Abstract Real Algebraic Linear Algebra

D B M

Contents

A Few Words

These notes start from the beginning of vector spaces over **R**; very occasionally vector spaces over **Q** and **C** are discussed. They are written for readers that might not be familiar with proofs in the sense that all of the basic details are worked out from the beginning (I tried my best to include everything but quite honestly some things should be left out). I have dispensed with a preliminary review of sets, functions, and the basics of logic. Aspects of these topics are used sometimes without discussion (e.g. intersection/union of subsets and methods of proof like induction/proof by contradiction).

Chapter 1

Vector Spaces

The main objects of study in these notes are real vector spaces. We will start with the definition of a real vector space and establish several basic results from these definitions. Several examples of real vector spaces are given and help serve as a starting point for concrete examples of real vector spaces.

1.1 Definition and Basic Properties

For the definition of a vector space, we will make use of functions on product sets. Recall that the (**Cartesian**) **product** of two sets X,Y is the set $X \times Y$ of pairs (x,y) such that $x \in X$ and $y \in Y$. That is,

$$X \times Y \overset{\text{def}}{=} \{(x,y) \; : \; x \in X, \; y \in Y\}.$$

We also will use the rational numbers **Q**, the real numbers **R**, and the complex numbers **C**. We first give the definition of a real vector space.

Definition 1.1 (Real Vector Space)**.** A **real vector space** is a set V with a pair of functions

$$+ : V \times V \longrightarrow V, \quad \cdot : \mathbf{R} \times V \to V$$

satisfying the following conditions:

(a)

$$v + w = w + v$$

for all $v, w \in V$.

(b) There exists $v_0 \in V$ such that
$$v_0 + v = v$$
for all $v \in V$.

(c)
$$v + (w + u) = (v + w) + u$$
for all $v, w, u \in V$.

(d) For each $v \in V$, there exists $w_v \in V$ such that
$$v + w_v = v_0.$$

(e)
$$\alpha \cdot (\beta \cdot v) = (\alpha\beta) \cdot v$$
for all $v \in V$ and $\alpha, \beta \in \mathbf{R}$.

(f)
$$1 \cdot v = v$$
for all $v \in V$.

(g)
$$(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$$
for all $v \in V$ and $\alpha, \beta \in \mathbf{R}$.

(h)
$$\alpha \cdot (v + w) = \alpha \cdot v + \alpha \cdot w$$
for all $v, w \in V$ and $\alpha \in \mathbf{R}$.

The function
$$+ : V \times V \longrightarrow V$$
is called **vector addition** and the function
$$\cdot : \mathbf{R} \times V \to V$$
is called **scalar multiplication**. The vector $v_0 \in V$ is called the **zero vector** (“the” because of Lemma 1.2).

Remark 1. Real vector spaces generalize $\mathbf{R}^n$. The definition above axiomatizes $\mathbf{R}^n$ with vector addition and scalar multiplication but with a general set V in place of $\mathbf{R}^n$. Given a set V, a vector space structure on V is the existence of functions as above that satisfy all of the conditions above. Not every set V can be given the structure of a real vector space. The empty set for instance and any finite set with at least two elements. Another is $\mathbf{Z}$, the integers and $\mathbf{Q}$, the rationals. All of these sets cannot be given a real vector space structure because they are too small. Specifically, if V is a real vector space, then $|V| = 1$ or $|V| \geq |\mathbf{R}|$ with $|\cdot|$ denotes the cardinality of a set. As we will see much later, if V admits a real vector space structure, then any other vector space structure on V is the "same" in the sense that they are isomorphic (a concept defined later in these notes).

We will now establish some of the most basic results via Definition 1.1. The reader unfamiliar with vector spaces might opt to review the examples section that proceeds this one before pushing forward as the forthcoming material is abstract.

Lemma 1.2 (Uniqueness: Zero Vector)**.** *If V is a real vector space and $w \in V$ satisfies*

$$v + w = v$$

for all $v \in V$, then $w = v_0$.

Proof. By definition of v_0, we know that $v + v_0 = v$ for all $v \in V$. Taking $v = w$, we see that $w + v_0 = w$. By assumption, we know that $v + w = v$ for all $v \in V$. Taking $v = v_0$, we see that $v_0 + w = v_0$. Finally, as vector addition is commutative, we have

$$w = w + v_0 = v_0 + w = v_0.$$

♠

We will denote the zero vector in V by 0_V as to not confuse it with the real number or scalar $0 \in \mathbf{R}$.

Lemma 1.3. *If V is a vector space and $v + w = v$ for some $v \in V$, then $w = 0_V$.*

Proof. If $v + w = v$, then

$$\begin{aligned} v + w &= v \\ (v + w) + w_v &= v + w_v \\ (v + w_v) + w &= 0_V \\ 0_V + w &= 0_V \\ w &= 0_V. \end{aligned}$$

♠

Lemma 1.4 (Uniqueness: Additive Inverses)**.** *If V is a real vector space with $v, w \in V$ such that $v + w = 0_V$, then $w = w_v$.*

Proof. By definition of w_v, we know that $w_v + v = 0_V$. Adding w_v to both sides of the equation

$$w + v = 0_V,$$

we obtain

$$(w + v) + w_v = 0_V + w_v. \tag{1.1}$$

Working with the left hand side of Equation 1.1, we have

$$(w + v) + w_v = w + (v + w_v) = w + 0_V = w.$$

Working with the right hand side of Equation 1.1, we know that

$$0_V + w_v = w_v$$

by definition of 0_V. Hence by Equation 1.1, we have $w = w_v$. ♠

We call w_v the **additive inverse** of v.

Lemma 1.5. *If V is a vector space, then $w_{0_V} = 0_V$.*

Proof. By definition of additive inverses, we have

$$0_V + w_{0_V} = 0_V.$$

Hence $w_{0_V} = 0_V$ by Lemma 1.3. ♠

Lemma 1.5 says that the additive inverse of the zero vector is the zero vector.

Lemma 1.6 (Zero Scales Like Zero Should)**.** *If V is a vector space and $v \in V$, then $0 \cdot v = 0_V$.*

Proof. By definition of the zero vector, we have

$$0 \cdot v + 0_V = 0 \cdot v.$$

Therefore,

$$0 \cdot v + 0_V = 0 \cdot v + (0 \cdot v + w_{0 \cdot v}) = (0 \cdot v + 0 \cdot v) + w_{0 \cdot v} = (0+0) \cdot v + w_{0 \cdot v} = 0 \cdot v + w_{0 \cdot v}.$$

Hence

$$0 \cdot v = 0 \cdot v + w_{0 \cdot v},$$

and so $w_{0 \cdot v} = 0_V$ by Lemma 1.3. Thus, by Lemma 1.5, we have $0 \cdot v = 0_V$. ♠

Lemma 1.7 (-1 scales like -1 should)**.** *If V is a vector space and $v \in V$, then $w_v = (-1) \cdot v$.*

Proof. Note that

$$(1 + (-1)) \cdot v = 1 \cdot v + (-1) \cdot v = v + (-1) \cdot v$$

and that

$$(1 + (-1)) \cdot v = 0 \cdot v = 0_V$$

by Lemma 1.6. Hence

$$v + (-1) \cdot v = 0_V,$$

and so $(-1) \cdot v = w_v$. ♠

For notational simplicity, we denote the additive inverse of v simply by $-v$ in all that follows.

Lemma 1.8 (The Zero Vector is Unmovable)**.** *If V is a vector space, then $\alpha \cdot 0_V = 0_V$ for all $\alpha \in \mathbf{R}$.*

Proof. The validity of the claim follows from the string of equalities below:

$$\begin{aligned}
\alpha \cdot 0_V + 0_V &= \alpha \cdot 0_V \\
\alpha \cdot 0_V + (\alpha \cdot 0_V + (-(\alpha \cdot 0_V))) &= \alpha \cdot 0_V \\
-(\alpha \cdot 0_V) + (\alpha \cdot 0_V + \alpha \cdot 0_V) &= \alpha \cdot 0_V \\
-(\alpha \cdot 0_V) + \alpha(0_V + 0_V) &= \alpha \cdot 0_V \\
-(\alpha \cdot 0_V) + \alpha \cdot 0_V &= \alpha \cdot 0_V \\
0_V &= \alpha \cdot 0_V.
\end{aligned}$$

♠

Lemma 1.9 (You are zero only when you should be zero)**.** *If V is a vector space, $v \in V$, and $\alpha \in \mathbf{R}$, then $\alpha \cdot v = 0_V$ if and only if $\alpha = 0$ or $v = 0_V$.*

Proof. For the reverse implication, if $\alpha = 0$ or $v = 0_V$, then $\alpha \cdot v = 0_V$ by Lemma 1.6 and Lemma 1.7. For the direct implication, if $\alpha \cdot v = 0_V$, then we must show that $\alpha = 0$ or $v = 0_V$. If $\alpha = 0$, then we are trivially done. In the event $\alpha \neq 0$, we see that

$$v = 1 \cdot v = \left(\frac{\alpha}{\alpha}\right) \cdot v = \left(\frac{1}{\alpha}\right) \cdot (\alpha \cdot v) = \left(\frac{1}{\alpha}\right) \cdot 0_V = 0_V,$$

as needed. ♠

We now define complex and rational vector spaces. Note that we use our notation 0_V and $-v$ instead of v_0 and w_v from our real definition as it is both more traditional and notationally simpler.

Definition 1.10 (Complex Vector Space)**.** A **complex vector space** is a set V with a pair of functions

$$+ : V \times V \longrightarrow V, \quad \cdot : \mathbf{C} \times V \to V$$

satisfying the following conditions:

(a)
$$v + w = w + v$$
for all $v, w \in V$.

(b) There exists $0_V \in V$ such that
$$0_V + v = v$$
for all $v \in V$.

(c)
$$v + (w + u) = (v + w) + u$$
for all $v, w, u \in V$.

(d) For each $v \in V$, there exists $-v \in V$ such that
$$v + (-v) = 0_V.$$

(e)
$$\alpha \cdot (\beta \cdot v) = (\alpha\beta) \cdot v$$
for all $v \in V$ and $\alpha, \beta \in \mathbf{C}$.

(f)
$$1 \cdot v = v$$
for all $v \in V$.

(g)
$$(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$$
for all $v \in V$ and $\alpha, \beta \in \mathbf{C}$.

(h)
$$\alpha \cdot (v + w) = \alpha \cdot v + \alpha \cdot w$$
for all $v, w \in V$ and $\alpha \in \mathbf{C}$.

Definition 1.11 (Rational Vector Space)**.** A **rational vector space** is a set V with a pair of functions
$$+ : V \times V \longrightarrow V, \quad \cdot : \mathbf{Q} \times V \to V$$
satisfying the following conditions:

(a)
$$v + w = w + v$$
for all $v, w \in V$.

(b) There exists $0_V \in V$ such that
$$0_V + v = v$$
for all $v \in V$.

(c)
$$v + (w + u) = (v + w) + u$$
for all $v, w, u \in V$.

(d) For each $v \in V$, there exists $-v \in V$ such that
$$v + (-v) = 0_V.$$

(e)
$$\alpha \cdot (\beta \cdot v) = (\alpha\beta) \cdot v$$
for all $v \in V$ and $\alpha, \beta \in \mathbf{Q}$.

(f)
$$1 \cdot v = v$$
for all $v \in V$.

(g)
$$(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$$
for all $v \in V$ and $\alpha, \beta \in \mathbf{Q}$.

(h)
$$\alpha \cdot (v + w) = \alpha \cdot v + \alpha \cdot w$$
for all $v, w \in V$ and $\alpha \in \mathbf{Q}$.

1.2 Examples

1.2.1 The Standard Ones

In this subsection, we give some standard examples of vector spaces. We start with the most basic example, the real line.

The Real Line

Let $V = \mathbf{R}$ and let
$$+ : \mathbf{R} \times \mathbf{R} \longrightarrow \mathbf{R}, \qquad \cdot : \mathbf{R} \times \mathbf{R} \longrightarrow \mathbf{R}$$
be given by addition and multiplication operations on the real numbers. Note that when we write the multiplication of $a, b \in \mathbf{R}$, we will write ab and not $a \cdot b$; this is because of the unique role $\mathbf{R}$ plays in this note as the field of scalars. The zero vector is the real number 0 in this case. We write down the vector conditions in this case where $a, b, c \in \mathbf{R}$:

- $a + b = b + a$.
- $a + 0 = a$.
- $a + (b + c) = (a + b) + c$.

- $a+(-a)=0$.
- $a(bc)=(ab)c$.
- $1a=a$.
- $(a+b)c=ac+bc$.
- $a(b+c)=ab+ac$.

One notes that these are all properties of the real numbers that were learned years ago. Commutativity of addition and multiplication. Associativity of addition and multiplication. The two distributive laws. Zero is zero and one is one.

Euclidean n–space

We define $\mathbf{R}^n$ to be

$$\mathbf{R}^n \stackrel{\text{def}}{=} \underbrace{\mathbf{R}\times\mathbf{R}\times\cdots\times\mathbf{R}}_{n\text{ times}} = \{(x_1,\dots,x_n)\ :\ x_1,\dots,x_n\in\mathbf{R}\}.$$

We define

$$+\ :\ \mathbf{R}^n\times\mathbf{R}^n\longrightarrow\mathbf{R}^n, \qquad \cdot\ :\ \mathbf{R}\times\mathbf{R}^n\longrightarrow\mathbf{R}^n$$

by

$$(x_1,x_2,\dots,x_n)+(y_1,y_2,\dots,y_n)\stackrel{\text{def}}{=}(x_1+y_1,x_2+y_2,\dots,x_n+y_n)$$

and

$$\alpha\cdot(x_1,\dots,x_n)\stackrel{\text{def}}{=}(\alpha x_1,\alpha x_2,\dots,\alpha x_n).$$

We call these operations, **coordinate-wise addition** and **coordinate-wise scalar multiplication** as the operations are done on the level of the fixed coordinate. We will see this type of "independence" behavior throughout these notes. The zero vector is given by $0_{\mathbf{R}^n}=(0,0,\dots,0)$. We will verify that $\mathbf{R}^n$ with these operations satisfies the conditions needed to be a real vector space. This verification is formal, relying only on the properties of real numbers and the definition of the operators $+,\cdot$. We will verify each property to illustrate the nature of such a check.

Commutativity. For $x,y\in\mathbf{R}^n$ with $x=(x_1,\dots,x_n)$, $y=(y_1,\dots,y_n)$, we have

$$\begin{aligned}x+y&=(x_1,\dots,x_n)+(y_1,\dots,y_n)=(x_1+y_1,\dots,x_n+y_n)\\&=(y_1+x_1,\dots,y_n+x_n)=(y_1,\dots,y_n)+(x_1,\dots,x_n)\\&=y+x.\end{aligned}$$

Zero Vector. For $x \in \mathbf{R}^n$, we have

$$x + 0_{\mathbf{R}^n} = (x_1, \dots, x_n) + (0, \dots, 0) = (x_1 + 0, \dots, x_n + 0) = (x_1, \dots, x_n) = x.$$

Associativity (additive). For $x, y, z \in \mathbf{R}^n$, we have

$$\begin{aligned} x + (y + z) &= (x_1, \dots, x_n) + ((y_1, \dots, y_n) + (z_1, \dots, z_n)) \\ &= (x_1, \dots, x_n) + (y_1 + z_1, \dots, y_n + z_n) \\ &= (x_1 + (y_1 + z_1), \dots, x_n + (y_n + z_n)) \\ &= ((x_1 + y_1) + z_1, \dots, (x_n + y_n) + z_n) \\ &= (x_1 + y_1, \dots, x_n + y_n) + (z_1, \dots, z_n) \\ &= ((x_1, \dots, x_n) + (y_1, \dots, y_n)) + (z_1, \dots, z_n) \\ &= (x + y) + z. \end{aligned}$$

Additive Inverse. For $x \in \mathbf{R}^n$, we have

$$x + (-x) = (x_1, \dots, x_n) + (-x_1, \dots, x_n) = (x_1 - x_1, \dots, x_n - x_n) = (0, \dots, 0) = 0_{\mathbf{R}^n}.$$

Associativity (scalar). For $x \in \mathbf{R}^n$ and $\alpha, \beta \in \mathbf{R}$, we have

$$\begin{aligned} \alpha \cdot (\beta \cdot x) &= \alpha \cdot (\beta \cdot (x_1, \dots, x_n)) \\ &= \alpha \cdot (\beta x_1, \dots, \beta x_n) \\ &= (\alpha(\beta x_1), \dots, \alpha(\beta x_n)) \\ &= ((\alpha\beta) x_1, \dots, (\alpha\beta) x_n) \\ &= (\alpha\beta) \cdot (x_1, \dots, x_n) = (\alpha\beta) \cdot x. \end{aligned}$$

1 acts like 1. For $x \in \mathbf{R}^n$, we have

$$1 \cdot x = 1 \cdot (x_1, \dots, x_n) = (1x_1, \dots, 1x_n) = (x_1, \dots, x_n) = x.$$

Distributive 1. For $\alpha, \beta \in \mathbf{R}$ and $x \in \mathbf{R}^n$, we have

$$\begin{aligned} (\alpha + \beta) \cdot x &= (\alpha + \beta) \cdot (x_1, \dots, x_n) \\ &= ((\alpha + \beta) x_1, \dots, (\alpha + \beta) x_n) \\ &= (\alpha x_1 + \beta x_1, \dots, \alpha x_n + \beta x_n) \\ &= (\alpha x_1, \dots, \alpha x_n) + (\beta x_1, \dots, \beta x_n) \\ &= \alpha \cdot (x_1 \dots, x_n) + \beta \cdot (x_1 \dots, x_n) \\ &= \alpha \cdot x + \beta \cdot x. \end{aligned}$$

Distributive 2. For $\alpha \in \mathbf{R}$ and $x, y \in \mathbf{R}^n$, we have

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

Spaces Arising From Matrices

Let $m, n \in \mathbf{N}$. We define an m by n real matrix A to be an m by n array of real numbers $(a_{j,k})$ with $j \in \{1, \ldots, m\}$ and $k \in \{1, \ldots, m\}$. Specifically,

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \ldots & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \ldots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & a_{m,3} & \ldots & a_{m,n} \end{pmatrix}.$$

We denote the set of m by n real matrices by $\mathrm{M}(m, n, \mathbf{R})$. Given $A, B \in \mathrm{M}(m, n, \mathbf{R})$, we define

$$\begin{aligned}
A + B &= \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \ldots & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \ldots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & a_{m,3} & \ldots & a_{m,n} \end{pmatrix} + \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} & \ldots & b_{1,n} \\ b_{2,1} & b_{2,2} & b_{2,3} & \ldots & b_{2,n} \\ b_{3,1} & b_{3,2} & b_{3,3} & \ldots & b_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & b_{m,3} & \ldots & b_{m,n} \end{pmatrix} \\
&= \begin{pmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} & a_{1,3} + b_{1,3} & \ldots & a_{1,n} + b_{1,n} \\ a_{2,1} + b_{2,1} & a_{2,2} + b_{2,2} & a_{2,3} + b_{2,3} & \ldots & a_{2,n} + b_{2,n} \\ a_{3,1} + b_{3,1} & a_{3,2} + b_{3,2} & a_{3,3} + b_{3,3} & \ldots & a_{3,n} + b_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} + b_{m,1} & a_{m,2} + b_{m,2} & a_{m,3} + b_{m,3} & \ldots & a_{m,n} + b_{m,n} \end{pmatrix}
\end{aligned}$$

We define

$$\alpha \cdot A \stackrel{\text{def}}{=} \alpha \cdot \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \dots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & a_{m,3} & \dots & a_{m,n} \end{pmatrix} = \begin{pmatrix} \alpha a_{1,1} & \alpha a_{1,2} & \alpha a_{1,3} & \dots & \alpha a_{1,n} \\ \alpha a_{2,1} & \alpha a_{2,2} & \alpha a_{2,3} & \dots & \alpha a_{2,n} \\ \alpha a_{3,1} & \alpha a_{3,2} & \alpha a_{3,3} & \dots & \alpha a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha a_{m,1} & \alpha a_{m,2} & \alpha a_{m,3} & \dots & \alpha a_{m,n} \end{pmatrix}.$$

The zero vector is given by the zero matrix $0_{m,n}$ defined by

$$0_{m,n} \stackrel{\text{def}}{=} \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

It is straightforward verify that $\mathrm{M}(m,n,\mathbf{R})$ is a real vector space. We list the needed properties below:

- $A+B = B+A$ for all $A,B \in \mathrm{M}(m,n,\mathbf{R})$.
- $A+0_{m,n} = A$ for all $A \in \mathrm{M}(m,n,\mathbf{R})$.
- $A+(B+C) = (A+B)+C$ for all $A,B,C \in \mathrm{M}(m,n,\mathbf{R})$.
- $A+(-A) = 0_{m,n}$ for all $A \in \mathrm{M}(m,n,\mathbf{R})$.
- $(\alpha\beta)\cdot A = \alpha\cdot(\beta\cdot A)$ for all $\alpha,\beta \in \mathbf{R}$ and $A \in \mathrm{M}(m,n,\mathbf{R})$.
- $1\cdot A = A$ for all $A \in \mathrm{M}(m,n,\mathbf{R})$.
- $(\alpha+\beta)\cdot A = \alpha\cdot A+\beta\cdot A$ for all $\alpha,\beta \in \mathbf{R}$ and $A \in \mathrm{M}(m,n,\mathbf{R})$.
- $\alpha\cdot(A+B) = \alpha\cdot A+\alpha\cdot B$ for all $\alpha \in \mathbf{R}$ and $A,B \in \mathrm{M}(m,n,\mathbf{C})$.

Exercise 1. Prove that $\mathrm{M}(m,n;\mathbf{R})$ is a vector space.

Complex Vector Spaces

Recall that the complex numbers can be represented as $x+iy$ where $x,y \in \mathbf{R}$ and $i^2 = -1$. Addition and multiplication are defined as

$$z+w = (z_1+iz_2)+(w_1+iw_2) \overset{\text{def}}{=} (z_1+w_1)+(z_2+w_2)i$$

and

$$zw = (z_1+iz_2)(w_1+iw_2) \overset{\text{def}}{=} (z_1w_1-z_2w_2)+(z_1w_2+z_2w_1)i$$

where $z = z_1+iz_2$ and $w = w_1+iw_2$. Zero is given by $0_{\mathbf{C}} = 0+i0$.

Exercise 2. Prove that $\mathbf{C}$ is a real vector space.

We can also define $\mathbf{C}^n = \underbrace{\mathbf{C}\times\mathbf{C}\times\cdots\times\cdots\times\mathbf{C}}_{n \text{ times}}$ with addition and scalar multiplication defined by

$$(z_1,\ldots,z_n)+(w_1,\ldots,w_n) \overset{\text{def}}{=} (z_1+w_1,\ldots,z_n+w_n), \quad \alpha(z_1,\ldots,z_n) \overset{\text{def}}{=} (\alpha z_1,\ldots,\alpha z_n)$$

where $z_1,\ldots,z_n,w_1,\ldots,w_n,\alpha \in \mathbf{C}$. With these operations, $\mathbf{C}^n$ is a complex vector space. We can also define $\mathrm{M}(m,n,\mathbf{C})$ as in the real case and give it also the structure of a complex vector space.

The complex numbers are also a real vector space. However, the real numbers are not a complex vector space with the usual operations. The problem is that the real numbers are not closed under complex scaling.

Rational Vector Spaces

The rational numbers $\mathbf{Q}$ with the usual addition and multiplication operations is a rational vector space. We can also define $\mathbf{Q}^n$ and $\mathrm{M}(m,n,\mathbf{Q})$, and these sets with the operations defined as before become rational vector spaces as well. As before, $\mathbf{R}$ and $\mathbf{C}$ are also rational vector spaces but $\mathbf{Q}$ is not a real or complex vector space with the usual operations because it is not closed under real or complex scaling.

1.2.2 Spaces Arising From Functions

In this section, we will describe additional, more abstract examples of vector spaces coming from functions. The example $\mathrm{Fun}(X,\mathbf{R})$ will play an important role throughout these notes.

Set Theoretic Constructions

Given a set X, we define $\mathrm{Fun}(X,\mathbf{R})$ to be the set of function $f\colon X\to\mathbf{R}$. We define two operations on $\mathrm{Fun}(X,\mathbf{R})$:

$$+\ :\ \mathrm{Fun}(X,\mathbf{R})\times\mathrm{Fun}(X,\mathbf{R})\longrightarrow\mathrm{Fun}(X,\mathbf{R}),\quad \cdot : \mathbf{R}\times\mathrm{Fun}(X,\mathbf{R})\longrightarrow\mathrm{Fun}(X,\mathbf{R})$$

by

$$(f+g)(x)\stackrel{\text{def}}{=}f(x)+g(x),\quad (\alpha\cdot f)(x)\stackrel{\text{def}}{=}\alpha f(x).$$

The zero vector in $\mathrm{Fun}(X,\mathbf{R})$ is the constant function $0_{\mathrm{Fun}(X,\mathbf{R})}\colon X\to\mathbf{R}$ given by

$$0_{\mathrm{Fun}(X,\mathbf{R})}(x)\stackrel{\text{def}}{=}0.$$

It is straightforward to verify that $\mathrm{Fun}(X,\mathbf{R})$ is a real vector space with these operations. If we replace $\mathbf{R}$ with $\mathbf{C}$ or $\mathbf{Q}$, then the spaces $\mathrm{Fun}(X,\mathbf{C})$ and $\mathrm{Fun}(X,\mathbf{Q})$ are complex and rational vector spaces.

Exercise 3. Prove $\mathrm{Fun}(X,\mathbf{R})$ is a real vector space with the above operations.

This construction can be generalized. Given any real vector space V, the set of functions $\mathrm{Fun}(X,V)$ can be made into a real vector space by defining the operations

$$+\ :\ \mathrm{Fun}(X,V)\times\mathrm{Fun}(X,V)\longrightarrow\mathrm{Fun}(X,V),\quad \cdot : \mathbf{R}\times\mathrm{Fun}(X,V)\longrightarrow\mathrm{Fun}(X,V)$$

by

$$(f+g)(x)=f(x)+g(x),\quad (\alpha\cdot f)(x)=\alpha\cdot f(x).$$

Note that on the right hand side of the above equations, the operations are the vector addition and scalar multiplication operations on V. The zero vector is the constant function $0_{\mathrm{Fun}(X,V)}\colon X\to V$ defined by

$$0_{\mathrm{Fun}(X,V)}(x)\stackrel{\text{def}}{=}0_V.$$

The space $\mathrm{Fun}(X,V)$ is a real vector space with the above operation. **Remember,** Fun **is fun**.

Exercise 4. Prove $\mathrm{Fun}(X,V)$ is a real vector space with the above operations.

Spaces Arising From Calculus

Let $\mathrm{C}^0([0,1])$ denote the continuous functions $f\colon[0,1]\to\mathbf{R}$ where $[0,1]$ is the closed interval $[0,1]$ defined by

$$[0,1]\stackrel{\text{def}}{=}\{x\in\mathbf{R}\ :\ 0\le x\le 1\}.$$

We define addition and scalar multiplication as in the case of $\mathrm{Fun}(X,\mathbf{R})$ where $X = [0,1]$. Namely, if $f,g \in \mathrm{C}^0([0,1])$ and $\alpha \in \mathbf{R}$, we define

$$(f+g)(x) = f(x)+g(x), \quad (\alpha \cdot f)(x) = \alpha f(x).$$

Since $f+g$ is continuous when f,g are continuous, we see that $f+g \in \mathrm{C}^0([0,1])$. Likewise, αf is continuous when f is continuous, and so $\alpha f \in \mathrm{C}^0([0,1])$. We also see that $0_{\mathrm{Fun}([0,1],\mathbf{R})} \in \mathrm{C}^0([0,1])$ (i.e. the constant zero function is continuous). We will see when we discuss vector subspaces that this implies that $\mathrm{C}^0([0,1])$ is a real vector space with these operations. Indeed, it is a vector subspace of $\mathrm{Fun}([0,1],\mathbf{R})$.

We can also take instead the set of differentiable functions on $[0,1]$ or the set of Riemann integrable functions on $[0,1]$. These sets with the same operations defined above are real vector spaces.

Spaces Arising From Sequences

Recall that a sequence in $\mathbf{R}$ is a function $f\colon \mathbf{N} \to \mathbf{R}$ where we write $f(j) \stackrel{\text{def}}{=} x_j$. The set of sequence in $\mathbf{R}$ is just the set $\mathrm{Fun}(\mathbf{N},\mathbf{R})$, which we saw is a real vector space. We define $\ell^1(\mathbf{R})$ to be the set of sequences $\{x_j\}$ such that

$$\sum_{j=1}^{\infty} |x_j|$$

is convergent. Under the operations of addition and scalar multiplication given in $\mathrm{Fun}(\mathbf{N},\mathbf{R})$, one can check that $\ell^1(\mathbf{R})$ is a real vector space. Given $p > 0$, we could also consider the sequences $\{x_n\}$ such that

$$\sum_{j=1}^{\infty} |x_n|^p$$

is convergent. This subset of $\mathrm{Fun}(\mathbf{N},\mathbf{R})$ is also a vector space and is denoted by $\ell^p(\mathbf{R})$. We can also consider the subset of bounded sequence, which is also a vector space, as well as the set of essentially bounded sequences (i.e. $\limsup_j |x_j| < \infty$).

Spaces Arising From Polynomials and Beyond

Recall that a **polynomial of degree** d is a function $P\colon \mathbf{R} \to \mathbf{R}$ given by

$$P(x) = \sum_{j=0}^{d} \alpha_j x^j$$

where $\alpha_0, \dots, \alpha_d \in \mathbf{R}$ and $\alpha_d \neq 0$. We denote the degree of the polynomial by $\deg(P)$. We define

$$\mathrm{Poly}_d(\mathbf{R}) \stackrel{\text{def}}{=} \{P\colon \mathbf{R} \to \mathbf{R} \ :\ P \text{ is a polynomial with } \deg(P) \leq d\}$$

and

$$\mathrm{Poly}(\mathbf{R}) \stackrel{\text{def}}{=} \{P\colon \mathbf{R} \to \mathbf{R} \ :\ P \text{ is a polynomial}\}.$$

Both $\mathrm{Poly}_d(\mathbf{R})$ and $\mathrm{Poly}(\mathbf{R})$ are vector spaces.

Exercise 5. Prove $\mathrm{Poly}_d(\mathbf{R})$ is a vector space.

Exercise 6. Prove that $\mathrm{Poly}(\mathbf{R})$ is a vector space.

We can also consider $\mathbf{R}[[x]]$ the set of **real formal power series in the variable** x. A formal power series is given by

$$\sum_{j=0}^{\infty} \alpha_j x^j$$

where $\alpha_j \in \mathbf{R}$. The set of formal power series is also a vector space under the (obvious) addition and scalar multiplication. That is

$$\sum_{j=0}^{\infty} \alpha_j x^j + \sum_{j=0}^{\infty} \beta_j x^j \stackrel{\text{def}}{=} \sum_{j=0}^{\infty} (\alpha_j + \beta_j) x^j$$

and

$$\alpha \cdot \left(\sum_{j=0}^{\infty} \alpha_j x^j \right) \stackrel{\text{def}}{=} \sum_{j=0}^{\infty} \alpha \alpha_j x^j.$$

Exercise 7. Prove $\mathbf{R}[[x]]$ is a vector space.

