Bo Li
Department of Mathematics
University of California, San Diego

Elements of Hilbert Space and Operator Theory

with Application to Integral Equations

May 30, 2005

Preface

These lecture notes present elements of Hilbert space and the theory of linear operators on Hilbert space, and their application to integral equations.

Chapter 1 reviews vector spaces: bases, subspaces, and linear transforms. Chapter 2 covers the basics of Hilbert space. It includes the concept of Banach space and Hilbert space, orthogonality, complete orthogonal bases, and Riesz representation. Chapter 3 presents the core theory of linear operators on Hilbert space, in particular, the spectral theory of linear compact operators and Fredholm Alternatives. Chapter 4 is the application to integral equations. Some supplemental materials are collected in Appendicies.

The prerequisite for these lecture notes is calculus. Occasionally, advanced knowledges such as the Lebesgue measure and Lebesgue integration are used in the notes. The lack of such knowledge, though, will not prevent readers from grasping the essence of the main subjects.

The presentation in these notes is kept as simple, concise, and self-complete as possible. Many examples are given, some of them with great details. Exercise problems for practice and references for further reading are included in each chapter.

The symbol $\square$ is used to indicate the end of a proof.

These lecture notes can be used as a text or supplemental material for various kinds of undergraduate or beginning graduate courses for students of mathematics, science, and engineering.

Bo Li
Spring, 2005

Contents

1 Vector Spaces 1
1.1 Definition 1
1.2 Linear Dependence 5
1.3 Bases and Dimension 9
1.4 Subspaces 17
1.5 Linear Transforms 17
Exercises 17
References 18

2 Hilbert Spaces 19
2.1 Inner Products 19
2.2 Normed Vector Spaces 20
2.3 Hilbert Spaces 20
2.4 Orthogonality 20
2.5 Complete Orthogonal Bases 20
2.6 Riesz Representation 20
Exercises 20
References 20

3 Operator Theory 21
3.1 Bounded Linear Operators 21
3.2 Self-Adjoint Operators 31
3.3 Compact Operators 31
Exercises 32
References 32

4 Integral Equations 33
4.1 Examples and Classification of Integral Equatons 33
4.2 Integral Equations of the Second Kind: Successive Iterations 33
4.3 Fredholm Theory of Integral Equations of the Second Kind 33
4.4 Singular Integral Equations 33

Problems 33
References 33

A Appendix 35
A.1 Zorn's Lemma 35
A.2 Cardinality 35

1

Vector Spaces

Notation:

$\mathbb{N}$: the set of all positive integers.
$\mathbb{Z}$: the set of all integers.
$\mathbb{R}$: the set of real numbers.
$\mathbb{C}$: the set of complex numbers.

1.1 Definition

We denote by $\mathbb{F}$ either $\mathbb{C}$ or $\mathbb{R}$.

Definition 1.1 (vector space). *A nonempty set X is called a* vector space (*or* linear space) *over $\mathbb{F}$, if there are two mappings $X \times X \to \mathbb{F}$, called* addition *and denoted by $(u, v) \to u + v$ for any $u, v \in X$, and $\mathbb{F} \times X \to X$, called* scalar multiplication *and denoted by $(\alpha, u) \to \alpha u$ or $\alpha \cdot u$ for any $\alpha \in \mathbb{F}$ and any $u \in X$, that satisfy the following properties:*

(1) *Commutative law for addition.*

$$u + v = v + u \qquad \forall u, v \in X; \tag{1.1}$$

(2) *Associative law for addition.*

$$(u + v) + w = u + (v + w) \qquad \forall u, v, w \in X; \tag{1.2}$$

(3) *A zero-vector. There exists $o \in X$, called a* zero-vector*, such that*

$$u + o = u \qquad \forall u \in X; \tag{1.3}$$

(4) *Negative vectors. For each $u \in X$, there exists $-u \in X$, called a* negative vector *of u, such that*

$$u + (-u) = o \qquad \forall u \in X; \tag{1.4}$$

(5) *Associative law for scalar multiplication.*

$$(\alpha\beta)u = \alpha(\beta u) \qquad \forall \alpha, \beta \in \mathbb{F},\ \forall u \in X; \tag{1.5}$$

(6) *Right distributive law.*

$$(\alpha + \beta)u = \alpha u + \beta u \qquad \forall \alpha, \beta \in \mathbb{F},\ \forall u \in X; \tag{1.6}$$

(7) *Left distributive law.*

$$\alpha(u + v) = \alpha u + \alpha v \qquad \forall \alpha \in \mathbb{F},\ \forall u, v \in X; \tag{1.7}$$

(8) *The unity of* $\mathbb{F}$.

$$1 \cdot u = u \qquad \forall u \in X. \tag{1.8}$$

Remark 1.2. Throughout these notes, vector spaces are always defined with respect to the scalar field $\mathbb{F}$, which is either $\mathbb{R}$ or $\mathbb{C}$.

Remark 1.3. Let X be a vector space.

(1) Suppose $o' \in X$ is also a zero-vector:

$$u + o' = u \qquad \forall u \in X. \tag{1.9}$$

Then, by (1.9) and (1.3),

$$o' = o + o' = o' + o = o'.$$

Hence, the zero-vector in X is unique.

(2) Let $u \in X$. If both $v_1 \in X$ and v_2 are negative vectors of u, then

$$v_1 = v_1 + o = v_1 + (u + v_2) = (v_1 + u) + v_2 = (u + v_1) + v_2 = o + v_2 = v_2 + o = v_2.$$

Thus, the negative vector of $u \in X$ is unniue. We shall denote

$$u - v = u + (-v) \qquad \forall u, v \in X.$$

(3) Let $n \in \mathbb{N}$ with $n \geq 2$ and $u_1, \cdots, u_n \in X$. We define recursively

$$u_1 + \cdots + u_n = (u_1 + \cdots + u_{n-1}) + u_n.$$

It is easy to see that the sum $u_1 + \cdots + u_n$ is uniquely defined regardless of the order of vectors in the sum. We shall denote

$$\sum_{k=1}^{n} u_k = u_1 + \cdots + u_n.$$

For convenience, we also denote

$$\sum_{k=1}^{1} u_k = u_1.$$

Proposition 1.4. *Let X be a vector space. Let $n \in \mathbb{N}$ with $n \geq 2$.*

(1) *For any $\alpha_k \in \mathbb{F}$ $(k = 1, \cdots, n)$ and any $u \in X$,*

$$\left(\sum_{k=1}^{n} \alpha_k \right) u = \sum_{k=1}^{n} \alpha_k u. \tag{1.10}$$

(2) *For any $\alpha \in \mathbb{F}$ and any $u_k \in X$ $(k = 1, \cdots, n)$,*

$$\alpha \sum_{k=1}^{n} u_k = \sum_{k=1}^{n} \alpha u_k. \tag{1.11}$$

(3) *For any $\alpha \in \mathbb{F}$ and any $u \in X$,*

$$\alpha u = o \quad \text{if and only if} \quad \alpha = 0 \text{ or } u = o. \tag{1.12}$$

Proof. (1) If $n = 2$, then (1.11) follows from (1.6). Suppose (1.11) is true for a general $n \geq 2$. Let $\alpha_{n+1} \in \mathbb{F}$. Then, by (1.6) and (1.11),

$$\begin{aligned} \left(\sum_{k=1}^{n+1} \alpha_k \right) u &= \left[\left(\sum_{k=1}^{n} \alpha_k \right) + \alpha_{n+1} \right] u \\ &= \left(\sum_{k=1}^{n} \alpha_k \right) u + \alpha_{n+1} u \\ &= \left(\sum_{k=1}^{n} \alpha_k u \right) + \alpha_{n+1} u \\ &= \sum_{k=1}^{n+1} \alpha_k u. \end{aligned}$$

Thus, by the Principle of Mathematical Induction,[1] (1.11) holds true for any $n \geq 2$.

(2) This part can be proved similarly.

(3) Let $u \in X$ and $\alpha \in \mathbb{F}$. By (1.6),

$$0u = (0 + 0)u = 0u + 0u.$$

Adding $-0u$, we get by (1.4) that

$$0u = o.$$

Similarly, by (1.7) and (1.3),

$$\alpha o = \alpha(o + o) = \alpha o + \alpha o$$

[1] A statement $P = P(n)$ that depends on an integer $n \in \mathbb{N}$ is true for all $n \in \mathbb{N}$, if: (1) $P(1)$ is true; and (2) $P(n + 1)$ is true provided that $P(n)$ is true.

Adding $-\alpha o$, we obtain by (1.4) that

$$\alpha o = o. \tag{1.13}$$

This proves the "if" part of (1.12).

Now, assume that $\alpha u = o$ and $\alpha \neq 0$. By (1.8) and (1.13),

$$u = 1u = \left(\frac{1}{\alpha}\alpha\right) u = \frac{1}{\alpha}(\alpha u) = \frac{1}{\alpha} o = o.$$

This proves the "only if" part completes the proof. □

Example 1.5. Let $n \in \mathbb{N}$ and

$$\mathbb{F}^n = \left\{ \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} : \xi_k \in \mathbb{F}, k = 1, \cdots, n \right\}.$$

Define

$$\begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} = \begin{bmatrix} \xi_1 + \eta_1 \\ \vdots \\ \xi_n + \eta_n \end{bmatrix} \qquad \forall \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}, \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} \in \mathbb{F}^n,$$

$$\alpha \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} = \begin{bmatrix} \alpha\xi_1 \\ \vdots \\ \alpha\xi_n \end{bmatrix} \qquad \forall \alpha \in \mathbb{F}, \forall \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \in \mathbb{F}^n.$$

It is easy to verify that with these two operations the set $\mathbb{F}^n$ is a vector space. The zero-vector is

$$o = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix},$$

and the negative vector is given by

$$- \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} = \begin{bmatrix} -\xi_1 \\ \vdots \\ -\xi_n \end{bmatrix} \qquad \forall \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \in \mathbb{F}^n.$$

Example 1.6. Let

$$\mathbb{F}^\infty = \{\{\xi_k\}_{k=1}^\infty : \xi_k \in \mathbb{F}, k = 1, \cdots\}$$

be the set of ordered, infinite sequences in $\mathbb{F}$. Define for any $\{\xi_k\}_{k=1}^\infty, \{\eta_k\}_{k=1}^\infty \in \mathbb{F}^\infty$ and $\alpha \in \mathbb{F}$ the addition and scalar multiplication

$$\{\xi_k\}_{k=1}^{\infty} + \{\eta_k\}_{k=1}^{\infty} = \{\xi_k + \eta_k\}_{k=1}^{\infty},$$
$$\alpha\{\xi_k\}_{k=1}^{\infty} = \{\alpha\xi_k\}_{k=1}^{\infty}.$$

Then, the set $\mathbb{F}^{\infty}$ is a vector space. Its zero-vector is the unique infinite sequence with each term being zero. For $\{\xi_k\}_{k=1}^{\infty}$, the negative vector is

$$-\{\xi_k\}_{k=1}^{\infty} = \{-\xi_k\}_{k=1}^{\infty}.$$

Example 1.7. Let $m, n \in \mathbb{N}$. Denote by $\mathbb{F}^{m\times n}$ the set of of all $m \times n$ matrices with entries in $\mathbb{F}$. With the usual matrix addition and scalar multiplication, $\mathbb{F}^{m\times n}$ is a vector space. The zero-vector is the $m \times n$ zero matrix .

