

# Chapter 1

## Vector Spaces

### 1.1 Groups and Fields

**Definition.** A binary operation on a set  $G$  is a mapping from  $G \times G$  to  $G$ .

**Definition.** A binary operation  $\star$  on a set  $G$  is called *associative* if for all  $a, b, c \in G$ ,  $(a \star b) \star c = a \star (b \star c)$  holds.

**Definition.** Let  $G$  be a set and  $\star$  be a binary operation on  $G$ . An *identity* of  $G$  with respect to  $\star$  is an element  $e \in G$  such that  $a \star e = a$  and  $e \star a = a$  for all  $a \in G$ .

**Theorem 1.1.** The identity of  $G$  with respect to  $\star$  is unique if it exists.

*Proof.* If  $e$  and  $e'$  are identity of  $G$  with respect to  $\star$ , then  $e = e \star e' = e'$ . □

**Notation.** The identity of  $G$  is denoted by  $1_G$ . However, if the binary operation is written additively, the identity is denoted by  $0_G$  instead.

**Definition.** Let  $\star$  be a binary operation on  $G$  with identity  $e$ . Let  $a$  be an element of  $G$ . An element  $b \in G$  is called an *inverse* of  $a$  if  $a \star b = e$  and  $b \star a = e$ .

**Theorem 1.2.** For all  $a \in G$ , the inverse of  $a \in G$  is unique if it exists.

*Proof.* If both  $b$  and  $b'$  are inverses of  $a$ , then

$$b = b \star 1_G = b \star (a \star b') = (b \star a) \star b' = 1_G \star b' = b'. \quad \square$$

**Notation.** The inverse of  $a$  in  $G$  is denoted by  $a^{-1}$ . However, if the binary operation is written additively, the inverse of  $a$  is denoted by  $-a$  instead.

**Definition.** A set  $G$  and a binary operation  $\star$  on  $G$  form a *group*  $(G, \star)$  if the following conditions hold.

(G 1)  $\star$  is associative.

(G 2) The identity of  $G$  (with respect to  $\star$ ) exists.

(G 3) For all  $a \in G$ , the inverse of  $a$  (with respect to  $\star$ ) exists.

**Example.** Let  $S$  denote the set of permutations of  $\{1, 2, 3\}$  and let  $\circ$  denote the composition of permutations. That is,

$$S = \{(1)(2)(3), (1)(2\ 3), (2)(3\ 1), (3)(1\ 2), (1\ 2\ 3), (3\ 2\ 1)\}.$$

Then  $(S, \circ)$  is a group.

**Definition.** A binary operation  $\star$  on a set  $G$  is called *commutative* if for all  $a, b \in G$ ,  $a \star b = b \star a$  holds.

**Definition.** A group  $(G, \star)$  is called an *Abelian group* if the following condition holds.

(G 4)  $\star$  is commutative.

**Example.**  $(\mathbb{Z}, +)$  and  $(\mathbb{Q} \setminus \{0\}, \cdot)$  are Abelian groups.

**Theorem 1.3.** Let  $(G, \star)$  be a group. Then for all  $a \in G$ ,  $(a^{-1})^{-1} = a$ .

*Proof.* Since  $a \star a^{-1} = 1_G$ ,  $a$  is the inverse of  $a^{-1}$  in  $G$ . Thus,  $(a^{-1})^{-1} = a$ .  $\square$

**Theorem 1.4** (Cancellation Law). Let  $(G, \star)$  be a group. Then the following statements are true.

- (a) For all  $a, b, c \in G$ , if  $c \star a = c \star b$ , then  $a = b$ .
- (b) For all  $a, b, c \in G$ , if  $a \star c = b \star c$ , then  $a = b$ .

*Proof.*

- (a) We have

$$a = 1_G \star a = (c^{-1} \star c) \star a = c^{-1} \star (c \star a)$$

and

$$b = 1_G \star b = (c^{-1} \star c) \star b = c^{-1} \star (c \star b).$$

Because  $c \star a = c \star b$ , we have  $a = b$ .

- (b) The proof is similar to (a).  $\square$

**Definition.** Let  $F$  be a set. Let  $+$  and  $\cdot$  be binary operations on  $F$ .

- The operation  $\cdot$  is called *left-distributive* over  $+$  if  $a \cdot (b + c) = a \cdot b + a \cdot c$  for all  $a, b, c \in F$ .
- The operation  $\cdot$  is called *right-distributive* over  $+$  if  $(a + b) \cdot c = a \cdot c + b \cdot c$  for all  $a, b, c \in F$ .
- The operation  $\cdot$  is called *distributive* over  $+$  if it is both left-distributive and right-distributive.

**Definition.** A set  $F$  and two binary operations  $+$  and  $\cdot$  on  $F$  form a *field*  $(F, +, \cdot)$  if the following conditions hold.

- (F 1)  $(F, +)$  is an Abelian group.
- (F 2)  $(F \setminus \{0_F\}, \cdot)$  is an Abelian group.

(F 3) The operation  $\cdot$  is distributive over  $+$ .

**Example.**  $(\mathbb{Q}, +, \cdot)$ ,  $(\mathbb{R}, +, \cdot)$ , and  $(\mathbb{C}, +, \cdot)$  are fields.

**Example.**  $(\mathbb{Q}[\sqrt{2}], +, \cdot)$  is a field, where

$$\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}.$$

**Theorem 1.5.** Let  $(F, +, \cdot)$  be a field. Then the following statements are true.

- (a) For all  $a \in F$ ,  $a \cdot 0_F = 0_F = 0_F \cdot a$ .
- (b) For all  $a, b \in F$ ,  $(-a) \cdot b = -(a \cdot b) = a \cdot (-b)$ .
- (c) For all  $a, b \in F$ ,  $(-a) \cdot (-b) = a \cdot b$ .