We could also consider the set of formal Laurent series $\mathbf{R}(x)$. A formal Laurent series is given by

$$\sum_{j=-\infty}^{\infty} \alpha_j x^j$$

where $\alpha_j \in \mathbf{R}$.

Exercise 8. Prove $\mathbf{R}(x)$ is a vector space.

1.2.3 Spaces Arising From (Systems) of Equations

Given a matrix $A \in \mathrm{M}(m,n,\mathbf{R})$ with $A = (a_{j,k})$ where $j \in \{1,\dots,m\}$ and $k \in \{1,\dots,n\}$, we can form m equations with n variables

$$\mathrm{E}_j(A): \quad \sum_{\ell=1}^{n} a_{j,\ell} x_\ell = 0.$$

We view $x = (x_1,\dots,x_n)$ as a variable vector in $\mathbf{R}^n$. When $n = m = 2$,

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix},$$

and $x = (x_1, x_2)$, we have the two equations:

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 &= 0 \\ a_{2,1}x_1 + a_{2,2}x_2 &= 0. \end{aligned}$$

When we have a general m,n, expanding the sum, we see the equations are:

$$\begin{array}{rr} \mathrm{E}_1(A): & a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + \cdots + a_{1,n}x_n = 0 \\ \mathrm{E}_2(A): & a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 + \cdots + a_{2,n}x_n = 0 \\ \mathrm{E}_3(A): & a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 + \cdots + a_{3,n}x_n = 0 \\ \vdots & \vdots \qquad\quad \vdots \qquad\quad \vdots \quad = \vdots \\ \mathrm{E}_m(A): & a_{m,1}x_1 + a_{m,2}x_2 + a_{m,3}x_3 + \cdots + a_{m,n}x_n = 0. \end{array}$$

This system of equations is a homogenous system of linear equations. If we denote the set of $x \in \mathbf{R}^n$ such that $x = (x_1,\dots,x_n)$ is a solution to each of the equations $\mathrm{E}_1(A),\dots,\mathrm{E}_n(A)$ by $\mathscr{S}(A)$, then $\mathscr{S}(A)$ is a vector space with addition and scalar multiplication given by that in $\mathbf{R}^n$. To see this, we simply need to check that if $x, y \in \mathbf{R}^n$ are solutions to the equations $\mathrm{E}_1(A),\dots,\mathrm{E}_m(A)$, then $x + y$ is also is a solution to the equations $\mathrm{E}_1(A),\dots,\mathrm{E}_m(A)$. Likewise, we need to show that αx and $0_{\mathbf{R}^n}$ are solutions to the equations $\mathrm{E}_1(A),\dots,\mathrm{E}_m(A)$. For concreteness, we discuss the case when $m = n = 2$. In this case, if $x = (x_1, x_2)$ and $y = (y_1, y_2)$ are solutions to $\mathrm{E}_1(A)$ and $\mathrm{E}_2(A)$, then

$$\begin{aligned} a_{1,1}(x_1+y_1) + a_{1,2}(x_2+y_2) &= (a_{1,1}x_1 + a_{1,2}x_2) + (a_{1,1}y_1 + a_{1,2}y_2) = 0 + 0 = 0 \\ a_{2,1}(x_1+y_1) + a_{2,2}(x_2+y_2) &= (a_{2,1}x_1 + a_{2,2}x_2) + (a_{2,1}y_1 + a_{2,2}y_2) = 0 + 0 = 0. \end{aligned}$$

Likewise

$$a_{1,1}(\alpha x_1) + a_{1,2}(\alpha x_2) = \alpha(a_{1,1}x_1 + a_{1,2}x_2) = \alpha(0) = 0$$
$$a_{2,1}(\alpha x_2) + a_{2,2}(\alpha x_2) = \alpha(a_{2,1}x_1 + a_{2,2}x_2) = \alpha(0) = 0.$$

It is worth noting that if we change the right hand side of the above equations, the space of solutions is not a vector space. We made (essential) use of the fact that our equations were homogenous. The solutions spaces of non-homogenous systems of equations as above are always affine spaces (see Corollary 1.72) which is also a very nice property to have with regard to describing the solution space.

1.3 Subspaces

In this section, we introduce the concept of a vector subspace formally. Many of the examples given in the previous section arise (naturally) as vector subspace.

1.3.1 Vector Subspaces

Our focus will begin with the primary object, vector subspaces, of this section. We will briefly discuss a generalization of vector subspaces later called affine subspaces.

Definition 1.12 (Closed Under Vector Addition)**.** Given a vector space V and subset $S \subset V$, we say that S is **closed under vector addition** if given $v, w \in S$, then $v + w \in S$.

Definition 1.13 (Closed Under Scalar Multiplication)**.** Given a vector space V and subset $S \subset V$, we say that S is **closed under scalar multiplication** if given $v \in S$ and $\alpha \in \mathbf{R}$, then $\alpha \cdot v \in S$.

Definition 1.14 (Closed Under Linear Combinations)**.** Given a vector space V and subset $S \subset V$, we say that S is **closed under linear combinations** if given $v, w \in S$ and $\alpha, \beta \in \mathbf{R}$, then $\alpha v + \beta w \in S$.

Definition 1.15 (Vector Subspace)**.** A subset $S \subset V$ of a space is a **vector subspace** if S satisfies the following three conditions:

(1) $S \neq \emptyset$.

(2) S is closed under vector addition.

(3) S is closed under scalar multiplication

When $S \subset V$ is a vector subspace, we will write $S \leq V$.

Remark 2. If $S \leq V$ is a vector subspace, then restricting the functions

$$+ : V \times V \longrightarrow V, \quad \cdot : \mathbf{R} \times V \longrightarrow V$$

to $S \times S \subset V \times V$ and $\mathbf{R} \times S \subset \mathbf{R} \times V$, we see that the resulting functions satisfy

$$+ : S \times S \longrightarrow S, \quad \cdot : \mathbf{R} \times S \longrightarrow S.$$

If S is not closed under vector addition, then the codomain (i.e $\star$ is the codomain for $\rightarrow \star$) must be strictly larger than S. Likewise if S is not closed under scalar multiplication, the codomain for scalar multiplication will be strictly bigger than S. Under these operations when S is a vector subspace, we see that S is also a real vector space. Our next lemma will show that $0_S = 0_V$.

Lemma 1.16. *If $S \leq V$, then $0_V \in S$.*

Proof. As S is a vector subspace, we know that there is some $v \in S$ since $S \neq \emptyset$. As S is closed under scaling, we know that $-v \in S$. Finally, as S is closed under vector addition, we see that $v + (-v) = 0_V \in S$ as desired. ♠

Example 1. Every vector space has a vector subspace called the **trivial subspace** which is just the singleton set $\{0_V\}$. It follows from properties of 0_V that $\{0_V\}$ is closed under addition and scaling. If $V = \{0_V\}$, then this is the only subspace. If $V \neq \{0_V\}$, then the space V is also a vector subspace of V. Every vector subspace S of V sits between these two subspaces. That is, $\{0_V\} \leq S \leq V$ by Lemma 1.16.

Example 2. If $V = \mathbf{R}^n$ and $1 \leq m < n$, we define

$$W_m \stackrel{\text{def}}{=} \left\{ (x_1, \ldots, x_n) \ : \ x_j = 0 \text{ for all } j > m \right\}.$$

Then W_m is a vector subspace of $\mathbf{R}^n$. We see $0_{\mathbf{R}^n} \in W_m$. That is is closed under scalar multiplication and vector addition amount to $\alpha 0 = 0$ and $0 + 0 = 0$ coordinate-wise.

Exercise 9. Prove W_m is a vector subspace.

Example 3. In $\mathbf{R}^2$, we can take the subset

$$S \stackrel{\text{def}}{=} \left\{(x,y) \in \mathbf{R}^2 \;:\; x - y = 0\right\}.$$

To prove that S is a subspace, notice that if $(x,y) \in S$, then $x = y$. Hence $S = \left\{(x,x) \in \mathbf{R}^2 \;:\; x \in \mathbf{R}\right\}$. We see $(0,0) = 0_{\mathbf{R}^2} \in S$ and so $S \neq \emptyset$. Also, $(x,x) + (y,y) = (x+y, x+y) \in S$ and $\alpha \cdot (x,x) = (\alpha x, \alpha x) \in S$.

Example 4. If $\mathrm{M}(n,\mathbf{R}) \stackrel{\text{def}}{=} \mathrm{M}(n,n,\mathbf{R})$. We define ${}^T : \mathrm{M}(n,\mathbf{R}) \to \mathrm{M}(n,\mathbf{R})$ by $A^T \stackrel{\text{def}}{=} (a_{k,j})$ if $A = (a_{j,k})$ (this is called the **transpose**). For instance, if

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

then

$$A^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

We say $A \in \mathrm{M}(n,\mathbf{R})$ is **symmetric** if $A^T = A$. The subset S of $\mathrm{M}(n,\mathbf{R})$ of symmetric matrices is a vector subspace of $\mathrm{M}(n,\mathbf{R})$. For this, one simply needs to check that S is non-empty (it contains the zero matrix) and

$$(\alpha \cdot A)^T = \alpha \cdot A^T, \quad (A+B)^T = A^T + B^T$$

hold for all $A, B \in \mathrm{M}(n,\mathbf{R})$ and $\alpha \in \mathbf{R}$. We leave these tasks to the reader.

Exercise 10. Prove the subset of symmetric n by n real matrices is a vector subspace.

Example 5. The set of $A \in \mathrm{M}(n,\mathbf{R})$ such that $A^T = -A$ is also a vector subspace of $\mathrm{M}(n,\mathbf{R})$. We call $A \in \mathrm{M}(n,\mathbf{R})$ **skew-symmetric** when $A^T = -A$. Every $A \in \mathrm{M}(n,\mathbf{R})$ can be expressed uniquely as

$$A = A_+ + A_-$$

where A_+ is symmetric and A_- is skew-symmetric. For this, take

$$A = \frac{1}{2}\left(A + A^T\right) + \frac{1}{2}\left(A - A^T\right).$$

We see that

$$A_+ = \frac{1}{2}\left(A + A^T\right), \quad A_- = \frac{1}{2}\left(A - A^T\right).$$

Exercise 11. Prove the subset of skew-symmetric n by n real matrices is a vector subspace.

Example 6. Let

$$S \stackrel{\text{def}}{=} \left\{(x,y,z) \in \mathbf{R}^3 \;:\; x+y+z=0\right\}.$$

One can check that S is a subspace. Note that $0_{\mathbf{R}^3} \in S$. Also, if

$$x+y+z=0,$$

then

$$\alpha(x+y+z)=0.$$

Hence $\alpha \cdot v \in S$ when $v \in S$. Additionally, if $(x_1,y_1,z_1),(x_2,y_2,z_2) \in S$, then

$$(x_1+x_2)+(y_1+y_2)+(z_1+z_2)=(x_1+y_1+z_1)+(x_2+y_2+z_2)=0.$$

Hence S is a vector subspace.

Example 7. $\mathrm{Poly}(\mathbf{R})$ is a vector subspace of the vector space of formal power series $\mathbf{R}[[x]]$ and formal Laurent series $\mathbf{R}(x)$.

Example 8. $\mathrm{Poly}_d(\mathbf{R})$ is a vector subspace of the vector space of formal power series $\mathbf{R}[[x]]$ and formal Laurent series $\mathbf{R}(x)$.

Exercise 12. Prove $\mathrm{Poly}(\mathbf{R})$ is a vector subspace of the vector space of formal power series $\mathbf{R}[[x]]$ and formal Laurent series $\mathbf{R}(x)$.

Exercise 13. Prove the vector space of formal power series $\mathbf{R}[[x]]$ is a vector subspace of the vector space of formal Laurent series $\mathbf{R}(x)$.

Lemma 1.17. *If $S \leq V$, then S is a vector space with $+_{|S\times S}$ and $\cdot_{|\mathbf{R}\times S}$ where these functions are the restriction of the vector space operations for V*

$$+ : V \times V \longrightarrow V, \quad \cdot : \mathbf{R} \times V \longrightarrow V$$

to the subsets $S \times S \subset V \times V$ and $\mathbf{R} \times S \subset \mathbf{R} \times V$

$$+_{|S\times S} : S \times S \longrightarrow V, \quad \cdot_{|\mathbf{R}\times S} : \mathbf{R} \times S \longrightarrow V.$$

Proof. This is discussed in Remark 2 above. ♠

Lemma 1.18. *If V is a vector space and $S \subset V$, then the following are equivalent:*

(i) *S is closed under linear combinations.*

(ii) *S is closed under vector addition and scalar multiplication.*

Proof. If S is closed under linear combinations, we see that $v+w = 1\cdot v + 1\cdot w \in S$ and $\alpha\cdot v + 0\cdot w = \alpha\cdot v \in S$. Conversely, given $v,w\in S$ and $\alpha,\beta\in\mathbf{R}$, we know that $\alpha\cdot v, \alpha\cdot w\in S$ since it is closed under scalar multiplication. Hence $\alpha\cdot v+\beta\cdot w\in S$ since it is closed under vector addition. ♠

Lemma 1.19. *If $S\subset V$ is closed under linear combinations, then for every $v_1,\dots,v_n\in S$ and $\alpha_1,\dots,\alpha_n\in\mathbf{R}$, we have*

$$\sum_{j=1}^{n}\alpha_j\cdot v_j\in S.$$

Proof. We will prove this lemma via mathematical induction. That this holds for $n=2$ is immediate from the definition of being closed under linear combinations. For general n, we have $v_1,\dots,v_n\in S$ and $\alpha_1,\dots,\alpha_n\in\mathbf{R}$. We see that

$$w=\sum_{j=1}^{n}\alpha_j\cdot v_j=\alpha_1\cdot v_1+\sum_{j=2}^{n}\alpha_j\cdot v_j.$$

Setting

$$v=\sum_{j=2}^{n}\alpha_j\cdot v_j,$$

we know that $v,\alpha_1\cdot v_1\in S$ by the induction hypothesis. Finally, since S is closed under linear combinations, we see that $w=\alpha_1\cdot v_1+v\in S$ as desired. ♠

Definition 1.20 (Sum Sets)**.** Given a vector space V and subsets $A,B\subset V$, we define the **sum set of** A **and** B to be the subset

$$A+B\stackrel{\text{def}}{=}\{v+w\ :\ v\in A,\ w\in B\}.$$

Lemma 1.21. *If $S_1,S_2\le V$, then $S_1+S_2\le V$.*

Proof. As $0_V\in S_1,S_2$ and $0_V+0_V=0_V$, we see that $0_V\in S_1+S_2$. Hence, $S_1+S_2\neq\emptyset$. Given $v,w\in S_1+S_2$ and $\alpha,\beta\in\mathbf{R}$, we must show that $\alpha\cdot v+\beta\cdot w\in S_1+S_2$. To that end, we know that

$$v=v_1+v_2,\quad w=w_1+w_2$$

where $v_1,w_1\in S_1$ and $v_2,w_2\in S_2$. Now, we have

$$\alpha\cdot v+\beta\cdot w=\alpha\cdot(v_1+v_2)+\beta\cdot(w_1+w_2)=(\alpha\cdot v_1+\beta\cdot w_1)+(\alpha\cdot v_2+\beta\cdot w_2).$$

Since S_1,S_2 are vector subspaces, we see that $v'=\alpha\cdot v_1+\beta\cdot w_1\in S_1$ and $w'=\alpha\cdot v_2+\beta\cdot w_2\in S_2$. In particular, $v'+w'\in S_1+S_2$. Hence, S_1+S_2 is closed under linear combinations and thus is a vector subspace. ♠

Lemma 1.22. *If $S \subset V$ and $S+S \subset S$ if and only if S is closed under vector addition.*

Proof. First we assume that $S+S \subset S$ and must prove that S is closed under vector addition. Given $v,w \in S$, we know that $v+w \in S+S$ by Definition 1.20. Since $S+S \subset S$, we see that $v+w \in S$ and so S is closed under multiplication. Conversely, to see that $S+S \subset S$ when S is closed under multiplication, we argue as follows. Given $u \in S+S$, by definition of $S+S$, we have $u = v+w$ for some $v,w \in S$. Since S is closed under multiplication, $u = v+w \in S$ as desired. ♠

One case of sum sets that is particularly useful is a translate of a subset of a vector space.

Definition 1.23 (Translated Subset)**.** Given a vector space V, subset $A \subset V$, and $v \in V$, we call the sum set $v+A$ the **translate of** A **by** v.

Example 9. Take $V = \mathbf{R}^2$ and take S to be the x–axis. Specifically

$$S = \left\{(x,0) \in \mathbf{R}^2 \ : \ x \in \mathbf{R}\right\}.$$

Given $v_0 \in \mathbf{R}^2$ with $v_0 = (x_0, y_0)$, we see that

$$v_0 + S \overset{\text{def}}{=} \{(x+x_0, y_0) \ : \ x \in \mathbf{R}\}.$$

Making the substitution $z = x + x_0$, we see that

$$S + v_0 = \left\{(z, y_0) \in \mathbf{R}^2 \ : \ z \in \mathbf{R}\right\}.$$

If $S = \mathbf{R}^2$, then we can also “absorb” the y_0 and obtain the obvious fact that

$$\mathbf{R}^2 + v_0 = \mathbf{R}^2.$$

Definition 1.24 (Scaled Sets)**.** Given a vector space V, a subset $A \subset V$, and a subset $\Omega \subset \mathbf{R}$, we define the **$\Omega$–scaling of** A to be the subset

$$\Omega \cdot A \overset{\text{def}}{=} \{\alpha \cdot v \ : \ \alpha \in \Omega, \ v \in A\}.$$

Example 10. If $V = \mathbf{R}^2$, W is the x–axis, and $S = v_0 + W$ for $v_0 = (x_0, y_0)$, we see for $\alpha \in \mathbf{R}$ that

$$\alpha \cdot (v_0 + W) \alpha \cdot S = \left\{(\alpha z, \alpha y_0) \in \mathbf{R}^2 \ : \ z \in \mathbf{R}\right\} = \{(t, \alpha y_0) \ : \ t \in \mathbf{R}\} = \alpha \cdot v_0 + W.$$

Lemma 1.25. *If $S \subset V$ and $\mathbf{R} \cdot S \subset S$ if and only if S is closed under scalar multiplication.*

Proof. For the direct implication, given $v \in S$ and $\alpha \in \mathbf{R}$, we know that $\alpha \cdot v \in \mathbf{R} \cdot S \subset S$. Thus S is closed under scalar multiplication. If S is closed under scalar multiplication and $u \in \mathbf{R} \cdot S$, then there exists $v \in S$ and $\alpha \in \mathbf{R}$ such that $u = \alpha \cdot v$. Since S is closed under scalar multiplication, we see that $u = \alpha \cdot v \in S$ as needed. ♠

Remark 3. If $\mathscr{P}(V)$ is the power set of V (i.e. this is the set of all subsets of V), then we can define

$$+ : \mathscr{P}(V) \times \mathscr{P}(V) \longrightarrow \mathscr{P}(V), \quad \cdot : \mathbf{R} \times \mathscr{P}(V) \longrightarrow \mathscr{P}(V)$$

given by

$$T_1 + T_2 \stackrel{\text{def}}{=} T_1 + T_2 = \{t_1 + t_1 \ : \ t_1 \in T_1,\, t_2 \in T_2\}, \quad \alpha \cdot T \stackrel{\text{def}}{=} \alpha \cdot T = \{\alpha \cdot t \ : \ t \in T\}.$$

Note that the right hand sides are the sum set of T_1, T_2 and the subset T scaled by α. Notice that $\{0_V\}$ behaves like the zero vector under this addition operation as

$$T + \{0_V\} = T$$

holds for all $T \subset V$. However, if $S \subset T$, we see that

$$T + S = T.$$

In particular, $\mathscr{P}(V)$ cannot be a vector space with these operations (see Lemma 1.3).

Intersection, Union, and Products

We now consider when intersections and unions of vector subspaces are vector subspaces.

Lemma 1.26. *If $S_1, S_2 \leq V$, then $S_1 \cap S_2 \leq V$.*

Proof. By Lemma 1.16, we know that $0_V \in S_1$ and $0_V \in S_2$. In particular, $0_V \in S_1 \cap S_2$ and so $S_1 \cap S_2 \neq \emptyset$. Given $v_1, v_2 \in S_1 \cap S_2$ and $\alpha_1, \alpha_2 \in \mathbf{R}$, we know that $\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 \in S_1$ and $\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 \in S_2$ since both S_1, S_2 are closed under linear combinations by Lemma 1.18. Hence $\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 \in S_1 \cap S_2$. Therefore, $S_1 \cap S_2$ is a vector subspace as it is non-empty and closed under linear combinations. ♠

Lemma 1.27. *If $S_1, S_2 \leq V$, then $S_1 \cup S_2 \leq V$ if and only if $S_1 \subset S_2$ or $S_2 \subset S_1$.*

Proof. If $S_1 \subset S_2$ or $S_2 \subset S_1$, then $S_1 \cup S_2 = S_2$ or S_1. In either case, $S_1 \cup S_2$ is a vector subspace. If neither $S_1 \subset S_2$ or $S_2 \subset S_1$, then there exists vectors $v_1 \in S$ with $v_1 \notin S_2$ and $v_2 \in S_2$ with $v_2 \notin S_1$. We assert that $v_1 + v_2 \notin S_1 \cup S_2$. If $v_1 + v_2 \in S_1$, then $v_1 + v_2 + (-v_1) = v_2 \in S_1$ which we know is impossible. Likewise, if $v_1 + v_2 \in S_2$, then $v_1 + v_2 + (-v_2) = v_1 \in S_1$ which is also impossible. Hence $v_1 + v_2 \notin S_1$ and $v_1 + v_2 \notin S_2$. Therefore, $v_1 + v_2 \notin S_1 \cup S_2$. However, $v_1, v_2 \in S_1 \cup S_2$ and so we see that $S_1 \cup S_2$ cannot be closed under vector addition. Thus, $S_1 \cup S_2$ cannot be a vector subspace unless $S_1 \subset S_2$ or $S_2 \subset S_1$. ♠

Definition 1.28 (Product Space)**.** Given vector spaces V, W, we define the product vector space to be the set $V \times W$ with addition and scalar multiplication defined by

$$(v_1, w_1) + (v_2, w_2) \stackrel{\text{def}}{=} (v_1 + v_2, w_1 + w_2), \quad \alpha \cdot (v, w) \stackrel{\text{def}}{=} (\alpha \cdot v, \alpha \cdot w).$$

The zero vector in $V \times W$ is $0_{V \times W} = (0_V, 0_W)$.

Exercise 14. Prove that $V \times W$ is a vector space as defined in Definition 1.28.

Lemma 1.29. *If V, W are vector spaces with subspaces $S \leq V$ and $S' \leq W$, then $S \times S' \leq V \times W$.*

Proof. This is clear. ♠

Exercise 15. Prove Lemma 1.29.

1.3.2 Affine Subspaces

Definition 1.30 (Affine Subspace)**.** Given a vector space V and a subset $A \subset V$, we say A is a **affine subspace** of V if there exists $v_0 \in A$ such that the set

$$S_{A,v_0} \stackrel{\text{def}}{=} \{v - v_0 \;:\; v \in A\} = A - v_0.$$

is a vector subspace of V.

Example 11. If V is a vector space, $S \leq V$, and $v_0 \in V$, then $A = v_0 + S$ is an affine subspace. This amounts to

$$S_{A,v_0} \stackrel{\text{def}}{=} A - v_0 = (v_0 + S) - v_0 = S.$$

Lemma 1.31. *Given a vector space and a subset $A \subset V$, the following are equivalent:*

(i) A is an affine subspace.

(ii) *$A = v_0 + S$ for some $v_0 \in A$ and some vector subspace $S \leq V$. That is, A is the translate of S by v_0.*

Proof. First, we prove the direct implication. By definition of an affine subspace, there exists $v_0 \in A$ such that S_{A,v_0} is a vector subspace. Of course, we see that

$$\begin{aligned} v_0 + S_{A,v_0} &\stackrel{\text{def}}{=} \{v_0 + v \ : \ v \in S_{A,v_0}\} \\ &= v_0 + \{v - v_0 \ : \ v \in A\} \\ &= \{v_0 + v - v_0 \ : \ v \in A\} \\ &= \{v \ : \ v \in A\} = A. \end{aligned}$$

For the reverse implication, if $A = v_0 + S$ for some $v_0 \in A$ and $S \leq V$, then

$$\begin{aligned} S_{A,v_0} &\stackrel{\text{def}}{=} \{v - v_0 \ : \ v \in A\} \\ &= \{(v_0 + v) - v_0 \ : \ v \in S\} \\ &= \{v \ : \ v \in S\} = S. \end{aligned}$$

As S is a vector subspace, we see that $S_{A,v_0} = S$ is a vector subspace. ♠

Corollary 1.32. *If $S \leq V$ and $v \in V$, then $v + S$ is an affine subspace.*

Proof. This follows either from Example 11 or Lemma 1.31. ♠

Corollary 1.33 (Vector Subspaces are Affine)**.** *If $S \leq V$, then S is an affine subspace.*

Proof. Take $v_0 = 0_V$. ♠

The following is a restatement of Lemma 1.31.

Corollary 1.34 (Affine Subspaces are Vector Subspace Translates)**.** *If V is a vector space, then every affine subspace is a translate of a vector subspace and conversely.*

Lemma 1.35. *If $A \subset V$ is an affine subspace, then $S_{A,v} = S_{A,w}$ for all $v, w \in A$.*

Proof. Since A is an affine subspace, we know that there exists $v_0 \in V$ such that $S = S_{A,v_0}$ is a vector subspace. By Lemma 1.31, we know that $A = v_0 + S$. First, we will prove that $v + S = v_0 + S$ for any $v \in A$. For that, note that

$$v + S = (v - v_0 + v_0) + S = (v - v_0) + v_0 + S = v_0 + S$$

since $v - v_0 \in S$. Hence $A = v + S$.

Next, we will prove that $S = S_{A,v}$, which suffices for the proof of the lemma. Given $u \in S_{A,v}$, we know that
$$u = v - w$$
for some $w \in A$. As $A = v_0 + S$, we know that
$$v = v_0 + s_v, \quad w = v_0 + s_w$$
for some $s_v, s_w \in S$. In particular,
$$u = v - w = v_0 + s_v - v_0 - s_w = s_v - s_w \in S.$$
Hence, $S_{A,v} \subset S$. If $u \in S$, since $A = v + S$, we know that $v + u = w$ for some $w \in A$. Thus, $v - w = u$ and so $u \in S_{A,v}$. ♠

Though important mathematically, affine subspace will play a somewhat limited role in these notes.

1.3.3 Linear Hulls and Spans

We will need to discuss finite linear combinations of vectors over a fixed but arbitrary subset $S \subset V$. For this, we will view the coefficients in these sums over S as functions $\alpha\colon S \to \mathbf{R}$. For notational simplicity, we define
$$\alpha_v \stackrel{\text{def}}{=} \alpha(v)$$
for $v \in S$. We will write
$$\sum_{v \in S} \alpha_v \cdot v$$
for the associated function $\alpha\colon S \to \mathbf{R}$. As we are interested in finite linear combinations, we need $\alpha_v \neq 0$ for only finitely many $v \in S$.

Definition 1.36 (Finite Support)**.** Given set X and a function $f\colon X \to \mathbf{R}$, we say f has **finite support** if the subset
$$\operatorname{supp}(f) \stackrel{\text{def}}{=} \{x \in X \;:\; f(x) \neq 0\}$$
is finite. When $S \subset V$ and $\alpha\colon S \to \mathbf{R}$, we will say that α_v has finite support if
$$\operatorname{supp}(\alpha) \stackrel{\text{def}}{=} \{v \in S \;:\; \alpha_v \neq 0\}$$
is finite. We say α_v is **non-zero** if $\alpha_{v_0} \neq 0$ for some $v_0 \in S$. Otherwise, we say α_v is **zero**.

Remark 4. The subset $\mathrm{Fun}_{\mathrm{fin}}(S,\mathbf{R})$ of functions with finite support is a vector subspace of $\mathrm{Fun}(S,\mathbf{R})$.

Exercise 16. Prove the subset $\mathrm{Fun}_{\mathrm{fin}}(S,\mathbf{R})$ of functions with finite support is a vector subspace of $\mathrm{Fun}(S,\mathbf{R})$.

Definition 1.37 (Linear Span)**.** Given a vector space V and subset $S \subset V$, we define the **linear span of** S to be the subset of V of given by

$$\mathrm{Span}(S) \stackrel{\mathrm{def}}{=} \left\{ u \in V \ : \ u = \sum_{v \in S} \alpha_v \cdot v, \text{ where } \alpha_v \text{ has finite support} \right\}.$$

If $S = \emptyset$, then we define $\mathrm{Span}(\emptyset) = \{0_V\}$.

Lemma 1.38 (Linear Spans are Vector Subspaces)**.** *If V is a vector space and $S \subset V$, then* $\mathrm{Span}(S)$ *is a vector subspace of V.*

Proof. By definition of $\mathrm{Span}(S)$, we know that $\mathrm{Span}(S)$ is non-empty. Given $w, u \in \mathrm{Span}(S)$, we know that

$$w = \sum_{v \in S} \alpha_v \cdot v, \quad u = \sum_{v \in S} \beta_v \cdot v$$

for some α_v, β_v with finite support. Given $\theta, \lambda \in \mathbf{R}$, we see that

$$\begin{aligned} \theta \cdot w + \lambda \cdot u &= \theta \cdot \left(\sum_{v \in S} \alpha_v \cdot v \right) + \lambda \cdot \left(\sum_{v \in S} \beta_v \cdot v \right) \\ &= \left(\sum_{v \in S} (\theta \alpha_v + \lambda \beta_v) \cdot v \right) \in \mathrm{Span}(S). \end{aligned}$$

Hence $\mathrm{Span}(S)$ is closed under linear combinations and so a vector subspace. ♠

Remark 5. Note that we have a function $\mathrm{Fun}_{\mathrm{fin}}(S,\mathbf{R}) \to \mathrm{Span}(S)$ given by

$$\alpha_v \longmapsto \sum_{v \in S} \alpha_v \cdot v.$$

This is a surjective (linear) function by definition.

Lemma 1.39. *If V is a vector space and $S \subset T \subset V$, then* $\mathrm{Span}(S) \subset \mathrm{Span}(T)$.