Example 1.8. The set of all polynomials with coefficients in $\mathbb{F}$,

$$\mathcal{P} = \left\{ \sum_{k=0}^{n} a_k z^k : a_k \in \mathbb{F}, k = 0, \cdots, n; n \geq 0 \text{ is an integer} \right\},$$

is a vector space with the usual addition of polynomials and the scalar multiplication of a number in $\mathbb{F}$ and a polynomial in $\mathcal{P}$. The zero-vector is the zero polynomial.

For each integer $n \geq 0$, the set of all polynomials of degree $\leq n$,

$$\mathcal{P}_n = \left\{ \sum_{k=0}^{n} a_k z^k : a_k \in \mathbb{F}, k = 0, \cdots, n \right\},$$

is a vector space.

Example 1.9. Let $a, b \in \mathbb{R}$ with $a < b$. Denote by $C[a, b]$ the set of all continuous functions from $[a, b]$ to $\mathbb{F}$. For any $u = u(x), v = v(x) \in C[a, b]$ and $\alpha \in \mathbb{F}$, define $u + v$ and αu by

$$(u + v)(x) = u(x) + v(x) \qquad \forall x \in [a, b],$$
$$(\alpha u)(x) = \alpha u(x) \qquad \forall x \in [a, b].$$

Clearly, $u+v, \alpha u \in C[a, b]$. Moreover, with these two operations, the set $C[a, b]$ is a vector space over $\mathbb{F}$.

1.2 Linear Dependence

Definition 1.10 (linear combination and linear dependence). *Let X be a vector space over $\mathbb{F}$ and $n \in \mathbb{N}$.*

(1) *If $u_1, \cdots, u_n \in X$ and $\alpha_1, \cdots, \alpha_n \in \mathbb{F}$, then*

$$\sum_{k=1}^{n} \alpha_k u_k$$

is a linear combination *of $u_1, \cdots, u_n$ with* coefficients *$\alpha_1, \cdots, \alpha_n$ in $\mathbb{F}$.*

(2) *Finitely many vectors* $u_1, \cdots, u_n \in X$ *are* linearly dependent *in* $\mathbb{F}$*, if there exist* $\alpha_1, \cdots, \alpha_n \in \mathbb{F}$*, not all zero, such that the linear combination*

$$\sum_{k=1}^{n} \alpha_k u_k = o.$$

They are linearly independent*, if they are not linearly dependent.*

(3) *Let* Y *be a nonempty subset of* X*. If there exist finitely many vectors in* Y *that are linearly dependent in* $\mathbb{F}$*, then* Y *is* linearly dependent *in* $\mathbb{F}$*. If any finitely many vectors in* Y *are linearly independent in* $\mathbb{F}$*, then* Y *is* lienarly independent *in* $\mathbb{F}$*.*

Example 1.11. Let $n \in \mathbb{N}$,

$$\xi^{(k)} = \begin{bmatrix} \xi_1^{(k)} \\ \vdots \\ \xi_n^{(k)} \end{bmatrix} \in \mathbb{F}^n, \qquad k = 1, \cdots, n,$$

and

$$A = \left(\xi_j^{(k)}\right)_{j,k=1}^{n} = \left[\xi^{(1)} \cdots \xi^{(n)}\right] = \begin{bmatrix} \xi_1^{(1)} & \cdots & \xi_1^{(n)} \\ \vdots & & \vdots \\ \xi_n^{(1)} & \cdots & \xi_n^{(n)} \end{bmatrix} \in \mathbb{F}^{n \times n}.$$

The linear combination of the n vectors $\xi^{(1)}, \cdots, \xi^{(n)}$ in $\mathbb{F}^n$ with coefficients $u_1, \cdots, u_n \in \mathbb{F}$ is given by

$$\sum_{k=1}^{n} u_k \xi^{(k)} = Au \qquad \text{with } u = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}.$$

Now, the n vectors $\xi^{(1)}, \cdots, \xi^{(n)}$ are linearly dependent, if and only if the linear system $Au = o$ has a nontrivial solution, if and only if the matrix A is non-singular.

Example 1.12. By the Fundamental Theorem of Algebra, any polynomial of degree $n \geq 1$ has at most n roots. Thus, if for some $n \in \mathbb{N}$ and $a_0, \cdots, a_n \in \mathbb{F}$,

$$\sum_{k=0}^{n} a_k z^k = 0$$

identically, then all $a_k = 0$ $(k = 0, \cdots, n)$. Conseqently, the infinite set $\{p_k\}_{k=0}^{\infty}$, with

$$p_k(z) = z^k \qquad (k = 0, \cdots),$$

is linearly independent in the vector space $\mathcal{P}$.

Example 1.13. Let $a, b \in \mathbb{R}$ with $a < b$ and consider the vector space $C[a, b]$ over $\mathbb{F}$. For each $r \in (a, b)$, define $f_r \in C[a, b]$ by

$$f_r(x) = e^{rx}, \qquad a \leq x \leq b.$$

We show that the infinite subset $\{f_r\}_{a<r<b}$ is linearly independent in $C[a, b]$.

Let $n \in \mathbb{N}$ and $a < r_1 < \cdots < r_n < b$. Suppose

$$\sum_{k=1}^{n} c_k f_{r_k} = 0 \qquad \text{in } C[a, b]$$

for some $c_1, \cdots, c_n \in \mathbb{R}$. Let $m \in \mathbb{N}$ and take the mth derivative to get

$$\sum_{k=1}^{n} c_k r_k^m e^{r_k x} = 0 \qquad \forall x \in [a, b].$$

Fix $x = a$ and divide both sides of the equation by r_n^m to get

$$\sum_{k=1}^{n} c_k \left(\frac{r_k}{r_n}\right)^m e^{r_k a} = 0 \qquad \forall m = 1, \cdots .$$

Now, taking the limit as $m \to \infty$, we obtain $c_n = 0$. Since each $f_r \in C[a, b]$ $(a < r < b)$ is linearly independent in $C[a, b]$, a simple argument of induction thus implies that any finitely many members of the set $\{f_r\}_{a<r<b}$ is linearly independent in $C[a, b]$. Consequently, the infinite set $\{f_r\}_{a<r<b}$ itself is linearly independent.

Proposition 1.14. *Let X be a vector space and $\emptyset \neq Z \subseteq Y \subseteq X$.*

(1) *If Y is linearly independent, then Z is also linearly indenpendent.*
(2) *If Z is linearly dependent, then Y is also linearly dependent.*

The proof of this proposition is left as an exercise, cf. Problem 1.3.

Proposition 1.15. *Let X be a vector space.*

(1) *A single vector $u \in X$ is linearly independent if and only if $u \neq o$.*
(2) *Let $n \in \mathbb{N}$ with $n \geq 2$. Then, any n vectors in X are are linearly dependendet if and only if one of them if a linear combination of the others.*

Proof. (1) If $u = o$, then $1 \cdot u = o$. Hence, by Definition 1.10, u is linearly dependent. Conversely, if $u \neq o$, then, for any $\alpha \in \mathbb{F}$ such that $\alpha u = o$, we must have $\alpha = 0$ by Part (3) of Proposition 1.4. This means that u is linearly independent.

(2) Let $u_1, \cdots, u_n \in X$. Suppose they are linearly dependent. Then, there exist $\alpha_1, \cdots, \alpha_n \in \mathbb{F}$, not all zero, such that

$$\sum_{k=1}^{n} \alpha_k u_k = o.$$

Assume $\alpha_j \neq o$ for some $j : 1 \leq j \leq n$. Then,

$$u_j = \sum_{k=1, k\neq j}^{n} \left(-\frac{\alpha_k}{\alpha_j}\right) u_k,$$

i.e., u_j is a linear combination of the others.

Suppose now one of $u_1, \cdots, u_n$, say, u_j for some $j : 1 \leq j \leq n$, is a linear combination of the others. Then, there exist $\beta_1, \cdots, \beta_{j-1}, \beta_{j+1}, \cdots, \beta_n \in \mathbb{F}$ such that

$$u_j = \sum_{k=1, k\neq j}^{n} \beta_k u_k,$$

Setting $\beta_j = -1 \in \mathbb{F}$, we have the linear combination

$$\sum_{k=1}^{n} \beta_k u_k = o$$

in which at least one coefficient, $\beta_j = -1$, is not zero. Thus, $u_1, \cdots, u_n$ are lienarly dependent. □

Definition 1.16 (span). *Let X be a vector space and $E \subseteq X$. The* span *of E, denoted by span E, is the set of all the linear combinations of elements in E:*

$$\mathit{span}\, E = \left\{ \sum_{k=1}^{n} \alpha_k u_k : \alpha_k \in \mathbb{F}, u_k \in E, k = 1, \cdots, n; n \in \mathbb{N} \right\}.$$

Theorem 1.17. *Let $u_1, \cdots, u_n$ be n linearly independent vectors in a vector space X. Then, any $n+1$ vectors in span$\{u_1, \cdots, u_n\}$ are linearly dependent.*

Proof. Let

$$v_k \in \text{span}\,\{u_1, \cdots, u_n\}, \qquad k = 1, \cdots, n+1.$$

If the n vectors $v_1, \cdots, v_n$ are linearly dependent, then, by Proposition 1.14, the $n+1$ vectors $v_1, \cdots, v_n, v_{n+1}$ are also linearly dependent.

