Chapter 2
Linear Spaces

Abstract This chapter is a rigorous introduction to linear spaces, but with a strong emphasis on infinite dimensional linear spaces. No prior knowledge of linear spaces is assumed, so that all definitions and proofs are included, but some mathematical maturity is assumed, dictating the level of detail given. In particular, all results (e.g., existence of dimension) are given in full generality using Zorn's Lemma.

Keywords Linear space · Vector space · Linear transformation · Operator · Dimension · Quotient linear space · Product linear space · Inner product space · Normed space · Cauchy-Schwarz inequality

This chapter assumes a rudimentary understanding of the linear structure of $\mathbb{R}^n$, a very rich structure, both algebraically and geometrically. Elements in $\mathbb{R}^n$, when thought of as vectors, that is as entities representing direction and magnitude, can be used to form parallelograms, can be scaled, the angle between two vectors can be computed, and the length of a vector can be found. These geometric features are given algebraically by means of, respectively, vector addition, scalar multiplication, the inner product of two vectors, and the norm of a vector. In this chapter these notions are abstracted to give rise to the concepts of linear space, inner product space, and normed space.

The chapter gives a detailed presentation of all of the relevant notions of linear spaces, provides examples, and contains rigorous proofs of all of the results therein. Exploiting the assumption of a rudimentary understanding of the linear structure of $\mathbb{R}^n$, and thus of finite dimensional linear spaces, the chapter has a clear secondary goal, namely to explore the subtleties of infinite dimensional linear spaces. This is an absolute necessity with Hilbert spaces in mind, since virtually all interesting examples of Hilbert spaces are infinite dimensional.

A consequence of this infinite dimensional theme of the chapter is that some proofs are considerably more involved than their finite dimensional counterparts. In particular, the theorems establishing the existence of bases and the concept of dimension are sophisticated and resort to an application of Zorn's Lemma. Also, cardinality considerations are important, since one needs to be able to compute, at least a little bit, with infinite quantities. To facilitate an easier reading of this chapter,

C. Alabiso and I. Weiss, *A Primer on Hilbert Space Theory*,
UNITEXT for Physics, DOI 10.1007/978-3-319-03713-4_2

the Preliminaries contain an account of Zorn's Lemma (Sect. 1.2.17) and some basic cardinal arithmetic (Sects. 1.2.12 and 1.2.13). Moreover, the text below indicates which proofs can safely be skipped on a first reading, leaving it to the discretion of the reader when, and whether, to tackle the technicalities involved in mastering the more advanced techniques.

Section 2.1 introduces the axioms of linear spaces in full generality, establishes basic properties and explores examples, most of which will be revisited throughout the book. Section 2.2 is concerned with establishing the notion of dimension for an arbitrary linear space. In particular, the dimension need not be finite, which, at times, necessitates some more intricate proofs. Section 2.3 discusses linear operators, the natural choice of structure preserving functions between linear spaces, studies their basic properties, and discusses the notion of isomorphic linear spaces. Section 2.4 introduces standard constructions producing new spaces from given ones, and in particular the kernel and image of a linear operator are discussed. The final section is devoted to inner product spaces and normed spaces and presents several important examples such as function spaces and ℓ_p spaces.

2.1 Linear Spaces—Elementary Properties and Examples

In $\mathbb{R}^2$, the scalar product $\alpha \cdot x$ is between a real number $\alpha \in \mathbb{R}$ and a vector $x \in \mathbb{R}^2$, and yields again a vector in $\mathbb{R}^2$. However, in similar situations, such as in $\mathbb{C}^2$, the scalar multiplication $\alpha \cdot x$ is now defined between any complex number $\alpha \in \mathbb{C}$ and an arbitrary vector x. The most general situation is when α is allowed to vary over the elements of an arbitrary (but fixed) *field* K. Prominent examples of fields are the field $\mathbb{R}$ of real numbers and the field $\mathbb{C}$ of complex numbers (for a more detailed discussion of fields the reader is referred to Sect. 1.2.19 of the Preliminaries). The definition of linear space given below is the result of a distillation of certain key properties of vector addition and scalar multiplication in $\mathbb{R}^2$ or $\mathbb{R}^3$, and is formalized in the most general form, namely with arbitrary fields. If the reader is not familiar with any fields other than $\mathbb{R}$ or $\mathbb{C}$ (or $\mathbb{Q}$, the field of rational numbers), then it is perfectly safe to proceed and replace any occurrence of an arbitrary field K by either $\mathbb{R}$ or $\mathbb{C}$. In the context of this book these are in any case the most important fields.

Definition 2.1 *(Linear Space)* Let K be an arbitrary field (such as the field of real numbers or of complex numbers, for concrete examples). A set V of elements $x, y, z, \ldots$ of an arbitrary nature, together with an operation, called *vector addition*, or simply *addition*, associating with any two elements $x, y \in V$ an element $z \in V$, called the *sum* of x and y, and denoted by $z = x + y$, as well as an operation associating with any $x \in V$ and $\alpha \in K$ an element $w \in V$, called the *product* or *scalar product* of α and x, and denoted by $w = \alpha \cdot x$, is called a *linear space* if

1. For all $x, y, z \in V$, the operation of vector addition satisfies:
   - Associativity, i.e., $(x + y) + z = x + (y + z)$.

- Commutativity, i.e., $x + y = y + x$.
- Existence of a *neutral* element, i.e., there exists an element $0 \in V$, for which

$$x + 0 = x = 0 + x.$$

- Existence of *additive inverses*, i.e., there exists an element $x' \in V$ such that

$$x + x' = 0 = x' + x.$$

2. For all $x \in V$ and $\alpha, \beta \in K$, the scalar product operation satisfies:
   - Associativity, i.e., $\alpha \cdot (\beta \cdot x) = (\alpha\beta) \cdot x$.
   - Neutrality of $1 \in K$, i.e., $1 \cdot x = x$.
3. For all $x, y \in V$ and $\alpha, \beta \in K$, the scalar product and vector addition operations are compatible in the sense that
   - Scalar product distributes over vector addition

$$\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y.$$

   - Scalar product distributes over scalar addition

$$(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x.$$

A linear space is also called a *vector space*, its elements are called *vectors*, and the elements of the field K are called *scalars*. If we wish to emphasize the relevant field, then we say that V is a *linear space over* K. Otherwise, the assertion that V is a linear space includes the implicit introduction of a field K serving as the field of scalars for V.

Scalars will typically be denoted by lower-case Greek letters from the beginning of the alphabet, namely α, β, γ, and so on, while vectors will be denoted by x, y, z, etc. In either case, subscripts or superscripts may be used to enhance readability.

2.1.1 Elementary Properties of Linear Spaces

We now turn to establish several properties of linear spaces that immediately follow from the axioms.

Proposition 2.1 *In any linear space V the following statements hold.*

1. *The neutral element $0 \in V$ is unique.*
2. *For all $x \in V$, the additive inverse x' is unique.*
3. *For all $\alpha \in K$ and $x \in V$, the equation $\alpha \cdot x = 0$ holds if, and only if, $\alpha = 0$ or $x = 0$.*

Proof

1. Suppose that $0' \in V$ is a neutral element. That is
$$x + 0' = x$$
for all $x \in V$, and thus
$$0 = 0 + 0' = 0' + 0 = 0'.$$

2. Suppose that $x'' \in V$ satisfies that $x + x'' = 0$, then
$$x' = x' + 0 = x' + (x + x'') = (x' + x) + x'' = 0 + x'' = x''.$$

3. For $\alpha = 0$
$$\alpha \cdot x = 0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x \implies 0 \cdot x = 0.$$
For $x = 0$
$$\alpha \cdot x = \alpha \cdot 0 = \alpha \cdot (0 + 0) = \alpha \cdot 0 + \alpha \cdot 0 \implies \alpha \cdot 0 = 0.$$
In the other direction, if $\alpha \cdot x = 0$ and $\alpha \neq 0$, then upon multiplication by α^{-1}, one obtains
$$x = 1 \cdot x = (\alpha^{-1} \cdot \alpha) \cdot x = \alpha^{-1} \cdot (\alpha \cdot x) = \alpha^{-1} \cdot 0 = 0. \qquad \square$$

Remark 2.1 It similarly follows that for any vector x, the additive inverse x' is given by $x' = (-1) \cdot x$. It is further common to neglect the $\cdot$ denoting the scalar product, write $x' = -x$ and resort to the familiar conventions for algebraic manipulations on vectors and scalars commonly used for addition and multiplication, e.g., we write $x - y$ for $x + (-y)$ or $x + y + z$ for $(x + y) + z$, and so on. This convention, of course, considerably shortens proofs such as those given above and will be silently used throughout the text below.

Remark 2.2 Most of the linear spaces we shall be concerned with will be over the field $\mathbb{R}$ of real numbers, in which case they are called *real linear spaces*, or over the field $\mathbb{C}$ of complex numbers, in which case they are called *complex linear spaces*. In the absence of further specification, or as implied by context, a linear space is assumed to be either a real or a complex linear space.

Remark 2.3 We make no notational distinction between the *zero vector*, i.e., the neutral element with respect to vector addition, and the element 0 in the field K. This is a generally safe practice, since context typically immediately points to the correct interpretation. For instance, in the equation $0 \cdot x = 0$, context dictates that the 0 on the left-hand-side is the scalar $0 \in K$ while on the right-hand-side it is the zero vector $0 \in V$.

2.1.2 Examples of Linear Spaces

Since linear spaces are very common in mathematics, presenting an exhaustive list of linear spaces is a daunting task. The chosen examples below are meant to present some commonly occurring linear spaces, to explore some less common possibilities, and to familiarize the reader with some linear spaces of great importance in the context of this book. The latter refers to linear spaces of sequences and of functions, linear spaces that will be revisited throughout the rest of the text.

Example 2.1 In the familiar spaces $\mathbb{R}^2$ and $\mathbb{R}^3$, which are the mathematical models of the physical plane and space, vector addition has the geometric interpretation known as the parallelogram law, while scalar multiplication αx has a scaling effect on the vector x determined by the magnitude and sign of α.

Example 2.2 Given a natural number $n \geq 1$, let $\mathbb{R}^n$ be the set of n-tuples of real numbers. Thus if $x \in \mathbb{R}^n$, then it is of the form $x = (x_1, \ldots, x_n)$ where (for every $1 \leq k \leq n$) x_k, the k-th *component of* x, is a real number.

For all $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ defining

$$\begin{aligned} x + y &= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \\ \alpha x &= (\alpha x_1, \alpha x_2, \ldots, \alpha x_n), \end{aligned}$$

endows $\mathbb{R}^n$ with the structure of a linear space over the field $\mathbb{R}$, as is easy to verify. Obviously, the cases $n = 2$ and $n = 3$ recover the familiar linear spaces $\mathbb{R}^2$ and $\mathbb{R}^3$ (respectively). Analogously, the set $\mathbb{C}^n$ of all n-tuples of complex numbers, with similar coordinate-wise operations, is a linear space over the field $\mathbb{C}$.

Example 2.3 Let $\mathbb{R}^\infty$ be the set of infinite sequences of real numbers. Thus, a typical element $x \in \mathbb{R}^\infty$ is of the form $x = (x_1, x_2, \ldots, x_k, \ldots)$ where $x_k \in \mathbb{R}$, for each $k \geq 1$, is a real number, the k-th *component* of x. For all $x, y \in \mathbb{R}^\infty$ and $\alpha \in \mathbb{R}$ setting

$$\begin{aligned} x + y &= (x_1 + y_1, x_2 + y_2, \ldots, x_k + y_k, \ldots) \\ \alpha x &= (\alpha x_1, \alpha x_2, \ldots, \alpha x_k, \ldots) \end{aligned}$$

endows $\mathbb{R}^\infty$ with the structure of a linear space over the field $\mathbb{R}$, as is easily seen. Similarly, the set $\mathbb{C}^\infty$ of all infinite sequences of complex numbers, with similar coordinate-wise operations, is a linear space over the field $\mathbb{C}$.

Remark 2.4 The convention that for an element x in either $\mathbb{R}^n$, $\mathbb{C}^n$, $\mathbb{R}^\infty$, or $\mathbb{C}^\infty$ its k-th *component* is denoted by x_k (as illustrated in the preceding examples) will be used throughout this text. To refer to a sequence of such vectors we may thus use super scripts for the different vectors, e.g., $\{x^{(m)}\}_{m\geq 1}$, and then $x_k^{(m)}$ refers to the k-th component of the m-th vector.

Example 2.4 Consider the subset $c \subseteq \mathbb{C}^\infty$ consisting of all convergent sequences (here convergence is in the usual sense of convergence of sequences of complex numbers), let $c_0 \subseteq c$ be the subset consisting of all sequences that converge to 0, and let $c_{00} \subseteq c_0$ be the set

$$c_{00} = \{(x_1, \dots, x_m, 0, 0, 0, \dots) \mid m \geq 1, \quad x_1, \dots, x_m \in \mathbb{C}\}$$

of all sequences that are eventually 0. With the same definition of addition and multiplication as in Example 2.3, each of these sets is easily seen to be a linear space over $\mathbb{C}$. Obviously, replacing $\mathbb{C}$ throughout by $\mathbb{R}$ yields similar linear spaces over $\mathbb{R}$.

Example 2.5 The constructions given above of the linear spaces $\mathbb{R}^n$, $\mathbb{C}^n$, $\mathbb{R}^\infty$, and $\mathbb{C}^\infty$ easily generalize to any field K and to any cardinality. Indeed, consider an arbitrary set B and an arbitrary field K. Recall from the Preliminaries (Sect. 1.2.8) that the set K^B is the set of all functions $x : B \to K$. For all such functions

$$x : B \to K, \quad y : B \to K$$

and

$$\alpha \in K$$

define the functions $x + y : B \to K$ and $\alpha x : B \to K$ by

$$(x + y)(b) = x(b) + y(b), \quad (\alpha x)(b) = \alpha \cdot x(b).$$

With these notions of addition and scalar multiplication, the set K^B is easily seen to be a linear space over K. In particular, we obtain the linear spaces $\mathbb{R}^{\{1,2,\dots,n\}} = \mathbb{R}^n$, $\mathbb{C}^{\{1,2,\dots,n\}} = \mathbb{C}^n$, $\mathbb{R}^{\mathbb{N}} = \mathbb{R}^\infty$, and $\mathbb{C}^{\mathbb{N}} = \mathbb{C}^\infty$. Further, restricting attention to the subset $(K^B)_0$ consisting only of those functions $x : B \to K$ for which $x(b) = 0$ for all but finitely many $b \in B$ also yields a linear space, with exactly the same definition for addition and scalar multiplication as in K^B. In particular the linear space c_{00} from Example 2.4 is recovered by noticing that $c_{00} = (\mathbb{C}^{\mathbb{N}})_0$. Note however that for a general field K no analogue of the linear spaces c or c_0 need exist since K need not have any notion of convergence.

Example 2.6 For $n \geq 0$, let P_n be the set of all polynomial functions, that is functions of the form $p(t) = a_n t^n + \cdots + a_1 t + a_0$, with real coefficients and degree at most n. With the ordinary operations

$$(p + q)(t) = p(t) + q(t), \quad (\alpha p)(t) = \alpha \cdot p(t)$$

taken as addition and scalar multiplication, it is immediate to verify that P_n is a linear space over $\mathbb{R}$. Removing the restriction on the degrees of the polynomials one obtains the set P of all polynomial functions with real coefficients which, again with the obvious notions of addition and scalar multiplication, forms a linear space over $\mathbb{R}$. One may also consider polynomials with coefficients in the field $\mathbb{C}$ of complex numbers, and obtain similar linear spaces over $\mathbb{C}$.