Proof. Given $v \in \mathrm{Span}(S)$, by definition

$$v = \sum_{w \in S} \alpha_w \cdot w$$

where α_w has finite support. As $S \subset T$, we see that

$$v = \sum_{w \in T} \beta_w \cdot w$$

where

$$\beta_w \stackrel{\text{def}}{=} \begin{cases} \alpha_w, & w \in S, \\ 0, & w \notin S. \end{cases}$$

As α_v has finite support, β_w has finite support. Thus, $v \in \mathrm{Span}(T)$ and so $\mathrm{Span}(S) \subset \mathrm{Span}(T)$. ♠

Lemma 1.40. *If $S \subset V$, then*

$$\mathrm{Span}(\mathrm{Span}(S)) = \mathrm{Span}(S).$$

In particular, if $T \subset \mathrm{Span}(S)$, then $\mathrm{Span}(T) \subset \mathrm{Span}(S)$.

Proof. This is clear since finite linear combinations of finite linear combinations are still finite linear combinations. ♠

Definition 1.41 (Linear Hull)**.** Given a vector space V and subset $S \subset V$, we define the **linear hull of** S or also called the **linear closure of** S to be

$$\mathrm{Hull}(S) \stackrel{\text{def}}{=} \bigcap_{W \le V,\ S \subset W} W.$$

Because every vector subspace contains 0_V, we know $\mathrm{Hull}(S)$ is non-empty even when $S = \emptyset$.

Remark 6. $\mathrm{Hull}(S)$ is the smallest vector subspace of V that contains S (by definition).

Lemma 1.42 (Linear Hulls are Vector Subspaces)**.** *If V is a vector space and $S \subset V$, then* $\mathrm{Hull}(S)$ *is a vector subspace of V.*

Proof. This follows from Definition 1.41 and Lemma 1.26. ♠

Lemma 1.43. *If V is a vector space and $S \subset T \subset V$, then* $\mathrm{Hull}(S) \subset \mathrm{Hull}(T)$.

Proof. Given that $S \subset T$, it follows that

$$\mathrm{Hull}(S) = \bigcap_{W \leq V,\ S \subset W} W \subset \bigcap_{W \leq V,\ T \subset W} W = \mathrm{Hull}(T).$$

♠

Lemma 1.44. *If V is a vector space, then $S = \mathrm{Hull}(S)$ if and only if $S \leq V$.*

Proof. If $\mathrm{Hull}(S) = S$, then $S \leq V$ by Lemma 1.42. If $S \leq V$, then $S \leq V$ with $S \subset S$. Hence $\mathrm{Hull}(S) \subset S$ by Definition 1.41. However, $S \subset \mathrm{Hull}(S)$ by Definition 1.41 and so $S = \mathrm{Hull}(S)$. ♠

Proposition 1.45. *If V is a vector space and $S \subset V$, then $\mathrm{Span}(S) = \mathrm{Hull}(S)$.*

Proof. As $\mathrm{Span}(S)$ contains S and is a vector subspace by Lemma 1.38, by definition of $\mathrm{Hull}(S)$, we have $\mathrm{Hull}(S) \subset \mathrm{Span}(S)$. It remains to prove $\mathrm{Span}(S) \subset \mathrm{Hull}(S)$. If $w \in \mathrm{Span}(S)$, then

$$w = \sum_{v \in S} \alpha_v \cdot v$$

for some α_v with finite support. If $W \leq V$ and $S \subset W$, then $w \in W$ since W is closed under linear combinations by Lemma 1.18. Hence $w \in \mathrm{Hull}(S)$ as needed. Therefore, $\mathrm{Span}(S) \subset \mathrm{Hull}(S)$ and we conclude that $\mathrm{Span}(S) = \mathrm{Hull}(S)$. ♠

Corollary 1.46. *If V is a vector space, then $S = \mathrm{Span}(S)$ if and only if $S \leq V$.*

Proof. This follows from Lemma 1.44 and Proposition 1.45. ♠

1.4 Linear Functions

In this section we introduce another foundational concept in linear algebra called a linear function.

1.4.1 Definition and Examples

Definition 1.47 (Linear Function)**.** Given vector spaces V, W, we say that a function $L\colon V \to W$ is **linear** if

$$L(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2) = \alpha_1 \cdot L(v_1) + \alpha_2 \cdot L(v_2).$$

Lemma 1.48. *If $L\colon V \to W$ is linear, then $L(0_V) = 0_W$.*

Proof. We know that

$$L(0_V) = L(v + (-v)) = L(v) + (-L(v)) = 0_W.$$

♠

Lemma 1.49. *If $L_1\colon V_1 \to V_2$ and $L_2\colon V_2 \to V_3$ are linear functions, then $L_2 \circ L_1\colon V_1 \to V_3$ is linear.*

Proof. Given $v, w \in V_1$ and $\alpha, \beta \in \mathbf{R}$, we must prove that

$$L_2(L_1(\alpha \cdot v + \beta \cdot w)) = \alpha \cdot L_2(L_1(v)) + \beta \cdot L_2(L_1(w)).$$

That that end, we have the following:

$$\begin{aligned} L_2(L_1(\alpha \cdot v + \beta \cdot w)) &= L_2(\alpha \cdot L_1(v) + \beta \cdot L_1(w)) \\ &= \alpha \cdot L_2(L_1(v)) + \beta \cdot L_2(L_1(w)). \end{aligned}$$

♠

Definition 1.50 (Injective Function)**.** Given sets X, Y, we say a function $L\colon X \to Y$ is **injective** if given $x_1, x_2 \in X$ with $x_1 \neq x_2$, then $L(x_1) \neq L(x_2)$.

Lemma 1.51. *If $f_1\colon X_1 \to X_2$ and $f_2\colon X_2 \to X_3$ are injective functions, then $f_2 \circ f_1\colon X_1 \to X_3$ is injective.*

Proof. Given $x, y \in X_1$ with $x \neq y$, we must show that $f_2(f_1(x)) \neq f_2(f_1(y))$. Since f_1 is injective, $f_1(x) \neq f_1(y)$. Last, as f_2 is injective, $f_2(f_1(x)) \neq f_2(f_1(y))$. ♠

Definition 1.52 (Surjective Function)**.** Given sets X, Y, we say a function $L\colon X \to Y$ is **surjective** if each $y \in Y$, there exists $x \in X$ with $L(x) = y$.

Lemma 1.53. *If $f_1 : X_1 \to X_2$ and $f_2 : X_2 \to X_3$ are surjective functions, then $f_2 \circ f_1 : X_1 \to X_3$ is surjective.*

Proof. Given $z \in X_3$, we must find an $x \in X_1$ such that $f_2(f_1(x)) = z$. Since f_2 is surjective, there exists $y \in X_2$ such that $f_2(y) = z$. Finally, as f_1 is surjective, there exists $x \in X_1$ such that $f_2(f_1(x)) = z$. ♠

Definition 1.54 (Isomorphism)**.** Given vector spaces V, W, we say that a linear function $L\colon V \to W$ is an **isomorphism** if L is injective and surjective.

Definition 1.55 (Isomorphic Vector Spaces)**.** We say two vector spaces V, W are **isomorphic** if there exists an isomorphism $L\colon V \to W$. When V, W are isomorphic, we will denote this by $V \cong W$.

Definition 1.56 (Inclusion Map)**.** Given a vector space V and a subspace $S \leq V$, the **inclusion map** $\iota_{S,V} : S \to V$ is defined by $\iota_{S,V}(v) = v$.

Example 12. Let $V = \mathbf{R}^2$ and

$$S = \left\{(x, 0) \in \mathbf{R}^2 \ : \ x \in \mathbf{R}\right\}.$$

S is a vector space of $\mathbf{R}^2$ and $\iota_{S,\mathbf{R}^2}$ in this case is given by

$$\iota_{S,\mathbf{R}^2}(x, 0) = (x, 0).$$

Lemma 1.57. *If V is a vector space, $S \leq V$, and $\iota_{S,V}$ is the inclusion map, then $\iota_{S,V}$ is linear.*

Proof. This is true by definition of vector subspaces. It feels wrong to say more. ♠

Given a pair of vector spaces, we define

$$\mathrm{Hom}(V, W) \overset{\text{def}}{=} \{L\colon V \to W \ : \ L \text{ is linear}\}.$$

Notice that $\mathrm{Hom}(V, W) \subset \mathrm{Fun}(V, W)$. Recall that $\mathrm{Fun}(V, W)$ is a vector space under pointwise addition and pointwise scalar multiplication.

Lemma 1.58. $\mathrm{Hom}(V, W)$ *is a vector subspace of* $\mathrm{Fun}(V, W)$.

Proof. First, we will prove that $0_{\mathrm{Fun}(V,W)} \in \mathrm{Hom}(V,W)$. For that, we have

$$0_{\mathrm{Fun}(V,W)}(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2) = 0_W.$$

Likewise,

$$0_{\mathrm{Fun}(V,W)}(v_1) = 0_W = 0_{\mathrm{Fun}(V,W)}(v_2).$$

Hence

$$\alpha_1 \cdot 0_{\mathrm{Fun}(V,W)}(v_1) + \alpha_2 \cdot 0_{\mathrm{Fun}(V,W)}(v_2) = 0_W = 0_{\mathrm{Fun}(V,W)}(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2).$$

Hence, $0_{\mathrm{Fun}(V,W)} \in \mathrm{Hom}(V,W)$. Next, we will show that if $L_1, L_2 \in \mathrm{Hom}(V,W)$ and $\alpha_1, \alpha_2 \in \mathbf{R}$, then $\alpha_1 \cdot L_1 + \alpha_2 \cdot L_2$ is linear. For that, let $L = \alpha_1 \cdot L_1 + \alpha_2 \cdot L_2$. To show that L is linear, we need to prove that

$$L(\beta_1 \cdot v_1 + \beta_2 \cdot v_2) = \beta_1 \cdot L(v_1) + \beta_2 \cdot L(v_2).$$

$$\begin{aligned}
L(\beta_1 \cdot v_1 + \beta_2 \cdot v_2) &= (\alpha_1 \cdot L_1 + \alpha_2 \cdot L_2)(\beta_1 \cdot v_1 + \beta_2 \cdot v_2) \\
&= \alpha_1 \cdot L_1(\beta_1 \cdot v_1 + \beta_2 \cdot v_2) + \alpha_2 \cdot L_2(\beta_1 \cdot v_1 + \beta_2 \cdot v_2) \\
&= \alpha_1 \cdot (\beta_1 \cdot L_1(v_1) + \beta_2 \cdot L_1(v_2)) + \alpha_2 \cdot (\beta_1 \cdot L_2(v_1) + \beta_2 \cdot L_2(v_2)) \\
&= \alpha_1 \cdot (\beta_1 \cdot L_1(v_1)) + \alpha_2 \cdot (\beta_1 \cdot L_2(v_1)) + \alpha_1 \cdot (\beta_2 \cdot L_1(v_2)) + \alpha_2 \cdot (\beta_2 \cdot L_2(v_2)) \\
&= \beta_1 \cdot (\alpha_1 \cdot L_1(v_1)) + \beta_1 \cdot (\alpha_2 \cdot L_2(v_1)) + \beta_2 \cdot (\alpha_1 \cdot L_1(v_2)) + \beta_2 \cdot (\alpha_2 \cdot L_2(v_2)) \\
&= \beta_1 \cdot (\alpha_1 \cdot L_1(v_1) + \alpha_1 \cdot L_2(v_1)) + \beta_2 \cdot (\alpha_1 \cdot L_1(v_2) + \alpha_2 \cdot L_2(v_2)) \\
&= \beta_1 \cdot L(v_1) + \beta_2 \cdot L(v_2).
\end{aligned}$$

Hence $L \in \mathrm{Hom}(V,W)$ and so $\mathrm{Hom}(V,W)$ is a vector subspace of $\mathrm{Fun}(V,W)$. ♠

When $V = W$, we can consider the subset $\mathrm{Aut}(V) \subset \mathrm{Hom}(V,V)$ given by

$$\mathrm{Aut}(V) \stackrel{\text{def}}{=} \{L \colon V \to V \ : \ L \text{ is an isomorphism}\}.$$

Lemma 1.59. *If $L \colon V \to W$ is an isomorphism, then $L^{-1} \colon W \to V$ is an isomorphism.*

Proof. As $L \colon V \to W$ is bijective, the inverse function $L^{-1} \colon W \to V$ exists and satisfies

$$L^{-1}(L(v)) = v, \quad L(L^{-1}(w)) = w$$

for all $v \in V$ and $w \in W$. It remains to prove that L^{-1} is linear. Given $w_1, w_2 \in W$ and $\alpha_1, \alpha_2 \in \mathbf{R}$, we must prove that

$$L^{-1}(\alpha_1 \cdot w_1 + \alpha_2 \cdot w_2) = \alpha_1 \cdot L^{-1}(w_1) + \alpha_2 \cdot L^{-1}(w_2).$$

Since L^{-1} is bijective, there exist unique vectors $v_1, v_2 \in V$ with $L(v_1) = w_1$ and $L(v_2) = w_2$. Additionally, we see that $L^{-1}(w_1) = v_1$ and $L^{-1}(w_2) = v_2$. In particular, we have

$$\begin{aligned} L^{-1}(\alpha_1 \cdot w_1 + \alpha_2 \cdot w_2) &= L^{-1}(\alpha_1 \cdot L(v_1) + \alpha_2 \cdot L(v_w)) \\ &= L^{-1}(L(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2)) \\ &= \alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 \\ &= \alpha_1 \cdot L^{-1}(w_1) + \alpha_2 \cdot L^{-1}(w_2). \end{aligned}$$

♠

Lemma 1.60. *If $L_1 : V_1 \to V_2$ and $L_2 : V_2 \to V_3$ are isomorphisms, then $L_2 \circ L_1 : V_1 \to V_3$ is also an isomorphism. In particular, if $V_1 \cong V_2$ and $V_2 \cong V_3$, then $V_1 \cong V_3$.*

Proof. This follows from Lemma 1.48, Lemma 1.51, and Lemma 1.53. ♠

Lemma 1.61. *The identity function $\mathrm{Id}_V : V \to V$ defined by $\mathrm{Id}_V(v) = v$ is an isomorphism.*

Proof. That Id_V is bijective is clear. As

$$\mathrm{Id}_V(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2) = \alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 = \alpha_1 \cdot \mathrm{Id}_V(v_1) + \alpha_2 \cdot \mathrm{Id}_V(v_2),$$

we see that Id_V is also linear. ♠

Collecting the previous few results, we have (groups are awesome by the way):

Corollary 1.62. *$\mathrm{Id}_V \in \mathrm{Aut}(V)$ and $\mathrm{Aut}(V)$ is closed under composition of functions and inverses. That is, if $L \in \mathrm{Aut}(V)$, then $L^{-1} \in \mathrm{Aut}(V)$ and if $L_1, L_2 \in \mathrm{Aut}(V)$, then $L_2 \circ L_1 \in \mathrm{Aut}(V)$. This endows $\mathrm{Aut}(V)$ with a group structure.*

Not that we will speak about groups here but $\mathrm{Aut}(V)$ is a subgroup of the subset of $\mathrm{Fun}(V,V)$ of bijective functions. Every more, $\mathrm{Aut}(V)$ is also a subgroup of the subset of bijective functions of $\mathrm{Fun}(\mathrm{Aut}(V), \mathrm{Aut}(V))$ which is also a group (under composition of functions of course). Note that the subset of surjective functions of $\mathrm{Fun}(X,X)$ is closed under composition but inverses need not exist. The subset of injective functions is also closed under composition but inverses need not exist (at least not with the domain of the inverse being X). The issue is two sided. If $f : X \to X$ is surjective, then we can always find $g : X \to X$ such that $g \circ f = \mathrm{Id}_X$ is the identity function. Unfortunately, $f \circ g \neq \mathrm{Id}_X$ unless f is bijective. Similarly, if f is injective, then we can find $g : f(X) \to X$ with $g \circ f = \mathrm{Id}_X$. Unfortunately, we can not extend g to X unless $f(X) = X$.

Exercise 17. Let $f,g\colon X \to X$ and $g \circ f = \mathrm{Id}_X$.

(1) If f is surjective, then g is injective.

(2) If f is injective, then g is surjective.

Exercise 18. Recall that $\mathrm{Poly}(\mathbf{R})$ is the vector space of polynomials. Define

$$\mathrm{D}\colon \mathrm{Poly}(\mathbf{R}) \longrightarrow \mathrm{Poly}(\mathbf{R})$$

by

$$\mathrm{D}\left(\sum_{j=0}^{n} \alpha_j x^j\right) \stackrel{\text{def}}{=} \sum_{j=1}^{n} j\alpha_j x^{j-1}.$$

Prove that D is linear.

Exercise 19. Recall that $\mathrm{Poly}(\mathbf{R})$ is the vector space of polynomials. Define

$$\mathrm{I}\colon \mathrm{Poly}(\mathbf{R}) \longrightarrow \mathrm{Poly}(\mathbf{R})$$

by

$$\mathrm{I}\left(\sum_{j=0}^{n} \alpha_j x^j\right) \stackrel{\text{def}}{=} \sum_{j=0}^{n} \frac{\alpha_j x^{j+1}}{j+1}.$$

Prove that I is linear.

Exercise 20. Prove that $\mathrm{D} \circ \mathrm{I} = \mathrm{Id}_{\mathrm{Poly}(\mathbf{R})}$. That is, prove that if $P(x) \in \mathrm{Poly}(\mathbf{R})$, then

$$\mathrm{D}(\mathrm{I}(P(x))) = P(x).$$

Exercise 21. Given a vector space V and a subset $S \subset V$, prove that the function

$$L\colon \mathrm{Fun}_{\mathrm{fin}}(S,\mathbf{R}) \longrightarrow \mathrm{Span}(S)$$

given by

$$L(\alpha_v) \stackrel{\text{def}}{=} \sum_{v \in S} \alpha_v \cdot v$$

is a surjective linear function.

1.4.2 Image and Kernel

Definition 1.63 (Image)**.** Given vector spaces V,W and a linear function $L\colon V \to W$, we define the **image of** V **under** L to be the subset $L(V) \subset W$ defined by

$$L(V) \stackrel{\text{def}}{=} \{w \in W \ : \ \text{there exists } v \in V \text{ with } L(v) = w\}.$$

Definition 1.64 (Kernel)**.** Given vector spaces V,W and a linear function $L\colon V \to W$, we define the **kernel of** L to be the subset $\ker(L) \subset V$ defined by

$$\ker(L) \stackrel{\text{def}}{=} \{v \in V \ : \ L(v) = 0_W\}.$$

Definition 1.65 (Image of a Subset)**.** Given vector spaces V,W, a linear function $L\colon V \to W$, and a subset $S \subset V$, we define the **image of** S **under** L to be the subset $L(S) \subset W$ defined by

$$L(S) \stackrel{\text{def}}{=} \{w \in W \ : \ \text{there exists } v \in S \text{ with } L(v) = w\}.$$

Definition 1.66 (Preimage of a Subset)**.** Given vector spaces V,W, a linear function $L\colon V \to W$, and a subset $T \subset W$, we define the **preimage of** T **under** L to be the subset $L^{-1}(T) \subset V$ defined by

$$L^{-1}(T) \stackrel{\text{def}}{=} \{v \in V \ : \ L(v) \in T\}.$$

Exercise 22. Determine $\ker(\mathrm{D})$ and $\mathrm{D}(\mathrm{Poly}(\mathbf{R}))$ for D defined in Exercise 18.

Exercise 23. Determine $\ker(\mathrm{I})$ and $\mathrm{I}(\mathrm{Poly}(\mathbf{R}))$ for I defined in Exercise 19.

Exercise 24. Determine $\mathrm{D}^{-1}(1)$ and $\mathrm{D}^{-1}(x)$ for D defined in Exercise 18.

Exercise 25. Prove that $\mathrm{D}(\mathrm{Poly}_d(\mathbf{R})) \leq \mathrm{Poly}_{d-1}(\mathbf{R})$ and $\mathrm{I}(\mathrm{Poly}_d(\mathbf{R})) \leq \mathrm{Poly}_{d+1}(\mathbf{R})$.

Exercise 26. Define

$$\mathrm{D}^n \stackrel{\text{def}}{=} \underbrace{\mathrm{D} \circ \mathrm{D} \circ \cdots \circ \mathrm{D}}_{n \text{ times}}.$$

Prove that if $P(x) \in \mathrm{Poly}_d(\mathbf{R})$, then

$$\mathrm{D}^{d+1}(P(x)) = 0.$$

Lemma 1.67. *If $L\colon V \to W$ is a linear function, then $\ker(L)$ is a vector subspace of V.*

Proof. By Lemma 1.48, we know that $0_V \in \ker(L)$ and so $\ker(L)$ is non-empty. Given $v,w \in \ker(L)$ and $\alpha,\beta \in \mathbf{R}$, we see that

$$L(\alpha \cdot v + \beta \cdot w) = \alpha \cdot L(v) + \beta \cdot L(w) = \alpha \cdot 0_W + \beta \cdot 0_W = 0_W + 0_W = 0_W.$$

Hence, $\ker(L)$ is closed under linear combinations and hence a vector subspace by Lemma 1.18. ♠

Lemma 1.68. *If $L\colon V \to W$ is a linear function and $S \leq V$, then $L(S)$ is a vector subspace of W. In particular, $L(V)$ is a vector subspace of W.*

Proof. To begin, since $S \neq \emptyset$, we know that $L(S) \neq \emptyset$. Given $w_1, w_2 \in L(S)$ and $\alpha_1, \alpha_2 \in \mathbf{R}$, we need to show that $\alpha_1 \cdot w_1 + \alpha_2 \cdot w_2 \in L(S)$. To this end, by definition of $L(S)$, there exists $v_1, v_2 \in S$ such that $L(v_1) = w_1$ and $L(v_2) = w_2$. Define $v = \alpha_1 \cdot v_1 + \alpha_2 \cdot v_2$ and note that $v \in S$ since S is a vector subspace. Now, we have

$$L(v) = L(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2) = \alpha_1 \cdot L(v_1) + \alpha_2 \cdot L(v_2) = \alpha_1 \cdot w_1 + \alpha_2 \cdot w_2.$$

Therefore $\alpha_1 \cdot w_1 + \alpha_2 \cdot w_2 = L(v) \in L(S)$. Thus, $L(S)$ is a vector subspace by Lemma 1.18. ♠

Lemma 1.69. *If $L\colon V \to W$ is a linear function and $T \leq W$, then $L^{-1}(T)$ is a vector subspace of V.*

Proof. Since $0_W \in T$ by Lemma 1.16 and $L(0_V) = 0_W$ by Lemma 1.48, we know that $0_V \in L^{-1}(T)$. Given $v_1, v_2 \in L^{-1}(T)$ and $\alpha_1, \alpha_2 \in \mathbf{R}$, we must show that $v = \alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 \in L^{-1}(T)$. By definition of $L^{-1}(T)$, there exist $w_1, w_2 \in W$ such that $L(v_1) = w_1$ and $L(v_2) = w_2$. We see that $w = \alpha_1 \cdot w_1 + \alpha_2 \cdot w_2 \in T$ since T is a vector subspace. Additionally,

$$L(v) = L(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2) = \alpha_1 \cdot L(v_1) + \alpha_2 \cdot L(v_2) = \alpha_1 \cdot w_1 + \alpha_2 \cdot w_2 = w.$$

Hence $v \in L^{-1}(T)$ as needed. ♠

Proposition 1.70. *Given a linear function $L\colon V \to W$, the following are equivalent:*

(i) *L is injective.*

(ii) $\ker(L) = \{0_V\}$.

Proof. For the direct implication, we assume L is injective. We know by Lemma 1.48 that $L(0_V) = 0_W$. If $v \in V$ with $L(v) = 0_W$, then $v = 0_V$ since L is injective. Therefore, $\ker(L) = \{0_V\}$. For the reverse implication, we assume that $\ker(L) = \{0_V\}$. If $v_1, v_2 \in V$ with $L(v_1) = L(v_2)$, then

$$0_W = L(v_1) + (-L(v_2)) = L(v_1 + (-v_2)).$$

Thus, $v_1 - v_2 \in \ker(L)$ and so $v_1 - v_2 = 0_V$. Therefore $v_1 = v_2$ and L is injective. ♠

Proposition 1.71. *If $L\colon V \to W$ is a linear function and $L(v_0) = w_0$ for some $v_0 \in V$, then for each $v \in V$ with $L(v) = w_0$, there exists $v' \in \ker(L)$ such that $v = v_0 + v'$.*

Proof. Let $v \in V$ with $L(v) = w_0$. By assumption, $L(v_0) = w_0$ and so

$$0_W = w_0 - w_0 = L(v) - L(v_0) = L(v - v_0).$$

Hence, $v - v_0 \in \ker(L)$ and so $v - v_0 = v'$ for some $v' \in \ker(L)$. Thus, $v = v_0 + v'$ with $v' \in \ker(L)$ as needed. ♠

Corollary 1.72 (Solution Spaces to Linear Systems are Affine)**.** *If $L\colon V \to W$ is a linear function with $L(v_0) = w_0$, then $L^{-1}(w_0) = v_0 + \ker(L)$. In particular, $L^{-1}(w_0)$ is an affine subspace.*

Proof. By Proposition 1.71, we know that if $L(v) = w_0$ then $v \in \ker(L) + v_0$. Therefore, $L^{-1}(w_0) \subset \ker(L) + v_0$. For $v \in v_0 + \ker(L)$, there exists some $v' \in \ker(L)$ such that $v = v_0 + v'$. In particular,

$$L(v) = L(v_0 + v') = L(v_0) + L(v') = w_0 + 0_W = w_0.$$

Therefore, $v \in L^{-1}(w_0)$ and so $v_0 + \ker(L) \subset L^{-1}(w_0)$. In total, we now see that $L^{-1}(w_0) = v_0 + \ker(L)$. ♠

1.4.3 Linear Self-Maps and Eigenvectors

Linear functions from a vector space to itself are especially important. In this subsection, we briefly introduce some basic language in this area. Given $L\colon V \to V$, we define

$$L^n \overset{\text{def}}{=} \underbrace{L \circ L \circ \cdots \circ L}_{n \text{ times}} \colon V \longrightarrow V.$$

Alternatively, we have

$$\begin{tikzcd} V \arrow[r, "L"] \arrow[rrrr, bend left, "L^n"] & V \arrow[r, "L"] & \dots \arrow[r, "L"] & V \arrow[r, "L"] & L \end{tikzcd}$$

Definition 1.73 (Eigenvector)**.** Given a linear function $L\colon V \to V$, we say that $v \in V$ is an **eigenvector for** L if $L(v) = \lambda \cdot v$ for some $\lambda \in \mathbf{R}$.

Definition 1.74 (Eigenvalue)**.** Given a linear function $L\colon V \to V$, we say that $\lambda \in \mathbf{R}$ is an **eigenvalue for** L if $L(v) = \lambda \cdot v$ for some $v \in V - \{0_V\}$.

Recall that $\mathrm{Hom}(V,V)$ is a vector space under point-wise addition and scalar multiplication. In particular, for any $\lambda \in \mathbf{R}$, we can define $L_\lambda : V \to V$ by

$$L_\lambda(v) = L(v) - \lambda \cdot v.$$

Lemma 1.75. *If $L\colon V \to V$ and $\lambda \in \mathbf{R}$, then $L_\lambda = L - \lambda \cdot \mathrm{Id}_V$ is a linear function.*

Proof. We leave this for the reader. ♠

Exercise 27. Prove Lemma 1.75.

Lemma 1.76. *If $L\colon V \to V$, then λ is an eigenvalue for L if and only if $\ker(L_\lambda) \neq \{0_V\}$.*

Proof. We leave this for the reader. ♠

Exercise 28. Prove Lemma 1.76.

Corollary 1.77. *If $L\colon V \to V$, then λ is an eigenvalue for L if and only if L_λ is not injective.*

Proof. This follows from Proposition 1.70 and Lemma 1.76. ♠

Definition 1.78 (Eigenvalue Set of a Linear Self-Map)**.** Given a linear function $L\colon V \to V$, we define the **eigenvalue set of** L to be

$$\mathrm{E}(L) \overset{\text{def}}{=} \{\lambda \in \mathbf{R} \ : \ \lambda \text{ is an eigenvalue for } L\}.$$

Definition 1.79 (Eigenspace)**.** Given a linear function $L\colon V \to V$ and $\lambda \in \mathrm{E}(L)$, we define the **$\lambda$–eigenspace of** L to be the subset

$$\mathrm{E}_{L,\lambda} \overset{\text{def}}{=} \{v \in V \ : \ L(v) = \lambda \cdot v\}.$$

Lemma 1.80. *If $L\colon V \to V$ and $\lambda \in \mathrm{E}(L)$, then $\mathrm{E}_{L,\lambda} \leq V$.*

Proof. We leave this for the reader. ♠

Exercise 29. Prove Lemma 1.76.

Exercise 30. We say that a linear function $L\colon V \to V$ is **nilpotent** if there exists an n such that $L^n(v) = 0_V$ for all $v \in V$. Define

$$W_k \overset{\text{def}}{=} \ker(L^k), \quad U_k \overset{\text{def}}{=} L^k(V).$$

(a) Prove that $L^{k+1}(V) \leq L^k(V)$.

(b) Prove that $\ker(L^k) \leq \ker(L^{k+1})$.

(c) Prove that $L^k(V) \leq \ker(L^{n-k})$.

(d) Prove that $\ker(L^n) = V$.

1.5 Quotient Spaces

Given a vector space V and a vector subspace $W \leq V$, we will construct a vector space V/W and a surjective linear function $L_W : V \to V/W$ such that $\ker(L_W) = W$. We will define an equivalence relation $\sim_W$ and the set of equivalence classes will be the desired underlying set for V/W. The equivalence classes, being subsets of V, are natural elements of the power set $\mathscr{P}(V)$ of V. We will use the addition and scalar operations we defined on $\mathscr{P}(V)$ before (see Remark 2) to endow V/W with a vector space structure.

To begin, given a vector space V and a vector subspace $W \leq V$, we write $v \sim_W u$ if $v - u \in W$. We define

$$[v]_W \stackrel{\text{def}}{=} \{u \in V \ : \ v \sim_W u\}.$$

Lemma 1.81. *If V is a vector space, W is a vector subspace, and $\sim_W$ is given above, then the following hold for all $v, u, z \in V$:*

(1) $v \sim_W v$.

(2) If $v \sim_W u$, then $u \sim_W v$.

(3) If $v \sim_W u$ and $u \sim_W z$, then $v \sim_W z$.

In particular, $\sim_W$ is an equivalence relation.

Proof. This is straightforward. ♠

Exercise 31. Prove Lemma 1.81.

Example 13. Let $V = \mathbf{R}^2$ and

$$W = \left\{(x,0) \in \mathbf{R}^2 \,:\, x \in \mathbf{R}\right\}.$$

We see that if $v = (x_1, y_1)$ and $u = (x_2, y_2)$, then $v \sim_W u$ if $v - u \in W$. Now, $v - u = (x_1 - x_2, y_1 - y_2)$ and so if $v - u \in W$ then $y_1 - y_2 = 0$. Hence, $v \sim_W u$ if and only if v, u have the same second coordinate.

Lemma 1.82. *If V is a vector space, $W \leq V$, and $v_0 \in V$, then*

$$[v_0]_W = v_0 + W.$$

In particular, $[v_0]_W$ is an affine subspace.