Assume $v_1, \cdots, v_n$ are linearly independent. Since v_k $(1 \leq k \leq n)$ is a linear combination of $u_1, \cdots, u_n$, there exist $\alpha_j^{(k)} \in \mathbb{F}$ $(1 \leq j, k \leq n)$ such that

$$v_j = \sum_{k=1}^{n} \alpha_j^{(k)} u_k, \qquad j = 1, \cdots, n. \tag{1.14}$$

Thus, for any $\xi_1, \cdots, \xi_n \in \mathbb{F}$, we have

$$\sum_{j=1}^{n} \xi_j v_j = \sum_{k=1}^{n} \left(\sum_{j=1}^{n} \alpha_j^{(k)} \xi_j \right) u_k.$$

Since both the groups of vectors $v_1, \cdots, v_n$ and $u_1, \cdots, u_n$ are linearly independent, all $\xi_j = 0$ $(1 \leq j \leq n)$, if and only if

$$\sum_{j=1}^{n} \xi_j v_j = o,$$

if and only if

$$\sum_{j=1}^{n} \alpha_j^{(k)} \xi_j = 0, \qquad k = 1, \cdots, n.$$

Hence, the matrix

$$A = \left(\alpha_j^{(k)} \right)_{j,k=1}^{n}$$

is nonsingular. Therefore, by (1.14), each u_k $(1 \leq k \leq n)$ is a linear combination of $v_1, \cdots, v_n$. But, v_{n+1} is a linear combination of $u_1, \cdots, u_n$. Thus, v_{n+1} is also a linear combination of $v_1, \cdots, v_n$. By Proposition 1.15, this means that $v_1, \cdots, v_n, v_{n+1}$ are linearly dependent. □

1.3 Bases and Dimension

Definition 1.18 (basis). *Let X be a vector space. A nonempty subset $B \subseteq X$ is a* basis *of X if it is linearly independent and $X = \operatorname{span} B$.*

Lemma 1.19. *Let X be a vector space. Let T be a linearly independent subset of X. If $T \cup \{u\}$ is linearly dependent for any $u \in X \setminus T$, then $X = \operatorname{span} T$.*

Proof. Let $u \in X$. If $u = o$ or $u \in T$, then obviously $u \in \operatorname{span} T$. Assume $u \neq o$ and $u \in X \setminus T$. Then $T' := T \cup \{u\}$ is linearly dependent. Hence, there are finitely many, linearly dependent vectors $v_1, \cdots, v_m$ in T' with $m \in \mathbb{N}$.

The vector u must be one of those vectors, for otherwise $v_1, \cdots, v_m \in T$ would be linearly dependent, in contradiction with the fact that T is linearly independent. Without loss of generality, we may assume that $v_1 = u$. Since $u \neq o$, by Proposition 1.15 the single vector u is linearly indendent. Thus, we must have $m \geq 2$.

Since $v_1(= u), v_2, \cdots, v_m$ are linearly dependent, there exist $\alpha_1, \cdots, \alpha_m \in \mathbb{F}$, not all zero, such that

$$\sum_{k=1}^{m} \alpha_k v_k = o$$

We have $\alpha_1 \neq 0$, since $v_2, \cdots, v_m \in T$ are linearly independent. Therefore,

$$u = v_1 = \sum_{k=2}^{m} \left(-\frac{\alpha_k}{\alpha_1} \right) v_k \in \operatorname{span} T.$$

Thus, $X = \operatorname{span} T$. □

A vector space that consists of only a zero-vector is called a *trivial* vector space.

Theorem 1.20. *Let X be a nontrivial vector space. Suppose S and T are two subsets of X that satisfy:*

(1) $X = \mathit{span}\, S$;
(2) R *is linearly independent;*
(3) $R \subseteq S$.

Then, there exists a basis, T, of X such that

$$R \subseteq T \subseteq S.$$

Proof. [2] Denote by $\mathcal{A}$ the collection of all the linearly independent subsets $E \subseteq X$ such that $R \subseteq E \subseteq S$. Since $R \in \mathcal{A}$, $\mathcal{A} \neq \emptyset$. For any $E_1, E_2 \in \mathcal{A}$, define $E_1 \leq E_2$ if $E_1 \subseteq E_2$. It is obvious that $(\mathcal{A}, \leq)$ is partially ordered.

Let $\mathcal{C} \subseteq \mathcal{A}$ be a chain of $\mathcal{A}$. Define

$$U = \bigcup_{C \in \mathcal{C}} C. \tag{1.15}$$

Clearly, $R \subseteq U \subseteq S$. Consider any finitely many vectors $u_1, \cdots, u_n \in U$. By (1.15), there exist $C_1, \cdots, C_n \in \mathcal{C}$ such that $u_k \in C_k$, $k = 1, \cdots, n$.. Since $\mathcal{C}$ is a chain, there exists $j : 1 \leq j \leq n$ such that $C_k \subseteq C_j$ for all $k = 1, \cdots, n$. Thus, $u_k \in C_j$ for all $k = 1, \cdots, n$. Since $C_j \in \mathcal{A}$ is linearly independent, the n vectors $u_1, \cdots, u_n$ are therefore linearly independent. Consquently, U is linearly independent. Therefore, $U \in \mathcal{A}$. Obviously, $C \leq U$ for any $C \in \mathcal{C}$. Thus, $U \in \mathcal{A}$ is an upper bound of the chain $\mathcal{C}$. By Zorn's Lemma, $\mathcal{A}$ has a maximal element $T \in \mathcal{A}$.

By the definition of $\mathcal{A}$, T is linearly independent. Let $u \in X \setminus T$. Then, $T' := T \cup \{u\}$ must be linearly dependent. Otherwise, $T' = T$, since $T \subseteq T'$ and T is a maximal element in $\mathcal{A}$. This is impossible, since $u \notin T$. Now, by Lemma 1.19, $X = \operatorname{span} T$. Therefore, T is a basis of X. Since $T \in \mathcal{A}$, $R \subseteq T \subseteq S$. □

Corollary 1.21 (existence of basis). *Any nontrivial vector space X has a basis.*

Proof. Since X is nontrivial, there exists $x_0 \in X$ with $x_0 \neq o$. Let $R = \{x_0\}$ and $S = X$. Then, it follows from Theorem 1.20 that there exists a basis of X. □

Corollary 1.22. *Let X be a nontrival vector space.*

(1) *If $R \subset X$ is linearly independent, then R can be extended to a basis of X.*

---

[2] See Section A.1 for Zorn's Lemma and the related concepts that are used in the proof.

(2) *If $X = \operatorname{span} S$ for some subset $S \subseteq X$, then S contains a basis of X.*

Proof. (1) This follows from Theorem 1.20 by setting $S = X$.

(2) Since X is nontrivial and $X = \operatorname{span} S$, there exists $u \in S$ with $u \neq o$. Let $R = \{u\}$. The assertion follows then from Theorem 1.20. □

Theorem 1.23. *If B is a basis of a vector space X, then any nonzero vector in X can be expressed* uniquely *as a linear combination of finitely many vectors in B.*

Proof. Let $u \in X$ be a nonzero vector. Since B is a basis of X,

$$u = \sum_{k=1}^{p} \alpha_k v_k = \sum_{k=1}^{q} \beta_k w_k \tag{1.16}$$

for some $p, q \in \mathbb{N}$, $\alpha_k \in \mathbb{F}$ with $\alpha_k \neq 0$ $(1 \leq k \leq p)$, $\beta_k \in \mathbb{F}$ with $\beta_k \neq 0$ $(1 \leq k \leq q)$, and $v_k \in B$ $(1 \leq k \leq p)$ and $w_k \in B$ $(1 \leq k \leq q)$.

If

$$\{v_k\}_{k=1}^p \cap \{w_k\}_{k=1}^q = \emptyset,$$

then (1.16) implies that

$$\sum_{k=1}^{p} \alpha_k v_k - \sum_{k=1}^{q} \beta_k w_k = o,$$

and hence that $v_1, \cdots, v_p, w_1, \cdots, w_q$ are linearly dependent. This is impossible, since all these vectors are in B and B is linearly independent.

Assume now

$$\{u_1, \cdots, u_r\} = \{v_k\}_{k=1}^p \cap \{w_k\}_{k=1}^q$$

for some $r \in \mathbb{N}$ with $r \leq \min(p, q)$. By rearranging the indecies, we can assume, without loss of generality, that

$$v_k = w_k = u_k, \qquad k = 1, \cdots, r.$$

Thus, by (1.16),

$$\sum_{k=1}^{r} (\alpha_k - \beta_k) u_k + \sum_{k=r+1}^{p} \alpha_k v_k - \sum_{k=r+1}^{q} \beta_k v_k = o.$$

Since the distinct vectors $u_1, \cdots, u_r, v_{r+1}, \cdots, v_p, w_{r+1}, \cdots, w_q$ are all in the basis B, they are linearly independent. Consequently, by the assumption that all the coefficients $\alpha_k \neq 0$ $(1 \leq k \leq p)$ and $\beta_k \neq 0$ $(1 \leq k \leq q)$, we must have that $r = p = q$ and $\alpha_k = \beta_k$, $k = 1, \cdots, r$. □

The following theorem states that the "number of elements" in each basis of a vector space is the same. The *cardinal number* or *cardinality* of an infinite set is the generalizataion of the concept of the "number of elements" in a finite set. See Section A.2 for the concept and some results of the cardinality of sets.

Theorem 1.24 (cardinality of bases). *Any two bases of a nontrivial vector space have the same cardinal number.*

Proof. Let X be a nontrivial vector space. Let R and S be any two bases of X.

Case 1. R is finite. Let $\operatorname{card} R = n$ for some $n \in \mathbb{N}$. Then, by Theorem 1.17, all $n+1$ vectors in X are linearly dependent. Hence, S is finite and

$$\operatorname{card} S \le n = \operatorname{card} T.$$

A similar argument than leads to

$$\operatorname{card} T \le \operatorname{card} S.$$

Therefore,

$$\operatorname{card} T = \operatorname{card} S.$$

Case 2. R is infinite. By *Case 1*, S is also infinite. Let

$$\mathcal{S}_k = \{A \subset S : \operatorname{card} A = k\}, \qquad k = 1, \cdots,$$

and

$$\mathcal{G} = \bigcup_{k=1}^{\infty} \mathbb{F}^k \times \mathcal{S}_k,$$

where

$$\mathbb{F}^k \times \mathcal{S}_k = \left\{(\xi, A) : \xi \in \mathbb{F}^k, A \in \mathcal{S}_k\right\}, \qquad k = 1, \cdots.$$

Let $u \in R$. Since S is a bisis of X, by Theorem 1.23, there exist a unique $k \in \mathbb{N}$, a unique vector $\xi = (\xi_1, \cdots, \xi_k) \in \mathbb{F}^k$, and a unique set of k vectors $A = \{v_1, \cdots, v_k\} \subset S$, such that

$$u = \sum_{j=1}^{k} \xi_j v_j.$$

Let

$$\varphi(u) = (\xi, A).$$

This defines a mapping $\varphi : R \to \mathcal{G}$. By Theorem 1.23, this mapping is injective.