*Proof.*

- (a) We have

$$a \cdot 0_F + a \cdot 0_F = a \cdot (0_F + 0_F) = a \cdot 0_F = a \cdot 0_F + 0_F.$$

Thus,  $a \cdot 0_F = 0_F$  by cancelltaion law (Theorem 1.4). The proof of  $0_F \cdot a = 0_F$  is similar.

- (b) By (a), we have

$$a \cdot b + (-a) \cdot b = (a + (-a)) \cdot b = 0_F \cdot b = 0_F.$$

Thus,  $(-a) \cdot b = -(a \cdot b)$ . The proof of  $a \cdot (-b) = -(a \cdot b)$  is similar.

- (c) We have

$$(-a) \cdot (-b) = -(a \cdot (-b)) = -(-(a \cdot b)) = a \cdot b$$

by applying (b) twice. □

*Remark.* Let  $G = F \setminus \{0_F\}$  and  $1_G$  be the multiplicative identity of  $G$ . By Theorem 1.5 (a), we have  $1_G \cdot 0_F = 0_F = 0_F \cdot 1_G$ . Therefore,  $1_G$  is also the multiplicative identity of  $F$ , and thus we denote it by  $1_F$ .

*Remark.* Subtraction and division are defined in terms of addition and multiplication by using additive and multiplicative inverses.

## 1.2 Vector Spaces

**Definition.** Let  $F$  be a field and let  $V$  be a set on which two operations  $+: V \times V \rightarrow V$  and  $\cdot: F \times V \rightarrow V$  are defined. Then  $(V, +, \cdot)$  is a *vector space* over  $F$  if the following conditions hold.

(V 1)  $(V, +)$  is an Abelian group.

(V 2) For all  $x \in V$ ,  $1_F \cdot x = x$ .

(V 3) For all  $a, b \in F$  and for all  $x \in V$ ,  $(a \cdot b) \cdot x = a \cdot (b \cdot x)$ .

(V 4) For all  $a, b \in F$  and for all  $x \in V$ ,  $(a + b) \cdot x = a \cdot x + b \cdot x$ .

(V 5) For all  $a \in F$  and for all  $x, y \in V$ ,  $a \cdot (x + y) = a \cdot x + a \cdot y$ .

*Remark.* We also say that  $V$  is a vector space over  $F$  if both  $+$  and  $\cdot$  are “standard”.

**Example.**  $(\mathbb{C}, +, \cdot)$  is a vector space over  $\mathbb{R}$ , and  $(\mathbb{R}, +, \cdot)$  is a vector space over  $\mathbb{Q}$ .

**Example.** Let  $F$  be a field.

- $(F^n, +, \cdot)$  is a vector space over  $F$ .
- Let  $\mathcal{P}(F)$  denote the set of polynomials with coefficients in  $F$ . Then  $(\mathcal{P}(F), +, \cdot)$  is a vector space over  $F$ .
- Let  $\mathcal{F}(S, F)$  denote the set of functions from  $S$  to  $F$ . Then  $(\mathcal{F}(S, F), +, \cdot)$  is a vector space over  $F$ .

**Theorem 1.6.** Let  $(V, +, \cdot)$  be a vector space over  $F$ . Then the following statements are true.

(a) For all  $x \in V$ ,  $0_F \cdot x = 0_V$ .

(b) For all  $a \in F$ ,  $a \cdot 0_V = 0_V$ .

(c) For all  $a \in F$  and  $x \in V$ ,  $(-a) \cdot x = -(a \cdot x) = a \cdot (-x)$ .

*Proof.* It is similar to the proof of Theorem 1.5.

(a) We have

$$0_F \cdot x + 0_F \cdot x = (0_F + 0_F) \cdot x = 0_F \cdot x = 0_F \cdot x + 0_V.$$

Thus,  $0_F \cdot x = 0_V$  by cancelltaion law (Theorem 1.4).

(b) It is similar to the proof of (a).

(c) By (a), we have

$$a \cdot x + (-a) \cdot x = (a + (-a)) \cdot x = 0_F \cdot x = 0_V.$$

Thus,  $(-a) \cdot x = -(a \cdot x)$ . By (b), we have

$$a \cdot x + a \cdot (-x) = a \cdot (x + (-x)) = a \cdot 0_V = 0_V.$$

Thus,  $a \cdot (-x) = -(a \cdot x)$ . □

## 1.3 Subspaces

**Definition.** Let  $(V, +_V, \cdot_V)$  be a vector space over a field  $F$ . Let  $W$  be a subset of  $V$ . If  $+_W : W \times W \rightarrow W$  and  $\cdot_W : F \times W \rightarrow W$  satisfy

$$x +_W y = x +_V y \quad \text{and} \quad a \cdot_W x = a \cdot_V x$$

for all  $a \in F$  and  $x, y \in W$ , then we say that  $+_W$  and  $\cdot_W$  *inherit*  $+_V$  and  $\cdot_V$ , respectively.

**Definition.** Let  $(V, +_V, \cdot_V)$  be a vector space over  $F$ . A subset  $W$  of  $V$  is called a *subspace* of  $V$  if  $(W, +_W, \cdot_W)$  is a vector space over  $F$ , where  $+_W$  and  $\cdot_W$  inherit  $+_V$  and  $\cdot_V$ , respectively.

**Theorem 1.7.** Let  $(V, +_V, \cdot_V)$  be a vector space over  $F$ . Let  $W$  be a subset of  $V$ . Then  $W$  is a subspace of  $V$  if the following conditions hold.

- (a) For all  $x, y \in W$ ,  $x +_V y \in W$ .
- (b) For all  $a \in F$  and  $x \in W$ ,  $a \cdot_V x \in W$ .
- (c)  $0_V \in W$ .