Proof. If $u \in [v_0]_W$, then $v_0 - u \in W$. In particular, $v - u = w_0$ for some $w_0 \in W$. Thus, $u = v_0 - w_0 = v_0 + (-w_0) \in v_0 + W$. If $u \in v_0 + W$, then $u = v_0 + w_0$ for some $w_0 \in W$. Hence $v_0 - u = v_0 - v_0 + w_0 = w_0 \in W$. Thus $v_0 \sim_W u$. ♠

Lemma 1.83. *If V is a vector space and $W \leq V$, then $[0_V]_W = W$.*

Proof. If $v \in [0_V]_W$ then $0_V - v = -v \in W$, then $(-1)\cdot(-v) = v \in W$. If $v \in W$, then $-v = 0_V - v \in W$. Hence $v \in [0_V]_W$. ♠

By definition, each $[v]_W \in \mathscr{P}(V)$, where $\mathscr{P}(V)$ is the power set which is the set of all subsets of V. We already defined operations

$$+ \,:\, \mathscr{P}(V) \times \mathscr{P}(V) \longrightarrow \mathscr{P}(V), \quad \cdot \,:\, \mathbf{R} \times \mathscr{P}(V) \longrightarrow \mathscr{P}(V)$$

given by

$$S + T \stackrel{\text{def}}{=} \{s + t \,:\, s \in S,\, t \in T\}, \quad \alpha \cdot S \stackrel{\text{def}}{=} \{\alpha \cdot s \,:\, s \in S\}.$$

When $\alpha = 0$, the set $\alpha \cdot S = \{0_V\}$ for any subset $S \subset V$. In defining a vector space structure on a set, how 0 acts is unimportant in the sense that every vector when scaled by 0 must be the zero vector. In what follows, we will define

$$0 \cdot S \stackrel{\text{def}}{=} S.$$

It is important to remember that $\mathscr{P}(V)$ is not a vector space with these operations. However, it does satisfy many of the requisite conditions:

$$\begin{aligned} S+T &= T+S \\ S+(T+U) &= (S+T)+U \\ (\alpha\beta)\cdot S &= \alpha\cdot(\beta\cdot S) \\ \alpha\cdot(S+T) &= \alpha\cdot S+\alpha\cdot T \\ 1\cdot S &= S. \end{aligned}$$

Exercise 32. If $S,T,U \in \mathscr{P}(V)$ and $\alpha,\beta \in \mathbf{R}$, prove the following:

(1) $S+T = T+S$.

(2) $S+(T+U) = (S+T)+U$.

Exercise 33. If $S,T,U \in \mathscr{P}(V)$ and $\alpha,\beta \in \mathbf{R}$, prove the following:

(1) $(\alpha\beta)\cdot S = \alpha\cdot(\beta\cdot S)$.

(2) $\alpha\cdot(S+T) = \alpha\cdot S+\alpha\cdot T$.

(3) $1\cdot S = S$.

Remark 7. The condition $(\alpha+\beta)\cdot S = \alpha\cdot S+\beta\cdot S$ need not hold for a general subset of V. For example, take $V = \mathbf{R}^2$ and $S = \{(1,0),(0,1)\}$. Take $\alpha = \beta = 1$. Then

$$(1+1)\cdot S = 2\cdot S = \{(2,0),(0,2)\}$$

while

$$1\cdot S+1\cdot S = \{(2,0),(1,1),(0,2)\}.$$

If S is a subspace, then $\alpha\cdot S = S$ and $S+S = S$. Hence

$$(1+1)\cdot S = S+S = S = 2\cdot S.$$

We define $L_W : V \to \mathscr{P}(V)$ by

$$L_W(v) \stackrel{\text{def}}{=} [v]_W = v+W.$$

We set $V/W = L_W(V)$ to be the image of V under this function. The points of V/W are $[v]_W = v+W$. Even though $\mathscr{P}(V)$ is not a vector space with the above operations (see Remark 2), amazingly V/W is a vector space!

Theorem 1.84. *V/W is a vector space with the operations $+,\cdot$.*

We will prove a few lemmas that will collectively imply Theorem 1.84.

Lemma 1.85. $[v_1]_W + [v_2]_W = [v_1 + v_2]_W$.

Proof. First, we have

$$[v_1]_W + [v_2]_W = (v_1 + W) + (v_2 + W) \stackrel{\text{def}}{=} \{u_1 + u_2 \ : \ u_1 \in v_1 + W,\ u_2 \in v_2 + W\}.$$

By definition,

$$u_1 = v_1 + w_1, \quad u_2 = v_2 + w_2$$

for some $w_1, w_2 \in W$. If $u \in [v_1]_W + [v_2]_W$, then

$$u = u_1 + u_2 = (v_1 + w_1) + (v_2 + w_2) = (v_1 + v_2) + (w_1 + w_2) = (v_1 + v_2) + w_3$$

where $w_3 = w_1 + w_2 \in W$. Hence, $u \in (v_1 + v_2) + W = [v_1 + v_2]_W$. If $u \in [v_1 + v_2]_W$, then

$$u = (v_1 + v_2) + w$$

for some $w \in W$. However

$$(v_1 + v_2) + w = (v_1 + 0_V) + (v_2 + w).$$

As $v_1 + 0_V \in [v_1]_W$ and $v_2 + w \in [v_2]_W$, we see that $u \in [v_1]_W + [v_2]_W$. ♠

When $\alpha = 0$, we require additional comments on the definition of $\alpha \cdot [v]_W$. In this case, if $S \subset V$ is any subset, then $0 \cdot S = \{0_V\}$. On V/W, we define

$$0 \cdot [v]_W \stackrel{\text{def}}{=} [0_V]_W.$$

Lemma 1.86. $\alpha \cdot [v]_W = [\alpha \cdot v]_W$.

Proof. First, we have

$$\alpha \cdot [v]_W \stackrel{\text{def}}{=} \{\alpha \cdot z \ : \ z \in [v]_W\}.$$

Since $[v]_W = v + W$, we see that

$$\alpha \cdot [v]_W = \{\alpha \cdot (v + w) \ : \ w \in W\}.$$

Given $u \in \alpha \cdot [v]_W$, we see that

$$u = \alpha \cdot (v + w) = \alpha \cdot v + \alpha \cdot w = (\alpha \cdot v) + w'$$

for some $w \in W$ and with $w' = \alpha \cdot w$. Hence $u \in [\alpha \cdot v]$. If $u \in [\alpha \cdot v]$, then

$$u = \alpha \cdot v + w.$$

The case when $\alpha = 0$ is clear. Otherwise,

$$u = \alpha \cdot (v + \left(\frac{1}{\alpha}\right) \cdot w) = \alpha \cdot (v + w')$$

where $w' = \left(\frac{1}{\alpha}\right) \cdot w$. Hence, $u \in \alpha \cdot [v]_W$. ♠

Lemma 1.87. $[0_V]_W + [v]_W = [v]_W$.

Proof. Given $u \in [0_V]_W + [v]_W$, by definition and Lemma 1.83, we have

$$u = w_1 + (v + w_2) = v + (w_1 + w_2) = v + w_3$$

for some $w_1, w_2 \in W$ and $w_3 \in W$ given by $w_3 = w_2 + w_1$. Hence $u \in [v]_W$. If $u \in [v]_W$, then

$$u = v + w = (0_V + 0_V) + (v + w)$$

for some $w \in W$. Hence $u \in [0_V]_W + [v]_W$. ♠

Lemma 1.88. $[v]_W + [-v]_W = [0_V]_W$.

Proof. Given $u \in [v]_W + [-v]_W$, we know that

$$u = (v + w_+) + (-v + w_-) = (v - v) + (w_+ + w_-) = w_+ + w_-$$

for some $w_+, w_- \in W$. Hence by Lemma 1.83, $u \in [0_V]_W$. If $u \in [0_V]$, then

$$u = 0_V + w = (v - v) + w = (v - v) + w + 0_V = (v + w) + (-v + 0_V)$$

for some $w \in W$. Hence $u \in [v]_W + [-v]_W$. ♠

Lemma 1.89.

$$\alpha \cdot [v]_W + \beta \cdot [v]_W = (\alpha + \beta) \cdot [v]_W.$$

Proof. This amounts to proving

$$\alpha \cdot (v+W) + \beta \cdot (v+W) = (\alpha+\beta) \cdot (v+W).$$

For that, we have

$$\begin{aligned} \alpha \cdot (v+W) + \beta \cdot (v+W) &= (\alpha \cdot v + \alpha \cdot W) + (\beta \cdot v + \beta \cdot W) \\ &= (\alpha \cdot v + W) + (\beta \cdot v + W) \\ &= (\alpha+\beta) \cdot v + W \\ &= (\alpha+\beta) \cdot v + (\alpha+\beta) \cdot W \\ &= (\alpha+\beta) \cdot (v+W). \end{aligned}$$

♠

Remark 8. We used that W was a vector space when we replaced $\alpha \cdot W$ and $\beta \cdot W$ with W in the argument above. We also used it when we replaced W with $(\alpha+\beta) \cdot W$.

Collectively this proves V/W is a vector space.

Definition 1.90 (Quotient Space)**.** Given a vector space V and $W \le V$, we call the vector space V/W the **quotient vector space of V by W**.

We have a surjective function $L_W : V \to V/W$ given by $L_W(v) = [v]_W$.

Lemma 1.91. *L_W is a surjective linear function.*

Summary of Quotients: We call L_W the **quotient map**. The quotient space construction can be viewed as the following. Given a vector space V and subspace $W \le V$, there is a vector space V/W and a surjective linear function $L_W : V \to V/W$ such that $\ker(L_W) = W$.

Proof. This follows from Lemma 1.85 and Lemma 1.86. ♠

Corollary 1.92. *Given any vector space V and subspace W, there exists a surjective linear function $L\colon V \to U$ such that* $\ker(L) = W$.

Proof. Take $U = V/W$ and $L = L_W$. ♠

Lemma 1.93. *If V is a vector space with vector subspaces $U \le W \le V$, then $W/U \le V/U$.*

Proof. This is straightforward. ♠

Exercise 34. Prove Lemma 1.93.

1.6 The Isomorphism Theorems

In this section, we state and prove the three isomorphisms theorems involving quotient spaces.

1.6.1 First Isomorphism Theorem

We start from the first/main isomorphism theorem. The first isomorphism theorem is used in the proof of the second isomorphism theorem, and the first/second isomorphism theorems are used in the proof of the third isomorphism theorem.

Theorem 1.94 (First Isomorphism Theorem: Formal)**.** *If V is a vector space and $L\colon V \to W$ is a linear function, then there exists a bijective linear function*

$$L'\colon V/\ker(L) \longrightarrow L(V)$$

such that the following diagram commutes

$$\begin{tikzcd} V \arrow[r, "L"] \arrow[d, "L_{\ker(L)}"'] & L(V) \\ V/\ker(L) \arrow[ur, "L'"'] & \end{tikzcd}$$

Specifically,

$$L = L' \circ L_{\ker(L)}.$$

Equivalently,

$$L_{\ker(L)} = (L')^{-1} \circ L.$$

Proof. We define $L'\colon V/\ker(L) \to L(V)$ by $L'([v]) = L(v)$. In order for this to be well-defined, we need to prove that if $v' \in [v]$, then $L(v') = L(v)$. Since $v' \in [v]$, we know that $v \sim v'$. By definition of $\sim$, this implies that $v - v' \in \ker(L)$. Hence $L(v-v') = L(v) - L(v') = 0_W$ and so L' is well defined. Note that

$$L'(L_{\ker(L)}(v)) = L'([v]) = L(v)$$

and so

$$L' \circ L_{\ker(L)} = L.$$

It remains to prove that L' is a linear bijection. For linearity, we see that

$$L'([v_1]+[v_2]) = L'([v_1+v_2]) = L(v_1)+L(v_2) = L'([v_1])+L'([v_2])$$

and

$$L'(\alpha \cdot [v]) = L'([\alpha \cdot v]) = L(\alpha \cdot v) = \alpha \cdot L(v) = \alpha \cdot L'([v]).$$

For surjectivity, given $w \in L(V)$, by definition there exists $v \in V$ such that $L(v) = w$. Thus $L'([v]) = w$. For injectivity, given $[v_1],[v_2]$ with $L'([v_1]) = L'([v_2])$, we see that $L(v_1) = L(v_2)$. Hence $v_1 - v_2 \in \ker(L)$ and so $[v_1] = [v_2]$. ♠

Remark 9. One way to think about how we define L' is as follows. Since $L_{\ker(L)}$ is a surjective function, there exists an injective function $f\colon V/\ker(L) \to V$ such that

$$L_{\ker(L)} \circ f = \mathrm{Id}_{V/\ker(L)}.$$

That is

$$L_{\ker(L)}(f([v]) = [v].$$

Hence $f([v]) \in V$ satisfies $f([v]) \in [v]$ since $L_{\ker(L)}(f([v])) = [f([v])] = [v]$. We define L' via

$$\begin{tikzcd} V \arrow[r, "L"] & L(V) \\ V/\ker(L) \arrow[u, "f"] & \end{tikzcd}$$

The function $L' = L \circ f$ and does not depend on the choice of f; we proved that it does not when we proved that L' is well defined. So we have

$$\begin{tikzcd} V \arrow[r, "L"] & L(V) \\ V/\ker(L) \arrow[u, "f"] \arrow[ur, "L'"'] & \end{tikzcd}$$

The function f is sometimes called a **section of the function** $L_{\ker(L)}$. Diagrammatically, we have

$$\begin{tikzcd}
V \arrow[d, bend right=60, "L_{\ker(L)}"'] \\
V/\ker(L) \arrow[u, bend right=60, "f"'] \arrow[loop below, "\mathrm{Id}"]
\end{tikzcd}$$

Exercise. Prove that if $f\colon X \to Y$ is a surjective function, then there exists an injective function $g\colon Y \to X$ such that $f \circ g = \mathrm{Id}_Y$.

Corollary 1.95. *If $L\colon V \to W$ is a surjective linear function, then $W \cong V/\ker(L)$.*

Corollary 1.96. *If $L_1\colon V \to W_1$ and $L_2\colon V \to W_2$ are linear functions with $\ker(L_1) = \ker(L_2)$, then $L_1(V) \cong L_2(V)$.*

Corollary 1.97. *If $L\colon V \to W$ is a linear function with $U = \ker(L)$, then $V/U \cong L(V)$.*

1.6.2 Second Isomorphism Theorem

Theorem 1.98 (Second Isomorphism Theorem)**.** *If V is a vector space with $S \le U \le V$, then*

$$\frac{V}{U} \cong \frac{V/S}{U/S}.$$

Proof. We have

$$\begin{tikzcd}
V \arrow[r, "L_S"] \arrow[rr, bend left, "L_{U/S} \circ L_S"] & V/S \arrow[r, "L_{V/U}"] & \frac{V/S}{U/S}.
\end{tikzcd}$$

By the First Isomorphism Theorem, it suffices to prove that $U = \ker(L_U) = \ker(L_{U/S} \circ L_S)$.

Given $u \in U$, we see that $L_S(u) \in U/S$ since $u \in U$. Since $\ker(L_{U/S}) = U/S$, we see that $L_{U/S}(L_S(u)) = 0$. Hence $u \in \ker(L_{U/S} \circ L_S)$. Thus $U \subset \ker(L_{U/S} \circ L_S)$.

If $u \in \ker(L_{U/S} \circ L_S)$, then $L_{U/S}(L_S(u)) = 0$. Hence $L_S(u) \in \ker(L_{U/S}) = U/S$. Since $L_S(u) \in U/S$, then $u \in U$ since $S \subset U$. Thus $\ker(L_{U/S} \circ L_S) = \ker(L_U) = U$. ♠

Corollary 1.99. *If $L\colon V \to W$ is a linear function with $U \leq V$ and $\ker(L) \leq U$, then*

$$\frac{V}{U} \cong \frac{L(V)}{L(U)}.$$

Remark 10. We used the First Isomorphism Theorem to show that

$$\frac{V/S}{U/S} \cong \frac{V}{U}.$$

In particular, there exists an isomorphism

$$\Phi\colon \frac{V/S}{U/S} \longrightarrow \frac{V}{U}$$

such that the following diagram commutes

$$\begin{tikzcd}
 & & \frac{V/S}{V/U} \arrow[ddd, bend right, "\Phi"'] \\
 & \frac{V}{S} \arrow[ur, "L_{U/S}"] & \\
V \arrow[ur, "L_S"] \arrow[uurr, bend left, "L_{U/S}\circ L_S"] \arrow[drr, "L_U"'] & & \\
 & & \frac{V}{U} \arrow[uuu, bend right, "\Phi^{-1}"']
\end{tikzcd}$$

1.6.3 Third Isomorphism Theorem

Theorem 1.100 (Diamond Isomorphism Theorem aka DIT)**.** *If V is a vector space and $S,U \leq V$ are vector subspaces, then*

$$(S+U)/U \cong S/(S\cap U).$$

Recall that

$$S+U \stackrel{\text{def}}{=} \{s+u \;:\; s\in S,\ u\in U\}.$$

When S,U are vector subspaces, then $S+U$ is a vector subspace.

Remark 11. It is worth noting that in the actual isomorphism, the vector space V places no visible role. It does only depend on $S+U$. However, this operation does depend on V. One can view the DIT isomorphism theorem as measuring the difference between $S+U$ and $S\times U$. This will be more visible in our discussion below.

The main difficulty in proving DIT is that $(S+U)/U$ and $S/(S\cap U)$ are not obviously related. One general approach to proving two vectors spaces V_1,V_2 are isomorphic is to find a third vector space V_3 such that $V_1,V_2\cong V_3$. The relationship in the DIT goes through the product space $S\times U$.

General Setup: Given a vector space V, we can view the addition operation as a linear function

$$+\colon V\times V\longrightarrow V.$$

We also have two natural projection operations $\mathrm{P}_R,\mathrm{P}_L\colon V\times V\to V$ given by

$$\mathrm{P}_R(v,w)=w,\quad \mathrm{P}_L(v,w)=v.$$

In total, we have

$$\begin{tikzcd} & V\times V \arrow[dl, "\mathrm{P}_L"'] \arrow[d, "+"] \arrow[dr, "\mathrm{P}_R"] & \\ V & V & V \end{tikzcd}$$

and each map is surjective. The kernel of the linear function $V\times V\to V$ given by addition is

$$\Delta=\ker(+)=\{(v,-v)\ :\ v\in V\}.$$

Specialization: Given $S,U\leq V$, we can restrict addition from $V\times V$ to $S\times U$. This yields the diagram:

$$\begin{tikzcd} & S\times U \arrow[dl, "\mathrm{P}_L"'] \arrow[d, "+"] \arrow[dr, "\mathrm{P}_R"] & \\ S & S+U & U \end{tikzcd}$$

The kernel of the linear function given by restricting addition to $S\times U$ is

$$\Delta_{S,U}=\Delta\cap(S\times U)=\{(v,-v)\ :\ v\in S\cap U\}.$$

By the First Isomorphism Theorem, we know that

$$(S\times U)/\Delta_{S,U}\cong S+U.$$

Remark 12. Given $S,U \leq V$, the subspace $\Delta_{S,U}$ measures the failure of $S\times U \cong S+U$. However, it can be that $S\times U \cong S+U$ and $\Delta_{S,U} \neq \{0\}$.

Proof. We will prove that

$$\frac{S}{S\cap U} \cong \frac{S\times U}{(S\cap U)\times U} \cong \frac{S+U}{U}.$$

This will be done using two diagrams. Note that edges without arrows are inclusion maps where the arrow points up.

Diagram 1: We will first prove $\frac{S\times U}{(S\cap U)\times U} \cong \frac{S+U}{U}$.

$$\begin{tikzcd}
 & S\times U \arrow[dl, dash] \arrow[dr, dash] \arrow[dd, "L"] & \\
S\times(S\cap U) \arrow[dd, "L"] & & (S\cap U)\times U \arrow[dd, "L"] \\
 & S+U \arrow[dl, dash] \arrow[dr, dash] \arrow[dd, "L_U"] & \\
S \arrow[dr, "L_U"'] & & U \arrow[ddl, "L_U"] \\
 & \frac{S+U}{U} \arrow[d, dash] & \\
 & 0 &
\end{tikzcd}$$

We use $L\colon S\times U \to S+U$ denote the addition function. Since $\ker(L) \leq (S\cap U)\times U$, the Second Isomorphism Theorem implies that

$$\frac{S\times U}{(S\cap U)\times U} \cong \frac{(S\times U)/\ker(L)}{((S\cap U)\times U)/\ker(L)} \cong \frac{L(S\times U)}{L((S\cap U)\times U)}.$$

By the First Isomorphism Theorem, we have

$$\frac{S\times U}{\ker(L)} \cong S+U, \quad \frac{(S\cap U)\times U}{\ker(L)} \cong U.$$

Hence

$$\frac{(S\times U)/\ker(L)}{((S\cap U)\times U)/\ker(L)}\cong\frac{S+U}{U}$$

and so

$$\frac{S\times U}{(S\cap U)\times U}\cong\frac{S+U}{U}.$$

Diagram 2: Next we will prove $\frac{S}{S\cap U}\cong\frac{S\times U}{(S\cap U)\times U}$.

$$\begin{tikzcd}
 & & S\times U \arrow[dll, no head] \arrow[drr, no head] \arrow[dd, "\mathrm{P}_L"] & & \\
S\times(S\cap U) \arrow[drr, "\mathrm{P}_L"'] & & & & (S\cap U)\times U \arrow[dd, "\mathrm{P}_L"] \\
 & & S \arrow[drr, no head] \arrow[dd, "L_{S\cap U}"'] & & \\
 & & & & S\cap U \arrow[ddll, "L_{S\cap U}"] \\
 & & \frac{S}{S\cap U} \arrow[d, no head] & & \\
 & & 0 & &
\end{tikzcd}$$

Since $\ker(\mathrm{P}_L)\le(S\cap U)\times U$, by the Second Isomorphism Theorem, we have

$$\frac{S\times U}{(S\cap U)\times U}\cong\frac{(S\times U)/\ker(\mathrm{P}_L)}{((S\cap U)\times U)/\ker(\mathrm{P}_L)}\cong\frac{\mathrm{P}_L(S\times U)}{\mathrm{P}_L((S\cap U)\times U)}.$$

By the First Isomorphism, we have

$$\frac{S\times U}{\ker(\mathrm{P}_L)}\cong S,\quad \frac{(S\cap U)\times U}{\ker(\mathrm{P}_L)}\cong S\cap U.$$

Hence

$$\frac{(S\times U)/\ker(\mathrm{P}_L)}{((S\cap U)\times U)/\ker(\mathrm{P}_L)}\cong\frac{S}{S\cap U}$$

and so

$$\frac{S\times U}{(S\cap U)\times U}\cong\frac{S}{S\cap U}.$$

♠

Remark 13. The main point in the proof is that we can prove the following string of isomorphisms

$$\frac{S+U}{U} \cong \frac{L(S\times U)}{L((S\cap U)\times U)} \cong \frac{S\times U}{(S\cap U)\times U} \cong \frac{\mathrm{P}_L(S\times U)}{\mathrm{P}_L((S\cap U)\times U)} \cong \frac{S}{S\cap U}.$$

Remark 14. The diamond in the diamond isomorphism theorem is

$$\begin{tikzcd}
 & S+U \arrow[dl, dash] \arrow[dr, dash] & \\
S \arrow[dr, dash] & & U \arrow[dl, dash] \\
 & S\cap U &
\end{tikzcd}$$

Note that the parallel edges in the above diamond are isomorphic. We have (the rather lovely) diagram that encapsulates everything:

$$\begin{tikzcd}
 & & S+U \arrow[rr, "L_U"] \arrow[dl, dash] \arrow[dr, dash] & & (S+U)/U \arrow[d, dash] \arrow[llll, bend right, "\cong"'] \\
S/(S\cap U) \arrow[d, dash] & S \arrow[l, "L_{S\cap U}"'] \arrow[dr, dash] & & U \arrow[r] \arrow[dl, dash] & 0 \\
0 & & S\cap U \arrow[ll] & &
\end{tikzcd}$$

Remark 15. Using the diagrams below, one can prove

$$\frac{S+U}{S} \cong \frac{U}{S\cap U}.$$

The details of the argument are the same. Specifically, using the diagrams below together with the First and Second Isomorphism Theorems, we can prove the string of isomorphisms

$$\frac{S+U}{S} \cong \frac{L(S\times U)}{L(S\times(S\cap U))} \cong \frac{S\times U}{S\times(S\cap U)} \cong \frac{\mathrm{P}_R(S\times U)}{\mathrm{P}_R(S\times(S\cap U))} \cong \frac{U}{S\cap U}.$$

Below are the two relevant diagrams for proving this.

Diagram 1:

$$\begin{tikzcd}
 & & S\times U \arrow[dll, dash] \arrow[drr, dash] \arrow[dd, "L"] & & \\
S\times(S\cap U) \arrow[dd, "L"'] & & & & (S\cap U)\times U \arrow[dd, "L"] \\
 & & S+U \arrow[dll, dash] \arrow[drr, dash] \arrow[dd, "L_S"] & & \\
S \arrow[ddrr, "L_S"'] & & & & U \arrow[dll, "L_S"] \\
 & & \frac{S+U}{S} \arrow[d, dash] & & \\
 & & 0 & &
\end{tikzcd}$$

Diagram 2:

$$
\begin{tikzcd}
 & & S \times U \arrow[dll, no head] \arrow[drr, no head] \arrow[dd, "\mathrm{P}_R"] & & \\
S \times (S \cap U) \arrow[dd, "\mathrm{P}_R"'] & & & & (S \cap U) \times U \arrow[dll, "\mathrm{P}_R"] \\
 & & U \arrow[dll, no head] \arrow[dd, "L_{S \cap U}"] & & \\
S \cap U \arrow[ddrr, "L_{S \cap U}"'] & & & & \\
 & & \frac{U}{S \cap U} \arrow[d, no head] & & \\
 & & 0 & &
\end{tikzcd}
$$

Chapter 2

Bases

In this chapter, we develop the foundational material on bases for vector spaces. Bases will provide use with an important invariant of vector spaces called the dimension. The existence of a basis and the independence of the cardinality of a basis are the two main technical results. The third main formal result is the universal mapping property for bases.

As this is algebraic, we will speak only of what are called **Hamel bases** for vector spaces.

2.1 Linear Dependence/Independence and Spanning

2.1.1 Linear Dependence/Independence

Recall that given a vector space V, $S \subset V$, and a function $\alpha\colon S \to \mathbf{R}$, we will write $\alpha_v \stackrel{\text{def}}{=} \alpha(v)$ and

$$\operatorname{supp}(\alpha) \stackrel{\text{def}}{=} \{v \in S \,:\, \alpha_v \neq 0\}\,.$$

In particular, α has **finite support** when $|\operatorname{supp}(\alpha)| < \infty$.

Definition 2.1 (Linear Dependence)**.** Given a vector space V and a subset $S \subset V$, we say S is **linearly dependent** if there exists a non-zero α_v with finite support such that

$$\sum_{v \in S} \alpha_v \cdot v = 0_V$$

Definition 2.2 (Linear Independence)**.** Given a vector space V and a subset $S \subset V$, we say S is **linearly independent** if when

$$\sum_{v \in S} \alpha_v \cdot v = 0_V,$$

for some α_v with finite support, then α_v is zero (i.e. $\alpha_v = 0$ for all $v \in S$).

We remark that if $S \subset V$ is linearly independent, then $0_V \notin S$. Indeed, if $0_V \in S$, then we see that

$$\lambda \cdot 0_V = 0_V$$

for $\lambda \neq 0$ which implies that S cannot be linearly independent.

Exercise 35. Given vectors $v, w \in \mathbf{R}^2$ with $v, w \neq 0_{\mathbf{R}^2}$, prove that $\{v, w\}$ is linearly dependent if and only if $v = \alpha \cdot w$ for some $\alpha \in \mathbf{R}$.

Exercise 36. Prove that if $v, w, u \in \mathbf{R}^3$ are linearly independent, then $v_1, w_1, u_1 \in \mathbf{R}^3$ are linearly independent where

$$v_1 = v - w, \quad w_1 = w - u, \quad u_1 = u.$$

Exercise 37. Prove that if $S_1, S_2 \subset V$ are linearly independent sets, then $S_1 \cap S_2$ is linearly independent (see also the more general Lemma 2.3 below).

Exercise 38. Let $v, w \in \mathbf{R}^2$ be linearly independent vectors. View $v = (v_1, v_2)$ and $w = (w_1, w_2)$. Define

$$S_1 = \{v, w\}, \quad S_2 = \{v + w, v - w\}.$$

Prove that S_2 is linearly independent. Prove that $S_1 \cup S_2$ is linearly dependent.

Exercise 39. Prove that $1, x, x^2, x^3, \dots, x^d$ are linearly independent in $\mathrm{Poly}_d(\mathbf{R})$.

Exercise 40. Prove that $A, B \in \mathrm{M}(2, \mathbf{R})$ given by

$$A = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{pmatrix}, \quad B = \begin{pmatrix} \beta_{1,1} & \beta_{1,2} \\ \beta_{2,1} & \beta_{2,2} \end{pmatrix}$$

are linearly dependent if and only if there exists $\lambda \in \mathbf{R}$ such that $\lambda \cdot A = B$.

Exercise 41. Prove that $A, B \in \mathrm{M}(2, \mathbf{R})$ given by

$$A = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{pmatrix}, \quad B = \begin{pmatrix} \beta_{1,1} & \beta_{1,2} \\ \beta_{2,1} & \beta_{2,2} \end{pmatrix}$$

with $AB \neq BA$ are linearly independent.

Lemma 2.3. *If $S \subset T \subset V$ and T is linearly independent, then S is linearly independent.*

Exercise 42. Prove Lemma 2.3.

Exercise 43. Given a vector space V and a subset $S \subset V$, define

$$L\colon \operatorname{Fun}_{\text{fin}}(S, \mathbf{R}) \longrightarrow \operatorname{Span}(S)$$

by

$$L(\alpha_v) \stackrel{\text{def}}{=} \sum_{v \in S} \alpha_v \cdot v.$$

Prove the following are equivalent:

(a) S is linearly independent.

(b) L is an isomorphism.

In particular, $\operatorname{Span}(S) \cong \operatorname{Fun}_{\text{fin}}(S, \mathbf{R})$ when S is linearly independent.

Definition 2.4 (Maximal Linearly Independent Subsets)**.** Given a vector space V and a subset $S \subset V$, we say S is a **maximal linearly independent set** if the following two conditions are satisfied:

(i) S is linearly independent.

(ii) If $S \subset S'$ and S' is linearly independent, then $S = S'$.

2.1.2 Spanning

Definition 2.5 (Spanning Set)**.** Given a vector space V and subset $S \subset V$, we say that S **spans** V or S **is a spanning set** if $\operatorname{Span}(S) = V$.

