) $(Tx|y) = (x|Ty)$ *for all* $x,y \in \mathfrak{H}$

(c) $(Tx|x) = (x|Tx)$ *for all* $x \in \mathfrak{H}$

(d) $(Tx|x)$ *is real, for all* $x \in \mathfrak{H}$.

Proof.

(a) *implies* (b): $(Tx|y) = (x|T^*y) = (x|Ty)$.

(c) *implies* (d): $(Tx|x)^* = (x|Tx) = (Tx|x)$.

(d) *implies* (a): $(Tx|x) = (Tx|x)^* = (x|Tx) = (T^*x|x)$. ∎

The next theorem follows easily from the properties of adjoints listed in § 1:

Theorem 2.

(1) *If S and T are self-adjoint, so is $S + T$.*

(2) *If T is self-adjoint, and α is real, αT is self-adjoint.*

(3) *If T is* **any** *operator, T^*T and $T + T^*$ are self-adjoint.*

(4) *If S and T are self-adjoint, then ST is self-adjoint if and only if $ST = TS$.*

Example 1. Notation as in *Theorem 3.1.* Thus, $Tx_k = \mu_k x_k$ and $T^*x_k = \mu_k{}^*x_k$. Clearly T is self-adjoint if and only if $\mu_k x_k = \mu_k{}^*x_k$ for all k, in other words, μ_k is *real* for all k. ▮

The decomposition of a complex number λ in the form $\lambda = \alpha + i\beta$, where α and β are real, has a generalization for operators:

Theorem 3. (Cartesian form) *If T is any operator, there exist self-adjoint operators A and B such that $T = A + iB$. Necessarily*

$$A = \frac{1}{2}(T + T^*)$$

$$B = \frac{1}{2i}(T - T^*).$$

Proof.

Define A and B by the above formulas; clearly A and B are self-adjoint, and $A + iB = T$. Suppose also $T = C + iD$, with C and D self-adjoint. Then, $T^* = C^* + i^*D^* = C - iD$, hence $T + T^* = 2C$ and $T - T^* = 2iD$. Thus, $C = A$ and $D = B$. ▮

The following result is important for the discussion of spectral theory in Chapter VIII:

Theorem 4. *If T is a self-adjoint operator,*

$$\| T \| = \text{LUB} \{ |(Tx|x)| : \| x \| \leq 1 \}.$$

Proof.

Define $\varphi(x,y) = (Tx|y)$; φ is a bounded sesquilinear form (see *Theorem V.7.1*), and $\| \varphi \| = \| T \|$. Since $\varphi(y,x) = (Ty|x) = (y|Tx) = (Tx|y)^* = [\varphi(x,y)]^*$, φ is Hermitian, hence $\| \varphi \| = \text{LUB} \{ |\varphi(x,x)| : \| x \| \leq 1 \}$ by *Theorem V.6.2.* ▮

Exercises

1. An operator T is said to be *skew-adjoint* in case $T^* = -T$. Show:
(i) T is skew-adjoint if and only if iT is self-adjoint.
(ii) For any operator T, $T - T^*$ is skew-adjoint.

2. If T is a self-adjoint operator, $\mathcal{S}$ and $\mathcal{T}$ are subsets of $\mathcal{H}$, and $T(\mathcal{S}) \subset \mathcal{T}$, then $T(\mathcal{T}^\perp) \subset \mathcal{S}^\perp$.

3. If T is self-adjoint, $Tx = \theta$ if and only if $TTx = \theta$. Give an example of an operator T such that $T \neq 0$ but $TT = 0$.

4. If T is self-adjoint, and S is any operator, then S^*TS is self-adjoint.

5. If T is self-adjoint, and S is unitarily equivalent to T, then S is self-adjoint.

6. (i) If φ is a bounded sesquilinear form on $\mathcal{H} \times \mathcal{H}$, there exist unique bounded Hermitian sesquilinear forms φ_1 and φ_2 such that $\varphi = \varphi_1 + i\varphi_2$.

(ii) If $\mathcal{U}$ is a vector space, and φ is a sesquilinear form on $\mathcal{U} \times \mathcal{U}$, there exist Hermitian sesquilinear forms φ_1, φ_2 such that $\varphi = \varphi_1 + i\varphi_2$.

(iii) Generalize (i) to bounded sesquilinear forms φ on $\mathcal{E} \times \mathcal{E}$, where $\mathcal{E}$ is a normed space.

7. An operator T is said to be *positive* in case $(Tx|x) \geq 0$ for all $x \in \mathcal{H}$. Notation: $T \geq 0$. Prove:

(i) If $T \geq 0$, then T is self-adjoint.
(ii) If $S \geq 0$ and $T \geq 0$, then $S + T \geq 0$.
(iii) If $T \geq 0$ and $\alpha \geq 0$, then $\alpha T \geq 0$.
(iv) If $T \geq 0$ and S is any operator, then $S^*TS \geq 0$.
(v) For any operator T, $T^*T \geq 0$.
(vi) If $S \geq 0$, $T \geq 0$, and $S + T = 0$, then $S = T = 0$.
(vii) If $T \geq 0$, then $|(Tx|y)|^2 \leq (Tx|x)(Ty|y)$ for all $x, y \in \mathcal{H}$.
(viii) If $T \geq 0$, then $Tx = \theta$ if and only if $(Tx|x) = 0$.

It can also be shown that:

*(ix) If $T \geq 0$, there exists a unique operator $S \geq 0$, such that $T = SS$; S is called the *positive square root* of T. Notation: $S = \sqrt{T}$.

*(x) Let $T \geq 0$, $S = \sqrt{T}$. For an operator R, $RT = TR$ if and only if $RS = SR$.

*(xi) If $S \geq 0$, $T \geq 0$, then $ST \geq 0$ if and only if $ST = TS$. In this case, $\sqrt{ST} = \sqrt{S}\sqrt{T}$.

*(xii) If T is self-adjoint, there exist positive operators A and B such that $T = A - B$ and $AB = 0$.

8. If S and T are self-adjoint operators, one writes $S \leq T$ (or $T \geq S$) in case $T - S \geq 0$ in the sense of *Exercise 7*. Prove:

(i) $T \leq T$ for every self-adjoint operator T.
(ii) If $S \leq T$ and $T \leq S$, then $S = T$.
(iii) If $R \leq S$ and $S \leq T$, then $R \leq T$.
(iv) If S and T are self-adjoint, then $S \leq T$ if and only if $(Sx|x) \leq (Tx|x)$ for all $x \in \mathfrak{C}$.
(v) If T is self-adjoint, $T \leq \| T \| I$, and $T \geq - \| T \| I$. Briefly, $- \| T \| I \leq T \leq \| T \| I$.
(vi) If $A \leq B$ and $C \leq D$, then $A + C \leq B + D$.
(vii) If $S \leq T$ and $\alpha \geq 0$, then $\alpha S \leq \alpha T$.
(viii) If $S \leq T$, then $-S \geq -T$.
(ix) If $S \leq T$, and R is any operator, then $R^*SR \leq R^*TR$.

9. An operator T is called a *contraction* in case $\| Tx \| \leq \| x \|$ for all $x \in \mathfrak{C}$. For an operator T, the following conditions are equivalent:
(a) T is a contraction
(b) $\| T \| \leq 1$
(c) $T^*T \leq I$ in the sense of *Exercise 8*
(d) $TT^* \leq I$
(e) T^* is a contraction
(f) T^*T is a contraction.

10. If T is a self-adjoint operator, there is a smallest number M, and a largest number m, such that $mI \leq T \leq MI$ in the sense of *Exercise 8*. One has $\| T \| = \max \{|m|, |M|\}$.

11. If T is an operator such that $T^*T \geq I$, then $T(\mathfrak{C})$ is a closed linear subspace of $\mathfrak{C}$. Such an operator is injective; it is not necessarily surjective.

12. Let T be a self-adjoint operator, and set $R = T + iI$. Prove:
(i) $\| Rx \|^2 = \| Tx \|^2 + \| x \|^2$.
(ii) R is injective, $\| Rx \| \geq \| x \|$.
(iii) $R(\mathfrak{C})$ is a closed linear subspace of $\mathfrak{C}$.
(iv) $R(\mathfrak{C}) = \mathfrak{C}$.
(v) R is bijective, R^{-1} is continuous, and $\| R^{-1} \| \leq 1$.
(vi) The operator $U = (T - iI)(T + iI)^{-1}$ is unitary; it is called the *Cayley transform* of T.

13. (i) If T is any operator, then the operator $R = T^*T + I$ is bijective, R^{-1} is continuous, and $\| R^{-1} \| \leq 1$.
(ii) If $T \geq 0$, then the operator $S = T + I$ is bijective, S^{-1} is continuous, and $\| S^{-1} \| \leq 1$.

14. Let $Uy_k = y_{k+1}$ be the two-sided shift operator, and P the projection whose range is the closed linear subspace generated by $y_0, y_1, y_2, \cdots$. Then, $P \geq 0$, $UPU^* \leq P$, but $UPU^* \neq P$.

§ 7. PROJECTION OPERATORS

Definition 1. *An operator T is called a* **projection** *in case $T^* = T = TT$.*

Thus, a projection operator is a self-adjoint operator which equals its "square."

Example 1. Let $\mathfrak{N}$ be a closed linear subspace of $\mathfrak{H}$, and $T = P_{\mathfrak{N}}$ the projection of $\mathfrak{H}$ on $\mathfrak{N}$. As shown in *Theorem III.8.1*, T is linear and $\| Tx \| \leq \| x \|$, hence T is an operator; also $T(Tx) = Tx$ and $(Tx \mid y) = (x \mid Ty)$, thus T is a projection operator in the sense of *Definition 1*. In fact, every projection operator can be obtained in this way from a suitable closed linear subspace:

Theorem 1. *If T is any projection operator, there is one and only one closed linear subspace $\mathfrak{N}$ such that $T = P_{\mathfrak{N}}$. Specifically, $\mathfrak{N}$ is the range of T, and $\mathfrak{N}^\perp$ is the null space of T.*

Proof.

Let $\mathfrak{N} = \{y \in \mathfrak{H} : Ty = y\}$; since $\mathfrak{N}$ is the null space of the operator $T - I$, it is a closed linear subspace of $\mathfrak{H}$.

The range of T is $\mathfrak{N}$. For, if $y \in \mathfrak{N}$, then $y = Ty \in T(\mathfrak{H})$; if conversely $y \in T(\mathfrak{H})$, say $y = Tx$, then $Ty = T(Tx) = Tx = y$ shows that $y \in \mathfrak{N}$. Thus, $\mathfrak{N} = T(\mathfrak{H})$.

By *Theorem V.8.6*, the null space of T is $[T^*(\mathfrak{H})]^\perp = [T(\mathfrak{H})]^\perp = \mathfrak{N}^\perp$.

One has $T = P_{\mathfrak{N}}$. For, given any $x \in \mathfrak{H}$, write $x = y + z$ with $y \in \mathfrak{N}$ and $z \in \mathfrak{N}^\perp$; quoting *Theorem III.8.1*, $P_{\mathfrak{N}}x = y = y + \theta = Ty + Tz = T(y + z) = Tx$.

Uniqueness: if $P_{\mathfrak{N}} = P_{\mathfrak{M}}$, then $\mathfrak{N} = P_{\mathfrak{N}}(\mathfrak{H}) = P_{\mathfrak{M}}(\mathfrak{H}) = \mathfrak{M}$. ∎

Example 2. Let y be a fixed vector of $\mathfrak{H}$, $\| y \| = 1$, and let $\mathfrak{N} = \{\lambda y : \lambda \text{ scalar}\}$ be the one-dimensional linear subspace generated by y. The operator T defined by $Tx = (x \mid y)y$ is the projection of $\mathfrak{H}$ on $\mathfrak{N}$. For, if $x \in \mathfrak{H}$, one can write $x = \lambda y + z$ with $z \perp \mathfrak{N}$ and λ a suitable scalar, and one has $Tx = (\lambda y + z \mid y)y = [\lambda(y \mid y) + (z \mid y)]y = [\lambda + 0]y = \lambda y = P_{\mathfrak{N}}x$.

Example 3. Notation as in *Theorem 3.1*. Let P_k be the projection whose range is the one-dimensional linear subspace generated by x_k; by *Example 2*, $P_k x = (x|x_k)x_k$. Define $T_n = \sum_1^n \mu_k P_k$, for $n = 1,2,3,\cdots$.

If $x \in \mathfrak{H}$, one has $x = \sum_1^\infty (x|x_k)x_k$, hence $Tx = \sum_1^\infty (x|x_k)\mu_k x_k = \sum_1^\infty \mu_k(P_k x) = \lim_n \sum_1^n \mu_k(P_k x) = \lim_n T_n x$. Thus, $T_n x \to Tx$, for each x; that is, $T_n \to T$ strongly, in the sense of *Exercise IV.8.2*.

Theorem 2. *Let $\mathfrak{M}$ and $\mathfrak{N}$ be closed linear subspaces of $\mathfrak{H}$, P the projection with range $\mathfrak{M}$, and Q the projection with range $\mathfrak{N}$. Then, $\mathfrak{M} \perp \mathfrak{N}$ if and only if $PQ = 0$.*

Proof.

If: Suppose $PQ = 0$. If $x \in \mathfrak{M}$ and $y \in \mathfrak{N}$, then $(x|y) = (Px|Qy) = (x|P^*Qy) = (x|PQy) = (x|\theta) = 0$.

Only if: Suppose $\mathfrak{M} \perp \mathfrak{N}$. If $x \in \mathfrak{H}$, then $Qx \in \mathfrak{N}$, hence $Qx \in \mathfrak{M}^\perp$, hence $P(Qx) = \theta$. ∎

Example 4. If $P_1,\cdots,P_n$ are projections, and $P_jP_k = 0$ whenever $j \neq k$, then $P = \sum_1^n P_k$ is a projection. For, $P^* = \sum_1^n P_k{}^* = \sum_1^n P_k = P$, and $PP = \sum_{j,k} P_jP_k = \sum_k P_kP_k = \sum_k P_k = P$. See also *Exercise 6*.