For each $k \in \mathbb{N}$, let

$$R_k = \{u \in R : u \text{ is a linear combination of exactly } k \text{ vectors in } S\}.$$

It is clear that

$$R = \bigcup_{k=1}^{\infty} R_k,$$

and

$$R_j \bigcap R_k = \emptyset \qquad \text{if } j \neq k.$$

Since both R and S are linearly independent, by Theorem 1.17, there exist at most k vectors in $\mathbb{F}^k$ such that

Thus, $\varphi(R_k)$.

to be completed

□

Definition 1.25 (dimension). *The dimension of a nontrivial vector space is the cardinality of any of its bases. The dimension of a trivial vector space is* 0.

The dimension of a vector space X is denoted by $\dim X$.

Remark 1.26. Let X be a vector space.

(1) If $\dim X < \infty$, then X is said to be *finitely dimensional.*
(2) If $\dim X = \infty$, then X is said to be *infinitely dimensional.*
  (a) If $\dim X = \aleph_0$, then X is said to be *countably infinitely dimensional.*
  (b) If $\dim X > \aleph_0$, then X is said to be *uncountably infinitely dimensional.*

Example 1.27. Let $n \in \mathbb{N}$ and

$$e_k = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow k\text{th-commponent}, \qquad k = 1, \cdots, n.$$

We call e_k the kth coordinate vector of $\mathbb{F}^n$ $(1 \le k \le n)$. It is easy to see that $e_1, \cdots, e_n$ are linearly independent. Moreover,

$$\xi = \sum_{k=1}^{n} \xi_k e_k \qquad \forall \xi = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \in \mathbb{F}^n.$$

Thus,

$$\mathbb{F}^n = \operatorname{span}\{e_1, \cdots, e_n\}.$$

Consequently, $\{e_1, \cdots, e_n\}$ is a basis of $\mathbb{F}^n$. Clearly,

$$\dim \mathbb{F}^n = n.$$

Example 1.28. For each integer $k \ge 0$, let $p_k(z) = z^k$. Then, it is easy to see that $\{p_0, p_1, \cdots, p_n\}$ is a basis of the vector space $\mathcal{P}_n$ of all polynomials of degrees $\le n$ and that the infinite set $\{p_0, p_1, \cdots\}$ is a basis of the vector space $\mathcal{P}$ of all polynomials. Obviously,

$$\dim \mathcal{P}_n = n + 1,$$
$$\dim \mathcal{P} = \aleph_0.$$

Example 1.29. Consider the vector space $C[a, b]$ over $\mathbb{F}$. By Example 1.13, there exist uncountably infinitely many vectors in $C[a, b]$ that are linearly independent. Thus, by Corollary 1.22, $C[a, b]$ is uncountably infinitely dimensional.

Definition 1.30. *Two vector spaces X and Y are* isomorphic*, if there exists a bijective mapping $\varphi : X \to Y$ such that*

$$(1) \qquad \varphi(u + v) = \varphi(u) + \varphi(v) \qquad \forall u, v \in X; \tag{1.17}$$

$$(2) \qquad \varphi(\alpha u) = \alpha\varphi(u) \qquad \forall \alpha \in \mathbb{F}, \forall u \in X. \tag{1.18}$$

Such a mapping $\varphi : X \to Y$ is called an isomorphism *between the vector spaces X and Y.*

Proposition 1.31. *Let $\varphi : X \to Y$ be an isomorphism between two vector spaces X and Y. Then,*

(1) *For any $\alpha_1, \cdots, \alpha_n \in \mathbb{F}$ and any $u_1, \cdots, u_n \in X$,*

$$\varphi\left(\sum_{k=1}^{n} \alpha_k u_k\right) = \sum_{k=1}^{n} \alpha_k \varphi(u_k);$$

(2) *If o is the zero-vector of X, then $\varphi(o)$ is the zero-vector of Y;*
(3) *For any $u \in X$,*

$$\varphi(-u) = -\varphi(u).$$

The proof of this proposition is left as an exercise problem, cf. Problem 1.4.

Theorem 1.32 (isomorphism). *Two vector spaces are isomorphic if and only if they have the same dimension.*

Proof. Let X and Y be two vector spaces.

(1) Assume X and Y are isomorphic. Then, there exists a bijective mapping $\varphi : X \to Y$ that satisfies (1.17) and (1.18).

If X is a trivial vector space, then

$$Y = \varphi(X) = \{\varphi(u) : u \in X\}$$

is also trivial. Thus, in this case,

$$\dim X = \dim Y = 0.$$

Assume now X is nontrivial. Then, there exist at least two vectors $u_1, u_2 \in X$ with $u_1 \neq u_2$. Thus, $\varphi(u_1), \varphi(u_2) \in Y$ and $\varphi(u_1) \neq \varphi(u_2)$. Hence, Y is also nontrivial. Let $A \neq \emptyset$ be a basis of X. Let

$$B = \varphi(A) = \{\varphi(u) : u \in A\} \subseteq Y.$$

Clearly, $B \neq \emptyset$. We show that B is in fact a basis of Y.

Let $\varphi(u_k)$ $(k = 1, \cdots, n)$ be n distinct vectors in B with $n \in \mathbb{N}$ and $n \leq \dim X$. Since $\varphi : X \to Y$ is bijective, $u_1, \cdots, u_n \in A$ are distinct. Let $\alpha_k \in \mathbb{F}$ $(k = 1, \cdots, n)$ be such that

$$\sum_{k=1} \alpha_k \varphi(u_k) = o,$$

the zero-vector in Y. Then, by Part (1) of Proposition 1.31,

$$\varphi\left(\sum_{k=1} \alpha_k u_k\right) = o,$$

the zero-vector in Y. By the fact that $\varphi : X \to Y$ is bijective and Part (2) of Proposition 1.31,

$$\sum_{k=1} \alpha_k u_k = o,$$

the zero-vector in X. Since $u_1, \cdots, u_n \in A$ are linearly independent, we have

$$\alpha_k = 0, \qquad k = 1, \cdots, n.$$

Thus, $\varphi(u_1), \cdots, \varphi(u_n)$ are linearly independent. Consequently, B is linearly independent.

Let $v \in Y$. Since $\varphi : X \to Y$ is bijective, there exists a unique $u \in X$ such that $v = \varphi(u)$. Since A is a basis of X, $u \in \operatorname{span} A$. By Proposition 1.31,

$$v = \varphi(u) \in \operatorname{span} \varphi(A) = \operatorname{span} B.$$

Thus,

$$Y = \operatorname{span} B.$$

Therefore, B is a basis of Y.

Clearly, the restriction $\varphi|_A : A \to B$ is a bijective mapping. Thus,

$$\operatorname{card} A = \operatorname{card} B,$$

cf. Section A.2. Thus,

$$\dim X = \dim Y. \tag{1.19}$$

(2) Assume (1.19) is true. If this common dimension is zero, then both X and Y are trivial, and are clearly isomorphic. Assume that this dimension is nonzero. Let A and B be bases of X and Y respectively. Then, since

$$\dim X = \operatorname{card} A = \operatorname{card} B = \dim Y,$$

there exists a bijective mapping $\psi : A \to B$.

We now construct a mapping $\Psi : X \to Y$. We define that Ψ maps the zero-vector of X to that of Y. For any nonzero vector $u \in X$, by Theorem 1.23, there exist a unique $n \in \mathbb{N}$, unique numbers $\alpha_1, \cdots, \alpha_n \in \mathbb{F}$ with each $\alpha_k \neq 0$ $(1 \leq k \leq n)$, and unique distinct vectors $u_1, \cdots, u_n \in A$ such that

$$u = \sum_{k=1}^{n} \alpha_k u_k. \tag{1.20}$$

Define

$$\Psi(u) = \sum_{k=1}^{n} \alpha_k \psi(u_k). \tag{1.21}$$

We thus have defined a mapping $\Psi : X \to Y$.

We show now the mapping $\Psi : X \to Y$ is bijective. Let $u \in X$ be a nonzero vector as given in (1.20). Notice that in (1.21) $\psi(u_1), \cdots, \psi(u_n) \in B$ are linearly independent and $\alpha_k \neq 0$ $(1 \leq k \leq n)$. Thus,

$$\Psi(u) \neq o.$$

Let $u' \in X$. If $u' = o$, then

$$\Psi(u') = o \neq \Psi(u),$$

Assume $u' \neq o$ and $u' \neq u$. Then, u' has the unique expression

$$u' = \sum_{k=1}^{m} \beta_k u'_k$$

for unique $m \in \mathbb{N}$, $\beta_k \in \mathbb{F}$ and $\beta \neq 0$ $(1 \leq k \leq m)$, and unique $u'_1, \cdots, u'_m \in A$. Consequently,

$$\Psi(u') = \sum_{k=1}^{m} \beta_k \psi(u'_k).$$

If

$$\Psi(u') = \Psi(u),$$

then

$$\sum_{k=1}^{n} \alpha_k \psi(u_k) = \sum_{k=1}^{m} \beta_k \psi(u'_k).$$

By Theorem 1.23, $m = n$, $\alpha_k = \beta_k$, and $u_k = u'_k$ for $k = 1, \cdots, n$. Thus, $u = u'$. Hence, $\Psi : X \to Y$ is injective.

Let $v \in Y$. If $v = o$ then

$$\Psi(o) = o.$$

Assume $v \neq o$. Then, by Theorem 1.23,

$$v = \sum_{k=1}^{l} \xi_k v_k \tag{1.22}$$

for some $l \in \mathbb{N}$, $\xi_k \in \mathbb{F}$ with $\xi_k \neq 0$ $(k = 1, \cdots, l)$, and $v_k \in B$ $(k = 1, \cdots, l)$. Since $\psi : A \to B$ is bijective, there exist $u_1, \cdots, u_l \in A$ such that

$$\psi(u_k) = v_k, \qquad k = 1, \cdots, l.$$

This, (1.22), and the definition of $\Psi : X \to Y$ imply that

$$v = \sum_{k=1}^{l} \xi_k \psi(u_k) = \Psi(\sum_{k=1}^{l} \xi_k u_k).$$

Hence, $\Psi : X \to Y$ is surjective. Consequently, it is surjective.