Exercise 44. If V is a vector space and $v \in V$, prove that

$$\operatorname{Span}(v) = \{\alpha \cdot v \;:\; \alpha \in \mathbf{R}\}.$$

Exercise 45. If V is a vector space and $v, w \in V$, prove that

$$\operatorname{Span}(v, w) = \operatorname{Span}(v)$$

if and only if v, w are linearly dependent.

Lemma 2.6. *If $S \subset T \subset V$, then* $\mathrm{Span}(S) \subset \mathrm{Span}(T)$. *In particular, if S is a spanning set, then T is a spanning set.*

Proof. Given $v_0 \in \mathrm{Span}(S)$, there exists α_v with finite support such that

$$v_0 = \sum_{v \in S} \alpha_v \cdot v.$$

Define β_v on T by

$$\beta_v \begin{cases} 0, & v \notin S \\ \alpha_v, & v \in S. \end{cases}$$

It follows that β_v has finite support and

$$v_0 = \sum_{v \in T} \beta_v \cdot v.$$

Hence $v_0 \in \mathrm{Span}(T)$ and so $\mathrm{Span}(S) \subset \mathrm{Span}(T)$. If S spans, then $V = \mathrm{Span}(S) \subset \mathrm{Span}(T) \subset V$. Hence T spans. ♠

Definition 2.7 (Minimal Spanning Subsets)**.** Given a vector space V and a subset $S \subset V$, we say S is a **minimal spanning set** if the following two conditions are satisfied:

(i) S is a spanning set.

(ii) If $S' \subset S$ and S' is a spanning set, then $S = S'$.

2.1.3 Relationship Between Dependence/Independence and Spanning

Lemma 2.8. *If V is a vector space, $S \subset V$, and $v_0 \in \mathrm{Span}(S)$, then*

$$\mathrm{Span}(S) = \mathrm{Span}(S \cup \{v_0\}).$$

Proof. As $v_0 \in \mathrm{Span}(S)$, there exists α_v with finite support such that

$$v_0 = \sum_{v \in S} \alpha_v \cdot v.$$

As $S \subset S \cup \{v_0\}$, we see that $\mathrm{Span}(S) \subset \mathrm{Span}(S \cup \{v_0\})$. It remains to prove that $\mathrm{Span}(S \cup \{v_0\}) \subset \mathrm{Span}(S)$. Given $u \in \mathrm{Span}(S \cup \{v_0\})$, there exists β_v with finite support and $\beta_{v_0} \in \mathbf{R}$ such that

$$u = \beta_{v_0} \cdot v_0 + \sum_{v \in S} \beta_v \cdot v = \beta_{v_0} \cdot \left(\sum_{v \in S} \alpha_v \cdot v \right) + \sum_{v \in S} \beta_v \cdot v = \sum_{v \in S} (\beta_{v_0} \alpha_v + \beta_v) \cdot v.$$

In particular, we see that $u \in \mathrm{Span}(S)$ by definition of spans. ♠

Lemma 2.9. *If V is a vector space and S is a linearly dependent set, then there exists $v_0 \in S$ such that*

$$\mathrm{Span}(S) = \mathrm{Span}(S - \{v_0\}).$$

Proof. As S is linearly dependent, there exists a non-zero α_v with finite support such that

$$\sum_{v \in S} \alpha_v \cdot v = 0_V$$

As α_v is non-zero, there exists $v_0 \in S$ with $\alpha_{v_0} \neq 0$. In particular,

$$-\alpha_{v_0} \cdot v_0 = \sum_{v \in S - \{v_0\}} \alpha_v \cdot v.$$

Scalar multiplying both sides by $-1/\alpha_{v_0}$, we see that

$$v_0 = \sum_{v \in S - \{v_0\}} \left(-\frac{\alpha_v}{\alpha_{v_0}} \right) \cdot v.$$

In particular, we see that $v_0 \in \mathrm{Span}(S - \{v_{j_0}\})$. Thus, $\mathrm{Span}(S) = \mathrm{Span}(S - \{v_{j_0}\})$ by Lemma 2.8. ♠

Lemma 2.10. *If V is a vector space and S is a linearly independent set, then $u \notin \mathrm{Span}(S - \{u\})$ for every $u \in S$. In particular, if $T_1, T_2 \subset S$ and $\mathrm{Span}(T_1) = \mathrm{Span}(T_2)$, then $T_1 = T_2$.*

Proof. If $u \in S$ and $v \in \mathrm{Span}(S - \{u\})$, then there exists α_v (defined on $S - \{u\}$) with finite support such that

$$u = \sum_{v \in S - \{u\}} \alpha_v \cdot v.$$

In particular, we see that

$$-u + \sum_{v \in S - \{u\}} \alpha_v \cdot v = 0_V.$$

If we define $\alpha_u = -1$, we can extend α_v to all of S. This function is non-zero, has finite support, and satisfies

$$\sum_{v \in S} \alpha_v \cdot v = 0_V.$$

However, since S is linearly independent, this is impossible. Hence, $u \notin \mathrm{Span}(S - \{u\})$. ♠

Lemma 2.11. *If V is a vector space, $S \subset V$ is a linearly independent set, and $u \notin \mathrm{Span}(S)$, then $S \cup \{u\}$ is a linearly independent set.*

Proof. Given

$$\sum_{v \in S \cup \{u\}} \alpha_v \cdot v = 0_V$$

where α_v has finite support, we must prove that α_v is zero. If $\alpha_u = 0$, then the linear independence of S implies α_v must be zero. Otherwise, we can assume that $\alpha_u \neq 0$. In this case, we see that

$$-\alpha_u \cdot u = \sum_{v \in S} \alpha_v \cdot v.$$

and so

$$u = \sum_{v \in S} \left(-\frac{\alpha_v}{\alpha_u} \right) \cdot v.$$

However, this is impossible since $u \notin \mathrm{Span}(S)$. Hence $S \cup \{u\}$ is linearly independent. ♠

2.1.4 Maximal Independent and Minimal Spannings Sets

Theorem 2.12. *If V is a vector space and $S \subset V$ is a linearly independent set, then there exists a maximal linearly independent set S' such that $S \subset S'$.*

Proof. We give a quasi-rigorous, intuitive proof of this claim. Any rigorous proof of this result in this generality requires the axiom of choice (or an equivalent axiom (e.g. Zorn's Lemma)).

We will build the set S' as follows. If $\mathrm{Span}(S) = V$, then S must be maximal. Indeed, if $\mathrm{Span}(S) = V$ and $v_0 \in V - S$, then since S spans, there exists α_v with finite support such that

$$v_0 = \sum_{v \in S} \alpha_v \cdot v.$$

However,

$$v_0 - \sum_{v \in S} \alpha_v \cdot S = 0_V.$$

Hence $S \cup \{v_0\}$ is not linearly independent. Hence, if S spans, then we take $S' = S$. If $\mathrm{Span}(S) \neq V$, we choose $v \in V - \mathrm{Span}(S)$. By Lemma 2.11, $S \cup \{v\}$ is linearly independent. If $\mathrm{Span}(S \cup \{v\}) = V$, then $S \cup \{v\}$ must be maximal by the argument above. Otherwise, $\mathrm{Span}(S \cup \{v\}) \neq V$ and we continue until we reach $S' \subset V$ that is linearly independent and $\mathrm{Span}(S') = V$. As above, we know that S' must be maximal linearly independent and $S \subset S'$ by construction. ♠

Remark 16. Running this process, we must choose some $v' \in V - \mathrm{Span}(S)$. The axiom of choice (by definition of the name of the axiom) allows us to select such an element.

Theorem 2.13. *If V is a vector space and $S \subset V$ is a spanning set, then there exists a minimal spanning set S' such that $S' \subset S$.*

Proof. As with the proof of Theorem 2.12, we will give a quasi-rigorous proof. We will build S' as follows. If S is linearly independent, then we assert S is a minimal spanning set. Indeed, if $S - \{v_0\}$ spans V, then there exists α_v (defined on $S - \{v_0\}$) with finite support such that

$$v_0 = \sum_{v \in S - \{v_0\}} \alpha_v \cdot v.$$

In particular,

$$v_0 - \sum_{v \in S - \{v_0\}} \alpha_v \cdot v = 0_V$$

and so S cannot be linearly independent. Hence, if S is linearly independent, we take $S' = S$. Otherwise, S is linearly dependent. By Lemma 2.9, there exists $v \in S$ such that $\mathrm{Span}(S) = \mathrm{Span}(S - \{v\})$. In particular, $S - \{v\}$ spans V. If $S - \{v\}$ is linearly independent, then $S - \{v\}$ must be a minimal spanning set. Otherwise, we continue until we reach S' as desired. ♠

Remark 17. Running this process requires that we be able to choose a non-zero point in the set

$$\left\{ \alpha_v \in \mathrm{Fun}_{\mathrm{fin}}(S, \mathbf{R}) \ : \ \sum_{v \in S} \alpha_v \cdot v = 0_V \right\}.$$

The axiom of choice allows us to select such an element.

Theorem 2.14. *If V is a vector space and $S \subset V$ is a maximal linearly independent set, then S is a spanning set.*

Proof. We will prove this via contradiction. If S is not a spanning set, then there exists $v \in V - \mathrm{Span}(S)$. By Lemma 2.11, we see that $S \cup \{v\}$ is a linearly independent set that contains S and is not equal to S. This contradicts the maximality of S. Hence, S must span. ♠

Theorem 2.15. *If V is a vector space and $S \subset V$ is a minimal spanning set, then S is linearly independent.*

Proof. We will prove this by contradiction. If S is linearly dependent, by Lemma 2.9, there exists $v_0 \in S$ such that $\mathrm{Span}(S) = \mathrm{Span}(S - \{v_0\})$. Since S spans, we see that $S - \{v_0\}$ spans. This contradicts the minimality of S and so S must be linearly independent. ♠

2.2 Basis

2.2.1 Definition and Basic Bases Examples

Definition 2.16 (Basis)**.** Given a vector space V and a subset $\mathscr{B} \subset V$, we say that $\mathscr{B}$ is a **basis for** V if the following two conditions hold:

(i) $\mathscr{B}$ is linearly independent.

(ii) $\mathscr{B}$ is a spanning set.

Example 14. Let $V = \mathbf{R}$. Then $\{\alpha\}$ is a basis for $\mathbf{R}$ for any $\alpha \in \mathbf{R}$ with $\alpha \neq 0$. For this, note that if $\alpha, \beta \in \mathbf{R}$ and $\alpha \neq 0$, then there exists a unique $\lambda \in \mathbf{R}$ such that $\lambda\alpha = \beta$ (i.e. $\lambda = \frac{\beta}{\alpha}$).

Example 15. Let $V = \mathbf{R}^2$. Take $e_1 = (1,0)$ and $e_2 = (0,1)$. Then $\{e_1, e_2\}$ is a basis. To see that $\{e_1, e_2\}$ is a spanning set, given $v = (x_1, x_2) \in \mathbf{R}^2$, we see that

$$x_1 \cdot e_1 + x_2 \cdot e_2 = v.$$

Hence, $\mathrm{Span}(\{e_1, e_2\}) = \mathbf{R}^2$ and so $\{e_1, e_2\}$ spans. If

$$\alpha_1 \cdot e_1 + \alpha_2 \cdot e_2 = 0_{\mathbf{R}^2} = (0,0)$$

Then $(\alpha_1, \alpha_2) = (0,0)$ and so $\alpha_1 = \alpha_2 = 0$. Thus, $\{e_1, e_2\}$ is linearly independent and spanning. Therefore, $\{e_1, e_2\}$ is a basis for $\mathbf{R}^2$.

Exercise 46. Given a linearly independent subset $\{v_1, v_2\} \subset \mathbf{R}^2$, prove that $\{v_1, v_2\}$ is a basis.

Exercise 47. Define

$$\delta_{j,k} \overset{\text{def}}{=} \begin{cases} j \neq k, & 0, \\ j = k, & 1. \end{cases}$$

Let $e_j \in \mathbf{R}^n$ be given by $e_j = (\delta_{j,1}, \delta_{j,2}, \dots, \delta_{j,n})$. For example, e_3 in $\mathbf{R}^6$ is $(0,0,1,0,0,0)$ or e_1 in $\mathbf{R}^{11}$ is $(1,0,0,0,0,0,0,0,0,0,0)$. Prove that $\{e_1, \dots, e_n\}$ is a basis for $\mathbf{R}^n$.

2.2.2 Existence of Bases

Theorem 2.17. *Given a vector space V and a subset $\mathscr{B} \subset V$, the following are equivalent:*

(i) *$\mathscr{B}$ is a basis for V.*

(ii) *$\mathscr{B}$ is a maximal linearly independent set.*

(iii) *$\mathscr{B}$ is a minimal spanning set.*

Proof. By Theorem 2.14 and Theorem 2.15, it suffices to prove that a maximal linearly independent set is a minimal spanning set and that a minimal spanning set is a maximal linearly independent set. If S is a maximal linearly independent set and there exists $v_0 \in S$ such that $\operatorname{Span}(S) = \operatorname{Span}(S - \{v_0\}) = V$, then there exists α_v (defined on $S - \{v_0\}$) with finite support such that

$$v_0 = \sum_{v \in S - \{v_0\}} \alpha_v \cdot v.$$

In particular,

$$-v_0 + \sum_{v \in S - \{v_0\}} \alpha_v \cdot v = 0_V.$$

However, this contradicts that S is linearly independent. Thus $S - \{v_0\}$ cannot span for any $v_0 \in S$ and so S is a minimal spanning set. Next, assume that S is a minimal spanning set and $S \subset S'$ where S' is linearly independent. Since S spans V, we know that $S \cup \{v\}$ also spans V. In particular, $\operatorname{Span}(S) = \operatorname{Span}(S \cup \{v\})$ and so $S = S'$ by Lemma 2.10. Therefore, S is a maximal linearly independent set. ♠

Corollary 2.18 (Existence of Bases)**.** *If V is a vector space, then there exists a subset $\mathscr{B} \subset V$ that is a basis for V.*

Proof. This follows from Theorem 2.12 (Theorem 2.13) and Theorem 2.14 (Theorem 2.15). ♠

Corollary 2.19. *If V is a vector space and $S \subset V$ is a linearly independent set, then there exists a basis $\mathscr{B}$ of V with $S \subset \mathscr{B}$.*

Proof. This follows from Theorem 2.12 and Theorem 2.17. ♠

Corollary 2.20. *If V is a vector space and $S \subset V$ is a spanning set, then there exists a basis $\mathscr{B}$ of V with $\mathscr{B} \subset S$.*

Proof. This follows from Theorem 2.13 and Theorem 2.17. ♠

Lemma 2.21. *If V is a vector space, $S \subset V$ is linearly independent, and $v_0 \in \mathrm{Span}(S)$, there exist unique $\alpha_v \in \mathbf{R}$ for each $v \in S$ such that*

$$v_0 = \sum_{v \in S} \alpha_v \cdot v.$$

Proof. Since $v_0 \in \mathrm{Span}(S)$, there exist α_v with finite support such that

$$v_0 = \sum_{v \in S} \alpha_v \cdot v.$$

If there exist β_v with finite support such that

$$v_0 = \sum_{v \in S} \beta_v \cdot v$$

then

$$v_0 - v_0 = 0_V = \sum_{v \in S} \alpha_v \cdot v - \sum_{v \in S} \beta_v \cdot v = \sum_{v \in S} (\alpha_v - \beta_v) \cdot v.$$

Since S is linearly independent, we see that $\alpha_v - \beta_v = 0$ for all $v \in S$. Thus, $\alpha_v = \beta_v$ for all $v \in S$ as desired. ♠

Theorem 2.22. *If V is a vector space, $\mathscr{B}$ is a basis for V, and $v_0 \in V$, then there exist unique α_v with finite support such that*

$$v_0 = \sum_{v \in \mathscr{B}} \alpha_v \cdot v.$$

Proof. As $\mathscr{B}$ spans V, we know that $\mathrm{Span}(\mathscr{B}) = V$. By Lemma 2.21, there exist unique α_v with finite support such that

$$v_0 = \sum_{v \in \mathscr{B}} \alpha_v \cdot v.$$

♠

Lemma 2.23. *If V is a vector space and $S \subset V$, then the following are equivalent:*

(i) S is linearly independent.

(ii) For each $v_0 \in \mathrm{Span}(S)$, then there exist unique α_v with finite support such that

$$v_0 = \sum_{v \in S} \alpha_v \cdot v.$$

Proof. (i) implies (ii) follows from Lemma 2.21. For (ii) implies (i), if

$$\sum_{v\in S}\alpha_v\cdot S=0_V$$

for some α_v with finite support, it follows that α_v must be zero. Indeed, if α_v is non-zero, then we would have

$$0_V=\sum_{v\in S}\alpha_v\cdot v=\sum_{v\in S}0\cdot v,$$

contradicting uniqueness. ♠

Corollary 2.24. *If V is a vector space and $\mathscr{B}\subset V$, then the following are equivalent:*

(i) $\mathscr{B}$ is a basis for V.

(ii) For every $v_0\in V$, then there exist unique α_v with finite support such that

$$v_0=\sum_{v\in\mathscr{B}}\alpha_v\cdot v.$$

Proof. (i) implies (ii) follows from Theorem 2.22. For (ii) implies (i), we see that $\mathscr{B}$ spans by assumption. That $\mathscr{B}$ is linearly independent follows from Lemma 2.23. ♠

2.2.3 Universal Mapping Property for Bases

The universal mapping property for bases reduces the study of linear functions from $V\to W$ to the the study of (general) functions $\mathscr{B}_V\to W$ where $\mathscr{B}_V$ is a fixed basis for V. Before stating and proving the universal mapping property for bases, we will prove that linear functions $L\colon V\to W$ are determined by their values on a fixed basis $\mathscr{B}_V$ for V.

Lemma 2.25. *If $L_1,L_2\colon V\to W$ are linear functions and $L_1(v)=L_2(v)$ for each $v\in\mathscr{B}$ and some basis $\mathscr{B}$ for V, then $L_1(v)=L_2(v)$ for all $v\in V$.*

Proof. If $L_1(v)=L_2(v)$ for every $v\in\mathscr{B}$ for some basis $\mathscr{B}$ for V, we must show that $L_1=L_2$. By Corollary 2.22, given $v_0\in V$, there exist unique α_v with finite support such that

$$v_0=\sum_{v\in\mathscr{B}}\alpha_v\cdot v.$$

Now, we have

$$L_1(v_0) = L_1\left(\sum_{v\in\mathscr{B}} \alpha_v \cdot v\right) = \sum_{v\in\mathscr{B}} \alpha_v \cdot L_1(v)$$
$$= \sum_{v\in\mathscr{B}} \alpha_v \cdot L_2(v) = L_2\left(\sum_{v\in\mathscr{B}} \alpha_v \cdot v\right) = L_2(v_0)$$

as desired. ♠

We will make extensive use of the following result.

Theorem 2.26 (Universal Mapping Property: Basis)**.** *If V,W are vector spaces, $\mathscr{B}$ is a basis for V, and $f\colon \mathscr{B} \to W$ is a function, then there exists a unique linear function $L = L_f\colon V \to W$ such that $L(v) = f(v)$ for all $v \in \mathscr{B}$.*

Proof. For $u \in V$, we know that

$$u = \sum_{v\in\mathscr{B}} \alpha_v \cdot v$$

for a unique β_v with finite support. Set $w_v = f(v)$ for $v \in \mathscr{B}$ and define $L\colon V \to W$ by

$$L(u) \overset{\text{def}}{=} \sum_{v\in\mathscr{B}} \alpha_v \cdot w_v.$$

The linearity of L follows from the definition of L while uniqueness follows from Lemma 2.25. ♠

Remark 18. The function f can be injective while the function L is not. Take $V = \mathbf{R}^2$ and $W = \mathbf{R}$. Define $f(e_1) = 0$ and $f(e_2) = 1$. The L associated to f is not injective since $L(0_{\mathbf{R}^2}) = 0 = L(e_1)$.

Remark 19. Given a function $f\colon \mathscr{B} \to W$, by Theorem 2.26, there exists a unique linear function $L_f \in \mathrm{Hom}(V,W)$ such that $L_f(v) = f(v)$ for all $v \in \mathscr{B}$. Since each linear function $L\colon V \to W$ induces a function $f_L\colon \mathscr{B} \to W$ via $f_L(v) = L(v)$ for $v \in \mathscr{B}$, we see that the function $\mathrm{Fun}(\mathscr{B},W) \to \mathrm{Hom}(V,W)$ induced by $f \mapsto L_f$ is an isomorphism. Note in the case $W = \mathbf{R}$, we have that $\mathrm{Hom}(V,\mathbf{R}) \cong \mathrm{Fun}(\mathscr{B},\mathbf{R})$.

Exercise 48. Prove that $f \mapsto L_f$ is an isomorphism between $\mathrm{Fun}(\mathscr{B},W)$ and $\mathrm{Hom}(V,W)$.

2.3 Dimension

In this section, we introduce a foundational concept in linear algebra. Namely, the concept of the dimension of a vector space. The dimension of a vector space will be defined to be the cardinality of a basis for the vector space. In order for this to be well defined, we must prove that the cardinalities of any two bases for a vector space are the same. For this, we will need the concept of the cardinality/size of a set.

Definition 2.27 (Cardinality)**.** Let X, Y be sets.

- We write $|X| = |Y|$ if there exists a bijective function $f: X \to Y$. We say X, Y have the same **cardinality** or **size** in this case.
- We write $|X| \leq |Y|$ if there exists an injective function $f: X \to Y$.
- We write $|X| \geq |Y|$ if there exists a surjective function $f: X \to Y$.
- We write $|X| < |Y|$ if there exists an injective function $f: X \to Y$ but no such bijective function.
- We write $|X| > |Y|$ if there exists a surjective function $f: X \to Y$ but no such bijective function.

Remark 20. If $|X| \leq |Y|$ and $|Y| \leq |X|$, then $|X| = |Y|$. This fact is non-trivial and requires the axiom of choice. We will assume this result without further discussion for sake of brevity.

To prove the main result of this section (i.e. Theorem 2.29), we require the following fact that we will not prove.

Fact 2.28. *If S_1, S_2 are infinite linearly subsets of V with $|S_1| < |S_2|$, then $|\mathrm{Span}(S_1)| < |\mathrm{Span}(S_2)|$.*

We are now ready to state and prove the main result of this section. This result is sometimes referred to as the Steinitz Exchange Lemma.

Theorem 2.29 (Steinitz Exchange Lemma)**.** *If V is a vector space, $S_1 \subset V$ is linearly independent, and $S_2 \subset V$ spans, then $|S_1| \leq |S_2|$.*

Proof. We have three cases to resolve.

Case 1: S_1 is finite and S_2 is infinite. Then $|S_1| < |S_2|$ and we win.

Case 2: S_1 and S_2 are finite. This case is classical. By Theorem 2.13, we know that S_2 contains a minimal spanning set S_4 and that S_1 is contained in a maximal linearly independent set S_3 by Theorem 2.12. As $S_1 \subset S_3 \subset S_4 \subset S_2$, it suffices to prove that $|S_3| \leq |S_4|$. To do this, we will produce a sequence of bases $S_3 = \mathscr{B}_0, \mathscr{B}_1, \ldots, \mathscr{B}_m \subset S_4$ such that $\left|\mathscr{B}_j\right| = |S_3|$ for all j. Write

$$S_3 = \{v_1, \ldots, v_r\}, \quad S_4 = \{w_1, \ldots, w_s\}.$$

As S_3 is a basis, by Theorem 2.22, there exist unique $\alpha_1, \ldots, \alpha_r \in \mathbf{R}$ such that

$$w_1 = \sum_{j=1}^{r} \alpha_j \cdot v_j.$$

Since $w_1 \neq 0$, $\alpha_{j_1} \neq 0$ for some $j_1 \in \{1, \ldots, r\}$. Relabeling the vectors $\{v_1, \ldots, v_r\}$, we can assume that $j_1 = 1$. Hence, we have

$$\begin{aligned} w_1 &= \sum_{j=1}^{r} \alpha_j \cdot v_j \\ w_1 &= \alpha_1 \cdot v_1 + \sum_{j=2}^{r} \alpha_j \cdot v_j \\ -\alpha_1 \cdot v_1 &= -w_1 + \sum_{j=2}^{r} \alpha_j \cdot v_j \\ v_1 &= \left(\frac{1}{\alpha_1}\right) \cdot w_1 - \sum_{j=2}^{r} \left(\frac{\alpha_j}{\alpha_1}\right) \cdot v_j. \end{aligned}$$

We set

$$\mathscr{B}_1 \stackrel{\text{def}}{=} \{v_2, \ldots, v_r, w_1\}.$$

We claim that $\mathscr{B}_1$ is a basis. First, note that $\mathscr{B}_1$ spans since $S_3 \subset \operatorname{Span}(\mathscr{B}_1)$ and S_3 spans by Lemma 1.40. If

$$\beta_{w_1} \cdot w_1 + \sum_{j=2}^{r} \beta_j \cdot v_j = 0_V$$

for some $\beta_2, \ldots, \beta_r, \beta_{w_1} \in \mathbf{R}$, we see that

$$\beta_{w_1} \cdot \left(\sum_{j=1}^{r} \alpha_j \cdot v_j\right) + \sum_{j=2}^{r} \beta_j \cdot v_j = 0_V.$$

$$(\beta_{w_1}\alpha_1) \cdot v_1 + \sum_{j=2}^{r} (\beta_{w_1}\alpha_j + \beta_j) \cdot v_j = 0_V.$$

Since S_3 is linearly independent, we see that $\beta_{w_1}\alpha_1 = 0$ and so $\beta_{w_1} = 0$ since $\alpha_1 \neq 0$. We also know that

$$\beta_{w_1}\alpha_j + \beta_j = \beta_j = 0$$

for all $j \in \{2,\dots,r\}$. Thus $\mathscr{B}_1$ is linearly independent. Now, since $\mathscr{B}_1$ is a basis, by Theorem 2.22, there exist unique $\alpha_2,\dots,\alpha_r,\beta_1$ such that

$$w_2 = \beta_1 \cdot w_1 + \sum_{j=2}^{r} \alpha_j \cdot v_j.$$

If $\alpha_2 = \cdots = \alpha_r = 0$, then $w_2 = \beta_1 \cdot w_2$. However, $w_1, w_2 \in S_4$ which is linearly independent. Thus some $\alpha_{j_2} \neq 0$ for some $j_2 \in \{2,\dots,r\}$. Relabeling the vectors v_j, we can assume that $j_2 = 2$. As before, we can solve for v_2 and obtain

$$v_2 = -\left(\frac{\beta_1}{\alpha_2}\right) \cdot w_1 + \left(\frac{1}{\alpha_2}\right) \cdot w_2 - \sum_{j=3}^{r} \left(\frac{\alpha_j}{\alpha_2}\right) \cdot v_j.$$

Define

$$\mathscr{B}_2 \stackrel{\text{def}}{=} \{v_3,\dots,v_r,w_1,w_2\}.$$

Since $\mathscr{B}_1 \subset \operatorname{Span}(\mathscr{B}_2)$, we see that $\mathscr{B}_2$ spans V by Lemma 1.40. If there exist $\theta_3,\dots,\theta_r,\lambda_1,\lambda_2 \in$ $\mathbf{R}$ such that

$$\lambda_1 \cdot w_1 + \lambda_2 \cdot w_2 + \sum_{j=3}^{r} \theta_j \cdot v_j = 0_V.$$

We see that

$$\lambda_1 \cdot w_1 + \lambda_2 \cdot w_2 + \sum_{j=3}^{r} \theta_j \cdot v_j = \lambda_1 \cdot w_1 + \lambda_2 \cdot \left(\beta_1 \cdot w_1 + \sum_{j=2}^{r} \alpha_j \cdot v_j\right) = 0_V.$$

In particular,

$$(\lambda_1 + \lambda_2\beta_2)w_1 + \lambda_2\alpha_2 \cdot v_2 + \sum_{j=3}^{r} (\lambda_2\alpha_j + \theta_j) \cdot v_j = 0_V.$$

Since $\mathscr{B}_1$ is linearly independent, we must have

$$\lambda_1 + \lambda_2\beta_2 = 0, \quad \lambda_2\alpha_2 = 0, \quad \lambda_2\alpha_j + \theta_j = 0$$

for all $j \in \{3,\dots,r\}$. We know that $\alpha_2 \neq 0$ and so $\lambda_2 = 0$. In this case, we see that

$$\lambda_1 \cdot w_1 + \sum_{j=3}^{r} \theta_j \cdot v_j = 0_V.$$

Since $\mathscr{B}_1$ is linearly independent, we see that this implies that $\lambda_1 = \theta_3 = \cdots = \theta_r = 0$. At the kth stage, we have a basis

$$\mathscr{B}_{j-1} = \{v_k, v_{k+1}, \ldots, v_r, w_1, \ldots, w_{k-1}\}.$$

By Theorem 2.22, there exist unique $\alpha_k, \ldots, \alpha_r, \beta_1, \ldots, \beta_{k-1} \in \mathbf{R}$ such that

$$w_k = \sum_{j=1}^{k-1} \beta_j \cdot w_j + \sum_{j=k}^{r} \alpha_j \cdot v_j.$$

Since S_4 is linearly independent, there must exist some $\alpha_{j_k} \neq 0$. Relabeling the vectors, we can assume that $j_k = k$. Hence,

$$v_k = \left(\frac{1}{\alpha_k}\right) \cdot w_k - \sum_{j=1}^{k-1} \left(\frac{\beta_j}{\alpha_k}\right) \cdot w_j - \sum_{k=j+1}^{r} \left(\frac{\alpha_j}{\alpha_k}\right) \cdot v_j.$$

Define

$$\mathscr{B}_k \stackrel{\text{def}}{=} \{v_{k+1}, \ldots, v_r, w_1, \ldots, v_r, w_1, \ldots, w_k\}.$$

As before, since $\mathscr{B}_{k-1} \subset \operatorname{Span}(\mathscr{B}_k)$, we see that $\mathscr{B}_k$ spans by Lemma 1.40. If

$$\sum_{j=1}^{k} \lambda_j \cdot w_j + \sum_{j=k+1}^{r} \theta_j \cdot v_j = 0$$

then

$$\sum_{j=1}^{k-1} \lambda_j \cdot w_j + \lambda_k \cdot w_k + \sum_{j=k+1}^{r} \theta_j \cdot v_j = \sum_{j=1}^{k-1} \lambda_j \cdot w_j + \lambda_k \cdot \left(\sum_{j=1}^{k-1} \beta_j \cdot w_j + \sum_{j=k}^{r} \alpha_j \cdot v_j \right) + \sum_{j=k+1}^{r} \theta_j \cdot v_j = 0_V.$$

Hence

$$\sum_{j=1}^{k-1} (\lambda_j + \lambda_k \beta_j) \cdot w_j + \lambda_k \alpha_k \cdot v_k + \sum_{j=k+1}^{r} (\lambda_k \alpha_j + \theta_j) \cdot v_j = 0_V.$$

Since $\mathscr{B}_{k-1}$ is linearly independent, we must have

$$\lambda_j + \lambda_k \beta_j = 0, \quad \lambda_k \alpha_k = 0, \quad \lambda_k \alpha_j + \theta_j = 0$$

for $j \in \{1, \ldots, k-1\}$ in the first equality and $j \in \{k+1, \ldots, r\}$ in the third inequality. Since $\alpha_k \neq 0$, we must have $\lambda_k = 0$. This implies $\lambda_j = 0$ for all j in the first equality and $\theta_j = 0$ for all j in the third equality. Hence $\mathscr{B}_j$ is a basis. Continuing until the rth stage, we see that $\mathscr{B}_r = \{w_1, \ldots, w_r\}$ and so $|S_3| \leq |S_4|$.