Exercises

1. An operator T is a projection if and only if $T = T^*T$.

2. If P is a projection, and Q is unitarily equivalent to P, then Q is a projection.

3. If P is a projection, then $P \geq 0$ in the sense of *Exercise 6.7*. If P is a non-zero projection, $\| P \| = 1$.

4. If P, Q are projections, and $PQ = QP$, then PQ is a projection. In this case, if $\mathfrak{M}$ is the range of P and $\mathfrak{N}$ is the range of Q, then the range of PQ is the intersection of $\mathfrak{M}$ and $\mathfrak{N}$ (see *Exercise I.5.4*).

5. If P is a projection, with range $\mathfrak{N}$, then $I - P$ is a projection, with range $\mathfrak{N}^\perp$.

6. Let P and Q be projections, with ranges $\mathfrak{M}$ and $\mathfrak{N}$, respectively. If $\mathfrak{M} \perp \mathfrak{N}$, then $P + Q$ is the projection with range $\mathfrak{M} \oplus \mathfrak{N}$ (see *Definition III.2.5*).

7. If P, Q are projections, and $PQ = QP$, then $P + Q - PQ$ is a projection. What is its range?

8. If P, Q, and $P + Q$ are projections, necessarily $PQ = 0$.

9. Let P and Q be projections, with ranges $\mathfrak{M}$ and $\mathfrak{N}$, respectively. The following conditions are equivalent:

(a) $P - Q$ is a projection.
(b) $P \geq Q$ in the sense of *Exercise 6.8*.
(c) $\| Px \| \geq \| Qx \|$ for all $x \in \mathfrak{H}$.
(d) $\mathfrak{M} \supset \mathfrak{N}$.

In this case, the range of $P - Q$ is the set of all vectors common to $\mathfrak{M}$ and $\mathfrak{N}^{\perp}$.

10. If P, Q are non-zero projections, and $PQ = 0$, then $\| P + Q \| < \| P \| + \| Q \|$.

11. If T is an isometric operator, then TT^* is a projection.

12. An operator T is said to be *partially isometric* in case T^*T is a projection. Suppose T is partially isometric, say $T^*T = P$. Prove:

(i) $TP = T$.

(ii) Set $Q = TT^*$. Then Q is a projection, and $QT = T$. Thus, T^* is partially isometric.

(iii) T maps $P(\mathfrak{H})$ isometrically onto $Q(\mathfrak{H})$, $[P(\mathfrak{H})]^{\perp}$ is the null space of T, and $[Q(\mathfrak{H})]^{\perp}$ is the null space of T^*.

(iv) For an operator S, $TS = 0$ if and only if $PS = 0$. Similarly, $ST = 0$ if and only if $SQ = 0$.

(v) If U is unitary, and R is any projection, then UR and RU are partially isometric. If S is isometric and T is partially isometric, then ST is partially isometric. If S and T are partially isometric, ST is not necessarily partially isometric.

13. Suppose $\mathfrak{H}$ is finite-dimensional.

(i) If T is partially isometric, and $TT^* \leq T^*T$, necessarily $TT^* = T^*T$.

(ii) If U is unitary, P is a projection, and $UPU^* \leq P$, necessarily $UP = PU$. Compare with *Exercise 6.14*.

14. If T is an operator with one-dimensional range, there exist vectors $y,z \in \mathfrak{H}$ such that $Tx = (x|z)y$ for all $x \in \mathfrak{H}$. Moreover,

(i) $T^*x = (x|y)z$ for all $x \in \mathfrak{H}$
(ii) $TT = \mu T$, μ a suitable scalar
(iii) $\| T \| = \| y \| \, \| z \|$

(iv) $T^* = T$ if and only if $y = \alpha z$ for some real number α. Equivalently, $(y|z) = \pm \| y \| \| z \|$.

(v) What are necessary and sufficient conditions on y and z, in order that T be a projection?

§ 8. NORMAL OPERATORS

Definition 1. *An operator T is said to be* **normal** *in case $T^*T = TT^*$.*

Every unitary operator is normal; so is every self-adjoint operator. An isometric operator is normal if and only if it is unitary.

Theorem 1. *The following conditions on an operator T are equivalent:*

(a) *T is normal*

(b) *T^* is normal*

(c) *$\| T^*x \| = \| Tx \|$ for all $x \in \mathfrak{H}$.*

Proof.

The equivalence of (a) and (b) results from $T^{**} = T$.

(a) *implies* (c): $\| T^*x \|^2 = (T^*x | T^*x) = (TT^*x | x) = (T^*Tx | x) = (Tx | Tx) = \| Tx \|^2$.

(c) *implies* (a): $(T^*Tx | x) = \| Tx \|^2 = \| T^*x \|^2 = (TT^*x | x)$ for all x. ∎

Theorem 2. *Let $T = A + iB$ be the Cartesian form of the operator T. Then, T is normal if and only if $AB = BA$.*

Proof.

Only if: Assuming T normal, it is clear from the formulas in *Theorem 6.3* that $AB = BA$.

If: Suppose $AB = BA$; since $T^* = A - iB$, clearly $T^*T = TT^* = AA + BB$. ∎

Exercises

1. If T is normal, and S is unitarily equivalent to T, then S is normal.

2. T is normal if and only if T and T^* are metrically equivalent.

3. (i) If T is normal, and $TT = 0$, then $T = 0$.
   (ii) Give an example of an operator $T \neq 0$ such that $TT = 0$.

4. If T is normal, and μ is a scalar, then $\| T^*x - \mu^*x \| = \| Tx - \mu x \|$ for all $x \in \mathcal{H}$.

5. If T is normal, then T and T^* have the same null space, and $\overline{T(\mathcal{H})} = \overline{T^*(\mathcal{H})}$. If $\mathfrak{N}$ is the null space of T, then $\mathfrak{N}^\perp = \overline{T(\mathcal{H})}$.

6. If T is normal, there exists a unitary operator U such that $T^* = UT$.

7. More generally, if S and T are metrically equivalent normal operators, there exists a unitary operator U such that $S = UT$.

8. Let T be a normal operator. Prove:
(i) T is surjective if and only if T^* is surjective.
(ii) T is injective if and only if T^* is injective.
(iii) T is bijective if and only if T^* is bijective.
(iv) If T is bijective, then T^*T^{-1} is unitary.

9. An operator T is said to be *invertible* if there exists an operator S such that $ST = TS = I$; in this case, it is clear that T is bijective, and $T^{-1} = S$. Prove:
(i) If S and T are invertible operators, then ST is invertible, and $(ST)^{-1} = T^{-1}S^{-1}$.
(ii) If T is an invertible operator, then so is T^*, and $(T^*)^{-1} = (T^{-1})^*$.
(iii) If T is an invertible normal (resp. self-adjoint) operator, then T^{-1} is normal (resp. self-adjoint).
(iv) If T is an invertible operator, then T^*T is invertible. The converse fails.
(v) If T is normal, then T is invertible if and only if T^*T is invertible.
*(vi) It can be shown that every bijective operator T is invertible (that is, T^{-1} is necessarily continuous). A proof of this is sketched in *Exercises 11* through *13*.

*10. If T is a bijective normal operator, then T^{-1} is continuous (see also *Exercise 13*). A brief proof results from *Exercise 8* and *Exercise V.8.6*.

11. An operator T is said to be *bounded below* in case there exists a constant $N > 0$ such that $\| Tx \| \geq N \| x \|$ for all $x \in \mathcal{H}$. Equivalently, there exists a constant $M > 0$ such that $T^*T \geq MI$ in the sense of *Exercise 6.8*. Prove:

(i) Every invertible operator is bounded below.

(ii) If an operator T is bounded below, then T is injective, and $T(\mathcal{H})$ is a closed linear subspace of $\mathcal{H}$. Hence, the mapping $T: \mathcal{H} \to T(\mathcal{H})$ has a continuous inverse.

(iii) An operator T is invertible if and only if, T is bounded below and $T(\mathcal{H})$ is dense in $\mathcal{H}$.

(iv) An operator T is invertible if and only if, T is bounded below and T^* is injective.

(v) An operator T is invertible if and only if, T is injective and T^* is bounded below.

(vi) An operator T is invertible if and only if both T and T^* are bounded below.

*12. If T is a surjective operator, then T^* is bounded below.

*13. If T is a bijective operator, then T^{-1} is continuous.

14. Let T be an operator.

(i) If $T^*T \geq I$ in the sense of *Exercise 6.8*, T is not necessarily invertible.

(ii) If $T^*T \geq I$ and $TT^* \geq I$, then T is invertible.

(iii) If T is normal and $T^*T \geq I$, then T is invertible.

(iv) If T is normal and bounded below, then T is invertible.

15. An operator S is said to be *similar* to the operator T in case there exists an invertible operator A such that $T = A^{-1}SA$. Notation: $S \sim T$. Then:

(i) $T \sim T$, for every operator T.

(ii) If $S \sim T$, then $T \sim S$.

(iii) If $R \sim S$ and $S \sim T$, then $R \sim T$. Moreover,

(iv) If $S \sim T$, then $S^* \sim T^*$.

(v) If $S \cong T$ in the sense of *Exercise 5.10*, then $S \sim T$.

*(vi) If S and T are normal operators, it can be shown that $S \cong T$ if and only if $S \sim T$.

§ 9. INVARIANT AND REDUCING SUBSPACES

Definition 1. *A closed linear subspace $\mathcal{M}$ is said to be* **invariant** *under the operator T in case $T(\mathcal{M}) \subset \mathcal{M}$.*

Examples

1. $\{\theta\}$ and $\mathfrak{H}$ are invariant under every operator T.

2. The null space of an operator T is invariant under T; for, $Tx = \theta$ implies $T(Tx) = \theta$.

3. More generally, if S and T are operators such that $ST = TS$, the null space of T is invariant under S. For, if $Tx = \theta$, then $T(Sx) = (TS)x = (ST)x = S(Tx) = S\theta = \theta$.

Theorem 1. *Let* $\mathfrak{s}$ *and* $\mathfrak{z}$ *be subsets of* $\mathfrak{H}$, $\mathfrak{M} = \mathfrak{s}^{\perp\perp}$, $\mathfrak{N} = \mathfrak{z}^{\perp\perp}$. *If* T *is an operator such that* $T(\mathfrak{s}) \subset \mathfrak{z}$, *then* $T(\mathfrak{M}) \subset \mathfrak{N}$.

Proof.

Incidentally, $\mathfrak{M}$ (resp. $\mathfrak{N}$) is the smallest closed linear subspace of $\mathfrak{H}$ which contains $\mathfrak{s}$ (resp. $\mathfrak{z}$). By *Theorem V.8.4*, $T^*(\mathfrak{z}^{\perp}) \subset \mathfrak{s}^{\perp}$, hence $T^{**}(\mathfrak{s}^{\perp\perp}) \subset \mathfrak{z}^{\perp\perp}$. ∎

Examples

4. Let $\mathfrak{s}$ be a subset of $\mathfrak{H}$, $\mathfrak{N}$ the smallest closed linear subspace containing $\mathfrak{s}$. If T is an operator such that $T(\mathfrak{s}) \subset \mathfrak{s}$, then $\mathfrak{N}$ is invariant under T; this follows at once from *Theorem 1*.

5. In particular, if $\mathfrak{s}$ is a linear subspace of $\mathfrak{H}$, $\mathfrak{N} = \bar{\mathfrak{s}}$, and T is an operator such that $T(\mathfrak{s}) \subset \mathfrak{s}$, then $\mathfrak{N}$ is invariant under T. See also *Exercise 1*.

6. For any operator T, $\overline{T(\mathfrak{H})}$ is invariant under T; this follows from the relation $T[T(\mathfrak{H})] \subset T(\mathfrak{H})$, and *Example 5*.

Theorem 2. *Let* T *be an operator,* $\mathfrak{M}$ *and* $\mathfrak{N}$ *closed linear subspaces of* $\mathfrak{H}$. *The following conditions are equivalent:*

(a) $T(\mathfrak{M}) \subset \mathfrak{N}$

(b) $T^*(\mathfrak{N}^{\perp}) \subset \mathfrak{M}^{\perp}$

(c) $P_{\mathfrak{N}} T P_{\mathfrak{M}} = T P_{\mathfrak{M}}$.

Proof.

(a) and (b) are equivalent by *Theorem V.8.5*.

(a) *implies* (c): Given any $x \in \mathcal{H}$, one has $P_{\mathfrak{M}}x \in \mathfrak{M}$ (see *Example 7.1*); by condition (a), $TP_{\mathfrak{M}}x \in \mathfrak{N}$, hence $P_{\mathfrak{N}}TP_{\mathfrak{M}}x = TP_{\mathfrak{M}}x$.

(c) *implies* (a): If $y \in \mathfrak{M}$, then $Ty = TP_{\mathfrak{M}}y = P_{\mathfrak{N}}TP_{\mathfrak{M}}y = P_{\mathfrak{N}}Ty$, hence $Ty \in \mathfrak{N}$. ∎

In particular, $\mathfrak{N}$ is invariant under T if and only if $\mathfrak{N}^{\perp}$ is invariant under T^*.