Example 2.7 Let I be a subset of $\mathbb{R}$ which is either an open interval (a, b), a closed interval $[a, b]$, or the entire real line $\mathbb{R}$. Consider the set $C(I, \mathbb{R})$ of all continuous real-valued functions $x : I \to \mathbb{R}$. The familiar definitions

$$(x + y)(t) = x(t) + y(t), \quad (\alpha x)(t) = \alpha \cdot x(t),$$

when applied to continuous functions $x, y : I \to \mathbb{R}$, are well-known to give continuous functions again, and it is easy to see that when these operations are taken as addition and scalar multiplication, the set $C(I, \mathbb{R})$ is a linear space over $\mathbb{R}$. One may also consider the set $C(I, \mathbb{C})$ of all continuous complex-valued functions $x : I \to \mathbb{C}$ to similarly obtain a linear space over $\mathbb{C}$. One may also consider, for each $k \geq 1$, the set $C^k(I, \mathbb{R})$ of all functions $x : I \to \mathbb{R}$ with a continuous k-th derivative, which is similarly a linear space over $\mathbb{R}$. It is then customary to equate $C^0(I, \mathbb{R})$ with $C(I, \mathbb{R})$. We may also allow $k = \infty$ so as to obtain $C^\infty(I, \mathbb{R})$, the linear space of all infinitely differentiable functions.

Example 2.8 Let K be any field and F a proper subfield of K, namely F is a proper subset of K, and with the induced operations from K it is itself a field. For instance, $\mathbb{Q}$ is a subfield of $\mathbb{R}$ which in turn is a subfield of $\mathbb{C}$. In such a situation any linear space V over the larger field K is also a *different* linear space over the smaller field F. The reason is that, of course, the additive structure of V is unaffected by the choice of field of scalars, whereas the axioms relating to the scalar product, if they hold for scalars ranging over K, then they certainly hold for scalars ranging over F (since all scalar axioms are universal equational quantifications). The process of considering a linear space over K as a linear space over F is named *restriction of scalars*.

In particular, each of the examples above of a linear space over $\mathbb{C}$ is also a linear space over $\mathbb{R}$, and also a linear space over $\mathbb{Q}$. Any linear space obtained by restriction of scalars $F \subset K$ is (except for trivial cases) very different than the original space. Another particular instance of restriction of scalars is the observation that since $\mathbb{R}$ is a linear space over itself (if this is confusing pause for a minute to realize that this is essentially the assertion that $\mathbb{R}^1$ is a linear space), restriction of scalars implies that $\mathbb{R}$ is also a linear space over $\mathbb{Q}$.

Exercises

Exercise 2.1 Let V be a linear space over K and $f : V \to X$ a bijection where X is a set with no a-priori extra structure. Let $f^{-1} : X \to V$ be the inverse function of f. Prove that the operations

$$x + y = f\left(f^{-1}(x) + f^{-1}(y)\right)$$

and

$$\alpha x = f\left(\alpha f^{-1}(x)\right),$$

defined for all $x, y \in X$ and $\alpha \in K$, turn X into a linear space over K.

Exercise 2.2 For any $\alpha \in \mathbb{C}$ let c_α be the set of all sequences $(x_1, x_2, \ldots) \in \mathbb{C}^\infty$ which converge to α. Prove that the linear structure of $\mathbb{C}^\infty$ restricts to a linear structure on c_α if, and only, if $\alpha = 0$.

Exercise 2.3 Let K be a field (such as $\mathbb{R}$ or $\mathbb{C}$ for familiarity) and let $M_{n,m}(K)$ be the set of $n \times m$ matrices with entries in the field K. Prove that with ordinary matrix addition and scalar product, the set $M_{n,m}(K)$ is a linear space over K.

Exercise 2.4 With the usual operations of addition and multiplication, is the set $(\mathbb{R} - \mathbb{Q}) \cup \{0\}$ a linear space over $\mathbb{Q}$?

Exercise 2.5 Let V be a linear space over K and fix some vector $w_0 \in V$. Define on the set V an addition operation by $x \oplus y = x + y - w_0$ and a scalar product operation by $\alpha \odot x = \alpha(x - w_0)$. Prove that with these operations V is a linear space over K whose zero vector is w_0.

Exercise 2.6 Prove that the set $\mathbb{R}$ of real numbers with addition given by $x \oplus y = xy$ and scalar multiplication given by $\alpha \odot y = y^\alpha$ is a linear space over the field $\mathbb{R}$.

Exercise 2.7 Let X be a set with one element. Prove that for any field K there is a unique choice of operations that turns X into a linear space over K.

Exercise 2.8 Prove that any linear space over $\mathbb{R}$ has either just a single vector or infinitely many.

Exercise 2.9 Let V be a linear space. Prove that $-x = (-1) \cdot x$ holds for all vectors $x \in V$.

Exercise 2.10 Let K be a field and $\mathscr{B}$ an arbitrary set. Consider the set $\langle \mathscr{B} \rangle_K$ of all formal expressions of the form

$$\sum_{b \in \mathscr{B}} \alpha_b \cdot b$$

where $\alpha_b \in K$ for each $b \in \mathscr{B}$ and at most finitely many of the α_b are non-zero. Verify that the obvious way to define addition and scalar multiplication on the set $\langle \mathscr{B} \rangle_K$ turns it into a linear space over K (known as the *free linear space generated by* $\mathscr{B}$).

2.2 The Dimension of a Linear Space

The familiar linear spaces $\mathbb{R}^n$ are all infinite sets (in fact, they all have the same cardinality, the cardinality c of the continuum). However, it is intuitively clear that $\mathbb{R}^3$ is considerably 'larger' than, say, $\mathbb{R}^2$. This fact is usually expressed in the claim that $\mathbb{R}^3$ has dimension 3, while $\mathbb{R}^2$ has dimension 2. The notion of dimension in general linear spaces is rather subtle, especially for the infinite dimensional ones. We now attend to investigate this situation in full generality, starting with the related concepts of linear independence and spanning sets.

2.2.1 Linear Independence, Spanning Sets, and Bases

By forming linear combinations of vectors from any given set of vectors in a linear space, one obtains a (potentially) large collection of vectors. The original set is said to be linearly independent if, intuitively, it does not allow for any redundancies when forming linear combinations while it is a spanning set if it is sufficiently large to generate any vector as a linear combination. The precise definitions follow.

Definition 2.2 Given any set $S \subseteq V$ of vectors in a linear space V (S may be finite or infinite), a *linear combination* of elements of S is any vector of the form

$$x = \sum_{k=1}^{m} \alpha_k x_k$$

with $x_1, \ldots, x_m$ vectors from S and $\alpha_1, \ldots, \alpha_m \in K$ arbitrary scalars. Equivalently,

$$x = \sum_{s \in S} \alpha_s \cdot s$$

with $\alpha_s = 0$ for all but finitely many $s \in S$. The two forms are essentially the same, differing only notationally. The *span* of S is then the set of all linear combinations of elements from S, and S is said to be a *spanning set* if its span is the entire linear space V. A spanning set S is a *minimal spanning set* if it is itself a spanning set but no proper subset of it is a spanning set. Further, S is said to be a set of *linearly independent* vectors if the only possibility of expressing the zero vector as a linear combination of elements from S is the trivial linear combination, i.e., where all the coefficients are 0. That is, S is linearly independent if whenever one has

$$0 = \sum_{k=1}^{m} \alpha_k x_k,$$

with $x_1, \ldots, x_m$ vectors in S, then necessarily $\alpha_k = 0$ for all $1 \leq k \leq m$. The set S is said to be a *maximal linearly independent* set if it is itself linearly independent but any set that properly contains it is not linearly independent. A set that is not linearly independent is also referred to as a *linearly dependent* set. Finally, a set that is both a spanning set and is linearly independent is said to be a *basis* of the linear space.

Remark 2.5 We speak of vectors $x_1, \ldots, x_m \in V$ as being either spanning or linearly independent, if the set $\{x_1, \ldots, x_m\}$ is spanning or linearly independent. Of course, we may also consider countably infinitely many vectors $x_1, x_2, \ldots$ as being spanning or linearly independent, in a similar fashion.

Example 2.9 The situation in $\mathbb{R}^n$ is probably very familiar to the reader. Any m vectors $x_1, \ldots, x_m$ in $\mathbb{R}^n$ are linearly independent if, and only if, the equation

$$\sum_{k=1}^{m} \alpha_k x_k = 0$$

admits the unique solution $\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$. It is well-known that linear independence implies $m \leq n$. Similarly, the given vectors are spanning if, and only if, for every vector $b \in \mathbb{R}^n$ the equation

$$\sum_{k=1}^{m} \alpha_k x_k = b$$

admits a solution. It is again a familiar fact that if the given vectors are spanning, then $m \geq n$. It thus follows that a basis for $\mathbb{R}^n$ must consist of precisely n vectors. In particular, all bases have the same size, namely n, which is referred to as the dimension of $\mathbb{R}^n$. Below we prove that every linear space has a dimension, if we allow infinite cardinalities into the picture. The result in that generality subsumes the properties of $\mathbb{R}^n$ just mentioned.

Example 2.10 In the space $\mathbb{C}^2$, considered as a linear space over $\mathbb{C}$, the vectors $(1, 0)$ and $(0, 1)$ are immediately seen to form a basis. However, if $\mathbb{C}^2$ is considered as a linear space over $\mathbb{R}$ (by the procedure of restriction of scalars from Example 2.8), then these two vectors are (of course) still linearly independent but they fail to span $\mathbb{C}^2$. Indeed, since only real scalars may now be used to form linear combinations of these vectors, the span will only be $\mathbb{R}^2$. To obtain a basis, the two vectors need to be augmented, for instance, by the vectors $(i, 0)$ and $(0, i)$. The four vectors together do form a basis of $\mathbb{C}^2$.

Example 2.11 In the linear space $\mathbb{R}^n$ or $\mathbb{C}^n$, the vectors $x_1, \ldots, x_n$ where

$$x_k = (0, \ldots, 0, 1, 0, \ldots, 0)$$

with 1 in the k-th position, are easily seen to be spanning and linearly independent, and thus form a basis, called the *standard basis* of $\mathbb{R}^n$, respectively $\mathbb{C}^n$. It is obvious that $\mathbb{R}^n$ and $\mathbb{C}^n$ have infinitely many bases.

In the examples presented so far it was quite straightforward to obtain a basis. The following example shows that this is not always the case. In fact, it is not even clear that the next linear space even has a basis.

Example 2.12 Let $\mathbb{C}^\infty$ be the linear space from Example 2.3. The vectors $x_1, x_2, \ldots$, given by

$$x_k = (0, \ldots, 0, 1, 0, \ldots)$$

with 1 in the k-th position are easily seen to be linearly independent, but they do not form a spanning set. Indeed, since the span consists only of the *finite* linear combinations of vectors from the set, the span in this case is the set of all vectors of the form $(a_1, \ldots, a_k, 0, \ldots, 0, 0, \ldots)$, namely those infinite sequences of complex numbers that are eventually 0. In other words, the span in $\mathbb{C}^\infty$ of the vectors $x_1, x_2, \ldots$ is the space c_{00} from Example 2.4, and thus we incidentally found a basis for c_{00}. It is now tempting to proceed as follows. Taking any vector $y_1 \in \mathbb{R}^\infty$ which is not spanned by $x_1, x_2, \ldots$, for instance the vector $y_1 = (1, 1, 1, \ldots, 1, \ldots)$, forming the set $\{y_1, x_1, x_2, \ldots\}$ must get us closer to obtaining a basis. Indeed, the new set is still linearly independent precisely because y_1 was not spanned by the rest of the vectors. But, this new set is still not a basis as there are still many vectors it fails to span, for instance the vector $y_2 = (1, 0, 1, 0, 1, 0, \ldots)$. Of course, we may now consider the larger set $\{y_1, y_2, x_1, x_2, \ldots\}$, but it too fails to be a basis. One may attempt to resolve the argument once and for all by claiming that proceeding in this way to infinity will eventually result in a basis. However, this is a very vague statement, and even if this process can be carried out mathematically (which it can, as will be shown below), it is entirely unclear as to which vectors will end up in the basis and which will not. In other words, even if this linear space has a basis, it is unlikely we can ever present one.

Naturally, similar observations hold true for $\mathbb{R}^\infty$ instead of $\mathbb{C}^\infty$, and in fact to most of the linear spaces in this book, and in analysis in general.

Example 2.13 Recall (see Example 2.8) that one can consider the space $\mathbb{R}$ as a linear space over the field $\mathbb{Q}$ of rational numbers (as a particular case of restriction of scalars). A real number α is said to be *transcendental* if it is not the root of a polynomial with rational coefficients. Examples of transcendental numbers include e and π, though the proofs are far from trivial. For any transcendental number x the set $\{1, x, x^2, x^3, \ldots\}$ is linearly independent (this is basically the definition of x being transcendental), but it is not a spanning set.

Example 2.14 Recall the space P_n from Example 2.6 of all polynomial functions with real coefficients of degree at most n. For every $k \geq 0$ let p_k be the vector $p_k(t) = t^k$. Clearly, the vectors $p_0, p_1, p_2, \ldots, p_n$ form a basis for P_n, but again there are infinitely many other choices of bases. The space P of all polynomial functions with real coefficients has a countable basis given by $p_0, p_1, p_2, \ldots$, a fact that is easily verified. Noticing that polynomials are continuous functions, no matter on which interval they are defined, we see by the above that in the linear space $C(I, \mathbb{R})$ of all continuous functions $x : I \to \mathbb{R}$, where I is a non-degenerate interval (i.e., not reducing to a point) the vectors $p_0, p_1, p_2, \ldots$ are linearly independent. However, they do not form a basis since any linear combination of these vectors is again a polynomial function, but not all continuous functions are polynomial functions. Once more, it is not at all clear that $C(I, \mathbb{R})$ even has a basis (what would one look like?).

From the discussion above we see that in some linear spaces, such as $\mathbb{R}^n$, P_n, or P, it is quite easy to find bases while in other linear spaces, such as $\mathbb{R}^\infty$, $\mathbb{R}$ over $\mathbb{Q}$, or $C(I, \mathbb{R})$, it is a highly non-trivial task. In more detail, we saw that it is rather simple to exhibit a large set of linearly independent vectors, but it is not so easy to have these vectors also span the entire space (the spaces c_{00} and P, as discussed above, are the exception to the rule). In fact, it is not at all clear that one can find bases for V in the case of $\mathbb{R}^\infty$ or $\mathbb{C}^\infty$, as well as for $\mathbb{R}$ as a linear space over $\mathbb{Q}$, or for $C(I, \mathbb{R})$.

2.2.2 *Existence of Bases*

In order to establish that every linear space does have a basis, we observe some immediate facts. It should be noted at once that the existence proof uses Zorn's Lemma in an essential way. That is, it can be shown that if every linear space has a basis, then the Axiom of Choice holds. Consequently, the existence proof is not constructive.