Case 3: S_1 and S_2 are both infinite. In this case, assume that $|S_1| > |S_2|$. Then there exists an injective function $f \colon S_2 \to S_1$ which is not surjective. By Fact 2.28, we know that $|\operatorname{Span}(S_2)| < |\operatorname{Span}(S_1)|$. However, since S_2 spans, we know that $|\operatorname{Span}(S_2)| \geq |\operatorname{Span}(S_1)|$ which is a contradiction. Hence $|S_1| \leq |S_2|$. ♠

The following is the main application needed to define the dimension of a vector space. This result is sometimes referred to as the Dimension Theorem.

Corollary 2.30 (Dimension Theorem)**.** *If V is a vector space with bases $\mathscr{B}_1$ and $\mathscr{B}_2$, then* $|\mathscr{B}_1| = |\mathscr{B}_2|$.

Proof. Applying Theorem 2.29 twice, we see that $|\mathscr{B}_1| \leq |\mathscr{B}_2|$ and $|\mathscr{B}_2| \leq |\mathscr{B}_1|$. Hence $|\mathscr{B}_1| = |\mathscr{B}_2|$. ♠

Definition 2.31 (Dimension)**.** Given a vector space V, we define the **dimension of** V to be the cardinality $|\mathscr{B}|$ of a basis for V. We denote this cardinality by $\dim(V)$ and occasionally by $\dim_{\mathbf{R}}(V)$ when emphasizing that V is a real vector space.

Example 16. $\dim(\mathbf{R}^n) = n$ since we saw that $\{e_1, \ldots, e_n\}$ is a basis for $\mathbf{R}^n$.

Example 17. $\dim_{\mathbf{Q}}(\mathbf{R}) = |\mathbf{R}|$.

Example 18. $\dim_{\mathbf{C}}(\mathbf{C}) = 1$, $\dim_{\mathbf{R}}(\mathbf{C}) = 2$, and $\dim_{\mathbf{Q}}(\mathbf{C}) = |\mathbf{C}| = |\mathbf{R}|$.

Lemma 2.32. *If V, W are vector spaces, then*

$$\dim(V \times W) = \dim(V) + \dim(W).$$

Proof. Take a basis $\mathscr{B}_V$ for V and $\mathscr{B}_W$ for W. Define $\mathscr{B}$ to be

$$\mathscr{B} \stackrel{\text{def}}{=} \{(v, 0_W) \in V \times W \ : \ v \in \mathscr{B}_V\} \cup \{(0_V, w) \in V \times W \ : \ w \in \mathscr{B}_W\}.$$

It is straightforward to check that $\mathscr{B}$ is a basis. By definition of $\mathscr{B}$ and cardinalities, we have

$$\dim(V \times W) = |\mathscr{B}| = |\mathscr{B}_V| + |\mathscr{B}_W| = \dim(V) + \dim(W).$$

♠

Lemma 2.33. *If $L\colon V \to W$ is an injective linear function and $S \subset V$ is linearly independent, then $L(S)$ is linearly independent.*

Proof. Assume that

$$\sum_{w \in L(S)} \alpha_w \cdot w = 0_W$$

where α_w has finite support. By definition, we know for each $w \in L(S)$, there exists $v \in S$ such that $L(v) = w$. In particular, we see that

$$\sum_{w \in L(S)} \alpha_w \cdot w = \sum_{v \in S} \alpha_w L(v) = L\left(\sum_{v \in S} \alpha_w \cdot v\right) = 0_W.$$

Thus,

$$\left(\sum_{v \in S} \alpha_w \cdot v\right) \in \ker(L).$$

Since L is injective, we know that $\ker(L) = 0_V$ by Proposition 1.70. Hence

$$\sum_{v \in S} \alpha_w \cdot v = 0_V$$

and the linearly independence implies that $\alpha_w = 0$ for all $w \in L(S)$. Hence, $L(S)$ is linearly independent. ♠

Lemma 2.34. *If $L \colon V \to W$ is surjective linear function and $S \subset V$ is a spanning set, then $L(S)$ is a spanning set.*

Proof. Given $w_0 \in W$, we need to show that

$$w_0 = \sum_{w \in L(S)} \alpha_w \cdot w$$

for some α_w with finite support. Since L is surjective, there exists $v_0 \in V$ such that $L(v_0) = w_0$. Since S spans V, we know that

$$v_0 = \sum_{v \in S} \alpha_v \cdot v$$

for some α_v with finite support. Now, we have

$$L(v_0) = L\left(\sum_{v \in S} \alpha_v \cdot v\right) = \sum_{v \in S} \alpha_v \cdot L(v) = \sum_{w \in L(S)} \alpha_v \cdot w$$

where we substituted $w = L(v)$ in the last step. In particular, $w_0 \in \mathrm{Span}(L(S))$ and so $L(S)$ is a spanning set. ♠

Theorem 2.35. *Let V, W be vector spaces and let $L \colon V \to W$ be a linear function.*

(1) If L is injective, then $\dim(V) \leq \dim(W)$.

(2) *If L is surjective, then* $\dim(V) \geq \dim(W)$.

Proof. For (1), if $\mathscr{B}_V$ is a basis for V, then $L(\mathscr{B}_V)$ is a linearly independent set by Lemma 2.33. By Corollary 2.19, there exists a basis $\mathscr{B}_W$ for W with $L(\mathscr{B}_V) \subset \mathscr{B}_W$. Since L is injective, we see that $|\mathscr{B}_V| \leq |\mathscr{B}_W|$ and so $\dim(V) \leq \dim(W)$.

For (2), if $\mathscr{B}_V$ is a basis for V, then $L(\mathscr{B}_V)$ is a spanning set for W by Lemma 2.34. By Corollary 2.20, there exists a basis $\mathscr{B}_W$ for W with $\mathscr{B}_W \subset L(\mathscr{B}_V)$. Since L restricted to $\mathscr{B}_V$ surjects $L(\mathscr{B}_V)$ as well, we see that $|\mathscr{B}_W| \leq |\mathscr{B}_V|$ and so $\dim(V) \geq \dim(W)$. ♠

Corollary 2.36. *If $L\colon V \to W$ is injective and $\mathscr{B}_V \subset V$ is a basis for V, then there exists a basis $\mathscr{B}_{W,L}$ such that $L(\mathscr{B}_V) \subset \mathscr{B}_W$.*

Proof. Since L is injective and $\mathscr{B}_V$ is linearly independent, we know that $L(\mathscr{B}_V) \subset W$ is linearly independent by Lemma 2.33. By Corollary 2.19, there exists a basis $\mathscr{B}_{W,L}$ such that $L(\mathscr{B}_V) \subset \mathscr{B}_{W,L}$ as desired. ♠

Corollary 2.37. *If $L\colon V \to W$ is surjective and $\mathscr{B}_V \subset V$ is a basis for V, then there exists a basis $\mathscr{B}_{W,L}$ such that $\mathscr{B}_{W,L} \subset L(\mathscr{B}_V)$.*

Proof. Since L is surjective and $\mathscr{B}_V$ spans V, we know that $L(\mathscr{B}_V)$ spans W by Lemma 2.34. By Corollary 2.20, there exists a basis $\mathscr{B}_{W,L}$ such that $\mathscr{B}_{W,L} \subset L(\mathscr{B}_V)$ as needed. ♠

Corollary 2.38. *Given a linear function $L\colon V \to W$, the following are equivalent:*

(i) *L is injective.*

(ii) *For every linearly independent subset $S \subset V$, the subset $L(S) \subset W$ is linearly independent.*

Proof. For (i) implies (ii), we apply Lemma 2.33. For (ii) implies (i), we assume that $L(S) \subset W$ is linearly independent for every $S \subset V$ that is linearly independent. To show that L is injective, it suffice to prove that $\ker(L) = \{0_V\}$ by Proposition 1.70. If $v \in \ker(L)$ and $v \neq 0_V$, then $L(v) = 0_W$. However, the set $\{v\}$ is linearly independent but $\{L(v)\} = \{0_W\}$ is not. Thus, $v = 0_V$ and $\ker(L) = 0_V$. ♠

Corollary 2.39. *Given a linear function $L\colon V \to W$, the following are equivalent:*

(i) *L is surjective.*

(ii) *For every spanning set $S \subset V$, the subset $L(S) \subset W$ is a spanning set.*

Proof. For (i) implies (ii), we apply Lemmma 2.34. For (ii) implies (i), we assume that $L(S)$ spans W for every $S \subset V$ that spans. Given w_0, there exists α_w with finite support such that

$$w_0 = \sum_{w \in L(S)} \alpha_w \cdot w.$$

Setting $w = L(v)$, we see that

$$L\left(\sum_{v \in S} \alpha_w \cdot v\right) = \sum_{v \in S} \alpha_w \cdot L(v) = \sum_{w \in L(S)} \alpha_w \cdot w = w_0.$$

Hence L is surjective. ♠

Corollary 2.40. *Given a linear function $L \colon V \to W$, the following are equivalent:*

(i) *L is an isomorphism.*

(ii) *For every basis $\mathscr{B} \subset V$, the subset $L(\mathscr{B}) \subset W$ is a basis.*

Proof. If L is an isomorphism and $\mathscr{B}$ is a basis for V, then we know $L(\mathscr{B})$ is linearly independent by Corollary 2.38 and $L(\mathscr{B})$ spans W by Corollary 2.39. Hence $L(\mathscr{B})$ is a basis for W. If $L(\mathscr{B})$ is a basis for W for every basis $\mathscr{B}$ for V, then L is injective by Corollary 2.38 and L is surjective by Corollary 2.39. Hence L is a bijection and thus an isomorphism. ♠

Corollary 2.41. *If V, W are vector spaces and $V \cong W$, then $\dim(V) = \dim(W)$.*

Proof. This follows from Theorem 2.35. ♠

Theorem 2.42. *Given vector spaces V, W, the following are equivalent:*

(i) $\dim(V) = \dim(W)$.

(ii) $V \cong W$.

Proof. The reverse implication follows from Corollary 2.41. For the direct implication, if $\dim(V) = \dim(W)$, then $|\mathscr{B}_V| = |\mathscr{B}_W|$ for any basis $\mathscr{B}_V$ for V and any basis $\mathscr{B}_W$ for W. Hence, there exists a bijective function $f \colon \mathscr{B}_V \to \mathscr{B}_W \subset W$. By Theorem 2.26, there exists a unique linear function $L \colon V \to W$ such that $L(v) = f(v)$ for all $v \in \mathscr{B}_V$. As $L(\mathscr{B}_V) = \mathscr{B}_W$, by Corollary 2.40, we see that L is an isomorphism. ♠

Corollary 2.43. *If* $\dim(V) \leq \dim(W)$*, then there exists an injective linear function* $L\colon V \to W$.

Proof. If $\dim(V) \leq \dim(W)$ and $\mathscr{B}_V$, $\mathscr{B}_W$ are bases for V, W, then there exists an injective function $f\colon \mathscr{B}_V \to \mathscr{B}_W$. By Theorem 2.26, there exists a unique linear function $L\colon V \to W$ such that $L(v) = f(v)$ for all $v \in \mathscr{B}_V$. Finally, we see L is injective by Corollary 2.38. ♠

Corollary 2.44. *If* $\dim(V) \geq \dim(W)$*, then there exists a surjective linear function* $L\colon V \to W$.

Proof. If $\dim(V) \geq \dim(W)$ and $\mathscr{B}_V$, $\mathscr{B}_W$ are bases for V, W, then there exists an surjective function $f\colon \mathscr{B}_V \to \mathscr{B}_W$. By Theorem 2.26, there exists a unique linear function $L\colon V \to W$ such that $L(v) = f(v)$ for all $v \in \mathscr{B}_V$. Finally, we see L is surjective by Corollary 2.39. ♠

An important subclass of real vector spaces are the vector spaces with finite dimension. These vector spaces behave rather differently from the infinite dimensional vector spaces. We will record an important example (Theorem 2.46) after defining them.

Definition 2.45 (Finite Dimension)**.** We say V is finite dimensional if $\dim(V) < \infty$.

Theorem 2.46. *Let* V, W *be finite dimensional vector spaces with* $\dim(V) = \dim(W)$*. The following are equivalent for a linear function* $L\colon V \to W$.

(1) L is an isomorphism.

(2) L is injective.

(3) L is surjective.

Proof. (1) implies (2) is immediate.

For (2) implies (3), we assume L is injective and must prove that L is surjective. Let $n = \dim(V)$ and let $\mathscr{B} = \{v_1, \ldots, v_n\}$ be a basis for V. We know that

$$L(\mathscr{B}) = \{L(v_1), \ldots, L(v_n)\}$$

is linearly independent by Lemma 2.33. If $L(\mathscr{B})$ does not span, then there exists $w \in W$ with $w \notin \mathrm{Span}(L(\mathscr{B}))$. By Lemma 2.11, we see that $\{L(v_1), \ldots, L(v_n), w\}$ is a linearly independent set. By Corollary 2.19, there exists a basis $\mathscr{B}'$ for W with

$$\{L(v_1), \ldots, L(v_n), w\} \subset \mathscr{B}'. \tag{2.1}$$

By assumption, $\dim(V) = \dim(W)$ and so $|\mathscr{B}'| = n$. However, by (2.1), we see that $|\mathscr{B}'| \geq n+1$ which is impossible. Hence, we see that $L(\mathscr{B})$ spans. In particular, L is surjective by Corollary 2.39.

For (3) implies (1), we must show that L is injective. Given a basis $\mathscr{B} = \{v_1, \ldots, v_n\}$ for V, we know that $L(\mathscr{B})$ spans W by Lemma 2.34. If $L(\mathscr{B})$ is linearly dependent, then by Lemma 2.9, there exists $L(v_j)$ such that $\mathrm{Span}(L(\mathscr{B})) = \mathrm{Span}(L(\mathscr{B}) - \{L(v_j)\})$. By Corollary 2.20, there exists a basis $\mathscr{B}'$ for W with

$$\mathscr{B}' \subset \left\{L(v_1), \ldots, L(v_{j-1}), L(v_{j+1}), \ldots, L(v_n)\right\}. \tag{2.2}$$

By assumption, $\dim(V) = \dim(W)$ and so $|\mathscr{B}'| = n$. However, by (2.2), we see that $|\mathscr{B}'| \leq n-1$ which is impossible. Hence, we see that $L(\mathscr{B})$ is linearly independent. In particular, L is injective by Corollary 2.38. ♠

The following pair of corollaries can be used to reduce the amount of work required to prove a subset of a finite dimensional vector space is a basis.

Corollary 2.47. *If V is finite dimensional and S is linearly independent with $|S| = \dim(V)$, then S is a basis.*

Corollary 2.48. *If V is finite dimensional and S is a spanning set with $|S| = \dim(V)$, then S is a basis.*

Exercise 49. Let V be a finite dimensional vector space and $L\colon V \to V$ be a linear function such that $\dim(L(V)) < \dim(V)$. Prove that L is nilpotent (see Exercise 30). That is, prove that there exists n such that $L^n(v) = 0_V$ for all $v \in V$.

Exercise 50. Let V be a vector space with a countably infinite basis

$$\mathscr{B} = \{v_1, v_2, v_3, \ldots\}.$$

Define $f\colon \mathscr{B} \to V$ by $f(e_j) = e_{j+1}$ and let $L\colon V \to V$ to be the unique linear extension.

(a) Prove that $L^{k+1}(V) \leq L^k(V)$ and $L^{k+1}(V) \neq L^k(V)$ for all $k \geq 0$.

(b) Prove that L is injective.

2.4 Coordinate Systems Via Bases

If V is a vector space with a basis $\mathscr{B}$, we can define a **coordinate system** associated to $\mathscr{B}$. If $v_0 \in V$, by Theorem 2.22, there exists a unique α_v with finite support such that

$$v_0 = \sum_{v \in \mathscr{B}} \alpha_v \cdot v.$$

We call α_v the coefficient of the vth coordinate. Since α_v has finite support, there are only finitely many coordinates for v_0 that are non-zero. If $\dim(V) = n$, then

$$\mathscr{B} = \{v_1, \ldots, v_n\}.$$

Then

$$v_0 = \sum_{j=1}^{n} \alpha_j \cdot v_j.$$

We define the associated **coordinate vector for** v_0 to be the string with n real numbers given by

$$v_0 = (\alpha_1, \ldots, \alpha_n).$$

The vector operations in this coordinate system are simple. If $v, w \in V$ with

$$v = (\alpha_1, \ldots, \alpha_n), \quad w = (\beta_1, \ldots, \beta_n),$$

then

$$v + w = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n), \quad \alpha \cdot v = (\alpha\alpha_1, \ldots, \alpha\alpha_n).$$

Exercise 51. Let $V = \mathbf{R}^3$ and $v_1, v_2, v_3 \in \mathbf{R}^3$ given by

$$v_1 = (0,1,1), \quad v_2 = (1,0,1), \quad v_3 = (1,1,0).$$

(a) Prove $\{v_1, v_2, v_3\}$ is a basis for $\mathbf{R}^3$.

(b) Write $e_1 = (1,0,0)$ in coordinates for this basis.

Exercise 52. Prove that $\{1, x, x^2, \ldots, x^d\}$ is a basis for $\mathrm{Poly}_d(\mathbf{R})$.

Exercise 53. For each $k \in \{0, 1, \ldots, d\}$, define $P_k(x) \in \mathrm{Poly}_d(\mathbf{R})$ by

$$P_k(x) = \sum_{j=0}^{k} \alpha_{j,k} x^j$$

where $\alpha_{j,k} \in \mathbf{R}$ and $\alpha_{k,k} \neq 0$. Prove that $\{P_0(x), P_1(x), \ldots, P_d(x)\}$ is a basis for $\mathrm{Poly}_d(\mathbf{R})$.

Exercise 54. Prove that if V is finite dimension and $S \subset V$ is linearly independent with $|S| = \dim(V)$, then S is a basis.

Exercise 55. Prove that if V is finite dimension and $S \subset V$ spans V with $|S| = \dim(V)$, then S is a basis.

2.5 (Short) Exact Sequences of Vector Spaces

We will prove a well known theorem called the Rank–Nullity theorem. As it requires no additional effort, we will frame this material in terms of short exact sequences of vector spaces. Short exact sequences of vector spaces arise natural in many branches of mathematics.

Definition 2.49. Given vector spaces V_1, V_2, V_3 and linear functions $L_1 \colon V_1 \to V_2$, $L_2 \colon V_2 \to V_3$, we say that the sequence of linear functions is **exact** if $L_1(V_1) = \ker(L_2)$. We typically write

$$V_1 \xrightarrow{L_1} V_2 \xrightarrow{L_2} V_3$$

and say that this sequence is exact.

Definition 2.50 (Short Exact Sequence)**.** Given vector spaces V_1, V_2, V_3 and linear functions $L_1 \colon V_1 \to V_2$, $L_2 \colon V_2 \to V_3$, we say the sequence is **short exact** if

$$\begin{tikzcd}
\{0\} \arrow[r, bend right, "\iota_{0,V_1}"'] & V_1 \arrow[r, bend left, "L_1"] & V_2 \arrow[r, bend right, "L_2"'] & V_3 \arrow[r, bend left, "0_{\mathrm{Fun}(V_3,\{0\})}"] & \{0\}
\end{tikzcd}$$

that is exact at every triple. Namely

$$\{0\} \xrightarrow{\iota_{0,V_1}} V_1 \xrightarrow{L_1} V_2,$$

$$V_1 \xrightarrow{L_1} V_2 \xrightarrow{L_2} V_3,$$

$$V_2 \xrightarrow{L_2} V_3 \xrightarrow{0_{\mathrm{Fun}(V_3,\{0\})}} \{0\}$$

are all exact.

Exercise 56. Let $L_1 \colon \mathbf{R}^2 \to \mathbf{R}^3$ be given by $L_1(x,y) \stackrel{\text{def}}{=} (x,y,0)$ and $L_2 \colon \mathbf{R}^3 \to \mathbf{R}$ given by $L_2(x,y,z) \stackrel{\text{def}}{=} z$. Prove that

$$\begin{tikzcd}
\{0\} \arrow[r, bend right, "\iota_{0,\mathbf{R}^2}"'] & \mathbf{R}^2 \arrow[r, bend left, "L_1"] & \mathbf{R}^3 \arrow[r, bend right, "L_2"'] & \mathbf{R} \arrow[r, bend left, "0_{\mathrm{Fun}(\mathbf{R},\{0\})}"] & \{0\}
\end{tikzcd}$$

is a short exact sequence.

Lemma 2.51. *If the diagram below is a short exact sequence of vector spaces*

$$\begin{tikzcd} \{0\} \arrow[r, bend right, "\iota_{0,V_1}"'] & V_1 \arrow[r, bend left, "L_1"] & V_2 \arrow[r, bend right, "L_2"'] & V_3 \arrow[r, bend left, "0_{\mathrm{Fun}(V_3,\{0\})}"] & \{0\} \end{tikzcd}$$

then L_1 is injective and L_2 is surjective.

Proof. Since

$$\{0\} \xrightarrow{\iota_{0,V_1}} V_1 \xrightarrow{L_1} V_2,$$

is exact, we see that $\iota_{0,V_1}(0) = \ker(L_1)$. However, $\iota_{0,V_1}(0) = 0_{V_1}$ and so L_1 is injective by Proposition 1.70. Next, as

$$V_2 \xrightarrow{L_2} V_3 \xrightarrow{0_{\mathrm{Fun}(V_3,\{0\})}} \{0\}$$

is exact, we see that $L_2(V_2) = \ker(0_{\mathrm{Fun}(V_3,\{0\})})$. However, $\ker(0_{\mathrm{Fun}(V_3,\{0\})}) = V_3$ and so L_2 is surjective. ♠

Theorem 2.52 (Splitting Short Exact Sequences)**.** *If*

$$\begin{tikzcd} \{0\} \arrow[r, bend right, "\iota_{0,V_1}"'] & V_1 \arrow[r, bend left, "L_1"] & V_2 \arrow[r, bend right, "L_2"'] & V_3 \arrow[r, bend left, "0_{\mathrm{Fun}(V_3,\{0\})}"] & \{0\} \end{tikzcd}$$

is a short exact sequence of vector spaces, then $V_2 \cong V_1 \times V_3$.

Proof. By Lemma 2.51, we know that L_1 is injective and L_2 is surjective. By Corollary 2.36, there exists a basis $\mathscr{B}_1$ for V_1 such that $L_1(\mathscr{B}_1) \subset V_2$ is a linearly independent set. By Corollary 2.19, there exists a basis $\mathscr{B}_2$ for V_2 such that $L_1(\mathscr{B}_1) \subset \mathscr{B}_2$. Define

$$\mathscr{B}_3 \stackrel{\text{def}}{=} \{L_2(v) \in V_3 \ : \ L_2(v) \neq 0_V\}.$$

Notice that

$$\mathscr{B}_3 = L_2(\mathscr{B}_2 - \mathscr{B}_1).$$

We assert that $\mathscr{B}_3$ is a basis for V_3. As $\mathscr{B}_2$ spans and L_2 is surjective, we see that $\mathscr{B}_3$ spans V_3. Assume that we have α_w (defined on $\mathscr{B}_3$) with finite support such that

$$\sum_{w \in \mathscr{B}_3} \alpha_w \cdot w = 0_{V_3}.$$

By definition of $\mathscr{B}_3$, for each $w \in \mathscr{B}_3$, there exists $v \in \mathscr{B}_2 - \mathscr{B}_1$ such that $L_2(v) = w$. In particular, we have

$$\sum_{v \in \mathscr{B}_2 - \mathscr{B}_1} \alpha_w \cdot L_2(v) = L_2\left(\sum_{v \in \mathscr{B}_2 - \mathscr{B}_1} \alpha_w \cdot v \right) = 0_{V_3}.$$

This implies that

$$\left(\sum_{v \in \mathscr{B}_2 - \mathscr{B}_1} \alpha_w \cdot v \right) \in \ker(L_2).$$

By selection of $\mathscr{B}_1$, there exist β_u (defined on $\mathscr{B}_1$) with finite support such that

$$\sum_{u \in \mathscr{B}_1} \beta_u \cdot u = \sum_{v \in \mathscr{B}_2 - \mathscr{B}_1} \alpha_v \cdot v.$$

In particular, we have

$$\sum_{u \in \mathscr{B}_1} \beta_u \cdot u - \sum_{v \in \mathscr{B}_2 - \mathscr{B}_1} \alpha_v \cdot v = 0_{V_2}.$$

However, α_v, β_u must be zero by linear independence.

Define the function

$$L \colon V_2 \longrightarrow V_1 \times V_3$$

as follows. Every $v \in V_2$ can be expressed uniquely as

$$v = \sum_{u \in \mathscr{B}_1} \beta_u \cdot u + \sum_{w \in \mathscr{B}_2 - \mathscr{B}_1} \alpha_w \cdot w$$

where β_u, α_w have finite support. We define

$$L(v) \stackrel{\text{def}}{=} L\left(\sum_{u \in \mathscr{B}_1} \beta_u \cdot u + \sum_{w \in \mathscr{B}_2 - \mathscr{B}_1} \alpha_w \cdot w \right) = \left(\sum_{u \in \mathscr{B}_1} \beta_u \cdot L_1^{-1}(u), \sum_{w \in \mathscr{B}_2 - \mathscr{B}_1} \alpha_w \cdot L_2(w) \right).$$

Since $L(\mathscr{B}_2) \subset V_1 \times V_2$ is a basis, we see that L is an isomorphism by Corollary 2.40. ♠

Corollary 2.53 (Rank-Nullity Theorem)**.** *If*

$$\begin{tikzcd}
\{0\} \arrow[r, bend right, "\iota_{0,V_1}"'] & V_1 \arrow[r, bend left, "L_1"] & V_2 \arrow[r, bend right, "L_2"'] & V_3 \arrow[r, bend left, "0_{\mathrm{Fun}(V_3,\{0\})}"] & \{0\}
\end{tikzcd}$$

is a short exact sequence then

$$\dim(V_2) = \dim(V_1) + \dim(V_3).$$

In particular, if $\dim(V_1), \dim(V_3) < \infty$*, then* $\dim(V_2) < \infty$*.*

Proof. This follows from Lemma 2.32 and Theorem 2.52. ♠

Corollary 2.54 (Rank-Nullity Theorem: Classical)**.** *If* $L\colon V \to W$ *is a surjective linear map, then* $V \cong \ker(L) \times W$ *and* $\dim(V) = \dim(W) + \dim(\ker(L))$*.*

Proof. The sequence

$$\begin{tikzcd}
\{0\} \arrow[r, bend right, "\iota_{0,\ker(L)}"'] & \ker(L) \arrow[r, bend left, "\iota_{\ker(L),V}"] & V \arrow[r, bend right, "L"'] & W \arrow[r, bend left, "0_{\mathrm{Fun}(W,\{0\})}"] & \{0\}
\end{tikzcd}$$

is short exact. The proof is completed via Corollary 2.53. ♠

2.6 Complementary Subspaces and Decompositions

In this section, we discuss complementary subspaces. These vector subspaces allow us to decompose a vector space into two or more pieces.

Definition 2.55 (Complementary Subspaces)**.** Given a vector space V and subspaces U, V, we say that U, W are **complementary** if

(i) $U \cap W = \{0_V\}$.

(ii) $V = U + W$.

Lemma 2.56. *If $U, W \leq V$, then the following are equivalent:*

(i) $U \cap W = \{0_V\}$.

(ii) For each $v \in U + W$, there exist unique $u_v \in U$ and $w_v \in W$ such that $v = u_v + w_v$.

Exercise 57. Let $V = \mathbf{R}^3$ and define for

$$S_1 \stackrel{\text{def}}{=} \left\{(x,y,z) \in R^3 \ : \ \alpha_1 x + \beta_1 y + \theta_1 z = 0\right\}$$
$$S_2 \stackrel{\text{def}}{=} \left\{(x,y,z) \in R^3 \ : \ \alpha_2 x + \beta_2 y + \theta_2 z = 0\right\}$$
$$S_3 \stackrel{\text{def}}{=} \left\{(x,y,z) \in R^3 \ : \ \alpha_3 x + \beta_3 y + \theta_3 z = 0\right\}$$

for $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \theta_1, \theta_2, \theta_3 \in \mathbf{R}$. Find the set of all $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \theta_1, \theta_2, \theta_3 \in \mathbf{R}$ such that S_1 and $S_2 \cap S_3$ are complementary subspaces for $\mathbf{R}^3$

Proof. If $U \cap W \neq \{0_V\}$, then there exists $v \in U \cap W$ with $v \neq 0$. Now, $0_V \in U \cap W$ and

$$0_V = 0_V + 0_V = v - v.$$

Hence, 0_V does not have a unique representation. It remains to prove that each $v \in U + W$ has a unique representation when $U \cap W = \{0_V\}$. If

$$v = u_1 + w_1 = u_2 + w_2$$

for $u_1, u_2 \in U$ and $w_1, w_2 \in W$, then

$$v - v = (u_1 - u_2) + (w_1 - w_2) = 0_V$$

and so

$$u_1 - u_2 = w_2 - w_1.$$

However, we see that $u_1 - u_2, w_2 - w_1 \in U \cap W$ and so $u_1 - u_2 = w_2 - w_1 = 0_V$ as needed. ♠

Proposition 2.57. *If V is a vector space and U, W are complementary subspaces, then for each $v \in V$, there exist unique $u_v \in U$ and $w_v \in W$ such that*

$$v = u_v + w_v.$$

Proof. This follows immediately from Lemma 2.56. ♠

Theorem 2.58. *If V is a vector space and U, W are complementary subspaces, then for any basis $\mathscr{B}_U$ for U and any basis $\mathscr{B}_W$ for W, the set $\mathscr{B} = \mathscr{B}_U \cup \mathscr{B}_W$ is a basis for V.*

Proof. We see that $\mathscr{B}_U \cup \mathscr{B}_W$ by Proposition 2.57, Theorem 2.22, and Corollary 2.24. Specifically, for each $v_0 \in V$, by Proposition 2.57, there exists unique $u_{v_0} \in U$ and $w_{v_0} \in W$ such that $v = u_{v_0} + w_{v_0}$. By Theorem 2.22, there exist unique α_u, β_w with finite support such that

$$u_{v_0} = \sum_{u \in \mathscr{B}_U} \alpha_u \cdot u$$

and

$$w_{v_0} = \sum_{w \in \mathscr{B}_W} \beta_w \cdot w.$$

Hence

$$v_0 = \sum_{u \in \mathscr{B}_U} \alpha_u \cdot u + \sum_{w \in \mathscr{B}_W} \beta_w \cdot w$$

for unique α_u, β_w with finite support. Hence $\mathscr{B}_U \cup \mathscr{B}_W$ is a basis by Corollary 2.24. ♠

Corollary 2.59. *If V is a vector space and $U, W \leq V$ are complementary subspaces, then $V \cong U \times W$ and*

$$\dim(V) = \dim(U) + \dim(W).$$

Proof. This follows from Theorem 2.58. ♠

If $V = U \times W$, then we define

$$U_V \stackrel{\text{def}}{=} \{(u, 0_W) \in V \ : \ u \in U\}, \quad W_V \stackrel{\text{def}}{=} \{(0_U, w) \in V \ : \ w \in W\}.$$

Lemma 2.60. *If $V = U \times W$, then $V/U_V \cong W$ and $V/W_V \cong U$.*

Proof. It is clear that U_V, W_V are complementary in $V = U \times W$. Hence it follows by Corollary 2.59. ♠

Corollary 2.61. *If V is a vector space and $U, W \leq V$ are complementary subspaces, then $V/W \cong U$ and $V/U \cong W$.*

Proof. This follows from Theorem 2.58.