We now show that $\Psi : X \to Y$ is indeed a isomorphism.

let $u \in X$ be given as in (1.20) and let $\alpha \in \mathbb{F}$. Then,

$$\alpha u = \sum_{k=1}^{n} (\alpha \alpha_k) u_k.$$

Hence,

$$\Psi(\alpha u) = \sum_{k=1}^{n} (\alpha \alpha_k) \psi(u_k) = \alpha \sum_{k=1}^{n} \alpha_k \psi(u_k) = \alpha \Psi(u).$$

to be completed

□

1.4 Subspaces

Example 1.33. Let $n \in 0$ be an integer. Denote by $\mathcal{P}_n$ the set of polynomials of degreee $\leq n$ that hasve coefficients in $\mathbb{F}$. Then, $\mathcal{P}_n$ is a subspace of $\mathcal{P}$.

Example 1.34. Let $a, b \in \mathbb{R}$ with $a < b$. For each $n \in \mathbb{N}$, denote by $C^n[a, b]$ the set of all functions $[a, b] \to \mathbb{R}$ that have nth continuous derivative on $[a, b]$. It is clear that $C^n[a, b]$ is a subspace of $C[a, b]$

1.5 Linear Transforms

Exercises

1.1. Verfity that the set $\mathbb{F}^\infty$ with the addition and scalar multiplication defined in Example 1.6 is a vector space.

1.2. Consider the vector space $C[a, b]$ over $\mathbb{R}$.

(1) Show that the three functions $1, \sin^2 x, \cos^2 x \in C[a, b]$ are linearly dependent.

(2) Let $n \in \mathbb{N}$. Let $\lambda_1, \cdots, \lambda_n \in \mathbb{R}$ be n distinct real numbers. Show that the n functions $e^{\lambda_1 x}, \cdots, e^{\lambda_n x} \in C[a,b]$ are linearly independent.

1.3. Prove Proposition 1.14.

1.4. Prove Proposition 1.31.

1.5. Let $a, b \in \mathbb{R}$ with $a < b$. Denote by $C^2[a,b]$ the set of all real-valued functions defined on $[a,b]$ that have continuous second-order derivatives on $[a,b]$. Let S be the set of all functions $y \in C^2[a,b]$ that solves the different equation

$$\frac{d^2y}{dx^2} + p(x)\frac{dy}{dx} + q(x)y = r(x), \qquad x \in [a,b],$$

for some given $p = p(x), q = q(x), r = r(x) \in C[a,b]$. Prove that $C^2[a,b]$ is a subspace of $C[a,b]$ and S is a subspace of $C^2[a,b]$.

References

2

Hilbert Spaces

2.1 Inner Products

For any complex number $z \in \mathbb{C}$, we denote by $\bar{z}$ its conjugate. Note that $\bar{z} = z$, if $z \in \mathbb{R}$.

Definition 2.1 (inner product). *Let X be a vector space over $\mathbb{F}$.*

(1) *An inner product of X is a mapping $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{F}$ that satisfies the following properties:*

$$\begin{aligned}
&(i) && \langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle \qquad \forall u, v, w \in X; && (2.1)\\
&(ii) && \langle \alpha u, v\rangle = \alpha\langle u, v\rangle \qquad \forall u, v \in X,\ \forall \alpha \in \mathbb{F}; && (2.2)\\
&(iii) && \langle u, v\rangle = \overline{\langle v, u\rangle} \qquad \forall u, v \in X; && (2.3)\\
&(iv) && \langle u, u\rangle \geq 0 \qquad \forall u \in X, && (2.4)\\
& && \langle u, u\rangle = 0 \Longleftrightarrow u = 0. && (2.5)
\end{aligned}$$

(2) *If $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{F}$ is an inner product of X, then $(X, \langle \cdot, \cdot \rangle)$, or simply X, is an inner-product space.*

Remark 2.2. An inner product is also called a *scalar product* or a *dot product*.

Example 2.3. Let $n \in \mathbb{N}$. Define

$$\langle \xi, \eta \rangle = \sum_{j=1}^{n} \xi_j \bar{\eta}_j \qquad \forall \xi = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}, \ \eta = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} \in \mathbb{F}^n. \tag{2.6}$$

It is easy to see that this is an inner product of $\mathbb{F}^n$, cf. Problem (1.1).

Example 2.4. Let

$$l_2 = \left\{ (\xi_1, \cdots) \in \mathbb{F}^\infty : \sum_{k=1}^{\infty} |\xi_k|^2 < \infty \right\}.$$

Define for any $\xi = (\xi_1, \cdots), \eta = (\eta_1, \cdots) \in l_2$

$$(\xi_1, \cdots) + (\eta_1, \cdots) = (\xi_1 + \eta_1, \cdots,),$$
$$\alpha(\xi_1, \cdots) = (\alpha\xi_1, \cdots).$$

Proposition 2.5. *Let X be a vector space over $\mathbb{F}$ and denote by o its zero-vector. Let $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{F}$ be an inner product of a vector space X over $\mathbb{F}$. Then,*

$$\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle \qquad u, v, w \in X; \tag{2.7}$$
$$\langle u, \alpha v\rangle = \overline{\alpha}\langle u, v\rangle \qquad \forall u, v \in X,\ \forall \alpha \in \mathbb{F}. \tag{2.8}$$
$$\langle o, u\rangle = \langle u, o\rangle = 0 \qquad \forall u \in X. \tag{2.9}$$

Proof. Let $u, v, w \in X$ and $\alpha \in \mathbb{F}$. By (2.3) and (2.1), we have

$$\langle u, v + w\rangle = \overline{\langle v + w, u\rangle} = \overline{\langle v, u\rangle + \langle w, u\rangle} = \overline{\langle v, u\rangle} + \overline{\langle w, u\rangle} = \langle u, v\rangle + \langle u, w\rangle,$$

proving (2.7). Similarly, by (2.3) and (2.3),

$$\langle u, \alpha v\rangle = \overline{\langle \alpha v, u\rangle} = \overline{\alpha\langle v, u\rangle} = \overline{\alpha}\,\overline{\langle v, u\rangle} = \overline{\alpha}\langle u, v\rangle.$$

This is (2.8). Finally, by (2.2),

$$\langle o, u\rangle = \langle 0o, u\rangle = 0\langle o, u\rangle = 0,$$

and by (2.8),

$$\langle u, o\rangle = \langle u, 0o\rangle = 0\langle u, o\rangle = 0.$$

These imply (2.9). □

2.2 Normed Vector Spaces

2.3 Hilbert Spaces

2.4 Orthogonality

2.5 Complete Orthogonal Bases

2.6 Riesz Representation

Exercises

2.1. Show that the mapping $\langle\cdot,\cdot\rangle : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ defined by (2.6) is an inner product of $\mathbb{F}^n$.

References

3

Operator Theory

Throughout this chapter, we assume that H is a Hilbert space over $\mathbb{F}$. We also call a mapping from H to H an *operator* on H.

3.1 Bounded Linear Operators

Definition 3.1 (linear operators). *An operator $T : H \to H$ is* linear*, if*

$$T(u+v) = Tu + Tv \qquad \forall u, v \in H,$$
$$T(\alpha u) = \alpha Tu \qquad \forall \alpha \in \mathbb{F}, \forall u \in H.$$

Proposition 3.2. *Let $T : H \to H$ be a linear operator. Then,*

(1) $$To = o, \tag{3.1}$$

(2) $$T(-u) = -Tu \qquad \forall u \in H, \tag{3.2}$$

(3) $$T\left(\sum_{k=1}^{n} \alpha_k u_k\right) = \sum_{k=1}^{n} \alpha_k T u_k \qquad \forall n \in \mathbb{N}, \ \forall \alpha_k \in \mathbb{F}, \forall u_k \in H, \ k = 1, \cdots, n. \tag{3.3}$$

Proof. (1) We have

$$To = T(o + o) = To + To.$$

Adding $-To$, we obtain (3.1).

(2) By (3.1),

$$T(-u) + Tu = T(-u + u) = To = o.$$

This implies (3.2).

(3) Clearly, (3.3) is true for $n = 1$ and $n = 2$. Assume it is true for $n \in \mathbb{N}$ with $n \geq 2$. Then,

$$
\begin{aligned}
T\left(\sum_{k=1}^{n} \alpha_k u_k\right) &= T\left(\alpha_1 u_1 + \sum_{k=2}^{n} \alpha_k u_k\right) \\
&= T(\alpha_1 u_1) + T\left(\sum_{k=2}^{n} \alpha_k u_k\right) \\
&= \alpha_1 T u_1 + \sum_{k=2}^{n} \alpha_k T u_k \\
&= \sum_{k=1}^{n} \alpha_k T u_k.
\end{aligned}
$$

Therefore, (3.3) is true for all $n \in \mathbb{N}$. $\square$

Example 3.3 (the identity operator and the zero operator). The identity operator $I : H \to H$, defined by

$$Iu = u \qquad \forall u \in H,$$

is a linear operator.

The zero operator $O : H \to H$, defined by

$$Ou = o \qquad \forall u \in H,$$

is also a linear operator.

Example 3.4. Let $n \in \mathbb{N}$ and $H = \mathbb{F}^n$. Let $A \in \mathbb{F}^{n \times n}$. Define $T : \mathbb{F}^n \to \mathbb{F}^n$ by

$$Tu = Au \qquad \forall u \in \mathbb{F}^n. \tag{3.4}$$

Clearly, $T : \mathbb{F}^n \to \mathbb{F}^n$ is a linear operator. Thus, every $n \times n$ matrix determines a linear operator in $\mathbb{F}^n$.

Conversely, suppose $T : \mathbb{F}^n \to \mathbb{F}^n$ is a linear operator. Let

$$a^{(k)} = Te_k \in \mathbb{F}^n, \qquad k = 1, \cdots, n,$$

where $e_1, \cdots, e_n$ are all the n coordinate vectors in $\mathbb{F}^n$ as defined in Example 1.27. Define $A \in \mathbb{F}^{n \times n}$ to be the $n \times n$ matrix that has columns $a^{(1)}, \cdots, a^{(n)}$:

$$A = \left[a^{(1)} \cdots a^{(n)}\right]. \tag{3.5}$$

We have

$$Tu = T\left(\sum_{k=1}^{n} u_k e_k\right) = \sum_{k=1}^{n} u_k T e_k = \sum_{k=1}^{n} u_k a^{(k)} = Au \qquad \forall u = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \in \mathbb{F}^n.$$

Thus, any linear operator $T : \mathbb{F}^n \to \mathbb{F}^n$ is given by (3.4) with the matrix A that is defined in (3.5). We call A the *transformation matrix* of $T : \mathbb{F}^n \to \mathbb{F}^n$.