Proposition 2.2 *Let $S \subseteq V$ be a set of vectors in a linear space V. The following conditions are equivalent.*

1. *S is a maximal linearly independent set.*
2. *S is a minimal spanning set.*
3. *S is a basis.*

Proof First we show that if S is a maximal linearly independent set, then it is a basis. All that is needed is to show that S is a spanning set. To that end, let $x \in V$ be a vector in the ambient linear space. If $x \in S$, then it is certainly spanned by S. If $x \notin S$, then by virtue of S being a maximal linearly independent set, the set $S \cup \{x\}$ is linearly dependent. Thus, there exist vectors $x_1, \ldots, x_m \in S \cup \{x\}$ and non-zero scalars $\alpha_1, \ldots, \alpha_m \in K$ with

$$0 = \sum_{k=1}^{m} \alpha_k x_k.$$

However, in the expression above it must be that x itself appears as a summand since otherwise we would have expressed 0 as a non-trivial linear combination of vectors from the linearly independent set S. We may thus isolate x in the expression above to obtain it as a linear combination of elements from S. As x was arbitrary, we conclude that S is a spanning set.

Next we show that if S is a minimal spanning set, then it is a basis. All that is needed is to show that S is linearly independent, and indeed, if it were linearly dependent, then we would have some expression as above, giving 0 as a non-trivial linear combination of vectors from S. Using that expression we are able to isolate some vector $x \in S$ and exhibit it as a linear combination of other vectors from S. It is then easy to see that $S - \{x\}$ is still a spanning set (simply since any linear combination containing x can be replaced by one that does not). But this contradicts S being a minimal spanning set, and thus S must be linearly independent.

So far we have shown that each of conditions 1 and 2 implies condition 3. The proof will be completed by showing the converse of these implications. The details are very similar in spirit to those given so far, and thus the rest of the proof is left for the reader. □

We are now ready to establish that every linear space has a basis. The proof makes essential use of Zorn's Lemma (see Sect. 1.2.17 of the Preliminaries) and the reader may safely choose to skip the proof on a first reading. To the reader interested in the technique of Zorn's Lemma we remark that the proof below is actually a straightforward application with little technical difficulties, and is thus a fortunate first encounter with this important proof technique.

Theorem 2.1 *Every linearly independent set $A \subseteq V$ in a linear space V can be extended to a maximal linearly independent set. In particular, by considering the linearly independent set $\emptyset \subseteq V$ and by Proposition 2.2, it follows that every linear space V has a basis.*

Proof Consider the set P of all linearly independent subsets S of V that contain A, which we order by set inclusion. Evidently P is a poset, and to find a maximal linearly independent set that extends A amounts to finding a maximal element in P, and so we apply Zorn's Lemma (Sect. 1.2.17 of the Preliminaries). First we note that P is certainly not empty since clearly $A \in P$. Now, assume that $\{S_i\}_{i \in I}$ is a chain in P. We will show that $S = \bigcup_{i \in I} S_i$ is an upper bound for the chain. Clearly $S_i \subseteq S$ for all $i \in I$, thus the only thing to show is that $S \in P$, namely that S contains A (which is immediate) and that it is linearly independent. For that, assume that

$$0 = \sum_{k=1}^{m} \alpha_k x_k$$

is a non-trivial linear combination of the vectors $x_1, \ldots, x_m \in S$. Since S is the union of the S_i, it follows that $x_m \in S_{f(m)}$ for a suitable $f(m) \in I$, but since $\{S_i\}_{i \in I}$ is a chain it follows that there is a single index $i_0 \in I$ such that $x_1, \ldots, x_m \in S_{i_0}$. But then the equality above expresses the zero vector as a non-trivial linear combination of vectors from S_{i_0}, contradicting the fact that S_{i_0} is linearly independent. With that the conditions of Zorn's Lemma are satisfied, and so the existence of a maximal element in P is guaranteed. This maximal element is a set $S_M \in P$, namely S_M contains A and S_M is a maximal set of linearly independent vectors, as required. □

2.2.3 Existence of Dimension

Now that we know that every linear space has at least one basis, it is tempting to define the dimension of a linear space to be the cardinality of its basis. However, for this to make sense we need to know that all bases have the same cardinality. While this is a very plausible assertion (certainly for such familiar spaces as $\mathbb{R}^n$), it requires a careful proof, particularly in the infinite dimensional case. The important ingredient is the following lemma, which states that the cardinality of any linearly independent set is not greater than the cardinality of any spanning set. This result again uses Zorn's Lemma and its proof may safely be skipped on a first reading. A word of caution to the ambitious reader that this proof, compared to the proof of Theorem 2.1, is technically more demanding.

Lemma 2.1 *If V is a linear space, $I \subseteq V$ a linearly independent set of vectors, and $S \subseteq V$ a spanning set, then there exists an injective function $f : I \to S$.*

Proof Consider a pair (J, f) where $J \subseteq I$ and $f : J \to S$ is an injection. The idea will be (using Zorn's Lemma) to keep on extending the domain of f until the entire set I is exhausted. An important ingredient for achieving that is the following. Thinking of the injective function $f : J \to S$ as an instruction to replace the vectors in J by their images in S, we only consider such pairs (J, f) for which $(I - J) \cup f(J)$ is still a linearly independent set. Let us now form the set P of all such pairs and introduce an ordering on it declaring that $(J_1, f_1) \le (J_2, f_2)$ precisely when $J_1 \subseteq J_2$ and f_2 extends f_1 (the latter means that $f_1(x) = f_2(x)$ holds for all $x \in J_1$). It is immediate that P is a poset, and that a maximal element in it will furnish us with an injective function $f : J_M \to S$ for a very large subset $J_M \subseteq I$. We will then show that necessarily $J_M = I$, and the result will be established.

To verify the conditions of Zorn's Lemma, notice first that P is not empty. Indeed, the pair $(\emptyset, \emptyset)$ is in P. Next, let $\{(J_t, f_t)\}_{t \in T}$ be a chain in P, for which we must now find an upper bound. Let

$$J = \bigcup_{t \in T} J_t$$

and notice that we may define $f : J \to S$ as follows. Given $x \in J$ there is some $t \in T$ such that $x \in J_t$, and so let $f(x) = f_t(x)$. To see that this is a well-defined function, i.e., that it is independent of the choice of $t \in T$, note that if $x \in J_{t'}$, then either $J_t \subseteq J_{t'}$ or $J_{t'} \subseteq J_t$ and then either $f_{t'}$ extends f_t or f_t extends $f_{t'}$, and in either case $f_t(x) = f_{t'}(x)$. It is clear that $J_t \subseteq J$ and that f extends f_t, for all $t \in T$, so all we need to do in order to show that (J, f) is an upper bound for the chain is establish that $(J, f) \in P$. Clearly, $J \subseteq I$, and $f : J \to S$ is injective (since we saw that in fact for any two $x_1, x_2 \in J$ the function f agrees with some f_t, which is injective, so $f(x_1) = f(x_2) \implies f_t(x_1) = f_t(x_2) \implies x_1 = x_2$). So it remains to observe that $(I - J) \cup f(J)$ is a linearly independent set of vectors. Indeed, if 0 can be obtained as a non-trivial linear combination of the vectors $x_1, \dots, x_m \in (I - J) \cup f(J)$, then, using the chain condition again, it follows that there exists an element $t \in T$ such that $x_1, \dots, x_m \in (I - J_t) \cup f_t(J_t)$, which is a linearly independent set, yielding a contradiction.

With the conditions of Zorn's Lemma now verified, it follows that there exists a maximal element (J_M, f_M), with $J_M \subseteq I$ and $f_M : J_M \to S$ an injection. If $J_M = I$, then we are done, since $f_M : I \to S$ is the required injection. Assuming this is not the case, let $x_! \in I - J_M$. If we can find a vector $y_! \in S - f_M(J_M)$ such that

$$(I - (J_M \cup \{x_!\})) \cup (f_M(J_M) \cup \{y_!\})$$

is linearly independent, then we will have that $(J_M \cup \{x_!\}, f_!) \in P$, where $f_!$ is the extension of f_M given by $f_!(x_!) = y_!$. But that would contradict the maximality of (J_M, f_M), and we will have our contradiction. So, we proceed to prove the existence of such a vector $y_!$. If no such $y_!$ exists, then that means that the set

$$(I - (J_M \cup \{x_!\})) \cup (f_M(J_M) \cup \{y\})$$

is a linearly dependent set for every $y \in S - f_M(J_M)$. Now, since S is a spanning set we may write

$$x_! = \sum_{s \in S} \alpha_s \cdot s = \sum_{s \in f_M(J_M)} \alpha_s \cdot s + \sum_{s \in S - f_M(J_M)} \alpha_s \cdot s = x_1 + x_2$$

where the sum is a finite sum, so that $\alpha_s = 0$ for all but finitely many s, and we simply split the sum according to whether or not $s \in f_M(J_M)$. By our assumption, the set

$$(I - (J_M \cup \{x_!\})) \cup f_M(J_M)$$

is linearly dependent if any of the vectors $s \in S - f_M(J_M)$ is added to it. Thus, this set is linearly dependent if any linear combination of such vectors, such as x_2, is added to it. Further, since x_1 is in the span of $f_M(J_M)$ it follows that the set above is linearly dependent if $x = x_1 + x_2$ is added to it. But the latter is the set

$$(I - (J_M \cup \{x_!\})) \cup f_M(J_M) \cup \{x_!\} = (I - J_M) \cup f_M(J_M)$$

which is linearly independent. The proof is now complete. □

We are now in position to establish the following important result.

Theorem 2.2 *Let V be a linear space. If B_1 and B_2 are any two bases for V, then they have the same cardinality.*

Proof By definition of basis, B_1 is linearly independent and B_2 is a spanning set. By Lemma 2.1, there is an injective function $f : B_1 \to B_2$. By the same argument there is also an injective function $g : B_2 \to B_1$. It now follows from the Cantor-Schröder-Bernstein theorem (Theorem 1.1 in the Preliminaries) that the cardinalities of B_1 and B_2 are equal. □

Definition 2.3 The *dimension* of a linear space V, indicated by $\dim(V)$, is equal to the cardinality of a basis for it. Notice that by Theorem 2.1, at least one basis exists, and by Theorem 2.2 all bases have the same cardinality. Thus, the notion of dimension is well-defined. The linear space V is said to be *finite dimensional* if its dimension is finite, and *infinite dimensional* otherwise. A basis for an infinite dimensional linear space is also referred to as a *Hamel basis*.

Example 2.15 Finite dimensional linear spaces, by Examples 2.11 and 2.14, include the spaces $\mathbb{R}^n$ and $\mathbb{C}^n$, having dimension n, and the space P_n of polynomials (see Example 2.6), which is of dimension $n+1$. Infinite dimensional linear spaces include, due to Examples 2.12 and 2.14, the space c_{00} of sequences that are eventually 0, the space P of all polynomials with real coefficients, and the space $C(I, \mathbb{R})$ of continuous function $x : I \to \mathbb{R}$, as long as I does not reduce to a point.

We have already encountered linear spaces of very large dimension. To argue about the dimension of such spaces (and in general) the following result, interesting on its own, is useful. Recall from Example 2.5 the linear space $(K^B)_0$ of all functions $x : B \to K$ satisfying $x(b) = 0$ for all but finitely many $b \in B$.

Proposition 2.3 *Let V be a linear space and B a basis for it. Then every vector $x \in V$ can be expressed uniquely as a linear combination of elements from B. In other words, there is a bijective correspondence between the vectors $x \in V$ and the elements of $(K^B)_0$.*

Proof Since a typical element in $(K^B)_0$ is nothing but a function $x : B \to K$ for which $x(b) = 0$ for all but finitely many $b \in B$, we may associate with each such x the vector $\sum_{b \in B} x(b) \cdot b$ (as the summation is finite). We claim that this correspondence is the desired bijection. Indeed, it is a tautology that the surjectivity of this process is precisely the claim that B is a spanning set, while it is almost a tautology that the injectivity of the process is the claim that B is linearly independent. □

We thus see that a basis B endows a linear space with a notion of coordinates. This is certainly a useful thing to have, but for practical reasons it is only as useful as the ability to explicitly describe the basis B. In finite dimensional linear spaces it is very common to work with bases but, as we remarked earlier, in infinite dimensional linear spaces being able to explicitly describe a basis is the exception rather than the rule. Consequently, Hamel bases in infinite dimensional linear spaces are generally used for theoretical rather than practical purposes.

Example 2.16 Recall from Example 2.8 that $\mathbb{R}$ may be viewed as a linear space over $\mathbb{Q}$. Suppose that $\mathscr{B}$ is a Hamel basis for $\mathbb{R}$ over $\mathbb{Q}$. By Proposition 2.3, there is then a bijection between $\mathbb{R}$ and the set $(\mathbb{Q}^{\mathscr{B}})_0$ of all functions $\mathscr{B} \to \mathbb{Q}$ which attain 0 at all but finitely many arguments. In other words, $|\mathbb{R}| = |(\mathbb{Q}^{\mathscr{B}})_0|$ (see Sect. 1.2.11 of the Preliminaries for the basics of cardinalities). If $\mathscr{B}$ is countable, then it is easy to write $(\mathbb{Q}^{\mathscr{B}})_0$ as a countable union of countable sets (refer to Sect. 1.2.13 from the Preliminaries for the relevant material) and thus the set $(\mathbb{Q}^{\mathscr{B}})_0$ itself would be countable. But that would imply that $\mathbb{R}$ is a countable set while the reals are well-known to be uncountable. We conclude that $\mathbb{R}$, as a linear space over $\mathbb{Q}$, is infinite dimensional of uncountable dimension.

We close this section by illustrating a difference between finite dimensional linear spaces and infinite dimensional ones.

Proposition 2.4 *The cardinality of any linearly independent set A in a linear space V is a lower bound for the dimension of V. Moreover, if V is finite dimensional, then any linearly independent set $A \subseteq V$ whose cardinality is equal to the dimension of V is a basis.*

Proof The dimension of V is the cardinality of any basis B, and as B is in particular a spanning set, it follows from Lemma 2.1 that there is an injection $A \to B$, and so the cardinalities satisfy $|A| \leq |B|$ (refer to Sect. 1.2.11 of the Preliminaries, if needed). The claim now follows since the latter is the dimension of V.

Now, if V is finite dimensional, say of dimension n, and $A = \{x_1, \ldots, x_n\}$ is a set of n linearly independent vectors, then to show A is a basis we just need to prove that it is a spanning set. But if it were not, and $y \in V$ is any vector not in its span, then the set $\{x_1, \ldots, x_n, y\}$ is linearly independent and contains $n + 1$ vectors. But then, by the first part of the proposition, it would follow that $n + 1 \leq n$, an absurdity.

□

Remark 2.6 The finite dimensionality assumption is crucial. For instance, in the linear space c_{00} of sequences which are eventually 0 (Example 2.4), consider the vectors $\{x_k\}_{k \geq 1}$ where $x_k = (0, \ldots, 0, 1, 0, \ldots)$, with 1 in the k-th position. These vectors are easily seen to be linearly independent and spanning, thus they form a basis of countably many vectors. The dimension of c_{00} is thus infinitely countable. The set $\{x_2, x_3, \ldots\}$ is clearly also linearly independent, has the same cardinality as the dimension of c_{00}, but it is not spanning, and thus not a basis.

Exercises

Exercise 2.11 Let V be a linear space and $S, S' \subseteq V$ two arbitrary sets of vectors with $S \subseteq S'$. Prove that if S' is linearly independent, then so is S, and prove that if S is a spanning set, then so is S'.

Exercise 2.12 Let V be a finite dimensional linear space and S a spanning set. Prove that S can be sifted to give a basis, that is, show that there exists a subset $S' \subseteq S$ such that S' is a basis for V.

Exercise 2.13 Let V be a linear space (not necessarily finite dimensional) and S a spanning set. Prove that S can be sifted to give a basis, that is, show that there exists a subset $S' \subseteq S$ such that S' is a basis for V.