♠

Theorem 2.62. *If V is a vector space with $U, W \leq V$, then there exists bases $\mathscr{B}_{U \cap W}$, $\mathscr{B}_U$, and $\mathscr{B}_W$ of $U \cap W$, U, and W respectively such that*

(1) $\mathscr{B}_U \cup \mathscr{B}_W$ is a basis for V.

(2) $\mathscr{B}_{U \cap W} = \mathscr{B}_U \cap \mathscr{B}_W$.

Proof. We start with a basis $\mathscr{B}_{U \cap V}$ which exists by Corollary 2.18. By Corollary 2.19, there exists a basis $\mathscr{B}_U$ for U such that $\mathscr{B}_{U \cap W} \subset \mathscr{B}_U$. By Corollary 2.19, there exists basis $\mathscr{B}'_{U \cup W}$ for $U \cup W$. By Theorem 2.29, we can change $\mathscr{B}'_{U \cup W}$ to a desired basis $\mathscr{B}_{U \cup W}$ for $U \cup W$. Specifically, $\mathscr{B}_{U \cup W}$ it a basis for $U \cup W$ which contains the given $\mathscr{B}_{U \cap W}$ and also satisfies that $U \cap \mathscr{B}_{U \cup V}$ is a basis for U and $W \cap \mathscr{B}_{U \cup W}$ is is basis for W. ♠

Lemma 2.63. *If V is a vector space with $U, W \leq V$ and $V = U + W$, then $L_{U \cap W}(U), L_{U \cap W}(W)$ are complementary subspaces in $V/(U \cap W)$.*

Proof. This is clear. ♠

Given $U, W \leq V$, we define $\Delta_{U,W} \leq U \times W \leq V \times V$ to be

$$\Delta_{U,W} \overset{\text{def}}{=} \{(v, v) \in U \times W \ : \ v \in U \cap W\}.$$

Corollary 2.64. *If V is a vector space with subspaces $U, W \leq V$, then $V \cong (U \times W)/\Delta_{U,W}$ and*

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W).$$

Proof. This follows from Theorem 2.62. ♠

2.7 Matrix Representations of Linear Functions

In this section, we describe how one can associate a matrix to a linear function $L \colon V \to W$ provided we have bases for V, W. Our focus will rest largly with the case when V, W are both finite dimensional as this is the main classical interest.

2.7.1 Finite Dimensional

Given finite dimensional vector spaces V, W, a linear function $L\colon V \to W$, and bases $\mathscr{B}_V, \mathscr{B}_W$ for V, W, we can associate to L a matrix $A_L \in \mathrm{M}(\dim(W), \dim(V); \mathbf{R})$. We will set $n = \dim(V)$ and $m = \dim(W)$ so that $A_L \in \mathrm{M}(m, n; \mathbf{R})$. In particular, we have $\mathscr{B}_V = \{v_1, \ldots, v_n\}$ and $\mathscr{B}_W = \{w_1, \ldots, w_m\}$. For each $v_k \in \mathscr{B}_V$, we know that there exist unique $\beta_{1,k} \ldots, \beta_{m,k} \in \mathbf{R}$ such that

$$L(v_k) = \sum_{j=1}^{m} \beta_{j,k} \cdot w_j$$

Hence,

$$\begin{aligned}
L(v_1) &= \sum_{j=1}^{m} \beta_{j,1} \cdot w_j = \beta_{1,1} \cdot w_1 + \beta_{2,1} \cdot w_2 + \cdots + \beta_{m,1} \cdot w_m \\
L(v_2) &= \sum_{j=1}^{m} \beta_{j,2} \cdot w_j = \beta_{2,1} \cdot w_1 + \beta_{2,2} \cdot w_2 + \cdots + \beta_{m,2} \cdot w_m \\
&\vdots \\
L(v_k) &= \sum_{j=1}^{m} \beta_{j,k} \cdot w_j = \beta_{1,k} \cdot w_1 + \beta_{2,k} \cdot w_2 + \cdots + \beta_{m,k} \cdot w_m \\
&\vdots \\
L(v_n) &= \sum_{j=1}^{m} \beta_{j,n} \cdot w_j = \beta_{1,n} \cdot w_1 + \beta_{2,n} \cdot w_2 + \cdots + \beta_{m,n} \cdot w_m.
\end{aligned}$$

We define $A \in \mathrm{M}(m, n; \mathbf{R})$ by $A_{j,k} = \beta_{j,k}$. That is

$$A \stackrel{\text{def}}{=} \begin{pmatrix} \beta_{1,1} & \beta_{1,2} & \beta_{1,3} & \ldots & \beta_{1,n} \\ \beta_{2,1} & \beta_{2,2} & \beta_{2,3} & \ldots & \beta_{2,n} \\ \beta_{3,1} & \beta_{3,2} & \beta_{3,3} & \ldots & \beta_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta_{m,1} & \beta_{m,2} & \beta_{m,3} & \ldots & \beta_{m,n} \end{pmatrix}.$$

The matrix A depends on L, $\mathscr{B}_V$, and $\mathscr{B}_W$.

Definition 2.65 (Columns of a Matrix)**.** If $A \in \mathrm{M}(m, n; \mathbf{R})$, we define the j**th column of** A, which we denote by $C_{A,j}$ to be the vector $C_{A,j} \in W$ defined by

$$C_{A,j} \stackrel{\text{def}}{=} L(v_j) = \beta_{1,j} \cdot w_1 + \beta_{2,j} \cdot w_2 + \cdots + \beta_{m,j} \cdot w_m.$$

For simplicity, we will write $C_{A,j}$ in **coordinate form** as

$$C_{A,j} = (\beta_{1,j}, \beta_{2,j}, \dots, \beta_{m,j}).$$

Definition 2.66 (Rows of a Matrix)**.** If $A \in \mathrm{M}(m,n;\mathbf{R})$, we define the **$j$th row of** A, which we denote by $R_{A,j}$ to be the vector $R_{A,j} \in V$ defined by

$$R_{A,j} \stackrel{\text{def}}{=} \beta_{j,1} \cdot v_1 + \beta_{j,2} \cdot v_2 + \cdots + \beta_{j,n} \cdot v_n.$$

For simplicity, we will write $R_{A,j}$ in **coordinate form** as

$$R_{A,j} = (\beta_{j,1}, \beta_{j,2}, \dots, \beta_{j,n}).$$

Given $(\beta_1, \dots, \beta_n), (\alpha_1, \dots, \alpha_n)$, we define

$$\langle (\alpha_1, \dots, \alpha_n), (\beta_1, \dots, \beta_n) \rangle \stackrel{\text{def}}{=} \sum_{j=1}^{n} \alpha_j \beta_j.$$

Given $A \in \mathrm{M}(m,n;\mathbf{R})$ and $v = (\alpha_1, \dots, \alpha_n) \in V$ where

$$v = \sum_{k=1}^{n} \alpha_k \cdot v_k,$$

we define $Av \in W$ by

$$Av \stackrel{\text{def}}{=} \left(\left\langle R_{A,1}, v \right\rangle, \left\langle R_{A,2}, v \right\rangle, \dots, \left\langle R_{A,m}, v \right\rangle \right).$$

Specifically, if $Av = w$, then in coordinates w is given by

$$w = \left(\sum_{k=1}^{n} \beta_{1,k} \alpha_k, \sum_{k=1}^{n} \beta_{2,k} \alpha_k, \dots, \sum_{k=1}^{n} \beta_{m,k} \alpha_k \right).$$

Lemma 2.67. $Av = L(v)$.

Proof. As

$$v = \sum_{k=1}^{n} \alpha_j \cdot v_k$$

and

$$L(v_k) = \sum_{j=1}^{m} \beta_{j,k} \cdot w_j,$$

then

$$\begin{aligned} L(v) &= L\left(\sum_{k=1}^{n} \alpha_k \cdot v_k\right) = \sum_{k=1}^{n} \alpha_k \cdot L(v_k) \\ &= \sum_{k=1}^{n} \alpha_k \cdot \left(\sum_{j=1}^{m} \beta_{j,k} w_j\right) = \sum_{k=1}^{n}\sum_{j=1}^{m} (\alpha_k \beta_{j,k}) w_j. \end{aligned}$$

Expanding, we see that w_j appears n times and the total coefficient is given by

$$\alpha_1\beta_{j,1} + \alpha_2\beta_{j,2} + \cdots + \alpha_n\beta_{j,n}.$$

Hence $Av = L(v)$. ♠

We view A, the matrix associated to $L\colon V \to W$ as a coordinate representation of L in the coordinates associated to the bases $\mathscr{B}_V$ and $\mathscr{B}_W$.

Theorem 2.68. *If V, W are vector spaces with bases $\mathscr{B}_V$ and $\mathscr{B}_W$, then the function $\mathscr{M}\colon \mathrm{Hom}(V,W) \to \mathrm{M}(m,n;\mathbf{R})$ given by $\mathscr{M}(L) = A$, where A is the associated matrix, is an isomorphism.*

Proof. This is straightforward. ♠

Exercise 58. Prove Theorem 2.68.

If $L_1 : V \to W$ and $L_2 : W \to V$ with $\dim(V) = n_V$, $\dim(W) = n_W$, $\dim(U) = n_U$, then we have

$$\begin{tikzcd} V \arrow[r, bend right, "L_1"'] \arrow[rr, bend left, "L_2 \circ L_1"] & W \arrow[r, bend right, "L_2"'] & U. \end{tikzcd}$$

Given bases $\mathscr{B}_V = \{v_1, \ldots, v_{n_V}\}$, $\mathscr{B}_W = \{w_1, \ldots, w_{n_W}\}$, and $\mathscr{B}_U = \{u_1, \ldots, u_{n_U}\}$, we have matrices $A_1 \in \mathrm{M}(n_W, n_V; \mathbf{R})$, $A_2 \in \mathrm{M}(n_U, n_W; \mathbf{R})$, and $A_{1,2} \in \mathrm{M}(n_U, n_V; \mathbf{R})$ associated to L_1, L_2, and $L_2 \circ L_1$.

Definition 2.69. If $A \in \mathrm{M}(m,n;\mathbf{R})$ and $B \in \mathrm{M}(p,m;\mathbf{R})$, then we define $BA \in \mathrm{M}(p,n;\mathbf{R})$ to have (i,j)–coefficient $(BA)_{j,k}$ given by

$$(BA)_{j,k} \overset{\text{def}}{=} \langle R_{B,j}, C_{A,k} \rangle.$$

If we view $v \in \mathrm{M}(1,n;\mathbf{R})$ and $A \in \mathrm{M}(m,n;\mathbf{R})$, then we see that Av matches Definition 2.69. If $w \in \mathrm{M}(1,m;\mathbf{R})$ and $A \in \mathrm{M}(m,n;\mathbf{R})$, then we have $w^T \in \mathrm{M}(m,1;\mathbf{R})$ and $A^T \in \mathrm{M}(n,m;\mathbf{R})$. Hence $w^T A^T$ makes sense and $w^T A^T \in \mathrm{M}(n,1;\mathbf{R})$.

This next result is the motivation for the previous definition.

Theorem 2.70. *If $L_1 \colon V \to W$ and $L_2 \colon W \to V$ with $\dim(V) = n_V$, $\dim(W) = n_W$, $\dim(U) = n_U$, and bases $\mathscr{B}_V = \{v_1,\dots,v_{n_V}\}$, $\mathscr{B}_W = \{w_1,\dots,w_{n_W}\}$, and $\mathscr{B}_U = \{u_1,\dots,u_{n_U}\}$, then the associated matrices $A_1 \in \mathrm{M}(n_W,n_V;\mathbf{R})$, $A_2 \in \mathrm{M}(n_U,n_W;\mathbf{R})$, and $A_{1,2} \in \mathrm{M}(n_U,n_V;\mathbf{R})$ associated to L_1, L_2, and $L_2 \circ L_1$ satisfy*

$$A_2 A_1 = A_{1,2}.$$

Proof. This is one of those things that you should just check for yourself. ♠

Exercise 59. Prove Theorem 2.70.

Corollary 2.71. *If $L_1 \colon V \to W$ and $L_2 \colon W \to V$ with $\dim(V) = n_V$, $\dim(W) = n_W$, $\dim(U) = n_U$, and bases $\mathscr{B}_V = \{v_1,\dots,v_{n_V}\}$, $\mathscr{B}_W = \{w_1,\dots,w_{n_W}\}$, and $\mathscr{B}_U = \{u_1,\dots,u_{n_U}\}$, then the associated matrices $A_1 \in \mathrm{M}(n_W,n_V;\mathbf{R})$, $A_2 \in \mathrm{M}(n_U,n_W;\mathbf{R})$, and $A_{1,2} \in \mathrm{M}(n_U,n_V;\mathbf{R})$ associated to L_1, L_2, and $L_2 \circ L_1$ satisfy*

$$A_{1,2} v = A_2(A_1 v)$$

for all $v \in V$.

Proof. This follows immediately from Theorem 2.70. ♠

When $V = W$ and $L \colon V \to V$, then we can associate a matrix to L given any basis $\mathscr{B}$ of V. We simply proceed as above and take $\mathscr{B} = \mathscr{B}_V = \mathscr{B}_W$.

2.7.2 Infinite Dimensional

One can define something like a matrix for a linear function $L \colon V \to W$. Given a basis $\mathscr{B}_V$ for V and a basis $\mathscr{B}_W$ for W, for each $v \in \mathscr{B}_V$, there is a unique β_w^v with finite support on $\mathscr{B}_W$ such that

$$L(v) = \sum_{w \in \mathscr{B}} \beta_w^v \cdot w.$$

We can also define a function $\alpha_v^w \colon \mathscr{B}_V \to \mathbf{R}$ by $\alpha_v^w \stackrel{\text{def}}{=} \beta_w^v$. Note that unlike β_w^v, the function α_v^w need not have finite support. The β_w^v are generalizations of the column vectors of the associated matrix and are vectors in W. The α_v^w are generalizations of the row vectors of the associated matrix though they do not correspond to vectors in V unless the α_v^w have finite support. By Theorem 2.26, there is a unique linear function $L_{\alpha_v^w} : V \to \mathbf{R}$ and so we can view $\alpha_v^w \in \mathrm{Hom}(V, \mathbf{R})$ via identifying it with $L_{\alpha_v^w}$; recall this discussion in Remark 19. Hence, the column vector of L in the infinite dimensional setting are vectors in W and the row vectors for L are vectors in $V^* = \mathrm{Hom}(V, \mathbf{R})$. The space $V^* = \mathrm{Hom}(V, \mathbf{R})$ is called the dual space and is the subject of the next chapter.

2.8 Eigen-basis for Linear Self-Maps

In this section, we discuss the concept of an eigen-basis for a linear self-map and the concept of diagonalizing a linear function/matrix.

Definition 2.72 (Eigen-basis)**.** Given a linear function $L: V \to V$ and a basis $\mathscr{B}$ of V, we say that $\mathscr{B}$ is an **eigen-basis for** L **and** V if each $v \in \mathscr{B}$ is an eigenvector for L.

Definition 2.73 (Multiplicity of an Eigenspace)**.** Given a linear function $L: V \to V$ and $\lambda \in \mathrm{E}(L)$, we define the **multiplicity of** λ to be

$$m_{L,\lambda} \stackrel{\text{def}}{=} \dim(\mathrm{E}_{L,\lambda})$$

where $\mathrm{E}_{L,\lambda}$ is the λ–eigenspace of L.

Definition 2.74 (Simple Eigenvalue)**.** Given a linear function $L: V \to V$ and $\lambda \in \mathrm{E}(L)$, we say λ is a **simple eigenvalue** if $m_\lambda = 1$. That is, there is a λ–eigenvector v_0 such that for each $v \in \mathrm{E}_{L,\lambda}$, $v = \alpha \cdot v_0$ for some $\alpha \in \mathbf{R}$.

Definition 2.75 (Semisimple)**.** Given $L: V \to V$, we say L is **semisimple** if there exists an eigen-basis for L and V.

Lemma 2.76. *If $L: V \to V$ is semisimple and V is finite dimensional, then there exists a basis $\mathscr{B}$ of V such that the matrix associated to L with this basis satisfies*

$$A = \begin{pmatrix} \alpha_{1,1} & 0 & 0 & \dots & 0 \\ 0 & \alpha_{2,2} & 0 & \dots & 0 \\ 0 & 0 & \alpha_{3,3} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \alpha_{n,n} \end{pmatrix}.$$

Proof. Take an eigen-basis for L and V. The associated matrix will satisfy the desired conclusion. ♠

Remark 21. Given $L\colon V \to V$ with V finite dimensional, when a basis as in Lemma 2.76, one also says that L is **diagonalizable**.

Example 19. Take $V = \mathbf{R}^2$ with basis $\mathscr{B} = \{e_1, e_2\}$ and let $L\colon \mathbf{R}^2 \to \mathbf{R}^2$ be the unique linear function defined by

$$L(e_1) = e_1, \quad L(e_2) = e_1 + e_2.$$

The matrix associated to L in the basis $\mathscr{B}$ is

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

We claim that L is not semisimple. To see this, first notice that e_1 is an eigenvector for L with eigenvalue $\lambda = 1$. Assume that there is an eigenvector $v \in \mathbf{R}^2$ with

$$v = \alpha \cdot e_1 + \beta \cdot e_2$$

and eigenvalue λ'. Then

$$L(v) = L(\alpha \cdot e_1 + \beta \cdot e_2) = \alpha \cdot L(e_1) + \beta \cdot L(e_2) = \alpha \cdot e_1 + \beta \cdot e_1 + \beta \cdot e_2 = (\alpha + \beta) \cdot e_1 + \beta \cdot e_2$$

and

$$L(v) = \lambda' \cdot (\alpha \cdot e_1 + \beta \cdot e_2) = \lambda'\alpha \cdot e_1 + \lambda'\beta \cdot e_2.$$

Hence

$$\alpha + \beta = \lambda'\alpha, \quad \beta = \lambda'\beta.$$

We see that $\lambda' = 1$ from the second equality. Combining this observation with the first equality, we see that $\beta = 0$. Thus $v = \alpha \cdot e_1$. Therefore, L is not semisimple.

Example 20. Take $V = \mathbf{R}^2$ with basis $\mathscr{B} = \{e_1, e_2\}$ and let $L\colon \mathbf{R}^2 \to \mathbf{R}^2$ be the unique linear function defined by

$$L(e_1) = 2 \cdot e_1 + e_2, \quad L(e_2) = e_1 + e_2.$$

The matrix associated to L in the basis $\mathscr{B}$ is

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

We claim that L is semisimple. **Blah**.

Exercise 60. Let $L\colon \mathbf{R}^2 \to \mathbf{R}^2$ be given by

$$L(e_1) = a \cdot e_1 + c \cdot e_2, \quad L(e_2) = b \cdot e_1 + d \cdot e_2.$$

Assume that

$$|a+d| > 2, \quad ad - bc = 1.$$

Prove that L is semisimple.

Exercise 61. Let $L\colon \mathbf{R}^2 \to \mathbf{R}^2$ be given by

$$L(e_1) = a \cdot e_1 + c \cdot e_2, \quad L(e_2) = b \cdot e_1 + d \cdot e_2.$$

Prove that if $b = c$, then L is semisimple.

Exercise 62. Let $L\colon \mathbf{R}^2 \to \mathbf{R}^2$ be given by

$$L(e_1) = a \cdot e_1, \quad L(e_2) = b \cdot e_1 + d \cdot e_2.$$

Prove that L is not semisimple.

2.9 Flags In Vector Spaces

In this expository section, we briefly discuss the concept of flags on vector spaces. Unlike much of the material in this text, the main result of this section requires the vector space to be complex.

Definition 2.77 (Flag)**.** Given a finite dimensional vector space V, we say that a collection of vector subspaces $\mathscr{F} = \{V_0, V_1, \ldots, V_r\}$ is a **flag in** V if

$$\{0_V\} = V_0 \leq V_1 \leq \cdots \leq V_{r-1} \leq V_r = V$$

and

$$0 = \dim(V_0) < \dim(V_1) < \cdots < \dim(V_{r-1}) < \dim(V_r) = \dim(V).$$

The integer r is called the **length of the flag** $\mathscr{F}$.

Exercise 63. If V is finite dimensional and $\mathscr{F}$ is a flag in V of length r, then $r \leq \dim(V)$.

Definition 2.78 (Full Flag)**.** Given a finite dimensional vector space and a flag $\mathscr{F} = \left\{V_j\right\}$ of V, we say that $\mathscr{F}$ is a **full flag** if $\dim(V_j) = j$ for all j.

Exercise 64. Prove that if V is a finite dimensional vector space and $\mathscr{F}$ is a flag in V of length r that $\mathscr{F}$ is a full flag if and only if $r = \dim(V)$.

Definition 2.79. Given a finite dimensional vector space V, a flag $\mathscr{F}$ of V, and a linear function $L\colon V \to V$, we say that L **fixes** $\mathscr{F}$ if $L(V_j) \leq V_j$ for all $j \in \{0, 1, \dots, r\}$.

Lemma 2.80. *If V is a finite dimensional vector space with $n = \dim(V)$ and $L\colon V \to V$ is a linear map that fixes a full flag $\mathscr{F}$ of V, then there exists a basis $\mathscr{B}$ of V such that the matrix associated to L in $\mathscr{B}$ is of the form*

$$\begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} & \dots & \alpha_{1,n} \\ 0 & \alpha_{2,2} & \alpha_{2,3} & \dots & \alpha_{2,n} \\ 0 & 0 & \alpha_{3,3} & \dots & \alpha_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \alpha_{n,n} \end{pmatrix}.$$

Remark 22. A matrix as in the previous lemma is called **upper triangular**. Note the this matrix looks like a flag if you think of the zeroes as empty space. It looks like a $(\pi/4, \pi/4, \pi/2)$ triangle.

We will not prove the following result, in part, because it requires the vector space to be a complex vector space and we have not discussed this topic enough.

Theorem 2.81. *If V is finite dimensional complex vector space and $L\colon V \to V$ is a linear function, then there exists a full flag $\mathscr{F}$ of V such that L fixes $\mathscr{F}$.*

2.10 Vector Spaces Associated to Sets

Given a set X, we can associate to X a vector space that we will denote by $\mathbf{R}[X]$. We define

$$\mathbf{R}[X] \stackrel{\text{def}}{=} \mathrm{Fun}_{\mathrm{fin}}(X, \mathbf{R}).$$

Lemma 2.82. $\mathbf{R}[X]$ *is a vector space and* $\dim(\mathbf{R}[X]) = |X|$.

Proof. That $\mathrm{Fun}_{\mathrm{fin}}(X, \mathbf{R})$ is a vector space is clear as it is a vector subspace of $\mathrm{Fun}(X, \mathbf{R})$. For each $x_0 \in X$, we before

$$\chi x_0 \colon X \longrightarrow \mathbf{R}$$

by

$$\chi_{x_0}(x) = \begin{cases} 0, & x \neq x_0, \\ 1, & x = x_0. \end{cases}$$

It follows that $\chi_{x_0} \in \mathrm{Fun}_{\mathrm{fin}}(X, \mathbf{R})$ for every $x_0 \in X$ since

$$\mathrm{supp}(\chi_{x_0}) = \{x_0\}.$$

Given $L \in \mathrm{Fun}_{\mathrm{fin}}(X, \mathbf{R})$, we see that

$$L = \sum_{x \in \mathrm{supp}(L)} L(x) \cdot \chi_x.$$

Linearly independence is clear since χ_x, χ_y are simultaneously non-zero if and only if $x = y$. Hence $\{\chi_x\}_{x \in X}$ is a basis for $\mathbf{R}[X]$ and visibly $|X| = \left|\{\chi_x\}_{x \in X}\right|$. ♠

Definition 2.83 (Vector Space of a Set)**.** Given a set X, we call $\mathbf{R}[X]$ the **vector space associated to** X. The **associated basis** is $\mathscr{B}_X \stackrel{\text{def}}{=} \{\chi_x\}_{x \in X}$.

Corollary 2.84. $\mathbf{R}[X] \cong \mathbf{R}[Y]$ *if and only if* $|X| = |Y|$.

Proof. This follows from Theorem 2.42. ♠

Corollary 2.85. *If* $f\colon X \to Y$ *is a function, then there exists a unique linear function* $L\colon \mathbf{R}[X] \to \mathbf{R}[Y]$ *such that* $L(\chi_x) = \chi_{f(x)}$.

Proof. This follows from Theorem 2.26. ♠

The next two lemmas are left for the reader to prove.

Lemma 2.86. *If* $Y, Z \subset X$*, then* $\mathbf{R}[Y \cup Z] \cong (\mathbf{R}[Y] \times \mathbf{R}[Z])/\mathbf{R}[Y \cap Z]$.

Lemma 2.87. *If* $Y \subset X$*, then* $\mathbf{R}[X - Y] \cong \mathbf{R}[X]/\mathbf{R}[Y]$.

Exercise 65. Prove Lemma 2.86.

Exercise 66. Prove Lemma 2.87.

Chapter 3

The Dual Space

In this chapter, we introduce and study the dual space of a vector space V. The dual space and its precise relationship to V depends on whether V is finite dimensional or not. In the finite dimensional case, V will be isomorphic to its dual space. However, when V is infinite dimensional, it is only a vector subspace of the dual space.

3.1 Definition and Basic Concepts

We start with the definition of a dual vector and the dual vector space.

Definition 3.1 (Dual Vector)**.** Given a vector space V, we call $\mathscr{L} \in \mathrm{Hom}(V, \mathbf{R})$ a **dual vector**.

Definition 3.2 (Dual Space)**.** Given a vector space V, we call $\mathrm{Hom}(V, \mathbf{R})$ the **dual vector space** and denote it by V^*.

Example 21. Let $V = \mathbf{R}^n$ and $j \in \{1, \ldots, n\}$. We define $\mathrm{P}_j \colon \mathbf{R}^n \to \mathbf{R}$ by

$$\mathrm{P}_j(x_1, \ldots, x_n) = x_j.$$

It is straightforward to see that P_j is linear. Given $\mathscr{L} \colon \mathbf{R}^n \to \mathbf{R}$, we define

$$\alpha_j = \mathscr{L}(e_j).$$

Given $v \in \mathbf{R}^n$, since $\{e_1, \ldots, e_n\}$ is a basis for $\mathbf{R}^n$, we have

$$v = \sum_{j=1}^{n} \beta_j \cdot e_j.$$

We see that

$$\mathscr{L}(v) = \mathscr{L}\left(\sum_{j=1}^{n} \beta_j \cdot e_j\right) = \sum_{j=1}^{n} \beta_j \cdot \mathscr{L}(e_j) = \sum_{j=1}^{n} \beta_j \alpha_j.$$

Define

$$\mathrm{P}_{\mathscr{L}} \stackrel{\text{def}}{=} \sum_{j=1}^{n} \alpha_j \cdot \mathrm{P}_j.$$

By definition, we have $\mathrm{P}_{\mathscr{L}} \colon \mathbf{R}^n \to \mathbf{R}$. We see that

$$\begin{aligned}\mathrm{P}_{\mathscr{L}}(v) = \sum_{j=1}^{n} \alpha_j \cdot \mathrm{P}_j(v) &= \sum_{j=1}^{n} \alpha_j \mathrm{P}_j\left(\sum_{k=1}^{n} \beta_k \cdot e_k\right) \\ &= \sum_{j=1}^{n} \alpha_j \left(\sum_{k=1}^{n} \beta_k \cdot \mathrm{P}_j(e_k)\right) = \sum_{j=1}^{n} \alpha_j \beta_j.\end{aligned}$$

In particular, we see that

$$\mathscr{L} = \mathrm{P}_{\mathscr{L}}.$$

Thus, $\{\mathrm{P}_1, \ldots, \mathrm{P}_n\}$ span $(\mathbf{R}^n)^*$. Next, we will check that $\{\mathrm{P}_1, \ldots, \mathrm{P}_n\}$ is linearly independent. If

$$\sum_{j=1}^{n} \alpha_j \cdot \mathrm{P}_j = 0_{(\mathbf{R}^n)^*}$$

then

$$\sum_{j=1}^{n} \alpha_j \cdot \mathrm{P}_j(e_k) = 0$$

for all $k \in \{1, \ldots, n\}$. However

$$\sum_{j=1}^{n} \alpha_j \cdot \mathrm{P}_j(e_k) = \alpha_k = 0.$$

Thus, $\{\mathrm{P}_1, \ldots, \mathrm{P}_n\}$ is linearly independent. In total, we see that $\{\mathrm{P}_1, \ldots, \mathrm{P}_n\}$ is a basis for $(\mathbf{R}^n)^*$. The function P_j is sometimes called the **projection onto the j coordinate or factor function**. Note that $\dim(\mathbf{R}^n) = n = \dim((\mathbf{R}^n)^*)$ and so $\mathbf{R}^n \cong (\mathbf{R}^n)^*$. Indeed the unique linear extension of the function $e_j \mapsto \mathrm{P}_j$ is an isomorphism.