*Proof.* We can define operations  $+_W : W \times W \rightarrow W$  and  $\cdot_W : F \times W \rightarrow W$  such that

$$x +_W y = x +_V y \quad \text{and} \quad a \cdot_W x = a \cdot_V x$$

for all  $a \in F$  and  $x, y \in W$  due to (a) and (b). Then  $+_W$  and  $\cdot_W$  inherit  $+_V$  and  $\cdot_V$ , respectively.

Now we prove that  $(W, +_W, \cdot_W)$  is a vector space over  $F$ . Since a vector in  $W$  is also in  $V$ , (V 2), (V 3), (V 4) and (V 5) hold trivially for  $W$ . Thus, one only needs to prove (V 1), i.e.,  $(W, +_W)$  is an Abelian group.

Since  $+_W$  inherits  $+_V$ ,  $+_V$  is associative implies that  $+_W$  is associative. Furthermore, since

$$0_V \in W \quad \text{and} \quad -x = -(1_F \cdot x) = (-1_F) \cdot x \in W$$

hold for all  $x \in W$ , we have

$$0_V +_W x = x = x +_W 0_V \quad \text{and} \quad x +_W (-x) = 0_V = (-x) +_W x$$

hold for all  $x \in W$ . Thus,  $0_V \in W$  is an additive identity of  $W$ , and each vector in  $W$  also has an additive inverse in  $W$ , which complete the proof.  $\square$

**Example.** Let  $\mathcal{P}_n(F)$  denote the set of polynomials in  $\mathcal{P}(F)$  with degree less than or equal to  $n$ , where  $n \geq -1$  is an integer. Then it follows from Theorem 1.7 that  $\mathcal{P}_n(F)$  is a subspace of  $\mathcal{P}(F)$ .

**Theorem 1.8.** Let  $(V, +_V, \cdot_V)$  be a vector space over  $F$ . Let  $I$  be an index set such that  $W_i$  is a subspace of  $V$  for all  $i \in I$ . Then the intersection

$$W = \bigcap_{i \in I} W_i$$

is a subspace of  $V$ .

*Proof.* For all  $a \in F$  and for all  $x, y \in W$ , since

$$x +_V y \in W_i \quad \text{and} \quad a \cdot_V x \in W_i \quad \text{and} \quad 0_V \in W_i$$

hold for all indices  $i \in I$ , we have

$$x +_V y \in W \quad \text{and} \quad a \cdot_V x \in W \quad \text{and} \quad 0_V \in W.$$

Thus,  $W$  is a subspace of  $V$ . □

**Definition.** Let  $(V, +_V, \cdot_V)$  be a vector space over  $F$ . Let  $S_1$  and  $S_2$  be subsets of  $V$ . Then the *sum* of  $S_1$  and  $S_2$ , denoted  $S_1 + S_2$ , is defined as

$$S_1 + S_2 = \{x + y : x \in S_1 \text{ and } y \in S_2\}.$$

**Theorem 1.9.** Let  $(V, +_V, \cdot_V)$  be a vector space over  $F$ . If  $W_1$  and  $W_2$  be subspaces of  $V$ , then the following statements are true.

- (a)  $W_1 + W_2$  is a subspace of  $V$ .
- (b) If  $W$  is a subspace of  $V$  with  $W_1 \cup W_2 \subseteq W$ , then  $W_1 + W_2 \subseteq W$ .

*Proof.*

- (a) Suppose that  $a \in F$  and  $x, y \in W_1 + W_2$ . Then there exists  $x_1, y_1 \in W_1$  and  $x_2, y_2 \in W_2$  such that

$$x = x_1 +_V x_2 \quad \text{and} \quad y = y_1 +_V y_2.$$

Thus,

$$a \cdot_V x = a \cdot_V (x_1 + x_2) = a \cdot_V x_1 + a \cdot_V x_2 \in W_1 + W_2$$

and

$$x +_V y = (x_1 +_V x_2) + (y_1 +_V y_2) = (x_1 +_V y_1) + (x_2 +_V y_2) \in W_1 + W_2.$$

We also have  $0_V = 0_V +_V 0_V \in W_1 + W_2$ . Hence,  $W_1 + W_2$  is a subspace of  $V$ .

- (b) If  $x \in W_1 + W_2$ , then there exists  $x_1 \in W_1$  and  $x_2 \in W_2$  such that  $x = x_1 + x_2$ . Since  $W_1 \subseteq W$  and  $W_2 \subseteq W$ , we have  $x_1 \in W$  and  $x_2 \in W$ , which implies  $x \in W$ . □

## 1.4 Spanning Sets

**Definition.** Let  $(G, +)$  be an Abelian group. Then we define

$$\sum_{i=m}^n a_i = \begin{cases} \sum_{i=m}^{n-1} a_i + a_n & \text{if } m \leq n \\ 0_G & \text{if } m > n, \end{cases}$$

where  $a_i \in G$  for each integer  $i$  with  $m \leq i \leq n$ .

**Definition.** Let  $(V, +, \cdot)$  be a vector space over  $F$ . Let  $S$  be a subset of  $V$ . Then a vector  $x \in V$  is called a *linear combination* of  $S$  if there exist some nonnegative integer  $n$ , scalars  $a_1, \dots, a_n \in F$ , and vectors  $x_1, \dots, x_n \in S$  such that

$$x = \sum_{i=1}^n a_i x_i.$$

*Remark.* Since  $n$  can be zero,  $0_V$  is a linear combination for all  $S \subseteq V$ .

*Remark.* Although  $S$  can be infinite, the number of terms in the summation must be finite. For example, in the vector space  $\mathbb{R}$  over  $\mathbb{Q}$ , although we have

$$e = \sum_{k=0}^{\infty} \frac{1}{k!} = \frac{1}{1} + \frac{1}{1} + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \cdots,$$

$e$  is still not a linear combination of  $\mathbb{Q}$ .

**Definition.** Let  $(V, +, \cdot)$  is a vector space over  $F$ . The *span* of  $S$ , denoted  $\text{span}(S)$ , is the set that consists of all linear combinations of  $S$ .