Exercise 2.14 Consider $\mathbb{R}$ as a linear space over $\mathbb{Q}$ and let α be a transcendental number. Prove that the set $\{1, \alpha, \alpha^2, \alpha^3, \ldots\}$ is linearly independent. Prove that it is not a spanning set by showing that its span is a countable set.

Exercise 2.15 Consider $\mathbb{R}$ as a linear space over $\mathbb{Q}$. Prove that the dimension of $\mathbb{R}$ over $\mathbb{Q}$ is $|\mathbb{R}|$, the cardinality of the real numbers.

Exercise 2.16 Let V be a linear space and $S \subseteq V$ a set of vectors. For a scalar $\alpha \in K$, let us write

$$\alpha S = \{\alpha x \mid x \in S\}.$$

Assuming that $\alpha \neq 0$ is fixed, prove that S is linearly independent (respectively spanning, a basis) if, and only if, αS is linearly independent (respectively spanning, a basis).

Exercise 2.17 Let V be a linear space and $\{x_k\}_{k \in \mathbb{N}}$ countably many vectors in V. For all $m \in \mathbb{N}$, let

$$y_m = \sum_{k=1}^{m} x_k.$$

Prove that $\{x_k\}_{k \in \mathbb{N}}$ is linearly independent (respectively spanning, a basis) if, and only if, $\{y_k\}_{k \in \mathbb{N}}$ is linearly independent (respectively spanning, a basis).

Exercise 2.18 Let B be an arbitrary set and K a field. Consider the linear spaces K^B and $(K^B)_0$. For every $b_0 \in B$ let $x_{b_0} : B \to K$ be the function

$$x_{b_0}(b) = \begin{cases} 1 & \text{if } b = b_0, \\ 0 & \text{otherwise.} \end{cases}$$

Prove that the set $\{x_{b_0}\}_{b_0 \in B}$ is linearly independent in K^B. Is it a basis for K^B? Is it a basis for $(K^B)_0$?

Exercise 2.19 In the linear space $C(\mathbb{R}, \mathbb{R})$, find three vectors $x, y, z : \mathbb{R} \to \mathbb{R}$ such that $\{x, y, z\}$ is a linearly independent set while $\{x^2, y^2, z^2\}$ is linearly dependent. (Here x^2 refers to the function given by $x^2(t) = (x(t))^2$, and similarly for the other functions.)

Exercise 2.20 Consider the space c_0 of sequences of complex (or real, if you like) numbers that converge to 0. How many linearly independent vectors can you find in c_0?

2.3 Linear Operators

The definition and properties of the structure preserving functions between linear spaces form the topic of this section. The definitions (including that of a linear isomorphism) and the results are discussed and exemplified in the context of the linear spaces introduced above.

Definition 2.4 Let V and W be linear spaces over the same field K. A function $T : V \to W$ is said to be *additive* if

$$T(x + y) = T(x) + T(y)$$

for all vectors $x, y \in V$, and it is said to be *homogenous* if

$$T(\alpha x) = \alpha T(x)$$

for all vectors $x \in V$ and all scalars $\alpha \in K$. If T is both additive and homogeneous, then it is called a *linear operator*, a condition equivalent to the equality

$$T(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 T(x_1) + \alpha_2 T(x_2)$$

for all $x_1, x_2 \in V$ and $\alpha_1, \alpha_2 \in K$.

Remark 2.7 Synonymous terms for linear operator are *linear transformation* and *linear homomorphism*. Throughout the book we will adopt the convention that any reference to a linear operator $T : V \to W$ immediately implies, implicitly at times, that V and W are linear spaces over the same field K.

It is a straightforward proof by induction that a linear operator T preserves any linear combination, that is

$$T(\sum_{k=1}^{m} \alpha_k x_k) = \sum_{k=1}^{m} \alpha_k T(x_k)$$

for all scalars $\alpha_1, \ldots, \alpha_m \in K$ and vectors $x_1, \ldots, x_m \in V$. In fact, this property can equivalently be taken as the definition of linear operator. The next result shows that the linearity requirement of a linear operator forces any linear operator to also respect the zero vector, additive inverses, and subtraction.

Proposition 2.5 *For a linear operator $T : V \to W$:*

1. $T(0) = 0$.
2. $T(-x) = -T(x)$ *for all* $x \in V$.
3. $T(x - y) = T(x) - T(y)$ *for all* $x, y \in V$.

Proof

1. Notice that $T(0) + T(0) = T(0 + 0) = T(0)$, and now subtract $T(0)$ from both sides of the equation.
2. $T(x) + T(-x) = T(x + (-x)) = T(0) = 0$.
3. $T(x - y) = T(x + (-y)) = T(x) + T(-y) = T(x) - T(y)$. □

2.3.1 Examples of Linear Operators

Trivial, but important, examples of linear operators are the following. For any linear space V, the identity function $\mathrm{id} : V \to V$ is always a linear operator. On the other extreme, given linear spaces V and W over the same field, the function $T_0 : V \to W$ given by $T_0(x) = 0$ is also immediately seen to be a linear operator.

In the presence of a basis for the domain V, linear operators $T : V \to W$ are easily characterized by the following result, readily yielding an endless supply of examples.

Lemma 2.2 *Let V and W be linear spaces over the same field K, and let B be a basis for V. It then holds that any function $F : B \to W$ extends uniquely to a linear operator $T_F : V \to W$.*

Proof For a given function $F : B \to W$, suppose that $T_F : V \to W$ is a linear operator extending F. Then, given any vector $x \in V$, write

$$x = \sum_{b \in B} \alpha_b \cdot b$$

as a (finite) linear combination of basis elements, and then it follows that

$$T_F(x) = T_F(\sum_{b \in B} \alpha_b \cdot b) = \sum_{b \in B} \alpha_b \cdot T_F(b) = \sum_{b \in B} \alpha_b \cdot F(b).$$

Noticing that the computation above expresses $T(x)$ in terms of vectors of the form $F(b)$ we conclude that an extension T_F, if it exists, is unique. If we now take the equality above as a definition, then we obtain a function $T_F : V \to W$ (relying on the uniqueness of the linear combination expressing x in order to assure that T_F is well-defined). Verifying that this T_F is indeed a linear operator and that it extends F follow immediately. □

Remark 2.8 The linear operator $T_F : V \to W$ constructed from the given function $F : B \to W$ above is said to be obtained by *linearly extending* F to all of V.

Often enough the technique of the last result is inadequate, either because a basis is not readily available or because a more direct formula for the linear operator is obtainable. The following examples illustrate these possibilities.

Example 2.17 The derivation operation d/dt, taking a function f to its derivative df/dt, satisfies the well-known properties

$$\frac{d}{dt}(f+g) = \frac{d}{dt}(f) + \frac{d}{dt}(g)$$

and

$$\frac{d}{dt}(\alpha f) = \alpha \frac{d}{dt}(f),$$

namely, it is additive and homogenous, and thus, if we choose the domain and codomain correctly, we expect it to be a linear operator. To turn this observation into a precise statement, recall the linear spaces $C^k([a,b],\mathbb{R})$ from Example 2.7 of all functions $x : [a,b] \to \mathbb{R}$ with a continuous k-th derivative. Then, for every $k \geq 1$, the observation above can be stated by saying that $d/dt : C^k([a,b],\mathbb{R}) \to C^{k-1}([a,b],\mathbb{R})$ is a linear operator. Similarly, $d/dt : C^\infty([a,b],\mathbb{R}) \to C^\infty([a,b],\mathbb{R})$ is a linear operator on the linear space of infinitely differentiable functions.

Example 2.18 The integral operator $\int_a^b dt\ f(t)$ is also well known to be additive and homogenous, and, since every continuous function on a closed interval is integrable, we obtain that for every interval $I = [a,b]$ the function

$$f \mapsto \int_a^b dt\ f(t)$$

is a linear operator from the linear space $C(I,\mathbb{R})$ of Example 2.7 to $\mathbb{R}$ as a linear space over itself.

Example 2.19 Spaces of infinite sequences, such as $\mathbb{R}^\infty$ from Example 2.3 or the space c_0 from Example 2.4 admit the following operators (we use $\mathbb{R}^\infty$ just to fix one possibility). The *shift operator* $S : \mathbb{R}^\infty \to \mathbb{R}^\infty$, given, for $x = (x_1, x_2, \ldots, x_k \ldots)$, by $S(x) = (x_2, x_3, \ldots, x_k, \ldots)$, is easily seen to be a linear operator. It is equally easy to see that the function $T : \mathbb{R}^\infty \to \mathbb{R}^\infty$, given for $x = (x_1, x_2, \ldots, x_k \ldots)$ as above by $T(x) = (0, x_1, x_2, \ldots, x_k, \ldots)$ is a linear operator. Note that $S \circ T$ is the identity while $T \circ S$ is not, a phenomenon that is known to be impossible in finite dimensional linear spaces. These operators are the *creation* and *annihilation* operators $a^\dagger$ and a, widely used in Quantum Mechanics.

We see thus that even when a basis can be given explicitly (and certainly when it cannot) it may be much simpler to define a linear operator directly rather than by linear extension on a basis.

2.3.2 Algebra of Operators

With suitable domain and codomain, linear operators can be composed or added. We now investigate the resulting algebraic laws related to these operations.

Proposition 2.6 *The composition $S \circ T : U \to W$ of any two linear operators $U \xrightarrow{T} V \xrightarrow{S} W$ (where in particular all linear spaces are over the same field) is a linear operator.*

Proof The additivity of $S \circ T$ follows from the definition of the composition and the computation

$$S(T(x+y)) = S(T(x) + T(y)) = S(T(x)) + S(T(y)),$$

valid for all vectors $x, y \in U$, utilizing the additivity of T and S. The homogeneity of the composition follows similarly and the reader is invited to fill in the details. □

Next, any two functions $T, S : V \to W$ between linear spaces over the field K can be added to give rise to the function $T + S : V \to W$ given by

$$(T + S)(x) = T(x) + S(x)$$

for all $x \in V$. Further, given a scalar $\alpha \in K$, the function $\alpha T : V \to W$ is given by

$$(\alpha T)(x) = \alpha\,(T(x))\,.$$

Proposition 2.7 *For all linear operators $T, S : V \to W$ and scalars $\alpha \in K$, both $T + S$ and αT are linear operators.*

Proof For all vectors $x, y \in V$ and scalars $\beta, \gamma \in K$:

$$\begin{aligned}
(T + S)(\beta x + \gamma y) &= T(\beta x + \gamma y) + S(\beta x + \gamma y) \\
&= T(\beta x) + T(\gamma y) + S(\beta x) + S(\gamma y) \\
&= \beta T(x) + \gamma T(y) + \beta S(x) + \gamma S(y) \\
&= \beta(T(x) + S(x)) + \gamma(T(y) + S(y)) \\
&= \beta(T + S)(x) + \gamma(T + S)(y).
\end{aligned}$$

The proof for αT is similar and left for the reader. □

The result above shows that the set of all linear operators $T : V \to W$ has naturally defined notions of addition and scalar multiplication. It is a pleasant fact that with these operations one obtains a linear space.

Definition 2.5 Given linear spaces V and W over the same field K, we denote by $\mathrm{Hom}(V, W)$ the set of all linear operators $T : V \to W$.

Theorem 2.3 *For linear spaces V and W over the same field K, the set* $\mathrm{Hom}(V, W)$*, when endowed with the operations of addition and scalar multiplication as above, is a linear space over K.*

Proof The proof is a straightforward verification of the linear space axioms, so we only give the details for a few of the axioms. For instance, given linear operators $T_1, T_2 : V \to W$, to show that

$$T_1 + T_2 = T_2 + T_1$$

we note that

$$(T_1 + T_2)(x) = T_1(x) + T_2(x) = T_2(x) + T_1(x) = (T_2 + T_1)(x)$$

where the commutativity of vector addition in W was used.

To show the existence of an additively neutral element, recall that $T_0 : V \to W$ given by $T_0(x) = 0$, is always a linear operator, and thus $T_0 \in \mathrm{Hom}(V, W)$. Seeing that T_0 is additively neutral is simply the computation

$$(T + T_0)(x) = T(x) + T_0(x) = T(x) + 0 = T(x).$$

We leave the verification of the other axioms to the reader. □

2.3.3 Isomorphism

The linear spaces $\mathbb{R}^{n+1}$ and P_n, while consisting of very different elements, are, in a strong sense, essentially identical. It is obvious that one may think of a sequence of $n + 1$ real numbers as the coefficients of a polynomial, and, vice versa, one may identify a polynomial function with its list of coefficients and thus obtain $n + 1$ real numbers. Thus, one may rename the elements in one space to obtain the elements of the other, and, moreover, the linear structure under this renaming is respected. This situation is made precise, and generalized, by the concept of isomorphism.

Definition 2.6 A linear operator $T : V \to W$ which, as a function, is a bijection is called a *linear isomorphism* or (simply an *isomorphism* if the linear context is clear). When $T : V \to W$ is an isomorphism, the spaces V and W are said to be *isomorphic*, denoted by $V \cong W$.

The following result establishes some expected behaviour of isomorphisms.

Proposition 2.8 *Suppose that U, V, and W are linear spaces over the same field K and consider the linear operators $T : U \to V$ and $S : V \to W$. The following then hold.*

1. *The identity function $id : V \to V$ is an isomorphism.*
2. *If T and S are isomorphisms, then so is $S \circ T$.*
3. *If T is an isomorphism, then so is the inverse function T^{-1}.*

Proof

1. We already noted that the identity function is always a linear operator. It is clearly bijective and thus an isomorphism.
2. We already noted that the composition of linear operators is a linear operator, and thus $S \circ T$ is a linear operator. In general, the composition of bijective functions is again bijective, and so if T and S are isomorphisms, then so is $S \circ T$.
3. Since $T : U \to V$ is an isomorphism, the inverse function $T^{-1} : V \to U$ exists, and is clearly a bijection. To conclude the proof it remains to be shown that T^{-1} is a linear operator. Indeed, T^{-1} is additive since

$$\begin{aligned} T^{-1}(x+y) &= T^{-1}(T(T^{-1}(x)) + T(T^{-1}(y))) \\ &= T^{-1}(T(T^{-1}(x) + T^{-1}(y))) \\ &= T^{-1}(x) + T^{-1}(y) \end{aligned}$$

for all vectors $x, y \in V$. Further, T^{-1} is homogenous since

$$T^{-1}(\alpha x) = T^{-1}(\alpha T(T^{-1}(x))) = T^{-1}(T(\alpha T^{-1}(x))) = \alpha T^{-1}(x)$$

for all vectors $x \in V$ and scalars $\alpha \in K$. $\square$

Corollary 2.1 *For all linear spaces U, V, and W over the same field K:*

1. $V \cong V$.
2. If $U \cong V$, then $V \cong U$.
3. If $U \cong V$ and $V \cong W$, then $U \cong W$.

Isomorphic linear spaces are essentially identical, except (possibly) for the names of the elements in them and they thus possess exactly the same linear properties. This is a somewhat vague statement but it is almost always immediate how to turn it into a precise statement. For instance, the dimension of a linear space is a linear property and thus any two isomorphic linear spaces have the same dimension, as we now show.