Example 3.5 (the Hilbert-Schmidt operator). Let $H = L^2[a, b]$ and $k \in L^2([a, b] \times [a, b])$. Let $u \in L^2[a, b]$ and define

$$(Ku)(x) = \int_a^b k(x,t)u(t)\,dt \qquad x \in [a, b].$$

Using the Cauchy-Schwarz inequality, we have

$$|(Ku)(x)|^2 = \left| \int_a^b k(x,t)u(t)\,dt \right|^2 \leq \int_a^b |k(x,t)|^2 dt \int_a^b |u(t)|^2 dt$$

Integrating over $x \in [a, b]$, we obtain

$$\int_a^b |(Ku)(x)|^2 dx \leq \int_a^b \int_a^b |k(x,t)|^2 dx dt \int_a^b |u(t)|^2 dt < \infty.$$

Thus, $Ku \in L^2[a, b]$. Consequently, $K : L^2[a, b] \to L^2[a, b]$ is an operator. One can easily verify that this operator is linear. For each k, the operator K is called a Hilbert-Schmidt operator.

Example 3.6 (orthogonal projections). Let M be a nonempty, closed subspace of H. Then, for any $u \in H$, there exists unique $P_M u \in M$ such that

$$\langle u - P_M u, v \rangle = 0 \qquad \forall v \in M,$$

cf. Fig 3.6.

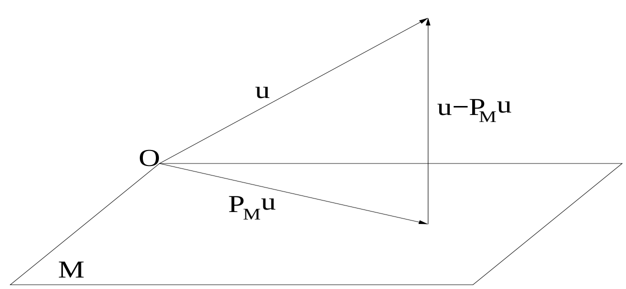

Fig. 3.1. The orthogonal projection $P_M : H \to M$

The operator $P_M : H \to M$ is a linear operator, called the *orthogonal projection operator* onto M. For $u \in H$, $P_M u \in M$ is called the orthogonal projection of u onto M.

We shall denote by $\mathcal{L}(H)$ the set of all linear operators from H to H.

Definition 3.7 (addition and scalar multiplication of operators). *For any $T, S \in \mathcal{L}(H)$ and any $\alpha \in \mathbb{F}$, we define the operator $T + S : H \to H$ and $\alpha T : H \to H$ by*

$$(T+S)u = Tu + Su \qquad \forall u \in H,$$
$$(\alpha T)u = \alpha Tu \qquad \forall \alpha \in \mathbb{F}, \forall u \in H.$$

The proof of the following proposition is left as an exercise problem, cf. Problem 3.1.

Proposition 3.8. *With respect to the operations defined in Definition 3.7, the set $\mathcal{L}(H)$ is a vector space over $\mathbb{F}$.*

Definition 3.9 (continuity and boundedness of operators). *Let $T \in \mathcal{L}(H)$.*

(1) *The operator T is* continuous, *if*

$$\lim_{k\to\infty} Tu_k = Tu \quad \textit{in } H, \quad \textit{whenever} \lim_{k\to\infty} u_k = u \quad \textit{in } H.$$

(2) *The operator T is* bounded, *if there exists a constant $M > 0$ such that*

$$\|Tu\| \le M\|u\| \qquad \forall u \in H. \tag{3.6}$$

Proposition 3.10. *Let $T \in \mathcal{L}(H)$. Then, the following are equivalent:*

(1) *The operator $T : H \to H$ is continuous;*
(2) *The operator $T : H \to H$ is continuous at o, i.e.,*

$$\lim_{k\to\infty} Tu_k = o \quad \textit{in } H, \quad \textit{whenever} \lim_{k\to\infty} u_k = o \quad \textit{in } H;$$

(3) *The operator $T : H \to H$ is bounded;*
(4) *For any bounded set $E \subseteq H$, the image $T(E) = \{Tu : u \in E\}$ is bounded in H.*

Proof. (1) $\Longrightarrow$ (2). This follows from setting $u = o$ in Part (1) of Definition 3.9.

(2) $\Longrightarrow$ (1). Suppose $u_k \to u$ in H. Then, $u_k - u \to o$ in H. By Proposition 3.2 and (2),

$$T(u_k - u) = Tu_k - Tu \to o. \qquad \text{in } H.$$

Hence,

$$Tu_k \to Tu \qquad \text{in } H.$$

This proves (1).

(3) $\Longrightarrow$ (2). If $u_k \to o$, then by (3)

$$\|Tu_k\| \le M\|u_k\| \to 0 \qquad \text{as } k \to \infty,$$

implying (2).

(2) $\Longrightarrow$ (3). Assume (2) is true. If (3) were not true, then, for any $k \in \mathbb{N}$ there exists $u_k \in H$ such that

$$\|Tu_k\| > M\|u_k\| \qquad k = 1, \cdots . \tag{3.7}$$

Clearly, by this and Part (1) of Proposition 3.2, each u_k $(k \geq 1)$ is nonzero. Moreover, setting

$$v_k = \frac{1}{k\|u_k\|}u_k, \qquad k = 1, \cdots ,$$

we have

$$\|v_k\| = \frac{1}{k} \to 0 \qquad \text{as } k \to \infty.$$

But, by (3.7),

$$\|Tu_k\| > 1, \qquad k = 1, \cdots ,$$

contradicting (2). Therefore, (3) is true.

(3) $\Longrightarrow$ (4). If $E \subseteq H$ is a bounded set, then there exists $C > 0$ such that

$$\|u \mid \leq C \qquad \forall u \in E.$$

Thus,

$$\|Tu\| \leq \|T\|\,\|u\| \leq C\|T\| \qquad \forall u \in E.$$

Hence, $T(E)$ is bounded.

(4) $\Longrightarrow$ (3). Setting

$$E = \{u \in H : \|u\| = 1\},$$

we have

$$M := \sup_{u\in H, \|u\|=1} \|Tu\| < \infty.$$

Since $(1/\|u\|)u \in E$ for any $u \in H$ with $u \neq o$,

$$\frac{1}{\|u\|}\|Tu\| = \left\|T\left(\frac{1}{\|u\|}u\right)\right\| \leq M \qquad \forall u \in H,\, u \neq o.$$

This implies (3.6). Hence, T is bounded. $\square$

We shall denote by $\mathcal{B}(H)$ the set of all the bounded linear operators in H.

Definition 3.11. *For any $T \in \mathcal{B}(H)$, we define*

$$\|T\| = \sup_{u\in H, u\neq o} \frac{\|Tu\|}{\|u\|}. \tag{3.8}$$

By Definition 3.9, $\|T\|$ is finite for any $T \in \mathcal{B}(H)$.

Proposition 3.12. *Let $T \in \mathcal{B}(H)$. Then,*

$$\|T\| = \sup_{u\in H, \|u\|\leq 1} \|Tu\| = \sup_{u\in H, \|u\|=1} \|Tu\|, \tag{3.9}$$

$$\|Tu\| \leq \|T\|\|u\| \qquad \forall u \in H. \tag{3.10}$$

Proof. By (3.8),

$$\sup_{u\in H, \|u\|=1} \|Tu\| \le \sup_{u\in H, \|u\|\le 1} \|Tu\| \le \|T\|. \tag{3.11}$$

If $u \in H$ and $u \ne o$, then $v = (1/\|u\|)u$ is a unit vector. Thus,

$$\frac{\|Tu\|}{\|u\|} = \|Tv\| \le \sup_{u\in H, \|u\|=1} \|Tu\|.$$

Hence, by (3.8),

$$\|T\| \le \sup_{u\in H, \|u\|=1} \|Tu\|.$$

This and (3.11) imply (3.9).

The inequality (3.10) follows from the definition (3.8). □

Theorem 3.13. *The set $\mathcal{B}(H)$ is a Banach space with respect to the operations in $\mathcal{L}(H)$ and $\|\cdot\| : \mathcal{B}(H) \to \mathbb{R}$.*

Proof. Step 1. The set $\mathcal{B}(H)$ is a vector subspace of the vector subspace $\mathcal{L}(H)$.

Obviously, the zero operator $O \in \mathcal{B}(H)$. So, $\mathcal{B}(H) \ne \emptyset$. Let $T, S \in \mathcal{B}(H)$. Then,

$$\begin{aligned} \|(T+S)u\| &= \|Tu + Su\| \le \|Tu\| + \|Su\| \\ &\le \|T\|\|u\| + \|S\|\|u\| = (\|T\| + \|S\|)\|u\| \qquad \forall u \in H. \end{aligned} \tag{3.12}$$

Hence, $T + S \in \mathcal{B}(H)$. Let $T \in \mathcal{B}(H)$ and $\alpha \in \mathbb{F}$. Then,

$$\|(\alpha T)u\| = \|\alpha(Tu)\| = |\alpha|\|Tu\| \le |\alpha|\|T\|\|u\| \qquad \forall u \in H.$$

Hence, $\alpha T \in \mathcal{B}(H)$. Therefore, by ?, $\mathcal{B}(H)$ is a vector subspace of $\mathcal{L}(H)$.

Step 2. The vector space $\mathcal{B}(H)$ is a normed vector space with respect to $\|\cdot\| : \mathcal{B}(H) \to \mathbb{R}$ defined in (3.8).

(1) Clearly,

$$\|T\| \ge 0 \qquad \forall T \in \mathcal{B}(H).$$

By (3.8) and (3.1), for $T \in \mathcal{B}(H)$,

$$\|T\| = 0 \qquad \text{if and only if} \qquad T = o.$$

(2) Let $T \in \mathcal{B}(H)$ and $\alpha \in \mathbb{F}$. By (3.8),

$$\begin{aligned} \|\alpha T\| &= \sup_{u\in H, u\ne o} \|(\alpha T)u\| = \sup_{u\in H, u\ne o} \|\alpha(Tu)\| \\ &= \sup_{u\in H, u\ne o} |\alpha|\|Tu\| = |\alpha| \sup_{u\in H, u\ne o} \|(\alpha T)u\| = |\alpha|\|T\|. \end{aligned}$$

(3) Let $T, S \in \mathcal{B}(H)$. It follows from (3.12) and (3.8) that

$$\|T+S\| \le \|T\| + \|S\|.$$

It follows from (1)–(3) that $\mathcal{B}(H)$ is a normed vector space.