**Theorem 1.10.** Let  $(V, +, \cdot)$  be a vector space over  $F$ . Let  $S \subseteq V$ . Then the following statements are true.

- (a)  $\text{span}(S)$  is a subspace of  $V$ .
- (b) If  $W$  is a subspace of  $V$  such that  $S \subseteq W$ , then  $\text{span}(S) \subseteq W$ .

*Proof.*

- (a) If  $c \in F$  and  $x, y \in \text{span}(S)$ , then there exist nonnegative integers  $m, n$ , scalars  $a_1, \dots, a_m, b_1, \dots, b_n \in F$  and vectors  $x_1, \dots, x_m, y_1, \dots, y_n \in S$  such that

$$x = \sum_{i=1}^m a_i x_i \quad \text{and} \quad y = \sum_{j=1}^n b_j y_j.$$

Thus, we have

$$\begin{aligned} cx &= c(a_1 x_1 + \cdots + a_m x_m) \\ &= c(a_1 x_1) + \cdots + c(a_m x_m) \\ &= (ca_1)x_1 + \cdots + (ca_m)x_m \in \text{span}(S) \end{aligned}$$

and

$$x + y = a_1 x_1 + \cdots + a_m x_m + b_1 y_1 + \cdots + b_n y_n \in \text{span}(S).$$

Also,  $0_V \in \text{span}(S)$ . Hence,  $\text{span}(S)$  is a subspace of  $V$ .

- (b) If  $x \in \text{span}(S)$ , then there exists a nonnegative integer  $n$ , scalars  $a_1, \dots, a_n \in F$  and vectors  $x_1, \dots, x_n \in S$  such that

$$x = \sum_{i=1}^n a_i x_i.$$

Thus, since  $x_1, \dots, x_n \in W$ , we have  $x = a_1 x_1 + \dots + a_n x_n \in W$ .  $\square$

**Definition.** A subset  $S$  of a vector space  $(V, +, \cdot)$  *spans*  $V$  if  $\text{span}(S) = V$ . In this case, we also say that  $S$  is a *spanning set* of  $V$ .

**Example.**  $\{(0, 1, 1), (1, 0, 1), (1, 1, 0)\}$  is a spanning set of  $\mathbb{R}^3$  since for all  $x, y, z \in \mathbb{R}$ ,

$$(x, y, z) = \frac{-x + y + z}{2} \cdot (0, 1, 1) + \frac{x - y + z}{2} \cdot (1, 0, 1) + \frac{x + y - z}{2} \cdot (1, 1, 0).$$



## 1.5 Linearly Independent Sets

**Definition.** Let  $(V, +, \cdot)$  be a vector space over  $F$ . Let  $S$  be a subset of  $V$ . For scalars  $a_1, \dots, a_n \in F$  and distinct vectors  $x_1, \dots, x_n \in S$ , we say that

$$\sum_{i=1}^n a_i x_i = 0_V$$

is a *trivial representation* of  $0_V$  as a linear combination of  $S$  if  $a_1 = \dots = a_n = 0_F$ .

**Definition.** Let  $(V, +, \cdot)$  be a vector space over  $F$ .

- A subset  $S$  of  $V$  is called *linearly dependent* if there exists a nontrivial representation of  $0_V$  as a linear combination of  $S$ .
- A subset  $S$  of  $V$  is called *linearly independent* if it is not linear dependent.

**Theorem 1.11.** Let  $(V, +, \cdot)$  be a vector space over  $F$  and let  $S \subseteq V$ . Then  $S$  is linearly independent if and only if there exists  $x \in S$  such that  $x \in \text{span}(S \setminus \{x\})$ .

*Proof.* ( $\Rightarrow$ ) Because  $S$  is linearly dependent, it follows that there exists a nontrivial representation

$$a_1 x_1 + a_2 x_2 + \dots + a_n x_n = 0_V$$

as a linear combination of  $S$ , where  $a_1, \dots, a_n \in F$  are scalars and  $x_1, \dots, x_n \in S$  are distinct vectors. Without loss of generality, let  $a_1 \neq 0_F$ . Then we have

$$\begin{aligned} x_1 &= (-a_1)^{-1}(a_2 x_2 + \dots + a_n x_n) \\ &= (-a_1)^{-1} a_2 x_2 + \dots + (-a_1)^{-1} a_n x_n \\ &\in \text{span}(S \setminus \{x_1\}). \end{aligned}$$

( $\Leftarrow$ ) Since  $x \in \text{span}(S \setminus \{x\})$ , there exists scalars  $a_1, \dots, a_n \in F$  and distinct vectors  $x_1, \dots, x_n \in S \setminus \{x\}$  such that

$$a_1 x_1 + \dots + a_n x_n = x.$$

Then

$$(-1_F)x + a_1 x_1 + \dots + a_n x_n = 0_V$$

is a nontrivial representation of  $0_V$  as a linear combination of  $S$ . □

**Theorem 1.12.** Let  $(V, +, \cdot)$  be a vector space over  $F$ . Let  $S$  be a subset of  $V$  and let  $x$  be an element of  $S$ . Then  $x \in \text{span}(S \setminus \{x\})$  if and only if  $\text{span}(S) = \text{span}(S \setminus \{x\})$ .

*Proof.* ( $\Rightarrow$ ) Since  $x \in \text{span}(S \setminus \{x\})$  and  $S \setminus \{x\} \subseteq \text{span}(S \setminus \{x\})$ , we have

$$S \subseteq \text{span}(S \setminus \{x\}) \quad \Rightarrow \quad \text{span}(S) \subseteq \text{span}(S \setminus \{x\})$$

by Theorem 1.10. Also,  $\text{span}(S \setminus \{x\}) \subseteq \text{span}(S)$  because  $S \setminus \{x\} \subseteq S$ . Thus, we can conclude that  $\text{span}(S \setminus \{x\}) = \text{span}(S)$ .