Proposition 2.9 *If $V \cong W$, then both linear spaces have the same dimension.*

Proof By assumption there exists a linear isomorphism $T : V \to W$. The dimension of a linear space is the cardinality of any of its bases, and thus the result will be established by exhibiting a bijection between a basis for V and a basis of W. Let $B \subseteq V$ be a basis and consider its image $T(B) \subseteq W$. Clearly, $T|_B$, the restriction of T to B, is a bijection between B and $f(B)$, so all that remains to be done is to show that $f(B)$ is a basis for W, namely that $f(B)$ is a spanning set of linearly independent vectors. Let $y \in W$ be an arbitrary vector. Since T is onto, we may write $y = T(x)$ for some $x \in V$. As B is a basis, we express x as a (finite!) linear combination of basis elements:

$$x = \sum_{b \in B} \alpha_b \cdot b.$$

But then

$$y = T(x) = \sum_{b \in B} \alpha_b \cdot T(b)$$

and thus y is in the span of $T(B)$. As y was arbitrary we conclude that $T(B)$ spans all of W. The proof that $T(B)$ is also linearly independent follows a similar pattern and is left for the reader. □

The converse to this result is also true, as the following result implies.

Theorem 2.4 *If V is a linear space and B is a basis for it, then $V \cong (K^B)_0$.*

Proof In Proposition 2.3 we established that the function $T : (K^B)_0 \to V$ given by

$$T(f) = \sum_{b \in B} f(b) \cdot b$$

is a bijection. We show now that it is in fact an isomorphism, thus completing the proof. We need to verify that $T(\alpha_1 f + \alpha_2 g) = \alpha_1 T(f) + \alpha_2 T(g)$, which amounts to showing that

$$\sum_{b \in B} (\alpha_1 f + \alpha_2 g)(b) \cdot b = \alpha_1 \sum_{b \in B} f(b) \cdot b + \alpha_2 \sum_{b \in B} g(b) \cdot b,$$

which follows by an immediate computation. □

Remark 2.9 Every linear space V over K is thus isomorphic to a linear space of the form $(K^B)_0$. The latter space is clearly only dependent on the cardinality of the set B, i.e., if X is any set with the same cardinality as B, then $(K^B)_0 \cong (K^X)_0$. In other words, the dimension of a linear space determines it up to an isomorphism. It should be noted however that generally there is no canonical choice for an isomorphism between V and $(K^X)_0$.

In the finite dimensional case we obtain the following corollary.

Corollary 2.2 *Let V be a finite dimensional linear space over the field K. Any choice of a basis $x_1, \ldots, x_n$ for V gives rise to an isomorphism $V \to K^n$. In this way we may identify every vector in V, by means of coordinates, with an n-tuple of scalars.*

In a more general context, and in particular in the infinite dimensional case, where bases are typically not explicitly available, we have the following results.

Corollary 2.3 *If V and W are linear spaces over the same field K and they have the same dimension, then they are isomorphic.*

Proof By Remark 2.9, if we let X be a set of cardinality equal to the common dimensions of V and W, then we obtain an isomorphisms $T_1 : (K^X)_0 \to V$ and an isomorphism $T_2 : (K^X)_0 \to W$. It follows that $T_2 \circ T_1^{-1} : V \to W$ is an isomorphism, showing that $V \cong W$, as claimed. □

Combining Proposition 2.9 and Corollary 2.3, the discussion above is summarized as follows.

Theorem 2.5 *Two linear spaces over the same field are isomorphic if, and only if, they have the same dimension.*

Remark 2.10 This result is an example of what is known as a rigidity phenomenon. Rigidity refers to a situation where two structures are essentially the same given that they have essentially the same substructure of some kind, typically of a much coarser nature than the original structures. In this case, the rigidity of linear spaces over a fixed field K is that the dimension, i.e., the cardinality of a basis, suffices to determine the linear space, up to an isomorphism.

Example 2.20 Recall the linear space P_n of polynomial functions from Example 2.6. We can easily show that $P_n \cong \mathbb{R}^{n+1}$ by constructing an isomorphism between the two spaces. Indeed, referring here to a typical element in $\mathbb{R}^{n+1}$ by $(a_0, \ldots, a_n)$, let

$$T : \mathbb{R}^{n+1} \to P_n$$

be given by

$$T(a_0, \ldots, a_n) = a_n t^n + \cdots + a_1 t + a_0.$$

It is a trivial matter to verify that T is a linear operator, clearly bijective, and thus an isomorphism. Similarly, the reader is invited to show that $c_{00} \cong P$. By the discussion above, the existence of an isomorphism (but not any explicit isomorphism) could have been deduced by noting that both $\mathbb{R}^{n+1}$ and P_n, as linear spaces over $\mathbb{R}$, have dimension $n + 1$. Similarly, both c_{00} and P have countably infinite dimension and are thus isomorphic.

Remark 2.11 Theorem 2.5 tells us that the dimension of a linear space essentially is all that one needs to know about a linear space in order to study it (since after all, isomorphic linear spaces are essentially the same). It is thus tempting to, once and for all, choose a single representative from each isomorphism class of linear spaces. This certainly suffices for the study of all linear spaces, and seems more economical than having all of this redundancy in the form of multiple isomorphic specimens of linear spaces. However, the example above clearly shows the danger in this approach. While $\mathbb{R}^{n+1}$ and P_n are isomorphic, they have different qualities (at least to us humans). For instance, it is very natural to consider the integral as an operator on polynomials, but not so much on elements in $\mathbb{R}^{n+1}$. The richness of having many different linear spaces, even if they are isomorphic, is a blessing that should not be given up for the sake of a more economical treatment.

Exercises

Exercise 2.21 Consider the linear space c of all convergent sequences of complex numbers (or real numbers, with the obvious adjustments). Prove that the assignment

$$\{x_m\}_{m\geq 1} \mapsto \lim_{m\to\infty} x_m$$

is a linear operator from c to $\mathbb{C}$.

Exercise 2.22 Finish the proof of Theorem 2.3 showing that $\mathrm{Hom}(V, W)$ is a linear space.

Exercise 2.23 Let V and W be finite-dimensional linear spaces over a field K, of dimensions n and m respectively, and let $B_1 = \{v_1, \ldots, v_n\}$ and $B_2 = \{w_1, \ldots, w_m\}$ be fixed bases for V and W respectively. For any vector $x \in V$, we write

$$[x]_{B_1} = (a_1, \ldots, a_n) \in K^n$$

where $a_1, \ldots, a_n$ are the unique scalars such that

$$x = \sum_{k=1}^{n} a_k v_k.$$

The tuple $[x]_{B_1}$ is called the vector of coordinates of x in the basis B_1. Similarly, one defines $[y]_{B_2}$ for all $y \in W$.

1. Given a linear operator $T : V \to W$, prove that there exists a unique matrix $[T]_{B_1,B_2} \in M_{n,m}(K)$ such that

$$[T(x)]_{B_2} = [T]_{B_1,B_2} \cdot [x]_{B_1}$$

where on the right-hand-side the '$\cdot$' stands for the ordinary product of a matrix by a vector. The matrix $[T]_{B_1,B_2}$ is called the *representative* matrix of the linear operator T in the given bases.
2. Prove that the assignment $T \mapsto [T]_{B_1,B_2}$, mapping a linear operator to the matrix representing it in the two given bases, is a linear isomorphism

$$\psi_{B_1,B_2} : \mathrm{Hom}(V, W) \to M_{n.m}(K).$$

3. Suppose now that U is a third linear space with its chosen basis B_3 and that $S : W \to U$ is a linear operator. Prove that $[S \circ T]_{B_1,B_3} = [S]_{B_2,B_3} \cdot [T]_{B_1,B_2}$, where the product on the right-hand-side is the ordinary product of matrices.

We remark that this exercise gives one justification for the definition of the algebraic operations on matrices the way they are defined.

Exercise 2.24 Prove that $\mathbb{R}^n$ and $\mathbb{C}^n$, $n \geq 1$, when considered as linear spaces over $\mathbb{R}$, are not isomorphic.

Exercise 2.25 Prove that $\mathbb{R}^\infty \cong \mathbb{C}^\infty$, when considered as linear spaces over $\mathbb{R}$.

Exercise 2.26 A linear operator $T : V \to V$ is *nilpotent* if there exists an $m \geq 1$ such that T^m, the m-fold composition of T with itself, is the zero operator $0 : V \to V$. Prove that if $T : V \to V$ is nilpotent, then $\mathrm{id}_V - T$ is an isomorphism.

Exercise 2.27 Let $T : V \to V$ be an operator with the property that for all $x \in V$ there exists an $m \geq 1$ such that $T^m(x) = 0$. Prove that if V is finite-dimensional, then T is nilpotent, but if V is infinite-dimensional, then T need not be nilpotent.

Exercise 2.28 Prove that for every set $\mathscr{B}$ there exists a linear space with $\mathscr{B}$ as a basis (and consequently for any cardinality κ there exists a linear space whose dimension is κ).

Exercise 2.29 Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function with $f(x + y) = f(x) + f(y)$ for all $x, y \in \mathbb{R}$. Prove that there exists some $a \in \mathbb{R}$ such that $f(x) = ax$ for all $x \in \mathbb{R}$.

Exercise 2.30 Use a Hamel basis for $\mathbb{R}$ as a linear space over $\mathbb{Q}$ to construct a function $f : \mathbb{R} \to \mathbb{R}$ satisfying $f(x + y) = f(x) + f(y)$ for all $x, y \in \mathbb{R}$, which is not of the form $f(x) = ax$ for any $a \in \mathbb{R}$.

2.4 Subspaces, Products, and Quotients

Subspaces arise naturally as portions of a given linear space that inherit a linear space structure from the ambient linear space. We show that any linear operator gives rise to certain subspaces, its kernel and its image, and we show how to combine two linear spaces to form their product space. We also show how to eliminate a subspace so as to obtain a quotient space. The quotient construction is related to the concept of complementary subspaces, which is also introduced and the connection made explicit.

2.4.1 Subspaces

Given a subset $A \subseteq V$ of a linear space V the operations of addition and scalar multiplication, when restricted to vectors in A always yield elements in the ambient space V. We are interested in the case where the results of these operations always yield elements in A.

Definition 2.7 A subset $A \subseteq V$ is said to be *closed under addition* if

$$x + y \in A$$

for all vectors $x, y \in A$. Similarly, A is said to be *closed under scalar multiplication* if

$$\alpha x \in A$$

for all vectors $x \in A$ and scalars $\alpha \in K$. The set A is called a *linear subspace* (or simply *subspace* if the linear context is evident) of V if $A \neq \emptyset$ and A is closed under addition and scalar multiplication. A subspace A is called a *proper subspace* if it is properly contained in the ambient space V.

Remark 2.12 If A is a linear subspace of V, then A, with vector addition and scalar multiplication induced from V, is a linear space. Indeed, the verification of each of the axioms follows the same pattern. For instance,

$$x + y = y + x$$

holds for all $x, y \in A$ since the same equality holds for all vectors x and y in the ambient space V. Moreover, notice that the zero vector 0 is always an element in A and is the zero vector of A. Indeed, $0 = 0 \cdot x$ holds for all $x \in A$, and since A is required to be non-empty, at least one $x \in A$ exists.

Notice that the subset $\{0\}$ consisting of just the zero vector is always a subspace, called the *trivial subspace* of V. Another immediate example of a subspace of V is V itself, i.e., a non-proper subspace. An immediate property of subspaces is their transitive property, that is if $A \subseteq B \subseteq V$, then if B is a subspace of V and A is a subspace of B, then A is a subspace of V.

Example 2.21 Consider the linear space $\mathbb{R}^2$ and its subset A consisting only of the vectors of the form $(x, 0) \in \mathbb{R}^2$. Clearly, A is non-empty, closed under addition, and closed under scalar multiplication, and hence is a subspace of $\mathbb{R}^2$. More generally, a line $l = \{tx \mid t \in \mathbb{R}\}$, with $x \in \mathbb{R}^2$ a non-zero vector, is a linear subspace of $\mathbb{R}^2$. These linear subspaces, together with the subspaces $\{0\}$ and $\mathbb{R}^2$, exhaust all of the linear subspaces of $\mathbb{R}^2$. Similarly, linear subspaces of $\mathbb{R}^n$, for larger values of n, correspond to the origin, to lines through the origin, to planes through the origin, and to hyper-planes through the origin. Notice the following subtle point. The field $\mathbb{R}$, when viewed as a linear space over itself, is obviously isomorphic to $\mathbb{R}^1$, that is to $\mathbb{R}^n$ where $n = 1$. However, even though the linear space $\mathbb{R} \cong \mathbb{R}^1$ may naturally be identified with either the **X**-axis or the **Y**-axis in $\mathbb{R}^2$, formally speaking, $\mathbb{R}$ itself (or $\mathbb{R}^1$) is not a subspace of $\mathbb{R}^2$. In fact, it is not even a subset of $\mathbb{R}^2$. More generally, $\mathbb{R}^n$ is a subspace of $\mathbb{R}^m$ if, and only if, $n = m$, in which case the two spaces coincide. When $n \leq m$ the space $\mathbb{R}^n$ may be identified in finitely many ways with various subspaces of $\mathbb{R}^m$, but that is a different story. An analogous discussion can be given for linear spaces over $\mathbb{C}$.

Example 2.22 Consider the linear space $\mathbb{C}^\infty$ from Example 2.3 and the linear spaces c of convergent sequences, c_0 of sequences that converge to 0, and c_{00} of sequences that are eventually 0 (see Example 2.4). One clearly has that

$$c_{00} \subset c_0 \subset c \subset \mathbb{C}^\infty$$

and in fact c_{00} is a linear subspace of c_0, which in turn is a linear subspace of c, which is a linear subspace of $\mathbb{C}^\infty$. The verification is immediate.

Other subspaces of $\mathbb{C}^\infty$ include, for each $n \geq 1$, the set of all vectors of the form $(x_1, \ldots, x_n, 0, 0, \ldots)$. This space is clearly isomorphic to $\mathbb{C}^n$, as are infinitely many other subspaces of $\mathbb{C}^\infty$, as the reader is asked to verify. In any event, $\mathbb{C}^n$ is not itself a subspace of $\mathbb{C}^\infty$.

Example 2.23 Consider the linear space P_n (Example 2.6) consisting of polynomial functions with real coefficients of degree not exceeding n, and P, the linear space of all polynomials with real coefficients. For all $n \leq m$, it holds that $P_n \subset P_m \subset P$, and it follows easily that P_n is a subspace of P_m, which in turn is a subspace of P. We thus obtain the tower of proper subspace inclusions

$$P_0 \subset P_1 \subset P_2 \subset \cdots \subset P_n \subset \cdots \subset P.$$

Example 2.24 With reference to Example 2.7, the subset of $C([a, b], \mathbb{R})$ consisting of the continuous functions $x : [a, b] \to \mathbb{R}$ which vanish at a and b, that is such that $x(a) = x(b) = 0$, constitutes a linear subspace. On the contrary, the subset of those continuous functions with $x(a) = x(b) = 1$ does not constitute a linear subspace (for instance since it fails to contain the zero vector).

The following family of linear spaces, which are subspaces of $\mathbb{C}^\infty$, is of significant importance.

Example 2.25 (The ℓ_p Spaces). Consider the linear space $\mathbb{C}^\infty$ from Example 2.3 and recall that a sequence $x = (x_1, \ldots, x_k, \ldots) \in \mathbb{C}^\infty$ is said to be *bounded* if there exists some $M \in \mathbb{R}$ such that

$$|x_k| < M$$

for all $k \geq 1$. Let $\ell_\infty \subseteq \mathbb{C}^\infty$ be the set of all bounded sequences. Clearly, the zero vector, namely the constantly zero sequence, is in ℓ_∞, the sum of two bounded sequences is again bounded, and thus in ℓ_∞, and if x is bounded by M, then $\alpha x = (\alpha x_1, \ldots, \alpha x_k, \ldots)$ is bounded by $|\alpha|M$, and thus is in ℓ_∞. It thus follows that ℓ_∞ is a linear subspace of $\mathbb{C}^\infty$.