Step 3. The normed vector space $\mathcal{B}(H)$ is a Banach space.

Let $\{T_n\}_{n=1}^\infty$ be a Cauchy sequence in $\mathcal{B}(H)$. Then, for any $u \in H$, $\{T_n u\}_{n=1}^\infty$ is a Cauchy sequence in H, since

$$\|T_m u - T_n u\| \le \|T_m - T_n\| \|u\| \to 0 \qquad \text{as } m, n \to \infty.$$

Since H is a Hilbert space, $\{T_n u\}_{n=1}^\infty$ converges in H. We define

$$Tu = \lim_{n\to\infty} T_n u \qquad \text{in } H. \tag{3.14}$$

Clearly, $T : H \to H$ is linear, since

$$\begin{aligned} T(u+v) &= \lim_{n\to\infty} T_n(u+v) = \lim_{n\to\infty} (T_n u + T_n v) \\ &= \lim_{n\to\infty} T_n u + \lim_{n\to\infty} T_n v = Tu + Tv \qquad \forall u, v \in H, \\ T(\alpha u) &= \lim_{n\to\infty} T_n(\alpha u) = \lim_{n\to\infty} \alpha(T_n u) \\ &= \alpha \lim_{n\to\infty} T_n u = \alpha T u \qquad \forall u, v \in H. \end{aligned}$$

By (3.14), we have

$$\|Tu\| = \lim_{n\to\infty} \|T_n u\| \le \left(\sup_{n\ge 1} \|T_n\| \right) \|u\| = M\|u\| \qquad \forall u \in H,$$

where

$$M = \sup_{n\ge 1} \|T_n\| < \infty,$$

since $\{T_n\}_{n=1}^\infty$ is a Cauchy sequence and hence a bounded sequence in $\mathcal{B}(H)$. Therefore, $T \in \mathcal{B}(H)$.

Finally, let $\varepsilon > 0$. Since $\{T_n\}_{n=1}^\infty$ is a Cauchy sequence in $\mathcal{B}(H)$, there exists $N \in \mathbb{N}$ such that

$$\|T_m - T_n\| < \varepsilon \qquad \forall m, n \ge N.$$

Let $u \in H$ with $\|u\| = 1$. We thus have

$$\|T_m u - T_n u\| \le \|T_m - T_n\| \cdot \|u\| < \varepsilon \qquad \forall m, n \ge N.$$

Sending $n \to \infty$, we obtain

$$\|T_m u - Tu\| \le \varepsilon \qquad \forall m \ge N.$$

This, together with (3.9), leads to

$$\|T_m - T\| \le \varepsilon \qquad \forall m \ge N,$$

i.e.,

$$\lim_{m\to\infty} T_m = T \qquad \text{in } \mathcal{B}(H).$$

Consequently, $\mathcal{B}(H)$ is a Banach space. $\square$.

Example 3.14. Let $H = \mathbb{F}^n$. Then, $\mathcal{B}(H) = \mathcal{L}(H)$.

Definition 3.15 (multiplication of operators). *Let $T, S \in \mathcal{B}(H)$. We define $TS : H \to H$ by*

$$(TS)u = T(Su) \qquad \forall u \in H.$$

We also define

$$T^0 = I \qquad \text{and} \qquad T^k = T^{k-1}T, \quad k = 1, \cdots .$$

Proposition 3.16. *If $T, S \in \mathcal{B}(H)$, then $TS \in \mathcal{B}(H)$. Moreover,*

$$\|TS\| \le \|T\| \, \|S\|. \tag{3.15}$$

In particular,

$$\left\|T^k\right\| \le \|T\|^k \qquad k = 0, 1, \cdots . \tag{3.16}$$

Proof. We have for any $u, v \in H$ and any $\alpha \in \mathbb{F}$ that

$$\begin{aligned} (TS)(u+v) &= T(S(u+v)) = T(Su+Sv) \\ &= T(Su) + T(Sv) = (TS)u + (TS)v, \\ (TS)(\alpha u) &= T(S(\alpha u)) = T(\alpha Su) = \alpha T(Su) = \alpha (TS)u. \end{aligned}$$

Thus, $TS : H \to H$ is linear.

The boundedness of TS and (3.15) follow from the following:

$$\|(TS)u\| = \|T(Su)\| \le \|T\| \, \|Su\| \le \|T\| \, \|S\| \, \|u\| \qquad \forall u \in H. \quad \square.$$

The following result is straightforward, and its proof is omitted.

Proposition 3.17. *Let $R, S, T \in \mathcal{B}(H)$ and $\alpha \in \mathbb{F}$. Then,*

$$(RS)T = R(ST), \tag{3.17}$$

$$R(S+T) = RS + RT, \tag{3.18}$$

$$(R+S)T = RS + ST, \tag{3.19}$$

$$(\alpha S)T = \alpha(ST) = S(\alpha T). \tag{3.20}$$

A *Banach algebra* over $\mathbb{F}$ is a normed vector space over $\mathbb{F}$ on which a multiplication is defined so that (3.17)–(3.20) and (3.15) hold true for all R, S, T in the space and all $\alpha \in \mathbb{F}$. By Proposition 3.16 and Proposition 3.17, we have the following corollary.

Corollary 3.18. *For any Hilbert space H over $\mathbb{F}$, the set $\mathcal{B}(H)$ of all bounded linear operators on H is a Banach algebra over $\mathbb{F}$.*

Proposition 3.19. (1) *If $T_n \to T$ in $\mathcal{B}(H)$, then*

$$T_n u \to Tu \qquad \forall u \in H.$$

(2) *If* $T_n \to T$ *and* $S_n \to S$ *in* $\mathcal{B}(H)$, *then*

$$T_n S_n \to TS.$$

Proof. (1) For any $u \in H$, we have by (3.10),

$$\|T_n u - Tu\| \le \|T_n - T\| \, \|u\| \to 0 \qquad \text{as } n \to \infty.$$

(2) Since $\{T_n\}_{n=1}^{\infty}$ converges in $\mathcal{B}(H)$, it is bounded. Thus, by the triangle inequality and (3.15), we have

$$\begin{aligned} \|T_n S_n - TS\| &= \|(T_n - T)S_n + T(S_n - S)\| \\ &\le \|(T_n - T)S_n\| + \|T(S_n - S)\| \\ &\le \|T_n - T\| \, \|S_n\| + \|T\| \, \|S_n - S\| \\ &\to 0 \qquad \text{as } n \to \infty. \end{aligned}$$

completing the proof. □

Example 3.20. Let H be a Hilbert space over $\mathbb{F}$ with a countable orthonormal basis $\{e_n\}_{n=1}^{\infty}$. Each $u \in H$ then has a unique Fourier expansion (cf. ?)

$$u = \sum_{k=1}^{\infty} \langle u, e_k \rangle e_k.$$

Define $T_n : H \to H$ for each $n \in \mathbb{N}$ by

$$T_n u = \sum_{k=1}^{n} \langle u, e_k \rangle e_k.$$

Clearly, T_n is the projection onto

$$M_n := \operatorname{span}\{e_1, \cdots, e_n\}.$$

Hence, $T_n \in \mathcal{B}(H)$. Moreover,

$$\|T_n\| = 1 \qquad \forall n \in \mathbb{N}.$$

In particular, $\{T_n\}_{n=1}^{\infty}$ does not converge to the zero operator $O : H \to H$ in $\mathcal{B}(H)$.

However, for any $u \in H$,

$$\|T_n u\|^2 = \sum_{k=1}^{n} |\langle u, e_k \rangle|^2 \to 0 \qquad \text{as } n \to \infty.$$

Hence, the converse of Part (1) in Proposition 3.19 is in general not true.

Definition 3.21 (invertible operators). *Let $S, T \in \mathcal{B}(H)$. If*

$$ST = TS = I, \tag{3.21}$$

then T is invertible, *and S is an* inverse *of T.*

Proposition 3.22. *If $T \in \mathcal{B}(H)$ is invertible, then its inverse is unique.*

Proof. Let $R, S \in \mathcal{B}(H)$ be inverses of T. Then,

$$R = RI = R(TS) = (RT)S = IS = S. \qquad \square$$

If $T \in \mathcal{B}(H)$ is invertible, we shall denote by T^{-1} its unique inverse.

Theorem 3.23 (The von Neumann series). *Let $T \in \mathcal{B}(H)$. Suppose $\|T\| < 1$. Then, $I - T \in \mathcal{B}(H)$ is invertible. Moreover,*

$$(I-T)^{-1} = \sum_{k=0}^{\infty} T^k.$$

Proof. For each $n \in \mathbb{N}$, set Let

$$S_n = \sum_{k=0}^{n} T^k \qquad n = 1, \cdots .$$

Since $\|T\| < 1$, the series $\sum_{k=0}^{\infty} \|T\|^k$ converges. Thus, for any $p \in \mathbb{N}$, we have by (3.16) that

$$\|S_{n+p} - S_n\| = \left\| \sum_{k=n+1}^{n+p} T^k \right\| \le \sum_{k=n+1}^{n+p} \left\|T^k\right\| \le \sum_{k=n+1}^{\infty} \|T\|^k \to 0 \qquad \text{as } n \to \infty.$$

Hence, $\{S_n\}_{n=1}^{\infty}$ is a Cauchy sequence in H, and therefore it converges in H. This shows that the series $\sum_{k=10}^{\infty} T^k$ converges in H, and

$$S_n \to \sum_{k=10}^{\infty} T^k \qquad \text{in } H. \tag{3.22}$$

One can easily verify that

$$(I-T)S_n = S_n(I-T) = I - T^{n+1}.$$

Sending $n \to \infty$, we obtain then by (3.22) and Part (2) of Proposition 3.19 that

$$(I-T)\sum_{k=10}^{\infty} T^k = \left(\sum_{k=10}^{\infty} T^k \right)(I-T) = I,$$

completing the proof. $\square$

3.2 Self-Adjoint Operators

3.3 Compact Operators

Definition 3.24 (compact operators). *An operator $T \in \mathcal{B}(H)$ is a* compact operator*, if for any bounded squence $\{u_n\}_{n=1}^{\infty}$ in H, there exists a subsequence $\{u_{n_k}\}_{k=1}^{\infty}$ such that $\{Tu_{n_k}\}_{k=1}^{\infty}$ converges in H.*

Proposition 3.25. (1) *If $T, S \in \mathcal{B}(H)$ are two compact operators and $\alpha \in \mathbb{F}$, then $T + S$ and αT are also compact operators on H.*
(2) *If $T \in \mathcal{B}(H)$ is a compact operator and $S \in \mathcal{B}(H)$, then TS and ST are also compact operators.*
(3) *If $\{T_n\}_{n=1}^{\infty}$ is a sequence of compact operators on H and*

$$T_n \to T \qquad \text{in } \mathcal{B}(H)$$

for some $T \in \mathcal{B}(H)$, then T is also a compact operator.