Corollary. *A closed linear subspace $\mathfrak{N}$ is invariant under the operator T if and only if $P_{\mathfrak{N}}TP_{\mathfrak{N}} = TP_{\mathfrak{N}}$.*

Definition 2. *Let $\mathfrak{N}$ be a closed linear subspace invariant under the operator T. The* **restriction** *of T to $\mathfrak{N}$ is the mapping $T/\mathfrak{N}\colon \mathfrak{N} \to \mathfrak{N}$ defined by $(T/\mathfrak{N})y = Ty$.*

Thus, $T/\mathfrak{N}$ is the mapping $y \to Ty$, restricted to $\mathfrak{N}$. The following is elementary:

Theorem 3. *Let $\mathfrak{N}$ be a closed linear subspace of $\mathcal{H}$.*

(1) *If $\mathfrak{N}$ is invariant under the operator T, then $T/\mathfrak{N}$ is an operator in the Hilbert space $\mathfrak{N}$, and $\| T/\mathfrak{N} \| \leq \| T \|$.*

(2) *If $\mathfrak{N}$ is invariant under the operators S and T, then it is invariant under $S + T$, ST, and λT; moreover,*

$$(S + T)/\mathfrak{N} = (S/\mathfrak{N}) + (T/\mathfrak{N})$$

$$(ST)/\mathfrak{N} = (S/\mathfrak{N})(T/\mathfrak{N})$$

$$(\lambda T)/\mathfrak{N} = \lambda(T/\mathfrak{N}).$$

Definition 3. *A closed linear subspace $\mathfrak{N}$ is said to* **reduce** *the operator T in case both $\mathfrak{N}$ and $\mathfrak{N}^{\perp}$ are invariant under T.*

Theorem 4. *If T is an operator, and $\mathfrak{N}$ is a closed linear subspace, the following conditions are equivalent:*

(a) *$\mathfrak{N}$ reduces T*

(b) *$\mathfrak{N}^{\perp}$ reduces T*

(c) *$\mathfrak{N}$ reduces T^**

(d) *$\mathfrak{N}$ is invariant under both T and T^**

(e) $TP_{\mathfrak{N}} = P_{\mathfrak{N}}T$.

Proof.

The equivalence of (a)–(d) is immediate from *Theorem 2* and the definitions. Let $P = P_{\mathfrak{M}}$.

(d) *implies* (e): By the Corollary of *Theorem 2*, $PTP = TP$ and $PT^*P = T^*P$. Then, $PT = P^*T = (T^*P)^* = (PT^*P)^* = PTP = TP$.

(e) *implies* (d): Assuming $PT = TP$, one has $PTP = TPP = TP$, hence $\mathfrak{M}$ is invariant under T by the cited corollary. Also, $PT^* = (TP)^* = (PT)^* = T^*P$, hence $\mathfrak{M}$ is invariant under T^*. ∎

Examples

7. If T is a normal operator, $\mathfrak{M} = \overline{T(\mathfrak{H})} = \overline{T^*(\mathfrak{H})}$ reduces T; for, $\mathfrak{M}^\perp$ is the null space of both T and T^*, hence is invariant under them both.

8. If $Tx_k = x_{k+1}$ is the one-sided shift operator (*Example 4.1*), the only closed linear subspaces which reduce T are $\{\theta\}$ and $\mathfrak{H}$. For, suppose $\mathfrak{M} \neq \{\theta\}$ is a closed linear subspace which reduces T. For a non-zero vector $y = \sum_1^\infty \lambda_k x_k$, define the *index* of y to be the smallest subscript k such that $\lambda_k \neq 0$. Let m be the smallest index of any non-zero vector in $\mathfrak{M}$, and choose any non-zero $y \in \mathfrak{M}$ with index m; clearly $y = \sum_m^\infty \lambda_k x_k$. Necessarily $m = 1$; otherwise, $\mathfrak{M}$ would contain the non-zero vector $T^*y = \sum_m^\infty \lambda_k x_{k-1} = \sum_{m-1}^\infty \lambda_{k+1} x_k$, contrary to the minimality of m. We may suppose $\lambda_1 = 1$, thus $y = x_1 + \sum_2^\infty \lambda_k x_k$. One has $T^*y = T^*x_1 + \sum_2^\infty \lambda_k T^*x_k = \theta + \sum_2^\infty \lambda_k x_{k-1}$, hence $TT^*y = \sum_2^\infty \lambda_k T x_{k-1} = \sum_2^\infty \lambda_k x_k = y - x_1$. Since $x_1 = y - TT^*y$, clearly $x_1 \in \mathfrak{M}$. Then, $\mathfrak{M}$ also contains $Tx_1 = x_2$, $Tx_2 = x_3$, and so on. Evidently $\mathfrak{M}$ is total, $\mathfrak{M}^\perp = \{\theta\}$, $\mathfrak{M} = \mathfrak{H}$. ∎

Theorem 5. *If $\mathfrak{M}$ reduces T, then $(T/\mathfrak{M})^* = T^*/\mathfrak{M}$.*

Proof.

Let $R = T/\mathfrak{M}$ and $S = T^*/\mathfrak{M}$; these are both defined, by condition (d) of *Theorem 4*. For all $x, y \in \mathfrak{M}$, $(R^*x|y) = (x|Ry) = (x|Ty) = (T^*x|y) = (Sx|y)$; since R^* and S are operators in $\mathfrak{M}$, $R^* = S$. ∎

Corollary 1. *If $\mathfrak{M}$ reduces T, and T is normal (resp. unitary), then $T/\mathfrak{M}$ is normal (resp. unitary).*

Proof.

By *Theorems 3* and *5*, $(T^*T)/\mathfrak{N} = (T^*/\mathfrak{N})(T/\mathfrak{N}) = (T/\mathfrak{N})^*(T/\mathfrak{N})$, and $I/\mathfrak{N}$ is the identity operator of $\mathfrak{N}$. ∎

Corollary 2. *If T is self-adjoint, and $\mathfrak{N}$ is invariant under T, then $T/\mathfrak{N}$ is self-adjoint.*

Proof.

Since $T^* = T$, $\mathfrak{N}$ reduces T by *Theorem 4*. Then, $(T/\mathfrak{N})^* = T^*/\mathfrak{N} = T/\mathfrak{N}$. ∎

Example 9. Suppose T is normal, and $\mathfrak{N}$ is invariant under T. Then, $T/\mathfrak{N}$ is normal if and only if $\mathfrak{N}$ reduces T.

For, if $\mathfrak{N}$ reduces T, $T/\mathfrak{N}$ is normal by Corollary 1 above.

Conversely, suppose $R = T/\mathfrak{N}$ is normal, that is, $R^*R = RR^*$, where R^* is an operator in $\mathfrak{N}$. Given $x \in \mathfrak{N}$, the problem is to show that $T^*x \in \mathfrak{N}$. For any $y \in \mathfrak{N}$, $(T^*x|y) = (x|Ty) = (x|Ry) = (R^*x|y)$; thus, $(T^*x - R^*x|y) = 0$ for all $y \in \mathfrak{N}$, $T^*x - R^*x \in \mathfrak{N}^\perp$. Since $R^*x \in \mathfrak{N}$, one has

$$\| T^*x \|^2 = \| (T^*x - R^*x) + R^*x \|^2 = \| T^*x - R^*x \|^2 + \| R^*x \|^2$$

by *Theorem II.6.2*. Since T and R are normal, $\| T^*x \| = \| Tx \|$ and $\| R^*x \| = \| Rx \| = \| Tx \|$, hence the above equation reduces to $\| Tx \|^2 = \| T^*x - R^*x \|^2 + \| Tx \|^2$. Thus, $T^*x - R^*x = \theta$, $T^*x = R^*x \in \mathfrak{N}$.

Exercises

1. Let $\mathfrak{s}$ and $\mathfrak{z}$ be subsets of the metric spaces $\mathfrak{X}$ and $\mathfrak{Y}$, respectively, and let $T: \mathfrak{X} \to \mathfrak{Y}$ be a continuous mapping such that $T(\mathfrak{s}) \subset \mathfrak{z}$. Then, $T(\bar{\mathfrak{s}}) \subset \bar{\mathfrak{z}}$.

2. If T is isometric, and $\mathfrak{N}$ is invariant under T, then $T/\mathfrak{N}$ is isometric.

3. If T has Cartesian form $T = A + iB$, and $\mathfrak{N}$ reduces T, then $T/\mathfrak{N}$ has Cartesian form $(A/\mathfrak{N}) + i(B/\mathfrak{N})$.

4. If $\mathfrak{N}$ is invariant under T, and P is the projection with range $\mathfrak{N}$, then $(T/\mathfrak{N})^* = (PT^*P)/\mathfrak{N}$.

5. Let $Uy_k = y_{k+1}$ be the two-sided shift operator (*Example 5.2*), $\mathfrak{N}$ the smallest closed linear subspace containing $y_1, y_2, y_3, \cdots$. Then $\mathfrak{N}$ is invariant under U, but does not reduce U.

6. If T is unitary, and $\mathfrak{N}$ is invariant under T, then $T/\mathfrak{N}$ is unitary if and only if $\mathfrak{N}$ reduces T.

7. If T is normal, $\mathfrak{N}$ is invariant under T, and $T/\mathfrak{N}$ is a scalar operator, then $\mathfrak{N}$ reduces T.

8. Suppose T is isometric, and $\mathfrak{N}$ is invariant under T. Then:
(i) If $T(\mathfrak{N}) = \mathfrak{N}$, $\mathfrak{N}$ reduces T.
(ii) If $T/\mathfrak{N}$ is normal, then $\mathfrak{N}$ reduces T and $T/\mathfrak{N}$ is unitary.

9. Call an operator T *hyponormal* in case $\| T^*x \| \leq \| Tx \|$ for all $x \in \mathfrak{H}$. Prove:
(i) T is hyponormal if and only if $TT^* \leq T^*T$ in the sense of *Exercise 6.8.*
(ii) If T is isometric, then T is hyponormal.
(iii) If T is hyponormal, $\mathfrak{N}$ is invariant under T, and $T/\mathfrak{N}$ is normal, then $\mathfrak{N}$ reduces T.

10. Let $\mathfrak{N}$ be a closed linear subspace invariant under the operator T. Prove:
(i) If $R = T/\mathfrak{N}$, $\| R^*x \| \leq \| T^*x \|$ for all $x \in \mathfrak{N}$.
(ii) If T is hyponormal, so is $T/\mathfrak{N}$.

11. An operator T is hyponormal if and only if there exists an operator V, $\| V \| \leq 1$, such that $T^* = VT$.

12. If T is an invertible hyponormal operator, then T^{-1} is hyponormal.

13. Let $Uy_k = y_{k+1}$ be the two-sided shift operator. Given a bounded sequence of complex numbers μ_n $(n = 0, \pm 1, \pm 2, \cdots)$, let R be the unique operator such that $Ry_k = \mu_k y_k$ for all k. Define $T = RU$. Then, T is hyponormal if and only if $|\mu_k| \leq |\mu_{k+1}|$ for all k. T is normal if and only if all the μ_k lie on the circle in the complex plane with center 0 and radius $|\mu_0|$; equivalently, T is a scalar multiple of a unitary operator.

14. *(i) It can be shown that every hyponormal operator on a finite-dimensional Hilbert space is normal.
(ii) If T is hyponormal, and $\mathfrak{N}$ is a finite-dimensional subspace invariant under T, then $\mathfrak{N}$ reduces T.

15. Let $T = A + iB$ be the Cartesian form of the operator T. Let $AB = C + iD$ be the Cartesian form of AB. Then, T is hyponormal if and only if $D \leq 0$.

16. Let $Tx_k = x_{k+1}$ be the one-sided shift operator. The only finite-dimensional subspace $\mathfrak{N}$ invariant under T is $\mathfrak{N} = \{\theta\}$.

17. If $Tx_k = x_{k+1}$ is the one-sided shift operator, there exists a linear subspace $\mathfrak{S}$, distinct from $\{\theta\}$ and $\mathfrak{H}$, such that $T(\mathfrak{S}) \subset \mathfrak{S}$ and $T^*(\mathfrak{S}) \subset \mathfrak{S}$. Such a linear subspace is necessarily dense in $\mathfrak{H}$. Among all such $\mathfrak{S}$, there is a smallest, that is, one which is contained in every other.

18. Suppose $\mathfrak{N}$ reduces the operator T. Then, T is an isometric operator if and only if both $T/\mathfrak{N}$ and $T/\mathfrak{N}^\perp$ are isometric. The word "isometric" can be replaced by "unitary," "normal," "self-adjoint," "projection," "hyponormal," "positive," "invertible."

**19. Given an arbitrary operator T. It is not known whether there exists a closed linear subspace $\mathfrak{N}$ invariant under T, other than $\{\theta\}$ and $\mathfrak{H}$.

Chapter

Proper Values VII

§ 1. *Proper vectors, proper values*
§ 2. *Proper subspaces*
§ 3. *Approximate proper values*

§ 1. PROPER VECTORS, PROPER VALUES

Definition 1. *A vector $x \in \mathfrak{X}$ is said to be a* **proper vector** *for the operator T in case:* (1) $x \neq \theta$, *and* (2) $Tx = \mu x$ *for a suitable scalar μ.*

If also $Tx = \nu x$, then $(\mu - \nu)x = \theta$, hence $\mu = \nu$. Thus, a proper vector x determines uniquely the associated scalar μ.

Definition 2. *A scalar μ is said to be a* **proper value** *for the operator T in case there exists a vector $x \neq \theta$ such that $Tx = \mu x$.*

Thus, μ is a proper value for T if and only if the null space of $T - \mu I$ is $\neq \{\theta\}$.

Examples

1. If $T = \mu I$, every non-zero vector of $\mathfrak{X}$ is a proper vector for T; the only proper value is μ.

2. Let P be the projection with range $\mathfrak{N}$. Assume $\mathfrak{N} \neq \{\theta\}$, $\mathfrak{N} \neq \mathfrak{X}$. If y is any non-zero vector of $\mathfrak{N}$, $Py = y = 1y$ shows that y is a proper vector, with associated proper value 1. Similarly, every non-zero $z \in \mathfrak{N}^{\perp}$ is a proper vector, with associated proper value 0. The only proper values of P are 0 and 1. Every proper vector of P belongs either to $\mathfrak{N}$ or to $\mathfrak{N}^{\perp}$.