( $\Leftarrow$ ) Since  $x \in S \subseteq \text{span}(S) = \text{span}(S \setminus \{x\})$ , we have  $x \in \text{span}(S \setminus \{x\})$ . □

**Example.** Let  $S = \{1, 1 + 2x, 1 + 2x + 3x^2, 1 + 2x + 3x^2 + 4x^3\}$  be a subset of  $\mathcal{P}_3(\mathbb{R})$ . Then  $S$  is linearly independent since the only solution to the following system of linear equations

$$\begin{aligned} a_1 &= 0 \\ a_1 + 2a_2 &= 0 \\ a_1 + 2a_2 + 3a_3 &= 0 \\ a_1 + 2a_2 + 3a_3 + 4a_4 &= 0 \end{aligned}$$

is  $a_1 = a_2 = a_3 = a_4 = 0$ .

**Theorem 1.13.** Let  $(V, +, \cdot)$  be a vector space, and let  $R \subseteq S \subseteq V$ . If  $R$  is linearly dependent, then  $S$  is linearly dependent.

*Proof.* If  $R$  is linearly dependent, then there exists  $x \in R$  such that  $x \in \text{span}(R \setminus \{x\})$ . By  $R \subseteq S$ , we have  $R \setminus \{x\} \subseteq S \setminus \{x\}$ . Since  $x \in S$  and  $x \in \text{span}(S \setminus \{x\})$ ,  $S$  is linearly dependent.  $\square$

**Corollary.** Let  $(V, +, \cdot)$  be a vector space, and let  $R \subseteq S \subseteq V$ . If  $S$  is linearly independent, then  $R$  is linearly independent.

*Proof.* Suppose that  $S$  is linearly independent. If  $R$  is linearly dependent, then so is  $S$  by Theorem 1.13, contradiction. Thus,  $R$  is linearly independent.  $\square$

**Theorem 1.14.** Let  $(V, +, \cdot)$  be a vector space. For each finite set  $S \subseteq V$ , there exists a linearly independent set  $Q \subseteq S$  such that  $\text{span}(Q) = \text{span}(S)$ .

*Proof.* The proof is by induction on  $n = |S|$ . The induction begins with  $n = 0$ , i.e.,  $S = \emptyset$ . Since  $\emptyset$  is linearly independent, we can choose  $R = \emptyset$ , and thus the theorem holds.

Now suppose that the theorem is true for some integer  $n \geq 0$ , and we prove that the theorem holds for  $n + 1$ . If  $S$  is linearly independent, then we can choose  $Q = S$ . Otherwise, there exists  $x \in S$  with  $\text{span}(S \setminus \{x\}) = \text{span}(S)$  because  $S$  is linearly dependent. Let  $S' = S \setminus \{x\}$ . Then there exists a linearly independent set  $Q \subseteq S'$  such that  $\text{span}(Q) = \text{span}(S')$  by induction hypothesis, implying  $Q \subseteq S$  and  $\text{span}(Q) = \text{span}(S)$ .  $\square$

## 1.6 Bases and Dimension

**Definition.** Let  $(V, +, \cdot)$  be a vector space. A subset  $S$  of  $V$  is a *basis* of  $V$  if  $S$  is not only a spanning set but also a linearly independent set of  $V$ .

**Example.** Following are some examples of bases.

- Since  $\text{span}(\emptyset) = \{0_V\}$  and  $\emptyset$  is linearly independent,  $\emptyset$  is a basis of  $\{0_V\}$ .
- Let  $S = \{x_1, \dots, x_n\}$  be a subset of  $F^n$  with  $(x_i)_j = \llbracket i = j \rrbracket$  for all  $i, j \in \{1, \dots, n\}$ . Then  $S$  is called the *standard basis* of  $F^n$ .
- The set  $S = \{1_F, x, x^2, \dots, x^n\}$  is called the *standard basis* of  $\mathcal{P}_n(F)$ .

**Theorem 1.15.** Let  $(V, +, \cdot)$  be a vector space over  $F$ . If there exists a finite set  $S$  that spans  $V$ , then there is a subset  $Q$  of  $S$  that is a finite basis of  $V$ .

*Proof.* By Theorem 1.14, there exists a linearly independent set  $Q \subseteq S$  such that  $\text{span}(Q) = \text{span}(S) = V$ . Thus,  $Q$  is a finite basis of  $V$ .  $\square$

**Theorem 1.16** (Replacement Theorem). Let  $(V, +, \cdot)$  be a vector space over  $F$ . Let  $S$  be a finite set that spans  $V$ , and let  $Q \subseteq S$  be a finite linearly independent set. Then  $|Q| \leq |S|$ , and there exists  $R \subseteq S \setminus Q$  such that both  $|Q \cup R| = |S|$  and  $\text{span}(Q \cup R) = V$  hold.

*Proof.* The proof is based on induction on  $|Q|$ . The induction begins with  $|Q| = 0$ , i.e.,  $Q = \emptyset$ . Choosing  $R = S$ , we have  $Q \cup R = S$ , and thus both  $|Q \cup R| = |S|$  and  $\text{span}(Q \cup R) = V$  hold.