Proof. (1) Let $\{u_n\}_{n=1}^{\infty}$ be a bounded sequence in H. Since $T \in \mathcal{B}(H)$ is compact, there exists a subsequence $\{u_{n_k}\}_{k=1}^{\infty}$, of $\{u_n\}_{n=1}^{\infty}$ such that $\{Tu_{n_k}\}_{k=1}^{\infty}$ converges in H. Clearly, $\{(\alpha T)u_{n_k}\}_{k=1}^{\infty}$ also converges. Hence, $\alpha T \in \mathcal{B}(H)$ is compact. Moreover, there exists a further subsequence, $\{u_{n_{k_j}}\}_{j=1}^{\infty}$, of $\{u_{n_k}\}_{=1}^{\infty}$ such that $\{Su_{n_{k_j}}\}_{j=1}^{\infty}$ converges in H, since $S \in \mathcal{B}(H)$ is compact. Thus, $\{(T+S)u_{n_{k_j}}\}_{j=1}^{\infty}$ converges in H. This implies that $T + S \in \mathcal{B}(H)$ is compact.

(2) Let $\{u_n\}_{n=1}^{\infty}$ be a bounded sequence in H. Since $S \in \mathcal{B}(H)$, $\{Su_n\}_{n=1}^{\infty}$ is also bounded in H. Thus, since $T \in \mathcal{B}(H)$ is compact and

$$(TS)u_n = T(Su_n), \qquad n = 1, \cdots,$$

there exists a subsequence, $\{u_{n_k}\}_{k=1}^{\infty}$, of $\{u_n\}_{n=1}^{\infty}$ such that $\{(TS)u_{n_k}\}_{k=1}^{\infty}$ converges. Hence, $TS \in \mathcal{B}(H)$ is compact.

Since T is compact, there exists a subsequence $\{u_{n_k}\}_{k=1}^{\infty}$, of $\{u_n\}_{n=1}^{\infty}$ such that $\{Tu_{n_k}\}_{k=1}^{\infty}$ converges. Since $S \in \mathcal{B}(H)$, $\{(ST)u_{n_k}\}_{k=1}^{\infty}$ also converges. Hence, ST is also compact.

(3) Let again $\{u_n\}_{n=1}^{\infty}$ be a bounded sequence in H. There exists $A > 0$ such that

$$\|u_n\| \le A, \qquad n = 1, \cdots. \tag{3.23}$$

Since T_1 is compact, there exists a subsequence, $\{u_n^{(1)}\}_{n=1}^{\infty}$, of $\{u_n\}_{n=1}^{\infty}$ such that $\{T_1 u_n^{(1)}\}_{n=1}^{\infty}$ converges in H. Suppose $\{u_n^{(k)}\}_{n=1}^{\infty}$ is a subsequence of $\{u_n\}_{n=1}^{\infty}$ such that $\{T_k u_n^{(k)}\}_{n=1}^{\infty}$ converges in H. Then, since T_{k+1} is compact, there exists a subsequence, $\{u_n^{(k+1)}\}_{n=1}^{\infty}$, of $\{u_n^{(k)}\}_{n=1}^{\infty}$ such that $\{T_{k+1}u_n^{(k+1)}\}_{n=1}^{\infty}$ converges in H. Thus, by induction, for each $k \in \mathbb{F}$, there exists a subsequence $\{u_n^{(k)}\}_{n=1}^{\infty}$ of $\{u_n^{(k-1)}\}_{n=1}^{\infty}$, with $\{u_n^{(0)}\}_{n=1}^{\infty}$ being $\{u_n\}_{n=1}^{\infty}$, such that $\{T_k u_n^{(k)}\}_{n=1}^{\infty}$ converges. Let

$$v_n = u_n^{(n)}, \qquad n = 1, \cdots .$$

Then, $\{v_n\}_{n=1}^{\infty}$ is a subsequence of $\{u_n\}_{n=1}^{\infty}$, and for each $k \in \mathbb{N}$, $\{T_k u_n\}_{n=1}^{\infty}$ converges in H.

We need now only to show that the sequence $\{Tv_n\}_{n=1}^{\infty}$ converges in H. Let $\varepsilon > 0$. Since $T_n \to T$ in $\mathcal{B}(H)$, there exists $N \in \mathbb{N}$ such that

$$\|T_k - T\| \le \varepsilon \forall k \ge N. \tag{3.24}$$

Since $\{T_N v_n\}_{n=1}^{\infty}$ converges, it is a Cauchy sequence. Thus, there exists $L \in \mathbb{N}$ such that

$$\|T_N v_p - T_N v_q\| \le \varepsilon \qquad \forall p, q \in \mathbb{N} \text{ with } p, q \ge L. \tag{3.25}$$

Consequently, by (3.10), (3.23), (3.24), and (3.25),

$$\begin{aligned} \|Tv_p - Tv_q\| &\le \|Tv_p - T_N v_p\| + \|T_N v_p - T_N v_q\| + \|T_N v_q - Tv_q\| \\ &\le \|T - T_N\| \, \|v_p\| + \|T_N v_p - T_N v_q\| + \|T_N - T\| \, \|v_q\| \\ &\le 2A\varepsilon + \varepsilon \qquad \forall p, q \in \mathbb{N} \text{ with } p, q \ge L. \end{aligned}$$

Hence, $\{Tv_n\}_{n=1}^{\infty}$ is a Cauchy sequence, and it thus converges in H. □

We shall denote by $\mathcal{C}(H)$ the set of all compact operators on H.

Proposition 3.26. *The set $\mathcal{C}(H)$ is a Banach algebra.*

Proof. Note that the zero operator $O : H \to H$ is a compact operator. Thus, $\mathcal{C}(H)$ is nonempty. Note also that $\mathcal{C}(H) \subseteq \mathcal{B}(H)$. By Proposition 3.25, $\mathcal{C}(H)$ is closed under addition, scalar multiplication, and multiplication. It is also closed with respect to the norm of $\mathcal{B}(H)$. Therefore, $\mathcal{C}(H)$ is a sub-Banach algebra of $\mathcal{B}(H)$. Hence, it is itself a Banach algebra. □

If $T \in \mathcal{L}(H)$, then the image of $T : H \to H$,

$$T(H) = \{Tu : u \in H\},$$

is a subspace of H. We call the dimension of $T(H)$ the *rank* of T.

Proposition 3.27. *If $T \in \mathcal{B}(H)$ has a finite rank, then T is compact.*

Proof. □

Exercises

3.1. Prove Proposition 3.8.

References

4

Integral Equations

4.1 Examples and Classification of Integral Equatons

4.2 Integral Equations of the Second Kind: Successive Iterations

4.3 Fredholm Theory of Integral Equations of the Second Kind

4.4 Singular Integral Equations

Exercises

References

A

Appendix

A.1 Zorn's Lemma

Given a set S. Let $\leq$ be a relation of S, i.e., for some $x, y \in S$, either $x \leq y$ or $y \leq x$.

Definition A.1. *A relation $\leq$ of a set S is a* partial order*, if the following are satisfied:*

(1) *For any $x \in S$, $x \leq x$;*
(2) *For any $x, y \in S$, $x \leq y$ and $y \leq x$ imply $x = y$;*
(3) *For any $x, y, z \in S$, $x \leq y$ and $y \leq z$ imply $x \leq z$.*

Definition A.2. *Let S be a partially ordered set with the partial order $\leq$.*

(1) *A subset $C \subseteq S$ is a* chain*, if either $x \leq y$ or $y \leq x$ for any $x, y \in C$.*
(2) *An element $m \in S$ is a* maximal element *of S, if $x \in S$ and $m \leq x$ imply that $x = m$.*

We hypothesize the following

Zorn's Lemma. *A partially ordered set in wich any chain has an upper bound contains a maximal element.*

A.2 Cardinality

Definition A.3. *Let S and T be two sets. A mapping $\varphi : S \to T$ is:*

(1) injective*, if $\varphi(s_1) \neq \varphi(s_2)$ for any $s_1, s_2 \in S$ with $s_1 \neq s_2$;*
(2) surjective*, if for any $t \in T$, there exists $s \in S$ such that $t = \varphi(s)$;*
(3) bijective*, if it is both injective and surjective.*

Two sets S and T are said to be *equipllent*, denoted by $S \sim B$, if there exists a bijective mapping from S to T. Clearly, equipllence is an equivalence relation on the class of all sets, i.e., for any sets R, S, T: $S \sim S$; $S \sim T$ implies $T \sim S$; and $R \sim S$ and $S \sim T$ imply $R \sim T$.

Any set S is therefore associated with a unique *cardinal nubmer*, or *cardinality*, denoted by $\operatorname{card} S$, that is characterized by

$$\operatorname{card} S = \operatorname{card} T \qquad \text{if and only if} \qquad S \sim T.$$

For two sets S and T, we say $\operatorname{card} S < \operatorname{card} T$, if $\operatorname{card} S \neq \operatorname{card} T$ and there exists $T_0 \subsetneq T$ such that $\operatorname{card} S = \operatorname{card} T_0$.

If S is a finite set, then $\operatorname{card} S$ is the number of elements in S. If S is infinite and $\operatorname{card} S = \operatorname{card} \mathbb{N}$, then S is said to be *countably infinite* or *countable*. The cardinal number of a countably infinite set is denoted by $\aleph_0$ ($\aleph$ is pronounced *aleph*). If S is infinite but $\operatorname{card} S \neq \operatorname{card} \mathbb{N}$, then S is said to be *uncountably infinite* or *uncountable*.

For example, the set of positive integers $\mathbb{N}$, the set of all integers $\mathbb{Z}$, and the set of all rational numbers $\mathbb{Q}$ are all countable. The set $\{x \in \mathbb{R} : 0 < x < 1\}$, the set of real numbers $\mathbb{R}$, the set of all complex numbers $\mathbb{C}$, the set $\mathbb{C}^n$ for any integer $n \in \mathbb{N}$ are all uncountable. They have the same cardinal number, denoted by $\aleph$. One can verify that $\aleph_0 < \aleph$.