3. Let x be a proper vector for T, say $Tx = \mu x$. If $\mathfrak{N}$ is the null space of $T - \mu I$, every non-zero vector of $\mathfrak{N}$ is also a proper vector. In particular, $\| x \|^{-1}x$ is a proper vector of norm 1.

4. Given a Hilbert space $\mathcal{H}$ of finite dimension n, and scalars $\mu_1, \cdots, \mu_n$ (not necessarily distinct). If $x_1, \cdots, x_n$ is any basis of $\mathcal{H}$, there is a unique linear mapping $T: \mathcal{H} \to \mathcal{H}$ such that $Tx_k = \mu_k x_k$ for all k (see *Theorem IV.1.3*). See *Exercise 1*.

Theorem 1. *Let μ be a proper value for T. Then:*

(1) $|\mu| \leq \| T \|$.

(2) *If T is self-adjoint, μ is real.*

(3) *If T is isometric, $|\mu| = 1$.*

Proof.

Let x be a non-zero vector such that $Tx = \mu x$; we may suppose $\| x \| = 1$.

(1): $|\mu| = \| \mu x \| = \| Tx \| \leq \| T \|$.

(2): $\mu = \mu(x|x) = (\mu x|x) = (Tx|x)$ is real by *Theorem VI.6.1*.

(3): $|\mu| = \| Tx \| = \| x \| = 1$. ∎

Example 5. Let $\mathcal{H}$ be classical Hilbert space, x_n an orthonormal basis for $\mathcal{H}$, μ_n a bounded sequence of scalars, and T the unique operator such that $Tx_k = \mu_k x_k$ for all k (see *Theorem VI.3.1*). Thus, the μ_k are proper values for T. In fact, these are the *only* proper values. For, suppose $Tx = \mu x$, $x \neq \theta$. Say $x = \sum_1^\infty \lambda_k x_k$; then $Tx = \sum_1^\infty \lambda_k \mu_k x_k$, hence $\lambda_k \mu_k = (Tx|x_k) = (\mu x|x_k) = \mu(x|x_k) = \mu\lambda_k$, $\lambda_k(\mu - \mu_k) = 0$ for all k. Since $x \neq 0$, there is an index m such that $\lambda_m \neq 0$; necessarily $\mu - \mu_m = 0$, thus $\mu = \mu_m$. Similarly, the $\mu_k{}^*$ are the proper values for T^*; this is a symptom of normality:

Theorem 2. *If T is a normal operator, x is a vector, and μ is a scalar, then $Tx = \mu x$ if and only if $T^*x = \mu^* x$. In particular:*

(1) *x is a proper vector for T if and only if it is a proper vector for T^*.*

(2) *μ is a proper value for T if and only if μ^* is a proper value for T^*.*

Proof.

By assumption, $T^*T = TT^*$. Since $(T - \mu I)^* = T^* - \mu^* I$, clearly $T - \mu I$ is normal, hence $\| (T - \mu I)x \| = \| (T - \mu I)^* x \|$ by *Theorem VI.8.1*. ∎

Examples

6. The one-sided shift operator $Tx_k = x_{k+1}$ (see *Example VI.4.1*) has *no* proper values. For, assuming $Tx = \mu x$, let us show that necessarily $x = \theta$. If $\mu = 0$, $Tx = \mu x = 0x = \theta$; since T is injective, $x = \theta$. Suppose $\mu \neq 0$. If $x = \sum_1^\infty \lambda_k x_k$, then $Tx = \sum_1^\infty \lambda_k x_{k+1} = \sum_2^\infty \lambda_{k-1} x_k$, hence $0 = (Tx|x_1) = (\mu x|x_1) = \mu\lambda_1$; thus $\lambda_1 = 0$. Also, for $k > 1$, $\lambda_{k-1} = (Tx|x_k) = (\mu x|x_k) = \mu\lambda_k$, hence $\lambda_k = \mu^{-1}\lambda_{k-1}$; since $\lambda_1 = 0$, this implies $\lambda_k = 0$ for all k.

7. If μ is a proper value for the operator S, it does not follow that μ^* is a proper value for S^*. For instance, let $S = T^*$, where T is the one-sided shift operator. As shown in *Example 6*, S^* has no proper values, but $Sx_1 = T^*x_1 = \theta = 0x_1$ shows that 0 is a proper value for S.

8. There exist normal operators having no proper values. For example, let $Uy_k = y_{k+1}$ be the two-sided shift operator (see *Example VI.5.2*). Suppose $Ux = \mu x$, and assume to the contrary that $x \neq \theta$. Since U is isometric, $|\mu| = 1$. If $x = \sum_{-\infty}^\infty \lambda_k y_k$, then

$$Ux = \sum_{-\infty}^\infty \lambda_k y_{k+1} = \sum_{-\infty}^\infty \lambda_{k-1} y_k.$$

For all k, $\lambda_{k-1} = (Ux|y_k) = \mu(x|y_k) = \mu\lambda_k$, hence $|\lambda_{k-1}| = |\lambda_k|$. It follows that $|\lambda_k| = |\lambda_0|$ for all k; since $x \neq \theta$, $\lambda_0 \neq 0$. But $\sum_{-\infty}^\infty |\lambda_k|^2 < \infty$, a contradiction.

Exercises

1. Notation as in *Example 4*. T is an operator, and a complex number μ is a proper value for T if and only if $\mu = \mu_k$ for some k.

2. Notation as in *Example 5*. If the μ_k are distinct, a non-zero vector x is a proper vector for T if and only if x is a scalar multiple of some x_k.

3. Let $Tx_k = x_{k+1}$ be the one-sided shift operator, and R the unique operator such that $Rx_k = (1/k)x_k$ for all k ($k = 1,2,3,\cdots$). Then $S = TR$ has no proper values.

4. If $Tx_k = x_{k+1}$ is the one-sided shift operator, then μ is a proper value for T^* if and only if $|\mu| < 1$.

5. There exists a self-adjoint operator T which has no proper values.

6. If T is a hyponormal operator, and $Tx = \mu x$, then $T^*x = \mu^* x$.

7. Let T be an operator. Then:

(i) A non-zero vector x is a proper vector for T if and only if $|(Tx|x)| = \| Tx \| \, \| x \|$.

(ii) In order that T have a proper value μ such that $|\mu| = \| T \|$, it is necessary and sufficient that there exist a vector x such that $\| x \| = 1$ and $|(Tx|x)| = \| T \|$.

(iii) If α_n is a sequence of scalars such that $\alpha_k \geq 0$ for all k, and $\sum_1^\infty \alpha_k{}^2 = 1$, then $\sum_1^\infty \alpha_k \alpha_{k+1} < 1$.

8. Proper vectors and values are definable for a linear mapping of a vector space into itself. If T is a continuous linear mapping in a normed space, and μ is a proper value for T, then $|\mu| \leq \| T \|$.

§ 2. PROPER SUBSPACES

Definition 1. *If T is an operator and μ is a scalar, the null space of the operator $T - \mu I$ is called the* **μ-th proper subspace** *of T, and is denoted $\mathfrak{N}_T(\mu)$. Thus,*

$$\mathfrak{N}_T(\mu) = \{x \in \mathfrak{H} : Tx = \mu x\}.$$

Briefly, $\mathfrak{N}_T(\mu)$ is the **μ-space** *of T.*

Thus, $\mathfrak{N}_T(\mu)$ is a closed linear subspace of $\mathfrak{H}$; it is different from $\{\theta\}$ if and only if μ is a proper value for T. A non-zero vector x is a proper vector for T if and only if x belongs to some μ-space of T.

Theorem 1. *If S and T are operators such that $ST = TS$, then the μ-spaces of T are invariant under S; that is,*

$$S(\mathfrak{N}_T(\mu)) \subset \mathfrak{N}_T(\mu) \text{ for all } \mu.$$

Proof.

If $x \in \mathfrak{N}_T(\mu)$, then $T(Sx) = (TS)x = (ST)x = S(Tx) = S(\mu x) = \mu(Sx)$ shows that $Sx \in \mathfrak{N}_T(\mu)$. ∎

Corollary. *The μ-spaces of T are invariant under T.*

Theorem 2. *If T is a* **normal** *operator,*

(1) *the μ-spaces of T reduce T;*

(2) $\mathfrak{N}_T(\mu) = \mathfrak{N}_{T^*}(\mu^*)$;

(3) $\mathfrak{N}_T(\mu) \perp \mathfrak{N}_T(\nu)$ *whenever* $\mu \neq \nu$.

Proof.

(1): Since $T^*T = TT^*$, $\mathfrak{N}_T(\mu)$ is invariant under T^* by *Theorem 1.* By the *Corollary*, $\mathfrak{N}_T(\mu)$ is also invariant under T, hence $\mathfrak{N}_T(\mu)$ reduces T by *Theorem VI.9.4.*

(2): See *Theorem 1.2.*

(3): Given $x \in \mathfrak{N}_T(\mu)$, $y \in \mathfrak{N}_T(\nu)$, and $\mu - \nu \neq 0$, the problem is to show that $x \perp y$. By (2), $T^*y = \nu^*y$. Then, $(Tx|y) = (x|T^*y)$, $(\mu x|y) = (x|\nu^*y)$, $\mu(x|y) = \nu(x|y)$, $(\mu - \nu)(x|y) = 0$, hence $(x|y) = 0$. ∎

Theorem 1 has a converse, provided T has "sufficiently many" proper vectors:

Theorem 3. *Suppose T is an operator such that the only vector x which is orthogonal to every μ-space of T is $x = \theta$. Then, the following conditions on an operator S are equivalent:*

(a) $ST = TS$

(b) *Every μ-space of T is invariant under S.*

Proof.

(b) *implies* (a): Let $\mathfrak{N}$ be the null space of $TS - ST$; the problem is to show that $\mathfrak{N} = \mathfrak{H}$, or, equivalently, that $\mathfrak{N}^\perp = \{\theta\}$.

One has $\mathfrak{N}_T(\mu) \subset \mathfrak{N}$; for, if $x \in \mathfrak{N}_T(\mu)$, by assumption $Sx \in \mathfrak{N}_T(\mu)$, hence $T(Sx) = \mu(Sx) = S(\mu x) = S(Tx)$, $(TS - ST)x = \theta$. It follows that if $x \perp \mathfrak{N}$, then $x \perp \mathfrak{N}_T(\mu)$ for all μ, hence $x = \theta$ by the hypothesis on T.

(a) implies (b) by *Theorem 1.* ∎

The hypothesis on the μ-spaces of T in *Theorem 3* is conveniently expressed in terms of the following:

Definition 2. *A family of closed linear subspaces is said to be* **total** *in case the only vector x which is orthogonal to every subspace $\mathfrak{N}$ belonging to the family is $x = \theta$.*

Thus, the condition on T in *Theorem 3* is that the μ-spaces are a total family. It will be shown in § 4 of Chapter VIII that the μ-spaces of any "completely continuous" normal operator are a total family; this is the essence of the "spectral theorem" for such operators.

Exercises

1. If T is any operator, and P_μ is the projection whose range is $\mathfrak{N}_T(\mu)$, then $P_\mu T P_\mu = TP_\mu = \mu P_\mu$.

2. Let $Tx_k = \mu_k x_k$, with notation as in *Example 1.5*.
(i) Describe the μ-spaces of T.
(ii) If S is an operator, $ST = TS$ if and only if every μ-space of T is invariant under S.
(iii) If S is an operator such that $ST = TS$, then $S^*T = TS^*$.
(iv) If S is an operator, $ST = TS$ if and only if the μ-spaces of T reduce S.

3. Suppose T is a normal operator whose μ-spaces are a total family. If S is an operator such that $ST = TS$, then $ST^* = T^*S$.

4. If the μ-spaces of an operator T are a total family, and $\mathfrak{N}_T(\mu) \subset \mathfrak{N}_T{}^*(\mu^*)$ for all μ, then T is normal.

5. Let T be a hyponormal operator.
(i) For every scalar μ, $\mathfrak{N}_T(\mu) \subset \mathfrak{N}_T{}^*(\mu^*)$.
(ii) For a fixed scalar μ, let $\mathfrak{N} = \mathfrak{N}_T(\mu)$. Then, $\mathfrak{N}$ reduces T, and $T/\mathfrak{N}$ is normal.
(iii) $\mathfrak{N}_T(\mu) \perp \mathfrak{N}_T(\nu)$ whenever $\mu \neq \nu$.
(iv) If the μ-spaces of T are a total family, then T is normal.

*6. If $\mathfrak{H}$ is finite-dimensional, every hyponormal operator T is normal.

7. *Definition 1* makes sense, and *Theorem 1* is true, for linear mappings in a vector space. For a continuous linear mapping in a normed space, the μ-spaces are closed linear subspaces.

8. Suppose $x_1, \cdots, x_n$ are proper vectors for T, $Tx_k = \mu_k x_k$. If $\mu_1, \cdots, \mu_n$ are distinct, then $x_1, \cdots, x_n$ are linearly independent.

9. If the μ-spaces of T reduce T, and are a total family, then T is normal.

§ 3. APPROXIMATE PROPER VALUES

A scalar μ is a proper value for the operator T if and only if there exists a vector x such that $\| x \| = 1$ and $\| Tx - \mu x \| = 0$. An operator may not have any proper values at all, even if it is normal (see

Example 1.8). A less stringent condition on μ would be the following: given any $\epsilon > 0$, there exists a vector x such that $\| x \| = 1$ and $\| Tx - \mu x \| < \epsilon$; equivalently,

Definition 1. *A scalar μ is said to be an* **approximate proper value** *for the operator T in case there exists a sequence of vectors x_n such that $\| x_n \| = 1$ and $\| Tx_n - \mu x_n \| \to 0$. Briefly, μ is an AP-***value** *for T.*

Examples

1. Every proper value for T is obviously also an AP-value.

2. Let $Tx_k = \mu_k x_k$, with notation as in *Example 1.5*. Assume moreover that the sequence μ_k converges to a limit μ, distinct from every μ_k. As shown in *Example 1.5*, μ is not a proper value for T; however, $\| Tx_k - \mu x_k \| = \| \mu_k x_k - \mu x_k \| = \| (\mu_k - \mu)x_k \| = |\mu_k - \mu| \to 0$, hence μ is an AP-value for T.