Now suppose that the theorem is true for  $|Q| = m$  with  $m \geq 0$ , and we prove that the theorem holds for  $|Q| = m + 1$ . Let  $Q = \{x_1, \dots, x_{m+1}\}$  and let  $Q' = Q \setminus \{x_{m+1}\}$ . By induction hypothesis, there exists  $R' = \{y_1, \dots, y_k\} \subseteq S \setminus Q'$  such that  $m + k = |S|$  and  $\text{span}(Q' \cup R') = V$ . Since  $Q' \cup R'$  spans  $V$ , there exists  $a_1, \dots, a_m, b_1, \dots, b_k \in F$  such that

$$x_{m+1} = \sum_{i=1}^m a_i x_i + \sum_{j=1}^k b_j y_j.$$

If  $b_j = 0_F$  for all  $j \in \{1, \dots, k\}$ , then  $x_{m+1}$  is a linear combination of  $Q$ , implying that  $Q$  is linearly dependent, contradiction. Thus, there must exist some  $j \in \{1, \dots, k\}$  such that  $b_j \neq 0_F$ . Without loss of generality let  $b_k \neq 0_F$ . Also, let  $R = \{y_1, \dots, y_{k-1}\}$ . Then  $|Q \cup R| = (m + 1) + (k - 1) = |S|$ . Since  $k \geq 1$ , we have  $|Q| \leq |S|$ . Note that  $(Q' \cup R') \setminus (Q \cup R) = \{y_k\}$ . By

$$y_k = (-b_k)^{-1} \left( \sum_{i=1}^m a_i x_i + (-1_F) x_{m+1} + \sum_{j=1}^{k-1} b_j y_j \right) \in \text{span}(Q \cup R),$$

we have

$$Q' \cup R' \subseteq Q \cup R \cup \{y_k\} \subseteq \text{span}(Q \cup R).$$

Thus, by Theorem 1.10 we have

$$V = \text{span}(Q' \cup R') \subseteq \text{span}(Q \cup R) \subseteq V,$$

implying  $\text{span}(Q \cup R) = V$ .  $\square$

**Corollary.** Let  $(V, +, \cdot)$  be a vector space over  $F$  that is spanned by a finite set. Then every linearly independent subset of  $V$  is finite.

*Proof.* Suppose that  $S$  is a finite spanning set of  $V$  and that  $Q$  is linearly independent. If  $Q$  is infinite, then there exists  $Q' \subseteq Q$  with  $|Q'| = |S| + 1$ . It follows that  $Q'$  is linearly independent by Theorem 1.13, and thus  $|Q'| \leq |S|$  by Theorem 1.16, contradiction to  $|Q'| = |S| + 1$ . Therefore,  $Q$  is finite.  $\square$

**Theorem 1.17.** Let  $(V, +, \cdot)$  be a vector space over  $F$ . If  $V$  has a finite basis, then all bases of  $V$  have the same size.

*Proof.* Let  $S$  be a finite basis of  $V$  and let  $Q$  be an arbitrary basis of  $V$ . Since  $V = \text{span}(S)$  and  $Q$  is linearly independent, it follows that  $Q$  is finite, and thus  $|Q| \leq |S|$  by replacement theorem (Theorem 1.16).

Also, since  $V = \text{span}(Q)$  and  $S$  is linearly independent, we have  $|S| \leq |Q|$  by replacement theorem (Theorem 1.16). Thus,  $|Q| = |S|$ .  $\square$

**Definition.** A vector space  $(V, +, \cdot)$  over  $F$  is called *finite-dimensional* if it has a finite basis. A vector space that is not finite-dimensional is called *infinite-dimensional*.

**Definition.** The number of vectors in each basis of a finite-dimensional vector space  $V$  is called the *dimension* of  $V$  and is denoted by  $\dim(V)$ .

**Example.** We have  $\dim(\{0_V\}) = 0$ ,  $\dim(F^n) = n$ , and  $\dim(\mathcal{P}_n(F)) = n + 1$ .

**Example.** The dimension of a vector space depends on its field of scalars.

- If  $V = \mathbb{C}$  is a vector space over  $\mathbb{R}$ , then  $\dim(V) = 2$  since  $\{1, i\}$  is a basis of  $V$ .
- If  $W = \mathbb{C}$  is a vector space over  $\mathbb{C}$ , then  $\dim(W) = 1$  since  $\{1\}$  is a basis of  $W$ .

**Theorem 1.18.** Let  $(V, +, \cdot)$  be a vector space over  $F$ . Then a subset of  $V$  of  $n = \dim(V)$  vectors is linearly independent if and only if it is a spanning set of  $V$ .

*Proof.* ( $\Rightarrow$ ) Suppose that  $Q$  is linearly independent with  $|Q| = n$ . By replacement theorem (Theorem 1.16), there exists  $R \subseteq S \setminus Q$  such that  $|Q \cup R| = |S|$  and  $\text{span}(Q \cup R) = V$ . Since  $|Q| = |S|$ , we have  $|R| = 0$ , i.e.,  $R = \emptyset$ . Thus,  $\text{span}(Q) = V$ .

( $\Leftarrow$ ) Suppose that  $S$  spans  $V$  with  $|S| = n$ . By Theorem 1.15, there is a subset  $Q$  of  $S$  that is a basis of  $V$ . Then we have  $|Q| = n$ , implying  $Q = S$ . Thus,  $S$  is a basis of  $V$ .  $\square$

**Theorem 1.19.** Let  $(V, +, \cdot)$  be a finite-dimensional vector space over  $F$ , and let  $V'$  be a subspace of  $V$ . Then the following statements hold.

- (a)  $\dim(V') \leq \dim(V)$ .
- (b) If  $\dim(V') = \dim(V)$ , then  $V' = V$ .

*Proof.* Let  $S$  be a basis of  $V$  and let  $S'$  be a basis of  $V'$ .

- (a) Since  $S'$  is linearly independent and  $V = \text{span}(S)$ , we have  $|S'| \leq |S|$  by replacement theorem (Theorem 1.16). Thus,  $\dim(V') \leq \dim(V)$ .
- (b) Since  $S'$  is linearly independent and  $|S'| = \dim(V)$ , we have  $\text{span}(S') = V$  by Theorem 1.18. Thus,  $V' = \text{span}(S') = V$ .  $\square$

# Chapter 2

## Linear Transformations

### 2.1 Linear Transformations, Null Spaces and Ranges

**Definition.** Let  $f : X \rightarrow Y$  be a function.

- $f$  is *injective* (i.e.,  $f$  is an *injection*) if  $T(x) = T(x')$  implies  $x = x'$  for  $x, x' \in X$ .
- $f$  is *surjective* (i.e.,  $f$  is a *surjection*) if for each  $y \in Y$ , there exists some  $x \in X$  with  $T(x) = y$ .
- $f$  is *bijective* (i.e.,  $f$  is a *bijection*) if  $f$  is injective and surjective.