Now, let $p \geq 1$ and consider the set ℓ_p of all sequences of complex numbers that are *absolutely p-power summable*, that is, all sequences $x \in \mathbb{C}^\infty$ such that

$$\sum_{k=1}^{\infty} |x_k|^p < \infty.$$

Clearly, the constantly zero sequence belongs to ℓ_p, and ℓ_p is evidently closed under scalar multiplication. To show that ℓ_p is a subspace of $\mathbb{C}^\infty$ it thus remains to be shown that ℓ_p is closed under addition, a task we leave to the reader.

We thus obtain a one-parameter family $\{\ell_p\}_{1\leq p\leq\infty}$ of linear subspaces of $\mathbb{C}^\infty$. In fact, it is easily seen that if $1 \leq p < q \leq \infty$, then $\ell_p \subset \ell_q$ (consider a suitable variation of the harmonic series), and so ℓ_p is a proper linear subspace of ℓ_q.

Example 2.26 Recall from Example 2.8 that $\mathbb{R}$ may be viewed as a linear space over $\mathbb{R}$, and that $\mathbb{C}$ can be viewed as a linear space over $\mathbb{C}$ or as a linear space over $\mathbb{R}$. The set inclusion $\mathbb{R} \subseteq \mathbb{C}$ is of course always valid, but $\mathbb{R}$ is a linear subspace of $\mathbb{C}$ only when both are considered to be linear spaces over $\mathbb{R}$ (or over $\mathbb{Q}$). A subspace relation between two linear spaces can only hold if both linear spaces are over the same field, and so, for instance, $\mathbb{C}^n$ as a linear space over $\mathbb{R}$ is not a linear subspace of $\mathbb{C}^n$ as a linear space over $\mathbb{C}$.

We now address the relationship between subspaces and dimension which, in the finite case, is as one would expect. But the infinite case, as usual, is more subtle.

Theorem 2.6 *For a linear space V of dimension n and a subspace $U \subseteq V$ with dimension m, the following hold.*

1. $m \leq n$.
2. If V is finite dimensional and $m = n$, then $U = V$.

Proof

1. The proof is given in full generality, thus n and m may be infinite cardinals. Let B_V be a basis for V and let B_U be a basis for U. By assumption, the cardinality of B_V is n while the cardinality of B_U is m. Now, by definition of basis, B_U is a linearly independent set of vectors in U, and hence is also linearly independent in V, while B_V is a spanning set in V. Applying Lemma 2.1, it follows that there exists an injection $B_U \to B_V$, and thus $n \leq m$.
2. Assume now that $m = n$ and that V is finite dimensional, namely that $n = m$ is a natural number. We may thus find a basis $\{x_1, \ldots, x_n\}$ of U. In particular these vectors are linearly independent in U, and thus also in V. By Proposition 2.4, it follows that the set $\{x_1, \ldots, x_n\}$ is a basis of V, so its span is V. But the span is also U, and the claim follows.

□

Remark 2.13 Much as in Remark 2.6, the finite dimensionality assumption cannot be avoided. Indeed, the linear space c_{00} from Example 2.4 is, as we saw, of countably infinite dimension but it does contain proper subspaces of the same dimension. For instance, it is easy to verify that the set c_{000} of all sequences in c_{00} whose first term is 0, is a proper subspace of c_{00} even though it is isomorphic to c_{00}, and thus has the same dimension as the ambient space.

2.4.2 Kernels and Images

With any linear operator $T : V \to W$ one can associate a subspace of the domain, called the kernel of T, and a subspace of the codomain, called the image of T. As we show below, the kernel suffices to detect whether or not T is injective.

Definition 2.8 Let $T : V \to W$ be a linear operator. The set

$$\mathrm{Ker}(T) = \{x \in V \mid T(x) = 0\}$$

is called the *kernel* of T.

Theorem 2.7 *The kernel of a linear operator $T : V \to W$ is a linear subspace of V.*

Proof In Proposition 2.5 we already noted that $T(0) = 0$ and thus $0 \in \mathrm{Ker}(T)$. All we need to do now is show that the kernel is closed under vector addition and scalar product. Indeed, if $x, y \in \mathrm{Ker}(T)$, then

$$T(x + y) = T(x) + T(y) = 0 + 0 = 0$$

and thus $x + y \in \mathrm{Ker}(T)$. Similarly, for all $\alpha \in K$, if $x \in \mathrm{Ker}(T)$, then

$$T(\alpha x) = \alpha T(x) = \alpha \cdot 0 = 0$$

and thus $\alpha x \in \mathrm{Ker}(T)$ and the proof is complete. □

Theorem 2.8 *A linear operator $T : V \to W$ is injective if, and only if, its kernel is trivial, i.e.,* $\mathrm{Ker}(T) = \{0\}$.

Proof If T is injective, then since we already know that $T(0) = 0$, it follows that $T(x) = 0$ implies $x = 0$, and thus $\mathrm{Ker}(T) = \{0\}$, and is thus trivial. Conversely, suppose that the kernel of T is trivial and suppose that

$$T(x) = T(y)$$

for some $x, y \in V$. Then

$$T(x - y) = T(x) - T(y) = 0$$

and thus

$$x - y \in \mathrm{Ker}(T).$$

But then $x - y = 0$, and so $x = y$, showing that T is injective. □

Another subspace naturally obtained from a linear operator is its image.

Definition 2.9 Let $T : V \to W$ be a linear operator. The set $\mathrm{Im}(T) = \{T(x) \mid x \in V\}$, namely the set-theoretic image of T, is called the *image* of T.

Proposition 2.10 *The image of a linear operator $T : V \to W$ is a linear subspace of W.*

Proof Obviously, $T(V)$ is not empty. Further, if $x, y \in \mathrm{Im}(T)$, then $x = T(x')$ and $y = T(y')$ for some vectors $x', y' \in V$. It then follows that

$$x + y = T(x') + T(y') = T(x' + y') \in \mathrm{Im}(T).$$

Similarly, one shows that $\alpha x \in \mathrm{Im}(T)$ for all $x \in \mathrm{Im}(T)$ and all scalars $\alpha \in K$. □

The following result, which the reader may recognize as the rank-nullity theorem, has important consequences for finite dimensional linear spaces.

Lemma 2.3 *The equality*

$$\dim(V) = \dim(\mathrm{Ker(T)}) + \dim(\mathrm{Im(T)})$$

holds for all linear operators $T : V \to W$ between finite dimensional linear spaces.

Proof We present the general strategy, inviting the reader to fill-in the details. Choose a basis $\{x_1, \ldots, x_k\}$ for $\mathrm{Ker}(T)$. Then augment this basis by (if needed) adding vectors to it, so as to obtain a basis $\{x_1, \ldots, x_k, y_1, \ldots, y_m\}$ of V and show that $\{T(y_1), \ldots, T(y_m)\}$ is a basis for $\mathrm{Im}(T)$. □

Remark 2.14 The proof above can be adapted to obtain a similar result for infinite dimensional linear spaces, provided one handles cardinal arithmetic with care, of course. However, such a result is not of great importance and we thus avoid its details.

Corollary 2.4 *For a linear operator $T : V \to W$ between finite dimensional linear spaces of equal dimension, the conditions*

- *T is injective*
- *T is surjective*
- *T is bijective*

are equivalent.

Proof We have seen that T is injective if, and only if, its kernel is trivial, namely precisely when $\dim(\mathrm{Ker}(T)) = 0$. It follows then that T is injective if, and only if, $\dim(V) = \dim(\mathrm{Im}(T))$. But since $\dim(V) = \dim(W)$, it follows by Theorem 2.6, that $\dim(V) = \dim(\mathrm{Im}(T))$ holds if, and only if, $\mathrm{Im}(T) = W$, in other words, precisely when T is surjective. This completes the verification of the non-trivial arguments in the proof. □

2.4.3 Products and Quotients

We now present the product construction for two linear spaces V and W over the same field K. We endow the set $V \times W$ with an addition operation and a scalar product that turn it into a linear space, as follows. Given $(x_1, x_2), (y_1, y_2) \in V \times W$, and $\alpha \in K$, the addition operation is given by

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$$

and the scalar product operation is given by

$$\alpha(x_1, x_2) = (\alpha x_1, \alpha x_2).$$

Theorem 2.9 *For linear spaces V and W over the same field K, the set $V \times W$, when endowed with addition and scalar multiplication as above, is a linear space over the field K.*

Proof The verification of the linear space axioms is straightforward. For instance, the additively neutral element is (0, 0) since

$$(x_1, x_2) + (0, 0) = (x_1 + 0, x_2 + 0) = (x_1, x_2)$$

for all $(x_1, x_2) \in V \times W$. We leave the rest of the verification to the reader. □

Obviously, one can similarly define the product $V_1 \times \cdots \times V_n$ of any finite number of linear spaces, and even the product of any collection of linear spaces. It is also evident, upon inspection of the linear structure of the linear space $\mathbb{R}^n$, that $\mathbb{R}^n$ is the n-fold product of $\mathbb{R}$ with itself, where $\mathbb{R}$ is viewed as a linear space over itself. Similarly, $\mathbb{C}^n$ is the n-fold product of $\mathbb{C}$ with itself.

We now turn to consider the quotient construction of a linear space by a linear subspace. This construction is, in some sense, the reversal of taking the product of linear spaces. We refer the reader to Sect. 1.2.15 of the Preliminaries for basic facts about equivalence relations.

Definition 2.10 Let V be a linear space and U a subspace of it. Two elements $x_1, x_2 \in V$ are said to be *equivalent modulo U* if

$$x_1 - x_2 \in U$$

that is, if

$$x_1 = x_2 + u$$

for some $u \in U$. This relation is denoted by $x_1 = x_2 (\mathrm{mod}\ U)$, by $x \equiv_U y$, or simply by $x \equiv y$ if U is evident.

We now show that $\equiv$ is an equivalence relation on V. Indeed, for all $x, y, z \in V$

$$x - x = 0 \in U,$$

and thus

$$x \equiv x,$$

so that $\equiv$ is reflexive. Further,

$$x \equiv y \implies x - y \in U \implies y - x = (-1) \cdot (x - y) \in U \implies y \equiv x,$$

and thus $\equiv$ is symmetric. Finally,

$$x - y, y - z \in U \implies x - z = (x - y) + (y - z) \in U \implies x \equiv z,$$

and so $\equiv$ is transitive.

Recall that the equivalence class of any vector $x \in V$ is the set

$$[x] = \{y \in V \mid x \equiv y\}.$$

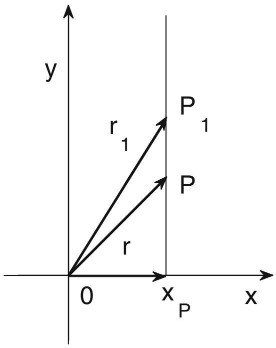
In other words, it is the set of all vectors of the form $x + u$, where $u \in U$. For that reason the equivalence class $[x]$ is also denoted by

$$x + U.$$

It follows from the general theory of equivalence relations that $\{x + U \mid x \in V\}$ is a partition of V.

Example 2.27 As a simple example, consider the usual 2-dimensional plane $\mathbb{R}^2$ and its subspace $\mathbf{Y}$ consisting of the ordinates. Let the element $\mathbf{r} \in \mathbb{R}^2$ be the vector $O - P$ from the origin to the point P in the plane. The equivalence class $[\mathbf{r}]_{\mathbf{Y}}$, formed by the vectors of $\mathbb{R}^2$ equivalent mod $\mathbf{Y}$ to $\mathbf{r}$, is clearly given by all vectors $O - P_i$ with P_i lying on the parallel line to the $\mathbf{Y}$-axis passing through the point P; all these vectors are equivalent *mod* $\mathbf{Y}$ to each other.

The fact that the entire 2-dimensional space is partitioned by all such equivalence classes is simply the fact that $\mathbb{R}^2$ is the disjoint union of all the lines parallel to the $\mathbf{Y}$ axis, which are precisely the translates $x + \mathbf{Y}$ of $\mathbf{Y}$.



We now turn to investigate the quotient set $V/U = \{x+U \mid x \in V\}$ of equivalence classes modulo U. It is natural to introduce an addition and a scalar product on V/U by the formulas

$$(y+U)+(z+U)=(y+z)+U \qquad \text{and} \qquad \alpha(y+U)=(\alpha y)+U,$$

but we must first verify that these operations are well-defined, namely that they do not depend on the chosen representatives. Indeed, for all $x, y, x', y' \in V$, we need to show that if

$$x+U=x'+U \quad \text{and} \quad y+U=y'+U,$$

then

$$(x+y)+U=(x'+y')+U.$$

But the former implies that

$$x'-x \in U \quad \text{and} \quad y'-y \in U$$

while the latter will be implied by showing that $(x'+y')-(x+y) \in U$. And indeed, since U is closed under addition,

$$(x'+y')-(x+y)=(x'-x)+(y'-y) \in U.$$

A similar argument shows that the scalar product above is well-defined and thus we see that the quotient set V/U of equivalence classes is naturally endowed with an addition operation and a scalar product operation.

Theorem 2.10 *Let V be a linear space and $U \subseteq V$ a subspace. The quotient set V/U of equivalence classes modulo U, with the operations of addition and scalar multiplication as given above, is a linear space whose zero vector is U. Moreover, the canonical projection $\pi : V \to V/U$ given by $\pi(x) = x+U$ is a surjective linear operator.*

Proof To see that $U = 0 + U$ behaves neutrally with respect to addition, note that for all $x + U \in V/U$

$$(x + U) + U = (x + U) + (0 + U) = (x + 0) + U = x + U.$$

The rest of the verification of the linear space axioms, as well as the claims about the canonical projection, follow similarly and are left for the reader. □

Definition 2.11 The linear space V/U constructed in Theorem 2.10 is called the *quotient space of* V *with respect to* U, or simply as V *modulo* U.

Continuing the example above, $\mathbb{R}^2/\mathbf{Y}$, which geometrically can be thought of as the space of all vertical lines in $\mathbb{R}^2$, is a linear space. Another, more complicated, example is c_0/c_{00}, the quotient of the space c_0 of all converging sequences of (say) real numbers, by the subspace of all eventually 0 sequences. The quotient may be thought of as the space of all sequences modulo finite changes. In analysis it is well known that the limit of a sequence is insensitive to finite changes in the sequence, and so it is precisely this quotient that is of interest already in elementary analysis.

2.4.4 Complementary Subspaces

Contemplating the quotient space associated to Example 2.27, one realizes that $\mathbb{R}^2/\mathbf{Y}$ is isomorphic to $\mathbf{X}$, and that under this isomorphism, the canonical projection $\pi : \mathbb{R}^2 \to \mathbb{R}^2/\mathbf{Y}$ is nothing but the projection of $\mathbb{R}^2$ onto the $\mathbf{X}$ axis. This observation generalizes to any quotient space construction, as we now show.

Definition 2.12 Two subspaces U and W of a given linear space V are said to be *complementary*

- if the only vector they have in common is the zero vector, in other words if $U \cap W = \{0\}$; and
- if $U + W = V$; that is, given $x \in V$, there exist $u \in U$ and $w \in W$ with $x = u + w$.