3. Let μ be an AP-value for T, and suppose $\| Tx_n - \mu x_n \| \to 0$, where $\| x_n \| = 1$. Then, $|(Tx_n|x_n) - \mu| = |(Tx_n|x_n) - \mu(x_n|x_n)| = |(Tx_n - \mu x_n|x_n)| \leq \| Tx_n - \mu x_n \| \to 0$, hence $(Tx_n|x_n) \to \mu$, $|(Tx_n|x_n)| \to |\mu|$. Clearly $|\mu| \leq \text{LUB}\,\{|(Tx|x)| : \| x \| \leq 1\} \leq \| T \|$.

Theorem 1. *The following conditions on an operator T are equivalent:*

(a) *T has an AP-value μ such that $|\mu| = \| T \|$.*

(b) $\text{LUB}\,\{|(Tx|x)| : \| x \| \leq 1\} = \| T \|$.

Proof.

(a) *implies* (b): Suppose $|\mu| = \| T \|$, $\| x_n \| = 1$, and $\| Tx_n - \mu x_n \| \to 0$. As shown in *Example 3*, $|(Tx_n|x_n)| \to |\mu| = \| T \|$. If M denotes the indicated LUB, one has $\| T \| \geq M \geq |(Tx_n|x_n)| \to \| T \|$ as $n \to \infty$, hence $M = \| T \|$.

(b) *implies* (a): Choose a sequence of vectors x_n such that $\| x_n \| = 1$ and $|(Tx_n|x_n)| \to \| T \|$. Passing to a subsequence, we may assume $(Tx_n|x_n) \to \mu$ for a suitable scalar μ, hence $|\mu| = \| T \|$. Then, $0 \leq \| Tx_n - \mu x_n \|^2 = \| Tx_n \|^2 - (Tx_n|\mu x_n) - (\mu x_n|Tx_n) + |\mu|^2 = \| Tx_n \|^2 - (Tx_n|\mu x_n) - (Tx_n|\mu x_n)^* + |\mu|^2 \leq \| T \|^2 - (Tx_n|\mu x_n) - (Tx_n|\mu x_n)^* + |\mu|^2 = |\mu|^2 - \mu^*(Tx_n|x_n) - \mu(Tx_n|x_n)^* + |\mu|^2 \to |\mu|^2 - \mu^*\mu - \mu\mu^* + |\mu|^2 = 0$ as $n \to \infty$. Thus, $\| Tx_n - \mu x_n \| \to 0$, and μ is an AP-value for T. ∎

There exist self-adjoint operators having no proper values at all (see *Exercise 1.5*). However, it can be shown that *every* operator has at least one AP-value. For our purposes, the following result will suffice:

Theorem 2. *If T is a self-adjoint operator, either $\| T \|$ or $-\| T \|$ is an AP-value for T.*

Proof.

Condition (b) of *Theorem 1* holds for T, by *Theorem VI.6.4*. Let $(Tx_n | x_n) \to \mu$, with notation as in the proof of *Theorem 1*. Since $(Tx_n | x_n)$ is real for all n, μ is real. Since $|\mu| = \| T \|$, one has $\mu^2 = \| T \|^2$. ∎

Exercises

1. AP-values are definable for a linear mapping in a normed space. If T is a continuous linear mapping, and μ is an AP-value for T, then $|\mu| \leq \| T \|$.

2. If μ is an AP-value for T, then $\mu + \lambda$ is an AP-value for $T + \lambda I$, and $\lambda\mu$ is an AP-value for λT.

3. If T is isometric, and μ is an AP-value for T, then $|\mu| = 1$.

4. If T is self-adjoint, every AP-value for T is real.

5. If T is normal, and μ is an AP-value for T, then μ^* is an AP-value for T^*.

6. Let T be a hyponormal operator.

(i) If μ is an AP-value for T, then μ^* is an AP-value for T^*.

(ii) Let μ and ν be AP-values for T, $\mu \neq \nu$. If x_n and y_n are sequences of unit vectors such that $\| Tx_n - \mu x_n \| \to 0$ and $\| Ty_n - \nu y_n \| \to 0$, then $(x_n | y_n) \to 0$.

7. If μ is an AP-value for the operator S, it does not follow that μ^* is an AP-value for S^*.

8. (i) μ is an AP-value for T if and only if $T - \mu I$ fails to be bounded below in the sense of *Exercise VI.8.11*.

(ii) If μ is an AP-value for T, $T - \mu I$ cannot be invertible.

(iii) If $Tx_k = x_{k+1}$ is the one-sided shift operator, then $T = T - 0I$ is not invertible, but 0 is not an AP-value for T.

*9. It can be shown that every normal operator T has an AP-value μ such that $|\mu| = \| T \|$, hence $\| T \| = \text{LUB } \{|(Tx|x)| : \| x \| \leq 1\}$. It would be nice to have a proof in the reverse order.

10. Let T be an operator, μ a scalar. The following conditions are equivalent:

(a) μ is an AP-value for T.

(b) There exists a sequence of operators S_n such that $\| S_n \| = 1$ and $\| (T - \mu I)S_n \| \to 0$.

*11. *Exercise 10* can be generalized to normed spaces.

12. The *spectrum* of an operator T is defined to be the set of all scalars μ such that $T - \mu I$ fails to be invertible in the sense of *Exercise VI.8.9;* it will be denoted $s(T)$. Then:

(i) If μ is an AP-value for T, $\mu \in s(T)$.

(ii) $\mu \in s(T)$ if and only if $\mu^* \in s(T^*)$.

(iii) $s(T + \lambda I) = \{\mu + \lambda : \mu \in s(T)\}$.

(iv) $s(\lambda T) = \{\lambda\mu : \mu \in s(T)\}$.

(v) T is invertible if and only if 0 is not in $s(T)$.

(vi) If T is invertible, $s(T^{-1}) = \{\mu^{-1} : \mu \in s(T)\}$.

(vii) If T is self-adjoint, every $\mu \in s(T)$ is real.

(viii) If T is positive, every $\mu \in s(T)$ is ≥ 0.

*(ix) It can be shown that for every operator T, $s(T)$ is a *non-empty* and *closed* set of complex numbers, and $|\mu| \leq \| T \|$ for all $\mu \in s(T)$.

(x) If T is normal, $\mu \in s(T)$ if and only if μ is an AP-value for T.

(xi) If T is unitary, then $|\mu| = 1$ for every $\mu \in s(T)$.

*(xii) If $Tx_k = x_{k+1}$ is the one-sided shift operator, $s(T)$ is the entire disc $|\mu| \leq 1$.

*13. If Λ is any non-empty closed and bounded set of complex numbers, and $\mathfrak{H}$ is classical Hilbert space, there exists a normal operator T such that (1) Λ is the spectrum of T, and (2) $\mu \in \Lambda$ if and only if μ is an AP-value for T.

Chapter

Completely Continuous Operators

VIII

§ 1. *Completely continuous operators*
§ 2. *An example*
§ 3. *Proper values of CC-operators*
§ 4. *Spectral theorem for a normal CC-operator*

§ 1. COMPLETELY CONTINUOUS OPERATORS

If the μ-spaces of an operator T are a total family, and S is any operator, $ST = TS$ if and only if the μ-spaces of T are invariant under S; this shows the power of proper values. However, the μ-spaces of an operator, even a self-adjoint one, may all be $\{\theta\}$; this shows their weakness.

Every self-adjoint operator T does have an AP-value μ, in fact, one for which $|\mu| = \| T \|$. Happily, there is an important class of operators for which AP-values are nearly always proper (see *Theorem 3.2*); these are the completely continuous operators:

Definition 1. *An operator T is said to be* **completely continuous** *in case: given any sequence of vectors x_n such that $\| x_n \|$ is bounded, Tx_n has a convergent subsequence. Briefly, T is a CC-*operator.

Clearly, an operator T is CC if and only if: $\| x_n \| \leq 1$ implies Tx_n has a convergent subsequence.

Examples

1. The zero operator is CC.

2. If y and z are fixed vectors, the operator T defined by the formula $Tx = (x|y)z$ is CC. For, suppose $\| x_n \| \leq 1$. Since $|(x_n|y)| \leq \| y \|$,

$(x_n|y)$ has a convergent subsequence, say $(x_{n_k}|y) \to \lambda$; then $Tx_{n_k} = (x_{n_k}|y)z \to \lambda z$.

3. If $\mathfrak{H}$ is infinite-dimensional, the identity operator is *not* CC. For, if $x_1, x_2, x_3, \cdots$ is any orthonormal sequence (see *Theorems I.7.3* and *II.6.4*), $\| Ix_m - Ix_n \|^2 = \| x_m - x_n \|^2 = 2$ whenever $m \neq n$, hence Ix_n cannot have a convergent subsequence. ∎

The results of this section and the next show that CC-operators exist abundantly; § 3 is preliminary to § 4, in which normal CC-operators are analyzed in detail.

Theorem 1. *If T is a CC-operator, and λ is a scalar, λT is a CC-operator.*

Proof.

If $Tx_{n_k} \to y$, then $(\lambda T)x_{n_k} \to \lambda y$. ∎

Theorem 2. *If S and T are CC-operators, $S + T$ is a CC-operator.*

Proof.

Given $\| x_n \| \leq 1$, the problem is to find a convergent subsequence of $Sx_n + Tx_n$. Passing to a subsequence, we may suppose $Sx_n \to u$. For a further subsequence, $Tx_{n_k} \to v$; since also $Sx_{n_k} \to u$, $(S + T)x_{n_k} \to u + v$. ∎

Definition 2. *An operator T is said to be* **finite-dimensional** *if its range $T(\mathfrak{H})$ is a finite-dimensional linear subspace.*

Theorem 3. *Every finite-dimensional operator T is CC.*

Proof.

Let $z_1, \cdots, z_n$ be an orthonormal basis for $T(\mathfrak{H})$. For each $x \in \mathfrak{H}$, $Tx = \sum_1^n f_k(x)z_k$, where $f_1(x), \cdots, f_n(x)$ are scalars uniquely determined by x; clearly $f_k(x) = (Tx|z_k) = (x|T^*z_k)$. Define $T_k x = f_k(x)z_k = (x|T^*z_k)z_k$; T_k is a CC-operator by *Example 2*, hence $T = \sum_1^n T_k$ is CC by *Theorem 2*. ∎

Theorem 4. *If T is a CC-operator, and S is* **any** *operator, then ST and TS are CC-operators.*

Proof.

Given $\| x_n \| \leq 1$. For a suitable subsequence, $Tx_{n_k} \to u$, hence $STx_{n_k} \to Su$; thus, ST is CC. Also, $\| Sx_n \| \leq \| S \|$, hence $T(Sx_n)$ has a convergent subsequence; thus, TS is CC. ∎

Theorem 5. *If T_n is a sequence of CC-operators, T is an operator, and $\| T - T_n \| \to 0$, then T is a CC-operator.*

Proof. ("diagonal procedure")

Given $\| x_n \| \leq 1$, the problem is to find a convergent subsequence of Tx_n.

Let x_n^1 be a subsequence of x_n for which $T_1 x_n^1$ is convergent. Say, $T_1 x_n^1 \to u_1$.

Let x_n^2 be a subsequence of x_n^1 for which $T_2 x_n^2$ is convergent. Say, $T_2 x_n^2 \to u_2$.

Continuing inductively, one obtains, for each k, a subsequence x_n^k of x_n^{k-1} such that $T_k x_n^k \to u_k$.

The x_n^k may be arranged in a rectangular array, with $x_1^k, x_2^k, x_3^k, \cdots$ as the k'th row. Then, the k'th row is by construction a subsequence of the $(k-1)$'th row. Consider the "diagonal sequence" $x_1^1, x_2^2, x_3^3, \cdots$. For each k, the sequence $x_k^k, x_{k+1}^{k+1}, x_{k+2}^{k+2}, \cdots$ is clearly a subsequence of the k'th row, hence

$$\lim_{n \to \infty} T_k x_n^n = u_k \quad \text{for } k = 1,2,3, \cdots.$$

We assert that Tx_n^n is a Cauchy sequence. For, given any $\epsilon > 0$, fix any index k such that $\| T - T_k \| \leq \epsilon$. For all m and n, $\| Tx_m^m - Tx_n^n \| \leq \| (T - T_k) x_m^m \| + \| T_k x_m^m - T_k x_n^n \| + \| (T_k - T) x_n^n \| \leq 2\epsilon + \| T_k x_m^m - T_k x_n^n \|$. Since $T_k x_n^n \to u_k$ as $n \to \infty$, there is an index N such that $\| T_k x_m^m - T_k x_n^n \| \leq \epsilon$ whenever $m,n \geq N$. Then, $\| Tx_m^m - Tx_n^n \| \leq 3\epsilon$ whenever $m,n \geq N$. This shows that the sequence Tx_n^n is Cauchy; since $\mathfrak{H}$ is complete, it is convergent. ∎

Corollary. *If T_n is a sequence of finite-dimensional operators, T is an operator, and $\| T_n - T \| \to 0$, then T is a CC-operator.*

Proof.

Each T_n is a CC-operator by *Theorem 3*, hence T is CC by *Theorem 5*. ∎

Exercises

1. Let T be an operator, $\mathfrak{N}$ a closed linear subspace of $\mathfrak{H}$.

(i) If $\mathfrak{N}$ is invariant under T, and T is CC, then $T/\mathfrak{N}$ is CC.

(ii) If $\mathfrak{N}$ reduces T, T is CC if and only if both $T/\mathfrak{N}$ and $T/\mathfrak{N}^\perp$ are CC.