*Remark.* If both domain and codomain of a function are vector spaces, then the function is usually said to be a *transformation*. Furthermore, it is said to be an *operator* if its domain and codomain are the same.

**Definition.** Let  $V$  and  $W$  be vector spaces over  $F$ . A transformation  $T : V \rightarrow W$  is *linear* if the following statements hold.

- (a)  $T(x + y) = T(x) + T(y)$  for all  $x, y \in V$ .
- (b)  $T(ax) = aT(x)$  for all  $a \in F$  and  $x \in V$ .

The set of all linear transformations from  $V$  to  $W$  is denoted by  $\mathcal{L}(V, W)$ . In the case that  $V = W$ , we write  $\mathcal{L}(V)$  for short.

**Example.** The *zero transformation* from  $V$  to  $W$  is the transformation  $O_{V,W} : V \rightarrow W$  that satisfies  $O_{V,W}(x) = 0_W$  for all  $x \in V$ . It is clear that  $O_{V,W} \in \mathcal{L}(V, W)$ .

**Example.** The *identity transformation* on  $V$  is the transformation  $I_V : V \rightarrow V$  that satisfies  $I_V(x) = x$  for all  $x \in V$ . It is clear that  $I_V \in \mathcal{L}(V)$ .

**Example.** Recall that  $\mathcal{P}(F)$  is the set of polynomials with coefficients in  $F$ .

- The differential operator  $D : \mathcal{P}(\mathbb{R}) \rightarrow \mathcal{P}(\mathbb{R})$  with  $D(f) = f'$  for  $f \in \mathcal{P}(\mathbb{R})$ , where  $f'$  is the derivative of  $f$ , is linear.
- The operator  $T : \mathcal{P}(\mathbb{R}) \rightarrow \mathcal{P}(\mathbb{R})$  such that for  $f \in \mathcal{P}(\mathbb{R})$ ,

$$(T(f))(x) = \int_0^x f(t)dt$$

for all  $x \in \mathbb{R}$ , is linear.

**Theorem 2.1.** If  $V$  and  $W$  are vector spaces over  $F$ , then  $\mathcal{L}(V, W)$  is also a vector space over  $F$ .

*Proof.*  $\mathcal{L}(V, W)$  is a vector space because it is a subspace of  $\mathcal{F}(V, W)$ , which is proved as follows.

(a) If  $T_1, T_2 \in \mathcal{L}(V, W)$ , then  $T_1 + T_2$  is linear because

$$\begin{aligned}(T_1 + T_2)(x + y) &= T_1(x + y) + T_2(x + y) \\ &= T_1(x) + T_1(y) + T_2(x) + T_2(y) \\ &= T_1(x) + T_2(x) + T_1(y) + T_2(y) \\ &= (T_1 + T_2)(x) + (T_1 + T_2)(y)\end{aligned}$$

and

$$\begin{aligned}(T_1 + T_2)(cx) &= T_1(cx) + T_2(cx) \\ &= cT_1(x) + cT_2(x) \\ &= c(T_1(x) + T_2(x)) \\ &= c(T_1 + T_2)(x)\end{aligned}$$

hold for  $x, y \in V$  and  $c \in F$ .

(b) If  $T \in \mathcal{L}(V, W)$  and  $a \in F$ , then  $aT$  is linear because

$$\begin{aligned}(aT)(x + y) &= aT(x + y) \\ &= a(T(x) + T(y)) \\ &= aT(x) + aT(y) \\ &= (aT)(x) + (aT)(y)\end{aligned}$$

and

$$(aT)(cx) = aT(cx) = a(cT(x)) = c(aT(x)) = c(aT)(x)$$

hold for  $x, y \in V$  and  $c \in F$ .

(c) It is clear that  $O_{V,W} \in \mathcal{L}(V, W)$ . □

**Theorem 2.2.** Let  $V$  and  $W$  be vector spaces over  $F$ , and let  $T : V \rightarrow W$  be linear. Let  $S$  be a subset of  $V$  and let  $U$  be a subspace of  $V$ . Then the following statements are true.

(a) If  $n$  is a nonnegative integer, then for  $a_1, \dots, a_n \in F$  and  $x_1, \dots, x_n \in V$ , we have

$$T\left(\sum_{i=1}^n a_i x_i\right) = \sum_{i=1}^n a_i T(x_i).$$

(b) If  $S$  spans  $U$ , then  $T(S)$  spans  $T(U)$ .

*Proof.*

- (a) The proof is by induction on  $n$ . For  $n = 0$ , it holds trivially. If the statement is true for some  $n \geq 0$ , then we have

$$\begin{aligned} T(a_1x_1 + \cdots + a_nx_n + a_{n+1}x_{n+1}) &= T(a_1x_1 + \cdots + a_nx_n) + T(a_{n+1}x_{n+1}) \\ &= a_1T(x_1) + \cdots + a_nT(x_n) + a_{n+1}T(x_{n+1}). \end{aligned}$$

Thus, the statement is true for nonnegative integer  $n$ .

- (b) We prove that  $\text{span}(T(S)) = T(U)$ . If  $y \in \text{span}(T(S))$ , then there exist  $a_i \in F$ ,  $x_i \in S$  for  $i \in \{1, \dots, n\}$  such that

$$y = \sum_{i=1}^n a_i T(x_i) = T\left(\sum_{i=1}^n a_i x_i\right) \in T(U),$$

so  $\text{span}(T(S)) \subseteq T(U)$ .