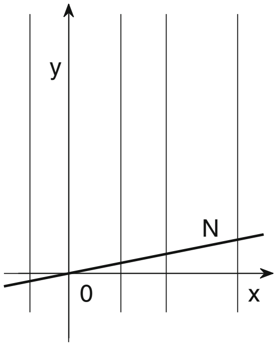
Remark 2.15 The decomposition above of an element $x \in V$ as the sum $x = u + w$ is in fact unique. Indeed, if $x = u + w$ and also $x = u' + w'$, with $u, u' \in U$ and $w, w' \in W$, then

$$0 = x - x = (u + w) - (u' + w') \implies u - u' = w' - w.$$

But $u - u' \in U$ and $w' - w \in W$, and since the only element in $U \cap W$ is the zero vector, it follows that $u - u' = 0 = w' - w$, and therefore, $u = u'$ and $w = w'$. This unique decomposition is said to express V as a *direct sum* of its subspaces U and W, a fact denoted by $V = U \oplus W$.

The following example shows that a given subspace $U \subseteq V$ may admit more than one complementary subspace.

Example 2.28 (Continuing Example 2.27) In $\mathbb{R}^2$ any straight line **N** passing through the origin and not coincident with **Y** is a complementary subspace, as can easily be verified by elementary means.



We may now present the main theorem relating complementary subspaces and the quotient construction.

Theorem 2.11 *Given a linear space V and two complementary subspaces U and W, the quotient space V/U is naturally isomorphic to W (the meaning of 'naturally' is discussed below).*

Proof Recall from Theorem 2.10 that the canonical projection $\pi : V \to V/U$, given by $\pi(x) = x + U$, is a linear operator. The result will be established by showing that the restriction of π to $W \subseteq V$ is an isomorphism between W and V/U. All we have to do is show that the restriction is a bijection. For injectivity we use Theorem 2.8 and show the kernel of the restriction is trivial. For that, suppose that $\pi(w) = 0$ for some $w \in W$, that is (recall that the zero vector in V/U is U)

$$w + U = U$$

and so $w \in U$. But since

$$U \cap W = \{0\},$$

it follows that $w = 0$, as needed. For surjectivity, let $y + U$ be an arbitrary element in V/U. Writing $y = u + w$, with $u \in U$ and $w \in W$, we get that

$$y + U = (u + w) + U = (u + U) + (w + U) = w + U = \pi(w). \qquad \square$$

Remark 2.16 The meaning of the isomorphism above being *natural* (or *canonical*) is that its construction does not depend on any choice of basis. This is an important observation since, as we saw, in infinite dimensional spaces it may be impossible to actually describe any basis.

Corollary 2.5 *Given a linear space V and a subspace $U \subseteq V$, any two subspaces W_1 and W_2 that are complementary to U are naturally isomorphic.*

Proof Each of W_1 and W_2 is naturally isomorphic to V/U. □

In the case of complementary spaces, the stated isomorphism is independent of a basis but it is dependent on the ability to express the ambient space as the direct sum of the given subspace and each of its complements. If the latter can be done explicitly, then an explicit formula for the isomorphism will emerge.

Exercises

Exercise 2.31 Let V be a linear space and $S \subseteq V$ a subset. Prove that S is a linear subspace of V if, and only if, S is closed under linear combinations, i.e., if

$$\sum_{k=1}^{m} \alpha_k s_k \in S$$

for all $\alpha_1, \dots, \alpha_m \in K$ and $s_1, \dots, s_m \in S$, and $m \geq 0$. Here we adopt the convention that the empty sum, i.e., the case $m = 0$, is equal to 0.

Exercise 2.32 Prove Lemma 2.3: If $T : V \to W$ is a linear operator between finite dimensional linear spaces, then

$$\dim(V) = \dim(\mathrm{Ker(T)}) + \dim(\mathrm{Im(T)}).$$

Exercise 2.33 Let V be a linear space and let $\{S_i\}_{i \in I}$ be a family of subspaces of V. Prove that the intersection

$$S = \bigcap_{i \in I} S_i$$

is a linear subspace of V. On the other hand show that if $S_1, S_2 \subseteq V$ are two linear subspaces of V, then $S_1 \cup S_2$ is a linear subspace of V if, and only if, either $S_1 \subseteq S_2$ or $S_2 \subseteq S_1$.

Exercise 2.34 (Refer to Example 2.25) Prove that ℓ_p, for $1 \leq p \leq \infty$, is a linear space and that if $1 \leq p < q \leq \infty$, then $\ell_p \subset \ell_q$.

Exercise 2.35 Let V and W be two linear spaces over the same field K. Prove that $V \times W \cong W \times V$. Can equality ever hold?

Exercise 2.36 Let V be a linear space and $U \subseteq V$ a linear subspace. Prove that the operation

$$\alpha(x + U) = \alpha x + U,$$

for all vectors $x \in V$ and scalars $\alpha \in K$, is well-defined. That is, show that it is independent of the choice of representative for the equivalence class $x + U$.

Exercise 2.37 Let V be a linear space. Prove that $V/V \cong 0$, where 0 denotes any linear space with a single element, necessarily its zero vector. Show also that $V/\{0\} \cong V$, where $\{0\}$ is the trivial subspace of V. Is it possible that any of these isomorphisms is an actual equality?

Exercise 2.38 Let V be a linear space and $U \subseteq V$ a linear subspace. Prove that the canonical projection $\pi : V \to V/U$ given by

$$\pi(x) = x + U$$

is a surjective linear operator.

Exercise 2.39 Let V be a linear space over K and $U \subseteq V$ a subspace. Consider the diagram

$$\begin{tikzcd} U \arrow[r, "i"'] \arrow[rr, bend left, "0"] & V \arrow[r, "\forall T"'] \arrow[d, "\pi"'] & V' \\ & V/U \arrow[ur, "\exists! T'"'] & \end{tikzcd}$$

where $i : U \to V$ is the inclusion function, $\pi : V \to V/U$ is the canonical projection to the quotient space, and V' is an arbitrary linear space over K. The rest of the diagram deciphers as follows: Prove that for all linear operators $T : V \to V'$ with the property that $T(x) = 0$ for all $x \in U$, there exists a unique linear operator $T' : V/U \to V'$ such that $T' \circ \pi = T$. (This exercise is establishing what is known as the universal property of the quotient construction)

Exercise 2.40 Let V and W be two linear spaces over the same field K, and consider the product $V \times W$. Show that

1. $U = \{(x, 0) \mid x \in V\}$ is a linear subspace of $V \times W$ which is isomorphic to V. We denote U more suggestively by $V \times \{0\}$.
2. Prove that $(V \times W)/(V \times \{0\}) \cong W$.

2.5 Inner Product Spaces and Normed Spaces

So far in our treatment of linear spaces two geometric aspects are missing, namely angles between vectors and the length of a vector. This situation is unavoidable since in general linear spaces it need not be possible to coherently define any of these notions. In the presence of extra structure, these notions become available, as we discuss below.

2.5.1 Inner Product Spaces

The reader is most likely familiar with the *standard inner product* on $\mathbb{R}^n$, i.e.,

$$\langle x, y \rangle = \sum_{k=1}^{n} x_k y_k.$$

The definition of a general inner product space is then a generalization of certain properties the standard inner product has. We will not pause here to motivate this definition any further, except for two comments. The first comment is that the results below will show that an inner product allows one (at least in the real case) to speak of angles between vectors, thus obtaining retrospective justification for the axioms. The second comment pertains to the non-arbitrary nature of the standard inner product; the reader is invited to discover it from elementary geometry, i.e., the law of cosines.

Definition 2.13 *(Inner Product Space)* Let V be a linear space over K, where K is either $\mathbb{R}$ or $\mathbb{C}$. Assume that a function $\langle -, - \rangle : V \times V \to K$ is given. The function $\langle -, - \rangle$ is said to be an *inner product* if for all $x, y, z \in V$ and $\alpha, \beta \in K$, the following conditions hold.

1. *Conjugate symmetry* or *Hermitian symmetry*, i.e., $\langle y, x \rangle = \overline{\langle x, y \rangle}$, where $\overline{w}$ denotes the complex conjugate of w.
2. Linearity in the second argument, i.e., $\langle x, \alpha y + \beta z \rangle = \alpha \langle x, y \rangle + \beta \langle x, z \rangle$.
3. The inner product is *positive definite*, i.e., $\langle x, x \rangle \geq 0$ for all $x \in V$, and $\langle x, x \rangle = 0$ implies that $x = 0$, the zero vector. In particular, we then introduce the *norm* of x to be $\|x\| = \sqrt{\langle x, x \rangle}$.

The linear space V together with the function $\langle -, - \rangle$ is then called an *inner product space*. To emphasize that $K = \mathbb{R}$ we use the term *real inner product space*, while the case $K = \mathbb{C}$ is given the term *complex inner product space*. The scalar $\langle x, y \rangle$ is the *inner product* of the given vectors (in that order!).

Remark 2.17 Notice that when $K = \mathbb{R}$, conjugate symmetry reduces to *symmetry*, that is

$$\langle y, x \rangle = \langle x, y \rangle,$$

and the inner product is also linear in its first argument, namely,

$$\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle.$$

In the complex case one generally has that

$$\langle \alpha x + \beta y, z \rangle = \overline{\alpha} \langle x, z \rangle + \overline{\beta} \langle y, z \rangle,$$

as follows easily from conjugate symmetry and linearity in the second argument. In both real and complex inner product spaces, the inner product is additive in each argument. We mention as well that in mathematical circles it is common to write (x, y) instead of $\langle y, x\rangle$, that is to define an inner product as being linear in the first argument rather than the second one. Obviously, the difference is only cosmetic.

Example 2.29 Consider the linear space $\mathbb{R}^n$ and the definition

$$x \cdot y = (x_1, \dots, x_n) \cdot (y_1, \dots, y_n) = \sum_{k=1}^{n} x_k y_k.$$

With this definition it is easily seen that $\mathbb{R}^n$ becomes a real inner product space. Similarly, considering $\mathbb{C}^n$, defining

$$x \cdot y = (x_1, \dots, x_n) \cdot (y_1, \dots, y_n) = \sum_{k=1}^{n} \overline{x_k} y_k$$

endows $\mathbb{C}^n$ with the structure of a complex inner product space. Notice that for $x, y \in \mathbb{R}^\infty$ the 'obvious definition'

$$x \cdot y = (x_1, \dots) \cdot (y_1, \dots) = \sum_{k=1}^{\infty} x_k y_k$$

fails to endow $\mathbb{R}^\infty$ with the structure of an inner product space, simply because the sum may fail to converge, and thus this is not even a function.

Example 2.30 Consider the space $C(I, \mathbb{R})$ of continuous real valued functions on the interval $I = [a, b]$. The familiar properties of the integral show at once that defining, for all $x, y \in C(I, \mathbb{R})$,

$$\langle x, y\rangle = \int_a^b dt\, x(t) y(t)$$

endows $C(I, \mathbb{R})$ with the structure of a real inner product space. With the proper conjugation in the integral above, the space $C(I, \mathbb{C})$ of continuous complex valued functions becomes a complex inner product space.

Before we present any further examples let us explore some elementary, and not so elementary, general results.

Proposition 2.11 *For all vectors x, y in an inner product space V (either real or complex)*

$$\|x + y\|^2 \leq \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2.$$

Proof Expanding $\|x + y\|^2 = \langle x + y, x + y\rangle$ we see that

$$\langle x + y, x + y\rangle = \langle x, x + y\rangle + \langle y, x + y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle$$

and thus the claim will follow by showing that

$$\langle x, y\rangle + \langle y, x\rangle \leq 2 \cdot |\langle x, y\rangle|,$$

which is immediate if V is a real inner product space. In the complex case suppose that $\langle x, y\rangle = u + iv$. Then

$$\langle x, y\rangle + \langle y, x\rangle = \langle x, y\rangle + \overline{\langle x, y\rangle} = 2 \cdot u$$

and since

$$|u| = \sqrt{u^2} \leq \sqrt{u^2 + v^2} = |u + iv| = |\langle x, y\rangle|,$$

the result follows. □

The geometric interpretation of the norm $\|x\|$ is that it is the length of the vector x. As for the geometric meaning of the inner product, notice that any non-zero vector $x \in V$ can be *normalized* by defining $\hat{x} = x/\|x\|$, a vector of *unit* length. For any vector x of unit length and an arbitrary vector y, the scalar $\langle x, y\rangle$ is interpreted as the *component* of the vector y in the direction x. Consequently, we make the following definition.

Definition 2.14 Let V be an inner product space. Two vectors $x, y \in V$ are said to be *orthogonal* or *perpendicular* if $\langle x, y\rangle = 0$.

Theorem 2.12 (Pythagoras' Theorem) *The equality*

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2$$

holds for all orthogonal vectors x and y in an inner product space.

Proof $\|x + y\|^2 = \langle x + y, x + y\rangle = \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2 = \|x\|^2 + \|y\|^2$.

□

2.5.2 The Cauchy-Schwarz Inequality

The following inequality is among the most important inequalities in mathematics. It has numerous uses, some of which we will see immediately and some later on.

Theorem 2.13 (Cauchy-Schwarz Inequality) *The inequality*

$$|\langle x, y\rangle| \leq \|x\|\|y\|$$

holds for all vectors x, y in an inner product space V.

Proof For simplicity let us assume V is a real inner product space (only a small adaptation is needed for the complex case). The inequality is trivial when $x = 0$ and so we proceed assuming that $x \neq 0$. Let $\hat{x} = x/\|x\|$ be the normalization of x, so in particular $\|\hat{x}\| = 1$. Dividing the desired inequality by $\|x\|$ we see that we need to establish the inequality

$$|\langle \hat{x}, y\rangle| \leq \|y\|$$

(and so the Cauchy-Schwarz Inequality acquires the interpretation that the absolute value of a component of y in a given unit direction sets a lower bound on the length of y). Rewriting y to identify its component in the $\hat{x}$ direction we obtain

$$y = y_{\hat{x}} + (y - y_{\hat{x}})$$

where $y_{\hat{x}} = \langle \hat{x}, y\rangle\hat{x}$. Since

$$\langle \hat{x}, y - y_{\hat{x}}\rangle = \langle \hat{x}, y\rangle - \langle \hat{x}, y_{\hat{x}}\rangle = \langle \hat{x}, y\rangle - \langle \hat{x}, y\rangle\langle \hat{x}, \hat{x}\rangle = 0$$

it follows that $y_{\hat{x}}$ and $y - y_{\hat{x}}$ are orthogonal. Pythagoras' Theorem (Theorem 2.12) now yields

$$\|y\|^2 = \|y_{\hat{x}}\|^2 + \|y - y_{\hat{x}}\|^2 \geq \|y_{\hat{x}}\|^2 = \langle \hat{x}, y\rangle^2.$$

Extracting square roots completes the proof. □

Corollary 2.6 *Since we now know that*

$$-1 \leq \frac{\langle x, y\rangle}{\|x\|\|y\|} \leq 1$$

for all non-zero vectors x and y in an arbitrary real inner product space, it follows that

$$\theta = \arccos \frac{\langle x, y\rangle}{\|x\|\|y\|}$$

is defined. Noting that

$$\langle x, y\rangle = \cos\theta\|x\|\|y\|$$

we define θ to be the angle between the vectors x and y.

We had just seen that a consequence of the Cauchy-Schwarz inequality is that, in a real inner product space, angles between vectors are a well-defined notion. There are many more consequences of the Cauchy-Schwarz inequality, among which we mention one. The Heisenberg uncertainty principle in Quantum Mechanics is the result of applying the Cauchy-Schwarz inequality in a certain Hilbert space, suitably constructed. Unfortunately, this result falls slightly short of the scope of this book.