2. If $\mathfrak{H}$ is finite-dimensional, every linear mapping $T: \mathfrak{H} \to \mathfrak{H}$ is a CC-operator.

3. (i) If S and T are metrically equivalent operators, and T is CC, then S is CC.

(ii) In particular, if T is normal and CC, then T^* is CC. See also *Exercise 6*.

4. If T is a CC-operator, and $S^*S \leq T^*T$, then S is a CC-operator. In particular, if T is a hyponormal CC-operator, T^* is CC. See also *Exercise 6*.

5. If $\mathfrak{H}$ is infinite-dimensional, a CC-operator cannot be invertible.

6. (i) An operator T is CC if and only if T^*T is CC.

(ii) If T is a CC-operator, then T^* is CC.

7. (i) If $\mathfrak{H}$ is infinite-dimensional, and the operator T is bounded below, then T cannot be CC.

*(ii) If $\mathfrak{H}$ is infinite-dimensional, a CC-operator cannot be surjective.

8. (i) If y and z are fixed vectors, and $y \neq \theta$, there exists a CC-operator T such that $Ty = z$.

(ii) More generally, the set of CC-operators is *n-fold transitive* in the following sense: if $y_1, \cdots, y_n$ are linearly independent, and $z_1, \cdots, z_n$ are arbitrary, there exists a CC-operator T such that $Ty_k = z_k$ for all k.

9. If $\mathfrak{A}$ is an algebra, an *ideal* of $\mathfrak{A}$ is a linear subspace $\mathfrak{I}$ such that: if $x \in \mathfrak{I}$, and $a \in \mathfrak{A}$ is arbitrary, both $ax \in \mathfrak{I}$ and $xa \in \mathfrak{I}$. Evidently, an ideal of $\mathfrak{A}$ is a subalgebra. The results of this section can be summarized as follows:

(i) The CC-operators form a closed ideal $\mathfrak{L}_{cc}(\mathfrak{H})$ of the Banach algebra $\mathfrak{L}_c(\mathfrak{H})$, and this ideal contains every finite-dimensional operator.

(ii) $\mathfrak{L}_{cc}(\mathfrak{H})$ is itself a Banach algebra.

(iii) One has $I \in \mathfrak{L}_{cc}(\mathfrak{H})$ if and only if $\mathfrak{H}$ is finite-dimensional, and in this case $\mathfrak{L}_{cc}(\mathfrak{H}) = \mathfrak{L}_c(\mathfrak{H})$.

(iv) In view of *Exercise 6*, $\mathfrak{L}_{cc}(\mathfrak{H})$ is a *-subalgebra of $\mathfrak{L}_c(\mathfrak{H})$, hence is a B^*-algebra.

10. Notation as in *Exercise V.8.11*, with $\mathfrak{K} = \mathfrak{H}$ a classical Hilbert space. Prove:

(i) The *-algebra $\mathfrak{L}_{hs}(\mathfrak{H},\mathfrak{H})$ of operators of Hilbert-Schmidt class is an ideal of $\mathfrak{L}_c(\mathfrak{H})$.

(ii) Every finite-dimensional operator is of Hilbert-Schmidt class.

*11. If $\mathcal{I}$ is an ideal of $\mathfrak{L}_c(\mathfrak{H})$, different from $\{0\}$, then $\mathcal{I}$ contains every finite-dimensional operator.

12. If $\mathcal{E}$ and $\mathcal{F}$ are normed spaces, complete continuity is definable for a linear mapping $T: \mathcal{E} \to \mathcal{F}$. If T is CC, it is necessarily continuous.

13. Let $\mathcal{E}$ and $\mathcal{F}$ be normed spaces, and denote by $\mathfrak{L}_{cc}(\mathcal{E},\mathcal{F})$ the set of all CC linear mappings $T: \mathcal{E} \to \mathcal{F}$. Then:

(i) $\mathfrak{L}_{cc}(\mathcal{E},\mathcal{F})$ is a linear subspace of $\mathfrak{L}_c(\mathcal{E},\mathcal{F})$.

(ii) $\mathfrak{L}_{cc}(\mathcal{E},\mathcal{E})$ is an ideal in the algebra $\mathfrak{L}_c(\mathcal{E})$, hence is a normed algebra.

(iii) If $\mathcal{F}$ is a Banach space, $\mathfrak{L}_{cc}(\mathcal{E},\mathcal{F})$ is a closed linear subspace of the Banach space $\mathfrak{L}_c(\mathcal{E},\mathcal{F})$, hence is itself a Banach space.

(iv) If $\mathcal{F}$ is a Banach space, $\mathfrak{L}_{cc}(\mathcal{F},\mathcal{F})$ is a closed ideal of the Banach algebra $\mathfrak{L}_c(\mathcal{F})$, hence is itself a Banach algebra.

*(v) If $T: \mathcal{E} \to \mathcal{F}$ is a continuous linear mapping such that $T(\mathcal{E})$ is finite-dimensional, then T is CC.

*(vi) It can be shown that the identity mapping $I: \mathcal{E} \to \mathcal{E}$ is CC if and only if $\mathcal{E}$ is finite-dimensional. In this case, every linear mapping $T: \mathcal{E} \to \mathcal{E}$ is CC.

14. If $\mathcal{E}$ and $\mathcal{F}$ are normed spaces, and $\mathcal{E}$ is finite-dimensional, then every linear mapping $T: \mathcal{E} \to \mathcal{F}$ is CC.

15. Let $\mathcal{E}$ be a normed space, and $\mathfrak{N}$ a dense linear subspace of $\mathcal{E}$. Prove:

(i) If $x \in \mathcal{E}$, there exists a sequence $y_n \in \mathfrak{N}$ such that $y_n \to x$, and $\| y_n \| = \| x \|$ for all n.

(ii) Let $\mathcal{B}$ be a Banach space, $S: \mathfrak{N} \to \mathcal{B}$ a continuous linear mapping. If $T: \mathcal{E} \to \mathcal{B}$ is the unique continuous linear mapping such that $Ty = Sy$ for all $y \in \mathfrak{N}$ (see *Theorem IV.7.6*), then T is CC if and only if S is CC.

*16. If $\mathcal{E}$ is a normed space, $\mathcal{B}$ is a Banach space, and $T: \mathcal{E} \to \mathcal{B}$ is a continuous linear mapping, it can be shown that T is CC if and only if T' is CC.

17. If $\mathcal{E}$ and $\mathcal{F}$ are normed spaces, consider the set of all linear mappings $T: \mathcal{E} \to \mathcal{F}$ satisfying the following condition: if $x_n \in \mathcal{E}$, $\| x_n \| \leq 1$, Tx_n has a Cauchy subsequence. Examine the theorems of this section, with complete continuity replaced by this condition.

§ 2. AN EXAMPLE

Let $\mathcal{H}$ be classical Hilbert space, x_n an orthonormal basis for $\mathcal{H}$, μ_n a bounded sequence of scalars, and T the unique operator such that $Tx_k = \mu_k x_k$ for all k (see *Theorem VI.3.1*). If P_k is the projection whose range is the one-dimensional subspace generated by x_k, and $T_n = \sum_1^n \mu_k P_k$, one has $T_n x \to Tx$ for each $x \in \mathcal{H}$ (see *Example VI.7.3*). It will be shown in § 4 that if T is *CC*, necessarily $\mu_n \to 0$; we are concerned here with the converse:

Theorem 1. *If* $\mu_n \to 0$, *then* $\| T - T_n \| \to 0$ *and* T *is completely continuous.*

Proof.

The range of T_n is contained in the n-dimensional subspace generated by $x_1, \cdots, x_n$, hence T_n is a finite-dimensional operator. In view of the Corollary of *Theorem 1.5*, it will suffice to show that $\| T - T_n \| \to 0$.

Given $\epsilon > 0$, let N be an index such that $|\mu_k| \leq \epsilon$ whenever $k \geq N$. Fix an index $n \geq N$; it will be shown that $\| T - T_n \| \leq \epsilon$. For any vector $x = \sum_1^\infty \lambda_k x_k$, one has $T_n x = \sum_1^n \lambda_k \mu_k x_k$, hence $\| Tx - T_n x \|^2 = \| \sum_1^\infty \lambda_k \mu_k x_k - \sum_1^n \lambda_k \mu_k x_k \|^2 = \| \sum_{n+1}^\infty \lambda_k \mu_k x_k \|^2 = \sum_{n+1}^\infty |\lambda_k \mu_k|^2 \leq \epsilon^2 \sum_{n+1}^\infty |\lambda_k|^2 \leq \epsilon^2 \sum_1^\infty |\lambda_k|^2 = \epsilon^2 \| x \|^2$. Thus, $\| (T - T_n)x \| \leq \epsilon \cdot \| x \|$; since x is arbitrary, $\| T - T_n \| \leq \epsilon$. ∎

Exercises

1. Let $\mathcal{H}$ be classical Hilbert space, x_n an orthonormal basis for $\mathcal{H}$, and (α_{jk}) an "infinite matrix" of scalars ($j = 1,2,3,\cdots$; $k = 1,2,3,\cdots$) such that $\sum_{j,k} |\alpha_{jk}|^2 < \infty$. Then, there exists one and only one operator T such that $Tx_k = \sum_j \alpha_{jk} x_j$ for all k; this operator is CC. Every operator of Hilbert-Schmidt class is completely continuous.

2. If $\mathcal{H}$ is classical Hilbert space, x_n is an orthonormal basis for $\mathcal{H}$, and β_{jk} are scalars such that $\sum_{j,k} |\beta_{jk}| < \infty$, there exists a unique operator T whose matrix is (β_{jk}) relative to the given basis. This operator is of Hilbert-Schmidt class, hence is CC.

§ 3. PROPER VALUES OF CC-OPERATORS

Recall that if T is an operator and μ is a scalar, $\mathfrak{N}_T(\mu)$ denotes the μ-space of T; it is the closed linear subspace $\{x \in \mathfrak{K}: Tx = \mu x\}$, that is, the null space of the operator $T - \mu I$.

Theorem 1. *If T is a CC-operator, and $\mu \neq 0$, then $\mathfrak{N}_T(\mu)$ is finite-dimensional.*

Proof.

Assume to the contrary; by *Theorem I.7.3*, $\mathfrak{N}_T(\mu)$ contains a linearly independent sequence x_n. By *Theorem II.6.4*, we may suppose the x_n are orthonormal. If $m \neq n$, $\| Tx_m - Tx_n \|^2 = \| \mu x_m - \mu x_n \|^2 = |\mu|^2 \| x_m - x_n \|^2 = 2 |\mu|^2 > 0$, hence Tx_n cannot have a convergent subsequence; since $\| x_n \| = 1$, this is contrary to the complete continuity of T. ∎

Definition 1. *Suppose μ is a proper value for the operator T. If $\mathfrak{N}_T(\mu)$ has finite dimension n, μ is said to be a proper value of* **finite multiplicity** *n. If $\mathfrak{N}_T(\mu)$ is infinite-dimensional, μ is said to have* **infinite multiplicity.** *If μ is not a proper value for T, that is, if $\mathfrak{N}_T(\mu) = \{\theta\}$, it is convenient to say that μ is a "proper value of* **multiplicity zero.**"

In this terminology, *Theorem 1* asserts that every non-zero proper value of a CC-operator has finite multiplicity. This result is not always helpful, for there exist CC-operators having no proper values at all:

Example 1. Let $Tx_k = x_{k+1}$ be the one-sided shift operator, R the unique operator such that $Rx_k = (1/k)x_k$ for all k, and let $S = TR$. Since R is CC by *Theorem 2.1*, S is CC by *Theorem 1.4;* however, S has no proper values (see *Exercise VII.1.3*). ∎

Every *normal* CC-operator does have at least one proper value, as is shown in *Theorem 3* below. This result is expedited by the following

Theorem 2. *If T is a CC-operator, and μ is a* **non-zero** *approximate proper value for T, then μ is a proper value.*

Proof.

By assumption, there exists a sequence of vectors x_n such that $\| x_n \| = 1$ and $\| Tx_n - \mu x_n \| \to 0$. Since T is CC, we may suppose, after passing to a subsequence, that Tx_n is convergent, say $Tx_n \to u$. Since $\| u - \mu x_n \| \leq \| u - Tx_n \| + \| Tx_n - \mu x_n \| \to 0$, one has $\mu x_n \to u$. Then, $T(\mu x_n) \to Tu$, thus $Tu = \lim T(\mu x_n) = \lim \mu(Tx_n)$

$= \mu \lim Tx_n = \mu u$. Since $\| u \| = \lim \| \mu x_n \| = \lim |\mu| = |\mu| > 0$, u is a proper vector. ∎

Lemma 1. *If S is any self-adjoint CC-operator, S has a proper value μ such that $|\mu| = \| S \|$.*

Proof.

If $S = 0$, $\mu = 0$ is a proper value (we are assuming $\mathfrak{H} \neq \{\theta\}$). If $S \neq 0$, then by *Theorem VII.3.2*, S has an approximate proper value μ such that $|\mu| = \| S \| > 0$; quote *Theorem 2*. ∎

Lemma 2. *If $\mathfrak{H}$ is finite-dimensional, every normal operator T has at least one proper value.*

Proof.

Every operator in $\mathfrak{H}$ is CC, by *Theorem 1.3*. Let $T = A + iB$ be the Cartesian form of T (see *Theorem VI.6.3*). By *Lemma 1*, there is a (real) scalar α such that the proper subspace $\mathfrak{N} = \mathfrak{N}_A(\alpha)$ is $\neq \{\theta\}$. Since T is normal, $AB = BA$ (see *Theorem VI.8.2*), hence $\mathfrak{N}$ is invariant under B (see *Theorem VII.2.1*). Let $R = B/\mathfrak{N}$ be the restriction of B to $\mathfrak{N}$; R is a self-adjoint operator in $\mathfrak{N}$ (see Corollary 2 of *Theorem VI.9.5*). By *Lemma 1*, there exists a scalar β, and a non-zero vector $x \in \mathfrak{N}$ such that $Rx = \beta x$, that is, $Bx = \beta x$. By the definition of $\mathfrak{N}$, $Ax = \alpha x$, hence $Tx = Ax + iBx = (\alpha + i\beta)x$. ∎

Theorem 3. *Every normal CC-operator T has at least one proper value; in fact, there exists a proper value μ such that $|\mu| = \| T \|$.*

Proof.