If  $y \in T(U)$ , then there exist  $a_i \in F$ ,  $x_i \in S$  for  $i \in \{1, \dots, n\}$  such that

$$y = T\left(\sum_{i=1}^n a_i x_i\right) = \sum_{i=1}^n a_i T(x_i) \in \text{span}(T(S)),$$

so  $T(U) \subseteq \text{span}(T(S))$ . Thus,  $\text{span}(T(S)) = T(U)$ .  $\square$

**Definition.** Let  $V$  and  $W$  be vector spaces over  $F$ , and let  $T : V \rightarrow W$  be linear.

- The *null space*  $\mathcal{N}(T)$  of  $T$  is the set of vectors  $x \in V$  with  $T(x) = 0_W$ ; that is,

$$\mathcal{N}(T) = \{x \in V : T(x) = 0_W\}.$$

- The *range*  $\mathcal{R}(T)$  of  $T$  is the image of  $V$  under  $T$ ; that is,

$$\mathcal{R}(T) = \{T(x) : x \in V\}.$$

**Example.** Let  $D : \mathcal{P}(\mathbb{R}) \rightarrow \mathcal{P}(\mathbb{R})$  be the differential operator. Then

$$\mathcal{N}(D) = \{a_0 : a_0 \in \mathbb{R}\} \quad \text{and} \quad \mathcal{R}(D) = \mathcal{P}(\mathbb{R}).$$

**Theorem 2.3.** Let  $V$  and  $W$  be vector spaces over  $F$ , and let  $T : V \rightarrow W$  be linear. Then  $\mathcal{N}(T)$  and  $\mathcal{R}(T)$  are subspaces of  $V$  and  $W$ , respectively.

*Proof.*

- (a) Let  $x, x' \in \mathcal{N}(T)$  and  $a \in F$ . Then we have  $T(x+x') = T(x) + T(x') = 0_W + 0_W = 0_W$ ,  $T(ax) = aT(x) = a0_W = 0_W$  and  $T(0_V) = 0_W$ . Thus,  $\mathcal{N}(T)$  is a subspace of  $V$ .
- (b) Let  $y, y' \in \mathcal{R}(T)$  and  $a \in F$ . There exist  $x, x' \in V$  with  $y = T(x)$  and  $y' = T(x')$ . Then we have  $y + y' = T(x) + T(x') = T(x + x')$ ,  $ay = aT(x) = T(ax)$  and  $0_W = T(0_V)$ . Thus,  $\mathcal{R}(T)$  is a subspace of  $W$ .  $\square$

**Definition.** Let  $V$  and  $W$  be vector spaces over  $F$ , and let  $T : V \rightarrow W$  be linear.

- The *nullity* of  $T$ , denoted by  $\text{nullity}(T)$ , is the dimension of  $\mathcal{N}(T)$ .

- The *rank* of  $T$ , denoted by  $\text{rank}(T)$ , is the dimension of  $\mathcal{R}(T)$ .

**Theorem 2.4** (Rank-nullity Theorem). Let  $V$  and  $W$  be vector spaces over  $F$ , and let  $T : V \rightarrow W$  be linear. If  $V$  is finite-dimensional, then  $\text{nullity}(T) + \text{rank}(T) = \dim(V)$ .

*Proof.* Let  $S$  be a basis for  $V$  and  $Q$  a basis for  $\mathcal{N}(T)$ . By corollary to replacement theorem (Theorem 1.16), there is  $R \subseteq S \setminus Q$  such that  $Q \cup R$  is a basis for  $V$ . Since  $|R| = |Q \cup R| - |Q| = \dim(V) - \text{nullity}(T)$ , the theorem holds if  $|R| = \dim(\mathcal{R}(T))$ .

If there exist different  $x, x' \in R$  with  $T(x) = T(x')$ , then we have  $T(x - x') = T(x) - T(x') = 0_W$ , and thus  $x - x' \in \mathcal{N}(T) = \text{span}(Q)$ . It follows that  $x \in \text{span}(Q \cup \{x'\})$ , contradiction to the fact that  $S$  is linearly independent. Thus,  $|R| = |T(R)|$ . We claim that  $T(R)$  is a basis for  $\mathcal{R}(T)$ .

First we prove that  $T(R)$  spans  $\mathcal{R}(T)$ . By Theorem 2.2 (b) and the fact that  $T(Q) = \{0_W\}$ , we have

$$\begin{aligned} \mathcal{R}(T) &= T(\text{span}(Q \cup R)) \\ &= \text{span}(T(Q \cup R)) \\ &= \text{span}(T(Q)) + \text{span}(T(R)) \\ &= \text{span}(T(R)). \end{aligned}$$

Then we prove that  $T(R)$  is linearly independent. Suppose that

$$a_1 T(x_1) + \cdots + a_n T(x_n) = 0_W$$

holds for some  $a_1, \dots, a_n \in F$  and some different  $x_1, \dots, x_n \in R$  with  $n \geq 1$ . Then by Theorem 2.2 we have  $T(a_1 x_1 + \cdots + a_n x_n) = 0_W$ , and thus  $a_1 x_1 + \cdots + a_n x_n \in \mathcal{N}(T)$ . Hence, there exist some  $b_1, \dots, b_m \in F$  and some different  $y_1, \dots, y_m \in Q$  such that

$$a_1 x_1 + \cdots + a_n x_n = b_1 y_1 + \cdots + b_m y_m.$$

That is,

$$a_1 x_1 + \cdots + a_n x_n + (-b_1) y_1 + \cdots + (-b_m) y_m = 0_V.$$

Since  $Q \cup R$  is linearly independent, we have  $a_1 = \cdots = a_n = b_1 = \cdots = b_m = 0_F$ , implying that  $T(R)$  is linearly independent.

Thus,  $T(R)$  is a basis for  $\mathcal{R}(T)$ , and we can conclude that  $\text{rank}(T) = |T(R)| = |R| = |Q \cup R| - |Q|$ , which completes the proof.  $\square$