2.5.3 Normed Spaces

In an inner product space every vector can be assigned a norm. A normed space is then an abstraction of certain key properties of this norm. This is useful since, as will be shown below, many linear spaces fail to admit an inner product but do admit the structure of a normed space. In such spaces one may still speak of the length of a vector, but not (necessarily) of angles between vectors.

Definition 2.15 *(Normed Space)* A linear space V with a function $x \mapsto \|x\|$, which associates with every vector $x \in V$ a real number $\|x\|$, called the *norm* of x, is said to be a *normed space* if for all vectors $x, y \in V$ and $\alpha \in K$ the following hold.

1. Positivity, i.e., $\|x\| > 0$ provided $x \neq 0$.
2. Homogeneity, i.e., $\|\alpha x\| = |\alpha|\|x\|$.
3. Triangle inequality, i.e., $\|x + y\| \leq \|x\| + \|y\|$.

The following result is an immediate consequence of the axioms.

Proposition 2.12 *Let V be a normed linear space. For all vectors $x, y \in V$:*

1. $\|x\| = 0$ *if, and only if,* $x = 0$.
2. $\| - x\| = \|x\|$.
3. $\|x - y\| \leq \|x\| + \|y\|$.
4. $|\|x\| - \|y\|| \leq \|x - y\|$.

Proof

1. Using homogeneity, $\|0\| = \|0 \cdot 0\| = 0 \cdot \|0\| = 0$, and together with positivity the claim follows.
2. Using homogeneity, $\| - x\| = \|(-1) \cdot x\| = 1 \cdot \|x\| = \|x\|$.
3. Using (2) and the triangle inequality, $\|x - y\| = \|x + (-y)\| \leq \|x\| + \|y\|$.
4. We need to show that $-\|x - y\| \leq \|x\| - \|y\| \leq \|x - y\|$. Using (3) and the triangle inequality, $\|x\| = \|(x + y) - y\| \leq \|x + y\| + \|y\|$ and similarly $\|y\| \leq \|x\| + \|x - y\|$, as needed. □

The definition of normed space was motivated by certain intuitive properties that the lengths of vectors, as modeled by the norm in an inner product space, satisfy. The next result shows that the inner product formalism indeed gives rise to a normed space, as expected.

Lemma 2.4 *Any inner product space, with its associated notion of norm, is a normed space.*

Proof Recall that the norm in an inner product space is given by $\|x\| = \sqrt{\langle x, x\rangle}$, and condition 1 in the definition of normed space is immediate. For condition 2, we have that

$$\|\alpha x\|^2 = \langle \alpha x, \alpha x\rangle = \alpha\overline{\alpha}\langle x, x\rangle = |\alpha|^2\|x\|^2$$

and the desired equality follows. As for the triangle inequality, by Proposition 2.11

$$\|x + y\|^2 = \langle x + y, x + y\rangle \leq \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2$$

and by applying the Cauchy-Schwarz inequality we obtain that

$$\|x + y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2,$$

and the result follows. □

We close the section, and the chapter, by introducing two very important families of normed spaces. The ℓ_p spaces, introduced already in Example 2.25, are spaces of sequences and each carries its own norm, the ℓ_p norm. The second family of spaces, the pre-L_p spaces, allude to the family of L_p spaces and the L_p norm. The pre-L_p spaces are, in a sense, the continuous version of the ℓ_p spaces and the L_p spaces are the completion of the pre-L_p spaces (completions are discussed in Chap. 4).

2.5.4 The Family of ℓ_p Spaces

For the reader's convenience, we repeat here the definition of the ℓ_p spaces, and augment it with the definition of the ℓ_p norm.

Definition 2.16 Let $1 \leq p < \infty$ be a real number. The set ℓ_p consists of all sequences $x = (x_1, \ldots, x_k, \ldots)$ of complex numbers for which

$$\sum_{k=1}^{\infty} |x_k|^p < \infty.$$

The ℓ_p-*norm* of x is

$$\|x\|_p = \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{1/p}.$$

The ℓ_∞ space is the space of all bounded sequences x of complex numbers and the ℓ_∞ norm is

$$\|x\|_\infty = \sup_k |x_k|.$$

Note that it is immediate that $\|x\|_p = 0$ if, and only if, $x = 0$. These spaces are over the field $\mathbb{C}$. By considering sequences of real numbers one obtains similar spaces over the field $\mathbb{R}$ which, ambiguously, are also referred to as ℓ_p spaces.

To further investigate the structure of ℓ_p, we need the following elementary result.

Proposition 2.13 (Young's Inequality) *The inequality*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

holds for all positive real numbers p and q satisfying $1/p + 1/q = 1$, and non-negative real numbers a and b, with equality if, and only if, $a^p = b^q$.

Proof Recall the weighted arithmetic-geometric means inequality

$$\sqrt[w]{x_1^{w_1} x_2^{w_2}} \leq \frac{w_1 x_1 + w_2 x_2}{w}$$

which holds for all non-negative real numbers x_1, x_2, and all positive weights w_1, w_2, with $w = w_1 + w_2$. Applying this inequality to the numbers $x_1 = a^p$ and $x_2 = b^q$, with weights $w_1 = 1/p$ and $w_2 = 1/q$, yields the desired inequality. □

The next inequality is used often when manipulating elements of the spaces ℓ_p. It is used below to establish the triangle inequality when proving that each ℓ_p space, with its ℓ_p norm, is a normed space. It is convenient to first introduce the following notation. For any two sequences $x = (x_1, \ldots, x_k, \ldots)$ and $y = (y_1, \ldots, y_k, \ldots)$, let

$$xy = (x_1 y_1, \ldots, x_k y_k, \ldots),$$

namely, the vector of component-wise products of the given vectors.

Theorem 2.14 (Hölder's Inequality) *Given positive real numbers p and q for which $1/p + 1/q = 1$, the inequality*

$$\|xy\|_1 \leq \|x\|_p \cdot \|y\|_q$$

holds for all $x \in \ell_p$ and $y \in \ell_q$.

Proof The case where either $\|x\|_p = 0$ or $\|y\|_q = 0$ is trivial and so we may assume this is not so. By component-wise division of x and y, respectively, by $\|x\|_p$ and $\|y\|_q$, we may assume that $\|x\|_p = \|y\|_q = 1$, namely that

$$\sum_{k=1}^{\infty} |x_k|^p = 1 = \sum_{k=1}^{\infty} |y_k|^q,$$

and we need to show that $\|xy\|_1 \leq 1$, or in other words that

$$\sum_{k=1}^{\infty} |x_k y_k| \leq 1.$$

By Young's Inequality (Proposition 2.13) we have that

$$|x_k y_k| = |x_k||y_k| \leq \frac{|x_k|^p}{p} + \frac{|y_k|^q}{q}$$

and therefore

$$\sum_{k=1}^{\infty} |x_k y_k| \leq \frac{1}{p} + \frac{1}{q} = 1,$$

as required. □

Theorem 2.15 (Minkowski's Inequality) *Let* $1 \leq p < \infty$. *The inequality*

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$

holds for all $x, y \in \ell_p$.

Proof Notice that

$$\|x + y\|_p^p = \sum_{k=1}^{\infty} |x_k + y_k|^p = \sum_{k=1}^{\infty} |x_k||x_k + y_k|^{p-1} + \sum_{k=1}^{\infty} |y_k||x_k + y_k|^{p-1}.$$

Applying Hölder's inequality (Theorem 2.14) with the given p and the associated $q = p/(p-1)$, we obtain

$$\sum_{k=1}^{\infty} |x_k||x_k + y_k|^{p-1} \leq (\sum_{k=1}^{\infty} |x_k|^p)^{\frac{1}{p}} (\sum_{k=1}^{\infty} |x_k + y_k|^p)^{1-\frac{1}{p}} = \|x\|_p \frac{\|x + y\|_p^p}{\|x + y\|_p}$$

and a similar inequality for the second summand. It follows that

$$\|x + y\|_p^p \leq (\|x\|_p + \|y\|_p) \frac{\|x + y\|_p^p}{\|x + y\|_p}$$

and, by simplifying, the claimed inequality is established. □

Theorem 2.16 *For all* $1 \leq p \leq \infty$ *the space* ℓ_p *is a normed linear space.*

Proof The case $p = \infty$ is left for the reader, so assume that $p < \infty$. The non-negativity of $\|x\|_p$ is clear from the definition of the norm and it was already noted that it is immediate that $\|x\| = 0$ implies $x = 0$. Homogeneity, is also immediate, since

$$\|\alpha x\|_p^p = \sum_{k=1}^{\infty} |\alpha x_k|^p = |\alpha|^p \|x\|_p^p$$

and finally the triangle inequality

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$

is precisely Minkowski's Inequality (Theorem 2.15). □

2.5.5 The Family of Pre-L_p Spaces

With the large family $\{\ell_p\}_{1 \leq p \leq \infty}$ of normed spaces at hand one can put the general theory to use for many sequences. Noticing that $\ell_p \subset \ell_q$ for all $1 \leq p < q \leq \infty$, and with sequences in ℓ_p converging faster to 0 than sequences in ℓ_q, given a bounded sequence one would typically try to identify a suitable p such that the sequence is found in ℓ_p, and proceed to apply the general theory to the problem at hand.

As long as sequences are concerned, the ℓ_p spaces are adequate but, quite often, when modeling a physical problem mathematically one obtains a function rather than a sequence. It is thus desirable to consider suitable normed spaces of functions. As explained above, obtaining the correct analogue of the ℓ_p spaces, namely the L_p spaces, requires some more sophisticated machinery than what we had presented so far, and so at this point we present what we call the pre-L_p spaces.

For the sake of simplicity, let us consider the ambient linear space $C([0, \infty), \mathbb{R})$ of all continuous function $x : [0, \infty) \to \mathbb{R}$ (we note that with the necessary modifications in the coming definitions and proofs, one may alter the domain, as well as consider complex-valued functions). The *pre-L_p* space, for $1 \leq p < \infty$, is the subset K_p of $C([0, \infty), \mathbb{R})$ whose elements are those functions x with

$$\int_0^{\infty} dt\ |x(t)|^p < \infty$$

where the integral is the usual Riemann integral (since we only consider continuous functions, the theory of Riemann integration is sufficient). The L_p *norm* of $x \in K_p$ is given by

$$\|x\|_p = \left(\int_0^{\infty} dt\ |x(t)|^p \right)^{1/p}$$

and we immediately note that $\|x\| = 0$ implies $x = 0$ (due to the continuity of x). Lastly, K_∞ is the space of all bounded continuous functions $x : [0, \infty) \to \mathbb{R}$ and the L_∞ norm is given by

$$\|x\|_\infty = \sup_{0 \le t < \infty} \{|x(t)|\}.$$

The rest of the section is devoted to showing that K_p with the L_p norm is a normed space, for each $1 \le p \le \infty$. In fact, the proof is formally identical to the case of the ℓ_p spaces. Indeed, the only ingredient one needs is the following version of Hölder's Inequality.

Theorem 2.17 (Hölder's Inequality) *Given positive real numbers p and q such that $1/p + 1/q = 1$, the inequality*

$$\|xy\|_1 \le \|x\|_p \cdot \|y\|_q$$

holds for all $x \in K_p$ and $y \in K_q$, where xy is the function $(xy)(t) = x(t)y(t)$.

Proof An inspection of the proof of Hölder's Inequality for sequences reveals that it was obtained by a point-wise application of Young's Inequality as well as properties of summation which are well-known to hold for integration as well, and so the same argument can be used to establish Hölder's Inequality for functions. □

The details of the proof of the following result are now formally identical to the proof of Theorem 2.16

Theorem 2.18 *For all $1 \le p \le \infty$ the pre-L_p space K_p is a normed linear space.*

Remark 2.18 Note again that the domain of the functions x in a pre-L_p space can be altered (sometimes with necessary extra care) and that the codomain can be replaced by $\mathbb{C}$ (with obvious adaptations to the definition of the L_p norm). All such spaces are still called pre-L_p spaces and denoted, ambiguously, by K_p. In particular, the space $C([a, b], \mathbb{R})$ of continuous real-valued functions on a closed interval can be endowed with an L_p norm and thus is a K_p space.

Exercises

Exercise 2.41 Prove the generalized Pythagoras' Theorem: In an inner product space, given m pairwise orthogonal vectors $x_1, \ldots, x_m$, the equality

$$\|x_1 + \cdots + x_m\|^2 = \|x_1\|^2 + \cdots + \|x_m\|^2$$

holds.

Exercise 2.42 Prove that ℓ_2 when endowed with the operation

$$\langle x, y \rangle = \sum_{k=1}^{\infty} \overline{x_k} y_k$$

is an inner product space.

Exercise 2.43 Prove that ℓ_∞ with the ℓ_∞ norm is a normed space and that K_∞ with the L_∞ norm is a normed space.

Exercise 2.44 Let V be an inner product space. Prove the parallelogram identity

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

for all $x, y \in V$.

Exercise 2.45 Prove that for $p \neq 2$ the space ℓ_p is not an inner product space.

Exercise 2.46 Does the equality

$$\ell_0 = \bigcap_{p>0} \ell_p$$

hold?

Exercise 2.47 For $x \in \ell_p$, with $1 \leq p < \infty$, does the equality

$$\lim_{q\to\infty} \|x\|_q = \|x\|_\infty$$

hold?

Exercise 2.48 Prove that K_2 is an inner product space.

Exercise 2.49 Let $1 \leq p \leq \infty$ and consider, for any $x \in K_p$ the sequence of samples $s(x) = (x(1), x(2), \ldots, x(k), \ldots)$. Prove that s is a linear operator $s : K_p \to \ell_p$. How do the norms $\|x\|_p$ and $\|s(x)\|_p$ compare?

Exercise 2.50 Prove Hölder's Inequality for functions and prove that K_p with the L_p norm is a normed space.

Further Reading

Introductory level texts on linear algebra typically treat linear spaces with a strong emphasis on techniques of matrices. When bases are not explicitly available, as in the infinite dimensional case, this approach must be replaced by a coordinate-free approach, as was done in this chapter. For an introductory level text treating linear algebra in a coordinate-free fashion see Chaps. 10–12 of Dummit and Foote [2]. A more advanced text with a coordinate-free approach, as well as an algebraically much deeper approach to linear algebra, making the connection between linear spaces and modules, is Roman [3]. Another text that bridges the gap between matrix-centric linear algebra and Hilbert space theory is Brown [1]. The reader seeking to enhance her intuition and ability with the Cauchy-Schwarz inequality is advised to consider Steele [4].

References

1. W.C. Brown, *A Second Course in Linear Algebra* (A Wiley-Interscience Publication, Wiley, New York, 1988) 264 p
2. D.S. Dummit, R.S. Foote, *Abstract Algebra*, 3rd edn. (Wiley, Hoboken, 2004) 932 p
3. S. Roman, *Advanced Linear Algebra*, 3rd edn., Graduate Texts in Mathematics 135 (Springer, New York, 2008) 522 p
4. J.M. Steele, *The Cauchy-Schwarz master class*, MAA Problem Books Series, Mathematical Association of America, Washington, DC (Cambridge University Press, Cambridge, 2004) 306 p

http://www.springer.com/978-3-319-03712-7

A Primer on Hilbert Space Theory
Linear Spaces, Topological Spaces, Metric Spaces, Normed Spaces, and Topological Groups
Alabiso, C.; Weiss, I.
2015, XVII, 255 p. 5 illus., Hardcover
ISBN: 978-3-319-03712-7