Let $S = T^*T$; S is self-adjoint, and is CC by *Theorem 1.4*. We may suppose $T \neq 0$, hence $\| S \| > 0$.

By the definition of $\| T \|$, there is a sequence of vectors x_n such that $\| x_n \| = 1$ and $\| Tx_n \| \to \| T \|$. Then, $\| S \| = \| T^*T \| = \| T \|^2 = \lim \| Tx_n \|^2 = \lim (T^*Tx_n | x_n) = \lim (Sx_n | x_n)$.

Passing to a subsequence, we may suppose, by the complete continuity of T, that $Tx_n \to y$ for a suitable vector y.

Let $\alpha = \| S \|$. We assert that α is a proper value for S. For, $\| Sx_n - \alpha x_n \|^2 = \| Sx_n \|^2 - 2\alpha(Sx_n | x_n) + \alpha^2 \leq \alpha^2 - 2\alpha(Sx_n | x_n) + \alpha^2 \to \alpha^2 - 2\alpha^2 + \alpha^2 = 0$, thus $Sx_n - \alpha x_n \to \theta$. Since $Tx_n \to y$, one has $Sx_n = T^*(Tx_n) \to T^*y$, hence $\alpha x_n = (\alpha x_n - Sx_n) + Sx_n \to \theta + T^*y$. Thus, $x_n \to \alpha^{-1}(T^*y)$. Setting $z = \alpha^{-1}(T^*y)$, one has $x_n \to z$, $\| z \| = \lim \| x_n \| = \lim 1 = 1$, and $Sz = \lim Sx_n = T^*y = \alpha z$. Thus, z is a proper vector for S, with associated proper value α. [One could

also have quoted *Lemma 1*, since $S \geq 0$; see the remarks following this theorem.]

Let $\mathfrak{N} = \mathfrak{N}_S(\alpha)$; it has just been shown that $\mathfrak{N} \neq \{\theta\}$. Since $\alpha > 0$, $\mathfrak{N}$ is finite-dimensional by *Theorem 1*. Since T is normal, clearly $TS = ST$ and $T^*S = ST^*$, hence $\mathfrak{N}$ reduces T (see *Theorems VII.2.1* and *VI.9.4*). Let $R = T/\mathfrak{N}$; R is a normal operator in $\mathfrak{N}$ (see Corollary 1 of *Theorem VI.9.5*). By *Lemma 2*, there exists a scalar μ, and a vector $x \in \mathfrak{N}$, such that $\| x \| = 1$ and $Rx = \mu x$, that is, $Tx = \mu x$. By *Theorem VII.1.2*, $T^*x = \mu^* x$, hence $Sx = T^*Tx = |\mu|^2 x$; since $Sx = \alpha x$ by the definition of $\mathfrak{N}$, $|\mu|^2 = \alpha$. Thus, $|\mu|^2 = \| S \| = \| T \|^2$. ∎

Remarks on Theorem 3

Suppose T is a normal CC-operator. Since $\| T^*x_m - T^*x_n \| = \| T^*(x_m - x_n) \| = \| T(x_m - x_n) \| = \| Tx_m - Tx_n \|$, it is clear that T^* is also CC. Let $T = A + iB$ be the Cartesian form; it is clear from the formulas in *Theorem VI.6.3* that A and B are CC-operators (see *Theorems 1.1* and *1.2*). It follows that the proof of *Lemma 2* is valid for T without the assumption of finite-dimensionality: T has at least one proper value.

Why mention *Lemma 2* at all? The point is that the existence of proper values for a linear mapping in finite-dimensional space is an essentially algebraic fact (see *Exercise 5*) which may be known to the reader from other contexts. Granted *Lemma 2*, one way or another, the proof of *Theorem 3* derives directly from *Lemma 2*, and one need not quote *Theorem 2* and *Lemma 1;* in this way, the concept of AP-value is circumvented, thus emphasizing the algebraic aspects of *Theorem 3* and the spectral theorem to follow.

Exercises

1. (i) If T is a normal CC-operator, there exists a vector x, $\| x \| = 1$, such that $|(Tx|x)| = \| T \|$. Compare this with *Exercise VII.1.7*.

*(ii) It can be shown that for every normal operator T, $\| T \| = \text{LUB}\, \{|(Tx|x)| : \| x \| = 1\}$.

2. Notation as in the proof of *Theorem 3*. Then, $\mathfrak{N}$ contains all vectors x such that $Tx = \lambda x$ for some scalar λ with $|\lambda| = \| T \|$. One has $\| Tx \| = \| T \| \, \| x \|$ for every $x \in \mathfrak{N}$.

3. Let $Rx_k = (1/k)x_k$, as in *Example 1*. R is a self-adjoint CC-operator, 0 is an AP-value for R, but 0 is not a proper value.

*4. It can be shown that every operator T has at least one AP-value μ. If moreover T is CC and $\mu \neq 0$, T has a proper value.

*5. It can be shown that if $\mathfrak{H}$ is finite-dimensional, every operator T has at least one proper value.

*6. *Theorem 1* holds for a CC linear mapping $T: \mathcal{E} \to \mathcal{E}$, where $\mathcal{E}$ is a normed space.

*7. In the wake of a result such as *Theorem 2*, the reader should be warned of the existence of a stubborn class of operators whose only AP-values are 0. Such an operator is called a *generalized nilpotent*. It can be shown that the following conditions on an operator T are equivalent:

(a) T is a generalized nilpotent.
(b) The spectrum of T is $\{0\}$.
(c) $\lim \| T^n \|^{1/n} = 0$.

[The operators T^n are defined inductively by the formulas $T^1 = T$, $T^{n+1} = T^n T$. Incidentally, T is called a *nilpotent* operator in case $T^n = 0$ for some n.] It can be shown that the only normal generalized nilpotent is the zero operator.

8. The operator S of *Example 1* is a generalized nilpotent.

§ 4. SPECTRAL THEOREM FOR A NORMAL CC-OPERATOR

Throughout this section, T is a fixed normal CC-operator.

Theorem 1. (Spectral Theorem) *The proper subspaces of T are a total family. That is, if $x \perp \mathfrak{N}_T(\mu)$ for every scalar μ, necessarily $x = \theta$.*

Proof.

Let $\mathfrak{S}$ be the smallest subset of $\mathfrak{H}$ such that $\mathfrak{N}_T(\mu) \subset \mathfrak{S}$ for all μ; clearly $\mathfrak{S} = \{x \in \mathfrak{H}: x \in \mathfrak{N}_T(\mu) \text{ for some } \mu\}$. Let $\mathfrak{N} = \mathfrak{S}^\perp$; the problem is to show that $\mathfrak{N} = \{\theta\}$.

Since each $\mathfrak{N}_T(\mu)$ reduces T (*Theorem VII.2.2*), $T(\mathfrak{S}) \subset \mathfrak{S}$ and $T^*(\mathfrak{S}) \subset \mathfrak{S}$; by *Theorem V.8.4*, $T^*(\mathfrak{N}) \subset \mathfrak{N}$ and $T(\mathfrak{N}) \subset \mathfrak{N}$, thus $\mathfrak{N}$ reduces T (see *Theorem VI.9.4*).

Assume to the contrary that $\mathfrak{N} \neq \{\theta\}$, and let $R = T/\mathfrak{N}$; R is a normal operator in $\mathfrak{N}$ (see Corollary 1 of *Theorem VI.9.5*). We assert that R is a CC-operator in the Hilbert space $\mathfrak{N}$. For, if $x_n \in \mathfrak{N}$ and $\| x_n \| \leq 1$, $Rx_n = Tx_n$ has a convergent subsequence, say $Rx_{n_k} \to u \in \mathfrak{H}$; since $\mathfrak{N}$ is closed, $u \in \mathfrak{N}$.

Thus, R is a normal CC-operator in the Hilbert space $\mathfrak{N} \neq \{\theta\}$. By *Theorem 3.3*, there exists a non-zero vector $x \in \mathfrak{N}$, and a scalar μ, such that $Rx = \mu x$, that is, $Tx = \mu x$. Then, $x \in \mathfrak{N}_T(\mu) \subset \mathfrak{S}$; but $x \in \mathfrak{N} = \mathfrak{S}^\perp$, hence $x \perp x$, $x = \theta$, a contradiction. ∎

Let $P(\mu)$ denote the projection with range $\mathfrak{N}_T(\mu)$. [More precisely, $P(\mu) = P_T(\mu)$; thus, $P_{T^*}(\mu) = P_T(\mu^*)$.] The following result is essentially an application of the spectral theorem:

Theorem 2. *The following conditions on an operator S are equivalent:*

(a) $ST = TS$

(b) *The μ-spaces of T are invariant under S.*

(c) $ST^* = T^*S$

(d) *The μ-spaces of T reduce S.*

(e) $SP(\mu) = P(\mu)S$ *for all* μ.

Proof.

(a) *implies* (b): See *Theorem VII.2.1.*

(b) *implies* (c): Since $\mathfrak{N}_{T^*}(\mu) = \mathfrak{N}_T(\mu^*)$, the μ-spaces of T^* are invariant under S; since they are a total family by the spectral theorem, $ST^* = T^*S$ by *Theorem VII.2.3.*

(c) *implies* (d): The μ-spaces of T^* are invariant under S by *Theorem VII.2.1;* in other words, the μ-spaces of T are invariant under S. Since $S^*T = TS^*$, the μ-spaces of T are also invariant under S^*.

(d) and (e) are equivalent by *Theorem VI.9.4.*

(d) implies (a) by *Theorem VII.2.3* and the spectral theorem. ∎

The next results will show that the proper values of T can be enumerated in a (finite or infinite) sequence.

Theorem 3. *For each $\epsilon > 0$, the annulus $\epsilon \leq |\lambda| \leq \| T \|$ contains at most finitely many proper values of T. Every proper value of T lies in the disc $|\lambda| \leq \| T \|$; that is, $\mathfrak{N}_T(\lambda) = \{\theta\}$ whenever $|\lambda| > \| T \|$.*

Proof.

If μ is a proper value of T, $|\mu| \leq \| T \|$ was shown in *Theorem VII.1.1;* this does not require any special hypotheses on T.

Given $\epsilon > 0$, assume to the contrary that there exists an infinite sequence μ_n of distinct proper values of T such that $\epsilon \leq |\mu_n| \leq \| T \|$.

Passing to a subsequence, we may assume $\mu_n \to \mu$, where $\epsilon \leq |\mu| \leq \| T \|$. Let $Tx_n = \mu_n x_n$, $\| x_n \| = 1$. Again passing to a subsequence; we may assume $Tx_n \to y$ for a suitable vector y. Thus, $\mu_n x_n \to y$; since $\mu_n^{-1} \to \mu^{-1}$, one has $x_n = \mu_n^{-1}(\mu_n x_n) \to \mu^{-1} y$, hence $\| x_m - x_n \| \to 0$ as $m,n \to \infty$. But x_n is an orthonormal sequence by *Theorem VII.2.2*, hence $\| x_m - x_n \|^2 = 2$ whenever $m \neq n$, a contradiction. ∎

Let us first dispose of a special case:

Theorem 4. *The following conditions on T are equivalent:*

(a) *T is finite-dimensional.*

(b) *T has only finitely many distinct proper values $\mu_1, \cdots, \mu_n$.*

In this case, $T = \sum_1^n \mu_k P(\mu_k)$.

Proof.

(a) *implies* (b): Assume to the contrary there is an infinite sequence $\mu_1, \mu_2, \mu_3, \cdots$ of distinct proper values. Let $Tx_k = \mu_k x_k$, $\| x_k \| = 1$; the sequence of vectors x_n is orthonormal, hence linearly independent. Since at most one of the μ_k is 0, $x_k = T(\mu_k^{-1} x_k) \in T(\mathfrak{H})$ for the remaining k, hence $T(\mathfrak{H})$ is infinite-dimensional.

(b) *implies* (a): Let $\mathfrak{M}$ be the smallest linear subspace containing $\mathfrak{N}_T(\mu_1), \cdots, \mathfrak{N}_T(\mu_n)$. Evidently, $\mathfrak{M}$ is the set of all vectors x of the form $x = \sum_1^n y_k$, with $y_k \in \mathfrak{N}_T(\mu_k)$ (see *Theorem I.5.1*). Since the $\mathfrak{N}_T(\mu_k)$ are mutually orthogonal, it is appropriate to write $\mathfrak{M} = \mathfrak{N}_T(\mu_1) \oplus \cdots \oplus \mathfrak{N}_T(\mu_n)$, in analogy with *Definition III.2.5*. We assert that $\mathfrak{M} = \mathfrak{H}$; since $\mathfrak{M}$ is closed (see the Corollary of *Theorem III.4.2*), it will suffice to show that $\mathfrak{M}^\perp = \{\theta\}$. Indeed, let

$$\mathfrak{S} = \{x \in \mathfrak{H} : x \in \mathfrak{N}_T(\mu) \text{ for some } \mu\};$$

clearly $\mathfrak{M}^\perp = \mathfrak{S}^\perp$. But $\mathfrak{S}^\perp = \{\theta\}$ was shown in the proof of *Theorem 1*.

Thus, given any vector $x \in \mathfrak{H}$, one can write $x = \sum_1^n y_k$, with $y_k \in \mathfrak{N}_T(\mu_k)$. If $j \neq k$, then $y_k \in [\mathfrak{N}_T(\mu_j)]^\perp$, hence $P(\mu_j) y_k = \theta$; it follows that $P(\mu_j) x = \sum_k P(\mu_j) y_k = y_j$, thus $x = \sum_1^n P(\mu_k) x$. This shows that $\sum_1^n P(\mu_k) = I$. Also,

$$Tx = \sum_1^n Ty_k = \sum_1^n \mu_k y_k = \sum_1^n \mu_k P(\mu_k) x,$$

hence $T = \sum_1^n \mu_k P(\mu_k)$.

Incidentally, if T is also injective, $\mathcal{H}$ is necessarily finite-dimensional. ∎

Assume henceforth that T is not finite-dimensional; that is, T has infinitely many proper values.

Theorem 5. *Assuming T is not a finite-dimensional operator,*

(1) *the distinct non-zero proper values of T can be enumerated in an infinite sequence $\mu_1,\mu_2,\mu_3,\cdots$;*

(2) *necessarily $\mu_n \to 0$;*

(3) *one can arrange to have $|\mu_1| \geq |\mu_2| \geq |\mu_3| \geq \cdots$.*

Proof.

(1): For each $n = 1,2,3,\cdots$, let Λ_n denote the set (possibly empty) of all proper values μ such that $1/n \leq |\mu| \leq \| T \|$. By *Theorem 3*, each Λ_n contains at most finitely many scalars, and every non-zero proper value belongs to some Λ_n. Since $\Lambda_1 \subset \Lambda_2 \subset \cdots$, the non-zero proper values can be enumerated in a sequence by first writing down those in Λ_1, those in Λ_2 not in Λ_1, and so on.

(2): Suppose λ_n is *any* sequence of distinct proper values of T. Given any $\epsilon > 0$. By *Theorem 3*, $|\lambda_n| \geq \epsilon$ for at most finitely many n, hence there is an index N such that $|\lambda_n| < \epsilon$ whenever $n \geq N$. Thus, $\lambda_n \to 0$.

(3): This is clear from the proof of (1); incidentally, $|\mu_1| = \| T \|$. ∎

Henceforth, $\mu_1,\mu_2,\mu_3,\cdots$ denotes the sequence of distinct non-zero proper values of T, enumerated as in *Theorem 5*. Then, $P(\mu_k)$ denotes the projection whose range is the proper subspace $\mathfrak{N}_T(\mu_k)$; $P(0)$ is the projection whose range is the null space of T. The next results will show that $T = \sum_1^\infty \mu_k P(\mu_k)$ in a suitable sense.

Lemma. *If $\mathfrak{N}_1,\mathfrak{N}_2,\mathfrak{N}_3,\cdots$ is a sequence of closed linear subspaces of $\mathcal{H}$ such that $\mathfrak{N}_j \perp \mathfrak{N}_k$ whenever $j \neq k$, the following conditions are equivalent:*

(a) *The $\mathfrak{N}_k$ are a total family.*

(b) *The smallest closed linear subspace of $\mathcal{H}$ which contains every $\mathfrak{N}_k$ is $\mathcal{H}$ itself.*

(c) *Every* $x \in \mathcal{H}$ *is uniquely expressible in the form* $x = \sum_1^\infty x_k$, *with* $x_k \in \mathfrak{N}_k$ *and* $\sum_1^\infty \| x_k \|^2 < \infty$.

Proof.

Let $\mathfrak{S} = \{x \in \mathcal{H}: x \in \mathfrak{N}_k \text{ for some } k\}$; as shown in Corollary 2 of *Theorem III.6.3*, $\mathfrak{M} = \mathfrak{S}^{\perp\perp}$ is the smallest closed linear subspace containing every $\mathfrak{N}_k$. Since $\mathfrak{M}^\perp = \mathfrak{S}^{\perp\perp\perp} = \mathfrak{S}^\perp$, clearly $\mathfrak{M} = \mathcal{H}$ if and only if $\mathfrak{S}^\perp = \{\theta\}$. This proves the equivalence of (a) and (b).

Let P_k be the projection with range $\mathfrak{N}_k$.

(a) *implies* (c): Suppose first that $x_k \in \mathfrak{N}_k$ and $\sum_1^\infty \| x_k \|^2 < \infty$. It is clear from the Lemma of *Theorem II.7.1* that one can form the vector $x = \sum_1^\infty x_k = \lim_n \sum_1^n x_k$. Since $P_j x_k = \theta$ when $k \neq j$, $P_j\left(\sum_1^n x_k\right) = x_j$ whenever $n \geq j$, hence $P_j x = \lim_n P_j\left(\sum_1^n x_k\right) = x_j$.

Now, given any $x \in \mathcal{H}$, define $x_k = P_k x$. For each n, $Q_n = \sum_1^n P_k$ is a projection (see *Example VI.7.4*); since the x_k are mutually orthogonal, $\sum_1^n \| x_k \|^2 = \| \sum_1^n x_k \|^2 = \| \sum_1^n P_k x \|^2 = \| Q_n x \|^2 \leq \| x \|^2$, hence $\sum_1^\infty \| x_k \|^2 < \infty$. Form the vector $y = \sum_1^\infty x_k$. For each j, $P_j y = x_j = P_j x$, $P_j(y - x) = \theta$, $y - x \in \mathfrak{N}_j{}^\perp$; since the $\mathfrak{N}_j$ are total, $y - x = \theta$, $x = y = \sum_1^\infty x_k$. Since $x_j = P_j\left(\sum_1^\infty x_k\right)$, the uniqueness of such a representation is clear.

(c) *implies* (a): Suppose $x \perp \mathfrak{N}_j$ for all j. By (c), write $x = \sum_1^\infty x_k$ with $x_k \in \mathfrak{N}_k$. Since $x \in \mathfrak{N}_j{}^\perp$, $\theta = P_j x = x_j$ for all j, thus $x = \theta$. ∎

Definition 1. *If the conditions of the Lemma are fulfilled, one writes*

$$\mathcal{H} = \sum_{\oplus 1}^\infty \mathfrak{N}_k = \mathfrak{N}_1 \oplus \mathfrak{N}_2 \oplus \mathfrak{N}_3 \oplus \cdots.$$

More generally, let $\mathfrak{N}_1, \mathfrak{N}_2, \mathfrak{N}_3, \cdots$ be any sequence of closed linear subspaces, and $\mathfrak{N}$ the smallest closed linear subspace containing every $\mathfrak{N}_k$. If $x \perp \mathfrak{N}_k$ for all k, then $x \perp \mathfrak{N}$; for, writing $\mathfrak{S} = \{x\}$, one has $\mathfrak{N}_k \subset \mathfrak{S}^\perp$ for all k, hence $\mathfrak{N} \subset \mathfrak{S}^\perp$. It follows that $\mathfrak{N}_k$ is a total family of subspaces in the Hilbert space $\mathfrak{N}$. If moreover $\mathfrak{N}_j \perp \mathfrak{N}_k$ whenever $j \neq k$, one has $\mathfrak{N} = \sum_{\oplus 1}^\infty \mathfrak{N}_k$ in the sense of *Definition 1*. It is easy to see that $\sum_{\oplus 1}^\infty \mathfrak{N}_k = \mathfrak{N}_1 \oplus \sum_{\oplus 2}^\infty \mathfrak{N}_k$.

Theorem 6. *If μ_n is the sequence of distinct non-zero proper values of T, then*

(1) $\mathfrak{K} = \mathfrak{N}_T(0) \oplus \sum_{\oplus 1}^{\infty} \mathfrak{N}_T(\mu_k)$

$= \mathfrak{N}_{T^*}(0) \oplus \sum_{\oplus 1}^{\infty} \mathfrak{N}_{T^*}(\mu_k^*)$;

(2) $\mathfrak{N}_T(0) = \mathfrak{N}_{T^*}(0)$ *is the null space of T and T^*, and*

$$\overline{T(\mathfrak{K})} = \overline{T^*(\mathfrak{K})} = \sum_{\oplus 1}^{\infty} \mathfrak{N}_T(\mu_k) = \sum_{\oplus 1}^{\infty} \mathfrak{N}_{T^*}(\mu_k^*);$$

(3) $\| T - \sum_1^n \mu_k P(\mu_k) \| \to 0,$

$\| T^* - \sum_1^n \mu_k^* P(\mu_k) \| \to 0.$

Proof.

(1) results from *Theorem 1*, the *Lemma*, and the fact that $\mathfrak{N}_{T^*}(\mu^*) = \mathfrak{N}_T(\mu)$.

(2): Let $\mathfrak{N} = \sum_{\oplus 1}^{\infty} \mathfrak{N}_T(\mu_k)$. It is clear from (1) that $\mathfrak{N} = [\mathfrak{N}_T(0)]^{\perp} = [\mathfrak{N}_{T^*}(0)]^{\perp}$; thus, $\mathfrak{N} = \overline{T^*(\mathfrak{K})} = \overline{T(\mathfrak{K})}$ by *Theorem V.8.6.*

(3): Given $x \in \mathfrak{K}$, one has $x = P(0)x + \sum_1^{\infty} P(\mu_k)x$ by the proof of the *Lemma*. Then, $Tx = \theta + \sum_1^{\infty} \mu_k P(\mu_k)x = \lim_n \left[\sum_1^n \mu_k P(\mu_k)\right] x$. Since $\mu_n \to 0$, $\| T - \sum_1^n \mu_k P(\mu_k) \| \to 0$ by essentially the same argument as in *Theorem 2.1*. The remaining assertion follows from $\| S^* \| = \| S \|$. ∎

Corollary 1. *If T is a normal CC-operator, there exists a sequence of finite-dimensional operators T_n such that $\| T - T_n \| \to 0$.*

Corollary 2. *Notation as in the theorem. Given any vector $y \in \overline{T(\mathfrak{K})}$, there exists an expansion $y = \sum_1^{\infty} y_k$, where the y_k are mutually orthogonal, $\sum_1^{\infty} \| y_k \|^2 < \infty$, and $Ty_k = \mu_k y_k$.*

Exercises

1. Let $Tx_k = \mu_k x_k$, with notation as in *Example VII.1.5*. Then, T is CC if and only if $\mu_n \to 0$.

2. An operator T is CC if and only if there exists a sequence of finite-dimensional operators T_n such that $\| T_n - T \| \to 0$.

3. Let T and S be operators.

*(i) It can be shown that if T is normal, and $ST = TS$, then $ST^* = T^*S$. Granted this result; if S and T are normal, and $ST = TS$, then $S + T$ and ST are normal.

(ii) If T and S are hyponormal, and $ST^* = T^*S$, then $T - S$ is hyponormal.

(iii) If T is normal, S is hyponormal, and $ST = TS$, then $T - S$ is hyponormal.

4. If S and T are normal operators, $ST = TS$, and T is CC, then $S + T$ and ST are normal.

5. Let S and T be normal CC-operators, and assume, for the sake of simplicity, that both S and T are injective. Then, S and T are unitarily equivalent if and only if: for each scalar μ, the multiplicity of μ for T equals the multiplicity of μ for S.

6. (i) If S and T are similar operators, and $\mathfrak{N}_S(\mu)$ has finite dimension n, then $\mathfrak{N}_T(\mu)$ also has finite dimension n.

(ii) Suppose S and T are normal CC-operators. Assume for the sake of simplicity that S and T are injective. If S and T are similar, there exists a unitary operator U such that $U^*SU = T$; that is, S and T are unitarily equivalent.

*(iii) One can prove (ii) without the hypotheses of complete continuity and injectivity.

7. The most general normal CC-operator in $\mathfrak{X}$ can be obtained as follows. Let $\mathfrak{N}_0, \mathfrak{N}_1, \mathfrak{N}_2, \cdots$ be a total sequence of mutually orthogonal closed linear subspaces, with $\mathfrak{N}_k$ finite-dimensional for $k \geq 1$; let $\mu_1, \mu_2, \cdots$ be a sequence of scalars such that $\mu_n \to 0$; let $\mu_0 = 0$. There is a unique operator T such that $Tx = \mu_k x$ when $x \in \mathfrak{N}_k$, $(k = 0,1,2,\cdots)$.

8. Let T_n be a sequence of normal CC-operators such that $T_jT_k = T_kT_j$ for all j and k. Then, there exists a vector z which is proper for every T_k.

*9. *Theorem 3* holds for a CC linear mapping in a normed space.

**10. The analog of *Exercise 2* for Banach space is an open question.

11. Does there exist a hyponormal CC-operator which is not normal?

References, and Hints to the Exercises

[1] BANACH, S., *Théorie des opérations linéaires*, Warsaw, 1932.
[2] HALMOS, P. R., *Finite-dimensional vector spaces*, New York, 1958.
[3] HALMOS, P. R., *Introduction to Hilbert space and the theory of spectral multiplicity*, New York, 1951.
[4] HILLE, E. and R. S. PHILLIPS, *Functional analysis and semigroups*, Providence, 1957.
[5] LANDAU, E., *Differential and integral calculus*, New York, 1951.
[6] LOOMIS, L. H., *An introduction to abstract harmonic analysis*, New York, 1953.
[7] NAIMARK, M. A., *Normed rings*, Groningen, 1959.
[8] RIESZ, F. and B. SZ.-NAGY, *Leçons d'analyse fonctionelle*, Budapest, 1952.
[9] STONE, M. H., *Linear transformations in Hilbert space and their applications to analysis*, New York, 1932.
[10] TAYLOR, A. E., *Introduction to functional analysis*, New York, 1958.

CHAPTER I

§ 1. 2. See *Example 2*.
3. See *Example 7*.
§ 5. 6. $e^t = \cosh t + \sinh t$.
8. Show that each vector in $\mathfrak{M} + \mathfrak{N}$ is in [S], and vice versa.
§ 6. 4. Factor theorem.
5. (ii) See *Exercise 4.4*.
§ 7. 6. See Section 1.72 of [10] for the existence of a basis, also called a *Hamel basis*.