3.2 The Dual of a Linear Function

Given a linear function $L\colon V \to W$, we will define a linear function $L^*\colon W^* \to V^*$. To do this, we must assign to each $\mathscr{L} \in \mathrm{Hom}(W, \mathbf{R})$ an element of $\mathrm{Hom}(V, \mathbf{R})$. We have the diagram below where the arrows are linear functions (the dashed arrow is the function we want):

$$\begin{tikzcd} V \arrow[r, "L"] \arrow[dr, dashed, "L^*(\mathscr{L})"'] & W \arrow[d, "\mathscr{L}"] \\ & \mathbf{R} \end{tikzcd}$$

Definition 3.3 (Dual Map)**.** Given a linear function $L\colon V \to W$, the dual linear function is the linear map

$$L^*\colon W^* \longrightarrow V^*$$

defined by

$$L(\mathscr{L})(v) = \mathscr{L}(L(v))$$

or

$$L^*(\mathscr{L}) = \mathscr{L} \circ L$$

where $\mathscr{L} \in W^*$. Namely,

$$\begin{tikzcd} V \arrow[r, "L"] \arrow[dr, "L^*(\mathscr{L})=\mathscr{L}\circ L"'] & W \arrow[d, "\mathscr{L}"] \\ & \mathbf{R} \end{tikzcd}$$

Lemma 3.4. *If $L\colon V \to W$ is an injective linear function, then $L^*\colon W^* \to V^*$ is a surjective linear function.*

Proof. Given $\mathscr{L} \in V^*$, we need to prove that there exists $\mathscr{L}' \in W^*$ such that $L^*(\mathscr{L}') = \mathscr{L}$. By definition of L^*, we see that

$$L^*(\mathscr{L}')(v) = \mathscr{L}'(L(v)).$$

In particular, we need

$$\mathscr{L}'(L(v)) = \mathscr{L}(v)$$

for all $v \in V$. Since L is injective, we know that $L(\mathscr{B}) \subset W$ is linearly independent by Lemma 2.33. By Corollary 2.18, there exists a basis $\mathscr{B}'$ with $L(\mathscr{B}) \subset \mathscr{B}'$. We define $f\colon \mathscr{B}' \to \mathbf{R}$ by

$$f(w) \stackrel{\text{def}}{=} \begin{cases} 0, & w \notin L(\mathscr{B}), \\ \mathscr{L}(L^{-1}(w)), & w \in L(\mathscr{B}). \end{cases}$$

By Theorem 2.26, there exists a unique linear function $\mathscr{L}'\colon W \to \mathbf{R}$ such that $\mathscr{L}'(w) = f(w)$ for all $w \in \mathscr{B}'$. Given $v_0 \in V$, there exists a unique α_v with finite support such that

$$v_0 = \sum_{v \in \mathscr{B}} \alpha_v \cdot v$$

by Theorem 2.22. We see that

$$\begin{aligned} L^*(\mathscr{L}')(v_0) = \mathscr{L}'(L(v_0)) &= \mathscr{L}'\left(L\left(\sum_{v\in\mathscr{B}} \alpha_v \cdot v\right)\right) \\ &= \mathscr{L}'\left(\sum_{v\in\mathscr{B}} \alpha_v \cdot L(v)\right) = \mathscr{L}'\left(\sum_{w\in L(\mathscr{B})} \alpha_v \cdot w\right) \\ &= \sum_{w\in L(\mathscr{B})} \alpha_v \cdot \mathscr{L}'(w) = \sum_{w\in L(\mathscr{B})} \alpha_v \cdot \mathscr{L}(L^{-1}(w)) \\ &= \sum_{v\in\mathscr{B}} \alpha_v \cdot \mathscr{L}(v) = \mathscr{L}\left(\sum_{v\in\mathscr{B}} \alpha_v \cdot v\right) \\ &= \mathscr{L}(v_0). \end{aligned}$$

Thus $L^*(\mathscr{L}') = \mathscr{L}$ and so L^* is surjective. ♠

Remark 23. We note that the function $\mathscr{L}'$ is not unique if $\mathscr{B}' \neq L(\mathscr{B})$ as we can define the function f in the above prove to be anything on $\mathscr{B}' - L(\mathscr{B})$ while maintaining $L^*(\mathscr{L}') = \mathscr{L}$.

Lemma 3.5. *If $L\colon V \to W$ is a surjective linear function, then $L^*\colon W^* \to V^*$ is an injective linear function.*

Proof. Given $\mathscr{L}_1, \mathscr{L}_2 \in W^*$ with $L^*(\mathscr{L}_1) = L^*(\mathscr{L}_2)$, we must prove that $\mathscr{L}_1 = \mathscr{L}_2$. Given $w \in W$, since L is surjective, there exists $v \in V$ with $L(v) = w$. By assumption, we have

$$\mathscr{L}_1(w) = \mathscr{L}_1(L(v)) = L^*(\mathscr{L}_1)(v) = L^*(\mathscr{L}_2)(v) = \mathscr{L}_2(L(v)) = \mathscr{L}_2(w).$$

Hence $\mathscr{L}_1(w) = \mathscr{L}_2(w)$ and since this holds for all $w \in W$, we see that $\mathscr{L}_1 = \mathscr{L}_2$. ♠

Corollary 3.6. *If $L\colon V \to W$ is an isomorphism, then $L^*\colon W^* \to V^*$ is an isomorphism.*

Proof. This follows from Lemma 3.4 and Lemma 3.5. ♠

3.3 Dual Vectors With Finite Support and Dual Bases

Given a basis $\mathscr{B}$ for V, we will define a subset $\mathscr{B}^*$ of V^*. Given $v \in \mathscr{B}$, we define $\mathscr{L}_v \colon V \to \mathbf{R}$ as follows. First, we define $f_v \colon \mathscr{B} \to \mathbf{R}$ by

$$f_v(w) \stackrel{\text{def}}{=} \begin{cases} 0, & w \neq v, \\ 1, & w = v. \end{cases}$$

By Theorem 2.26, there exists a unique linear function $\mathscr{L}_v \colon V \to \mathbf{R}$ such that $\mathscr{L}_v(w) = f(w)$ for all $w \in \mathscr{B}$. We define

$$\mathscr{B}^* \stackrel{\text{def}}{=} \{\mathscr{L}_v \in V^* \ : \ v \in \mathscr{B}\}.$$

Definition 3.7 (Finite Support: Dual Vectors)**.** Given a vector space V, basis $\mathscr{B}$, and $\mathscr{L} \in V^*$, we say that $\mathscr{L}$ has **finite support with respect to** $\mathscr{B}$ if

$$\operatorname{supp}(\mathscr{L}, \mathscr{B}) \stackrel{\text{def}}{=} \{v \in \mathscr{B} \ : \ \mathscr{L}(v) \neq 0\}.$$

Definition 3.8 (Finite Support Dual Space)**.** Given a vector space V and basis $\mathscr{B}$ for V, we define the **finite support dual space** to be the subset of V^* given by

$$V^*_{\text{fin}} \stackrel{\text{def}}{=} \{\mathscr{L} \in V^* \ : \ \mathscr{L} \text{ has finite support with respect to } \mathscr{B}\}.$$

Lemma 3.9. *If* $\dim(V) < \infty$*, then* $V^* = V^*_{\text{fin}}$.

Proof. If $\dim(V) < \infty$, then for any basis $\mathscr{B}$ and any $\mathscr{L} \in V^*$, we have

$$|\operatorname{supp}(\mathscr{L}, \mathscr{B})| \leq |\mathscr{B}| < \infty.$$

Hence $V^* = V^*_{\text{fin}}$. ♠

Lemma 3.10. *The subset* $\mathscr{B}^*$ *is a basis for* V^*_{fin}.

Proof. Given $\mathscr{L} \in V^*_{\text{fin}}$, we define

$$\alpha_v = \mathscr{L}(v)$$

for each $v \in \mathscr{B}$ and note that α_v has finite support since $\mathscr{L}$ has finite support with respect to $\mathscr{B}$. We assert that

$$\mathscr{L} = \sum_{v \in \mathscr{B}} \alpha_v \cdot \mathscr{L}_v.$$

Given $v_0 \in V$ there exists a unique β_v with finite support such that

$$v_0 = \sum_{v \in \mathscr{B}} \beta_v \cdot v.$$

Thus, we have

$$\begin{aligned}
\mathscr{L}(v_0) = \mathscr{L}\left(\sum_{v\in\mathscr{B}} \beta_v \cdot v\right) &= \sum_{v\in\mathscr{B}} \beta_v \cdot \mathscr{L}(v) = \sum_{v\in\mathscr{B}} \beta_v \alpha_v \\
&= \sum_{v\in\mathscr{B}} (\beta_v \alpha_v) \cdot \mathscr{L}_v(v) = \sum_{v\in\mathscr{B}} \alpha_v \cdot \mathscr{L}_v(\beta_v \cdot v) \\
&= \sum_{v\in\mathscr{B}} \alpha_v \cdot \mathscr{L}_v\left(\sum_{v\in\mathscr{B}} \beta_v \cdot v\right) = \sum_{v\in\mathscr{B}} \alpha_v \cdot \mathscr{L}_v(v_0).
\end{aligned}$$

Thus $\mathscr{L} \in \mathrm{Span}(\mathscr{B}^*)$ and so $\mathscr{B}^*$ spans V^*_{fin}.

Next, we prove that $\mathscr{B}^*$ is linearly independent. If

$$\sum_{v\in\mathscr{B}} \alpha_v \cdot \mathscr{L}_v = 0_{V^*} = 0_{\mathrm{Fun}(V,\mathbf{R})}$$

where α_v has finite support, then we must prove that α_v is zero . By assumption

$$\sum_{v\in\mathscr{B}} \alpha_v \cdot \mathscr{L}_v(v_0) = 0_V$$

for all $v_0 \in V$. Taking $v_0 = w \in \mathscr{B}$, we see that

$$\sum_{v\in\mathscr{B}} \alpha_v \cdot \mathscr{L}_v(w) = \alpha_w.$$

In particular, $\alpha_w = 0$ for all $w \in \mathscr{B}$. Hence, $\mathscr{B}^*$ is linearly independent. ♠

Definition 3.11 (Dual Basis)**.** Given a vector space V and a basis $\mathscr{B}$, we define the **dual basis of** $\mathscr{B}$ to be $\mathscr{B}^*$ which is a basis for V^*_{fin}.

Corollary 3.12. *If V is a vector space, then $V \cong V^*_{\mathrm{fin}}$.*

Proof. It is straightforward to see that $|\mathscr{B}| = |\mathscr{B}^*|$. Hence $\dim(V) = \dim(V^*_{\mathrm{fin}})$ and so $V \cong V^*_{\mathrm{fin}}$ by Theorem 2.42. We will give an alternative proof by constructing an isomorphism from a basis $\mathscr{B}$ for V. By Lemma 3.10, $\mathscr{B}^*$ is a basis for V^*_{fin}. We define $f\colon \mathscr{B} \to V^*_{\mathrm{fin}}$ by $f(v) = \mathscr{L}_v$. By Theorem 2.26, there exists a unique linear function $L\colon V \to V^*_{\mathrm{fin}}$ such that $L(v) = f(v)$ for all $v \in \mathscr{B}$. Since $L(\mathscr{B}) = \mathscr{B}^*$, we see that L is an isomorphism by Corollary 2.40. ♠

Corollary 3.13. *If* $\dim(V) < \infty$ *and* $\mathscr{B}$ *is a basis for* V*, then* $\mathscr{B}^*$ *is a basis for* V^**. Moreover,* $V \cong V^*$.

Proof. This follows from Lemma 3.9, Lemma 3.10 and Corollary 3.12. ♠

Remark 24. If V is not finite dimensional, then $V^*_{\text{fin}} \neq V^*$. To see this, it is enough to construct $\mathscr{L} \in V^*$ that does not have finite support with respect to $\mathscr{B}$. Given a basis $\mathscr{B}$, we define $f\colon \mathscr{B} \to \mathbf{R}$ by $f(v) = 1$ for all $v \in \mathscr{B}$. By Theorem 2.26, there exists a unique linear function $\mathscr{L}: V \to \mathbf{R}$ such that $\mathscr{L}(v) = f(v) = 1$ for all $v \in \mathscr{B}$. In particular, since

$$\operatorname{supp}(\mathscr{L}, \mathscr{B}) = \mathscr{B},$$

we see that $\mathscr{L}$ as finite support if and only if $\mathscr{B}$ is finite.

Remark 25. Let $\mathscr{B} = \{v_1, v_2, v_3, \dots\}$ be a countable basis for V with $\dim(V) = |\mathbf{N}|$. As we saw in the previous remark, the extension of the function $f\colon \mathscr{B} \to \mathbf{R}$ given by $f(v) = 1$ does not have finite support with respect to $\mathscr{B}$. However, if we define

$$\mathscr{B}' = \{v_1, v_1 - v_2, v_1 - v_3, \dots\}$$

then this is a basis for V and we see that $\operatorname{supp}(f, \mathscr{B}') = \{v_1\}$. In particular, linear functions $L\colon V \to R$ can have infinite support for some basis and finite support for another basis.

Exercise 67. Prove or disprove: If $\dim(V) = |\mathbf{N}|$, $\mathscr{B}$ is a basis for V, and $\mathscr{L} \in V^*$ with $\operatorname{supp}(\mathscr{L}, \mathscr{B})$ infinite, then there exists a basis $\mathscr{B}'$ such that $\operatorname{supp}(\mathscr{L}, \mathscr{B}')$ is finite.

Exercise 68. Prove or disprove: If V is a vector space with $\dim(V) = |\mathbf{N}|$ and $\mathscr{L} \in V^*$, then there exists a basis $\mathscr{B}$ for V such that $\operatorname{supp}(\mathscr{L}, \mathscr{B})$ is infinite.

Definition 3.14. We define $\operatorname{Fun}_{\text{fin}}(\mathscr{B}, \mathbf{R})$ to be the subset of $\operatorname{Fun}(\mathscr{B}, \mathbf{R})$ given by

$$\operatorname{Fun}_{\text{fin}}(\mathscr{B}, \mathbf{R}) \stackrel{\text{def}}{=} \{f \in \operatorname{Fun}(\mathscr{B}, \mathbf{R}) \ : \ f \text{ has finite support}\}.$$

Lemma 3.15. $\operatorname{Fun}_{\text{fin}}(\mathscr{B}, \mathbf{R})$ *is a vector subspace of* $\operatorname{Fun}(\mathscr{B}, \mathbf{R})$.

Proof. This is clear. ♠

Exercise 69. Prove Lemma 3.15.

Theorem 3.16. *If* V *is a vector space and* $\mathscr{B}$ *is a basis for* V*, then* $V \cong \operatorname{Fun}_{\text{fin}}(\mathscr{B}, \mathbf{R})$ *and* $V^* \cong \operatorname{Fun}(\mathscr{B}, \mathbf{R})$.

Proof. Given $v_0 \in V$, by Theorem 2.22, there exists a unique α_v with finite support such that

$$v_0 = \sum_{v \in \mathscr{B}} \alpha_v \cdot v.$$

By definition, $\alpha \colon \mathscr{B} \to \mathbf{R}$ is a function where $\alpha(v) = \alpha_v$ and so $\alpha \in \mathrm{Fun}_{\mathrm{fin}}(\mathscr{B}, \mathbf{R})$. We define $L \colon V \to \mathrm{Fun}_{\mathrm{fin}}(\mathscr{B}, \mathbf{R})$ by $L(v_0) = \alpha$. The linearity of this map is clear. Injectivity follows from the uniqueness of α_v. Given $\alpha \in \mathrm{Fun}_{\mathrm{fin}}(\mathscr{B}, \mathbf{R})$, we define

$$u = \sum_{v \in \mathscr{B}} \alpha(v) \cdot v$$

and note that $L(u) = \alpha$. The isomorphism for V^* and $\mathrm{Fun}(\mathscr{B}, \mathbf{R})$ is done similarly using Theorem 2.26. ♠

Corollary 3.17. *If V is not finite dimensional, then* $\dim(V) < \dim(V^*)$.

Let

$$X_n = \{1, 2, \dots, n\}.$$

Corollary 3.18. $\mathbf{R}^n \cong \mathrm{Fun}(X_n, \mathbf{R})$.

Exercise 70. Prove Corollary 3.18.

3.4 Eval and the Double Dual

In this section, we investigate the double dual (i.e. the dual of the dual). The double dual is another layer of abstraction and can be difficult to parse at the start. We will see that there is a canonical function called the **evaluation map** that relates V with its double dual. Again when V is finite dimensional, the evaluation map will provide us with a canonical isomorphism between V and its double dual.

Given a set X, the set $\mathrm{Fun}(X, \mathbf{R})$ is a vector space. We define a function

$$\mathrm{Eval} \colon X \times \mathrm{Fun}(X, \mathbf{R}) \longrightarrow \mathbf{R}$$

by

$$\mathrm{Eval}(x, f) = f(x).$$

Fixing $x \in X$, we have the function

$$\mathrm{Eval}_x \colon \mathrm{Fun}(X, \mathbf{R}) \longrightarrow \mathbf{R}$$

given by $\mathrm{Eval}_x(f) = f(x)$. The function Eval_x is the function above on $X \times \mathrm{Fun}(X, \mathbf{R})$ restricted to the subset $\{x\} \times \mathrm{Fun}(X, \mathbf{R})$ which we think of as just $\mathrm{Fun}(X, \mathbf{R})$. We see that

$$\mathrm{Eval}_x(\alpha_1 \cdot f_1 + \alpha_2 \cdot f_2) = \alpha_1 f_1(x) + \alpha_2 f_2(x) = \alpha_1 \mathrm{Eval}_x(f_1) + \alpha_2 \mathrm{Eval}_x(f_2).$$

Thus Eval_x is a linear function and so $\mathrm{Eval}_x \in (\mathrm{Fun}(X, \mathbf{R}))^* = \mathrm{Hom}(\mathrm{Fun}(X, \mathbf{R}), \mathbf{R})$.

When $X = V$ is a vector space, since $V^* = \mathrm{Hom}(V, \mathbf{R}) \subset \mathrm{Fun}(V, \mathbf{R})$, we can take the restriction of Eval_v to V^* for each $v \in V$. We see that $\mathrm{Eval}_v \in \mathrm{Hom}(V^*, \mathbf{R})$. Of course,

$$\mathrm{Hom}(V^*, \mathbf{R}) = \mathrm{Hom}(\mathrm{Hom}(V, \mathbf{R}), \mathbf{R}) = (V^*)^*.$$

Definition 3.19. Given a vector space V, we define the **double dual** of V to be $V^\bigstar = (V^*)^*$. Given a linear function $L\colon V \to W$, we denote the double dual map $(L^*)^*$ by $L^\bigstar$.

We define a function $\mathrm{Eval}_V : V \to V^\bigstar$ by

$$\mathrm{Eval}_V(v) \stackrel{\text{def}}{=} \mathrm{Eval}_v. \tag{3.1}$$

Remark 26. By Corollary 3.17, if V is not finite dimensional, then $\dim(V) < \dim(V^*) < \dim(V^\bigstar)$. Hence,$V$ and $V^\bigstar$ cannot be isomorphic.

Theorem 3.20. Eval_V *is a linear injection. If* $\dim(V) < \infty$*, then* Eval_V *is an isomorphism.*

Proof. To prove this, we will show that Eval_V is linear first. We must prove that

$$\mathrm{Eval}_V(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2) = \alpha_1 \cdot \mathrm{Eval}_V(v_1) + \alpha_2 \cdot \mathrm{Eval}_V(v_2). \tag{3.2}$$

Using (3.1) in (3.2), we obtain the functional equation:

$$\mathrm{Eval}_{\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2} = \alpha_1 \cdot \mathrm{Eval}_{v_1} + \alpha_2 \cdot \mathrm{Eval}_{v_2}. \tag{3.3}$$

In particular, the validity of (3.3) is equivalent to showing

$$\mathrm{Eval}_{\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2}(\mathscr{L}) = \alpha_1 \cdot \mathrm{Eval}_{v_1}(\mathscr{L}) + \alpha_2 \cdot \mathrm{Eval}_{v_2}(\mathscr{L}) \tag{3.4}$$

for all $\mathscr{L} \in V^*$. By definition of Eval_v, we see that

$$\mathrm{Eval}_{\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2}(\mathscr{L}) = \mathscr{L}(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2) = \alpha_1 \mathscr{L}(v_1) + \alpha_2 \mathscr{L}(v_2)$$

and

$$\alpha_1 \cdot \mathrm{Eval}_{v_1}(\mathscr{L}) + \alpha_2 \cdot \mathrm{Eval}_{v_2}(\mathscr{L}) = \alpha_1 \mathscr{L}(v_1) + \alpha_2 \mathscr{L}(v_2).$$

Hence (3.4) holds for all $\mathscr{L} \in V^*$ as needed and so Eval_V is linear.

For injectivity, it suffices to prove that $\ker(\mathrm{Eval}_V) = \{0_V\}$ by Proposition 1.70. If $v_0 \in \ker(\mathrm{Eval}_V)$, then we must have

$$\mathrm{Eval}_V(v_0)(\mathscr{L}) = 0$$

for all $\mathscr{L} \in V^*$. By definition of Eval_V, we have

$$\mathrm{Eval}_{v_0}(\mathscr{L}) = 0$$

for all $\mathscr{L} \in V^*$. If $v_0 \neq 0_V$, then there exists a unique α_v with finite support such that

$$v_0 = \sum_{v \in \mathscr{B}} \alpha_v \cdot v$$

and with α_v non-zero. In particular, $\alpha_u \neq 0$ for some $u \in \mathscr{B}$. Define $f \colon \mathscr{B} \to \mathbf{R}$ by

$$f(v) \stackrel{\text{def}}{=} \begin{cases} 0, & v \neq u, \\ 1, & v = u. \end{cases}$$

By Theorem 2.26, there exists a unique linear function $\mathscr{L}_u \in V^*$ such that $\mathscr{L}_u(v) = f(v)$ for all $v \in \mathscr{B}$. By construction, $\mathscr{L}_u(v_0) = \alpha_u \neq 0$. Thus, $\ker(\mathrm{Eval}_V) = 0_V$ and so Eval_V is injective.

For surjectivity (which requires $\dim(V) < \infty$), we know that $V \cong V^*$ and $V^* \cong V^\bigstar$ by Corollary 3.13. Hence $V \cong V^\bigstar$ by Lemma 1.60 and so $\dim(V) = \dim(V^\bigstar)$ by Theorem 2.42. Hence Eval_V is surjective by Theorem 2.46. ♠

Remark 27. One might object to using $V \cong V^\bigstar$ in the proof that Eval_V is an isomorphism. However, the string of results we used all rely on the choice of a basis. The function Eval_V is an isomorphism between V and $V^\bigstar$ that **does not require the choice of a basis to define**. Such isomorphisms are called **natural**; one also wants the conclusion of Theorem 3.21 below to hold when saying L is natural.

Theorem 3.21. *Given a linear function $L \colon V \to W$, the diagram below commutes:*

$$\begin{tikzcd} V \arrow[r, "L"] \arrow[d, "\mathrm{Eval}_V"'] & W \arrow[d, "\mathrm{Eval}_W"] \\ V^\bigstar \arrow[r, "L^\bigstar"'] & W^\bigstar. \end{tikzcd}$$

That is $\mathrm{Eval}_W \circ L = L^\bigstar \circ \mathrm{Eval}_V$ *or alternatively,*

$$\mathrm{Eval}_W(L(v)) = L^\bigstar(\mathrm{Eval}_V(v))$$

for all $v \in V$.

Proof. We must prove
$$\mathrm{Eval}_W(L(v)) = L^{\bigstar}(\mathrm{Eval}_V(v)) \tag{3.5}$$
holds for all $v \in V$. Note that this is a functional equation and is equivalent to prove for all $v \in V$ and all $\mathscr{L} \in W^*$ that
$$\mathrm{Eval}_W(L(v))(\mathscr{L}) = L^{\bigstar}(\mathrm{Eval}_V(v))(\mathscr{L}).$$
Given $v \in V$ and $\mathscr{L} \in W^*$, we know that
$$((\mathrm{Eval}_W \circ L)(v))(\mathscr{L}) = \mathrm{Eval}_{L(v)}(\mathscr{L}) = \mathscr{L}(L(v)).$$
Also, for $v \in V$ and $\mathscr{L} \in W^*$, we know that
$$((L^{\bigstar} \circ \mathrm{Eval}_V)(v))(\mathscr{L}) = (L^{\bigstar}(\mathrm{Eval}_v))(\mathscr{L}) \stackrel{!}{=} \mathrm{Eval}_v(L^*(\mathscr{L})) = L^*(\mathscr{L})(v) = \mathscr{L}(L(v)), \tag{3.6}$$
completing the proof. ♠

Remark 28. Understanding (3.6) can be difficult. The equality given by $\stackrel{!}{=}$ in (3.6) is the most "layered" but is just the definition of $L^{\bigstar}$.

Corollary 3.22. *If $L\colon V \to W$ is a linear function and $\dim(V) < \infty$, then*
$$L^{\bigstar} = \mathrm{Eval}_W \circ L \circ \mathrm{Eval}_V^{-1}.$$

Proof. This follows from Theorem 3.20 and Theorem 3.21. ♠

Corollary 3.23. *Given a linear function $L\colon V \to W$, we have the following:*

(i) L is injective if and only if $L^{\bigstar}$ is injective.

(ii) L is surjective if and only if $L^{\bigstar}$ is surjective.

Proof. This follows from Lemma 3.4 and Lemma 3.5 ♠

Corollary 3.24. *Given a linear function $L\colon V \to W$ with $\dim(V), \dim(W) < \infty$, we have the following:*

(i) L is injective if and only if L^* is surjective.

(ii) L is surjective if and only if L^* is injective.

Proof. The direct implications in (i) and (ii) follow from Lemma 3.4 and Lemma 3.5. The reverse implications in (i) and (ii) following from from Lemma 3.4, Lemma 3.5, and Corollary 3.23. ♠

3.5 Matrix Associated to the Dual Map

In this section, we will investigate the structure of the matrix associated to the dual map and how it relates to the matrix associated to the original linear function. When V is finite dimensional, we will see that the matrices are related by the matrix transpose.

3.5.1 Finite Dimensional

Given vector space V,W with $\dim(V)=n$ and $\dim(W)=m$, we will discuss the matrix associated to $L^*\colon W^* \to V^*$. We fix bases $\mathscr{B}_V$ and $\mathscr{B}_W$ for V,W. Specifically, we have

$$\mathscr{B}_V = \{v_1,\dots,v_n\}, \quad \mathscr{B}_W = \{w_1,\dots,w_m\}.$$

We have dual basis

$$\mathscr{B}_V^* = \{\mathscr{L}_{v_1},\dots,\mathscr{L}_{v_n}\}, \quad \mathscr{B}_W^* = \{\mathscr{L}_{w_1},\dots,\mathscr{L}_{w_m}\}.$$

Given $L\colon V \to W$, we have the associated matrix A given by $A = (\beta_{j,k})$ where

$$L(v_k) = \sum_{j=1}^{m} \beta_{j,k} \cdot w_j.$$

Specifically, we have

$$A \overset{\text{def}}{=} \begin{pmatrix} \beta_{1,1} & \beta_{1,2} & \beta_{1,3} & \dots & \beta_{1,n} \\ \beta_{2,1} & \beta_{2,2} & \beta_{2,3} & \dots & \beta_{2,n} \\ \beta_{3,1} & \beta_{3,2} & \beta_{3,3} & \dots & \beta_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta_{m,1} & \beta_{m,2} & \beta_{m,3} & \dots & \beta_{m,n} \end{pmatrix}.$$

By definition, $L^*\colon W^* \to V^*$ is given by $L^*(\mathscr{L}) = \mathscr{L} \circ L$. We see that

$$L^*(\mathscr{L}_{w_j})(v) = \mathscr{L}_{w_j}(L(v)).$$

Taking $v = v_k$, we see that

$$L^*(\mathscr{L}_{w_j})(v_k) = \mathscr{L}_{w_j}(L(v_k)) = \mathscr{L}_{w_j}\left(\sum_{\ell=1}^{m} \beta_{\ell,k} \cdot w_\ell\right) = \beta_{j,k}.$$

Hence, we see that

$$L^*(\mathscr{L}_{w_j}) = \sum_{\ell=1}^{n} \beta_{j,k} \cdot \mathscr{L}_{v_k}.$$

Hence, the matrix A^* associated to L^* is given by

$$A^* = \begin{pmatrix} \beta_{1,1} & \beta_{2,1} & \beta_{3,1} & \dots & \beta_{m,1} \\ \beta_{1,2} & \beta_{2,2} & \beta_{3,2} & \dots & \beta_{m,2} \\ \beta_{1,3} & \beta_{2,3} & \beta_{3,3} & \dots & \beta_{m,3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta_{1,n} & \beta_{2,n} & \beta_{3,n} & \dots & \beta_{m,n} \end{pmatrix}$$

This proves the following proposition.

Proposition 3.25. *If $L\colon V \to W$ is linear with associated matrix A in the bases $\mathscr{B}_V, \mathscr{B}_W$, then the matrix A^* associated to $L^*\colon W^* \to V^*$ in the bases $\mathscr{B}_V^*, \mathscr{B}_W^*$ satisfies*

$$A^* = A^T$$

where A^T is the transpose.

Exercise 71. Prove that $(A^*)^* = (A^T)^T = A$.

3.5.2 Infinite Dimensional

When $\dim(V), \dim(W)$ are infinite, we will review how to describe the "associated matrix" in terms of column and row vectors. Given a linear function $L\colon V \to W$ and basis $\mathscr{B}_V$, $\mathscr{B}_W$ for V, W, we can define two functions

$$\beta^v\colon \mathscr{B}_W \longrightarrow \mathbf{R}, \quad \alpha^w\colon \mathscr{B}_V \longrightarrow \mathbf{R}.$$

Namely, for each $v \in \mathscr{B}_V$, we have a unique β_w^v with finite support on $\mathscr{B}_W$ such that

$$L(v) = \sum_{w \in \mathscr{B}_W} \beta_w^v \cdot w.$$

We define $\alpha_v^w = \beta_w^v$. We can view $\beta_w^v \in W$ since β_w^v has finite support. However, α_v^w need not have finite support and thus $\alpha_v^w \in V^*$. The $\beta_w^v \in W$ represent the column vectors and the $\alpha_v^w \in V^*$ represent the row vectors.

Working more abstractly, if we view β^v as a function of $v \in \mathscr{B}_V$, we obtain

$$\mathrm{C}^L\colon \mathscr{B}_V \longrightarrow \mathrm{Fun}_{\mathrm{fin}}(\mathscr{B}_W, \mathbf{R}) \cong W.$$

The unique linear extension $L_{\mathrm{C}_L}\colon V \to W$ of C_L is L.

Exercise 72. Prove that $L_{\mathrm{C}_L} = L$.

Viewing α^w as a function of $w \in \mathscr{B}_W$, we obtain

$$\mathrm{R}^L \colon \mathscr{B}_W \longrightarrow \mathrm{Fun}(\mathscr{B}_V, \mathbf{R}) \cong V^*.$$

The unique linear extension $L_{\mathrm{R}^L} \colon W \to V^*$ is the restriction of L^* to W, where we view $W = \mathrm{Fun}_{\mathrm{fin}}(\mathscr{B}_W, \mathbf{R})$ inside of $W^* = \mathrm{Fun}(\mathscr{B}_W, \mathbf{R})$.

For the dual map $L^* \colon W^* \to V^*$ with basis $\mathscr{B}_{V^*}, \mathscr{B}_{W^*}$, for each $w^* \in \mathscr{B}_{W^*}$, there exists a unique $\beta_{v*}^{w^*}$ with finite support on $\mathscr{B}_{V^*}$ such that

$$L^*(w^*) = \sum_{v^* \in \mathscr{B}_{V^*}} \beta_{v^*}^{w^*} \cdot v^*.$$

We define $\alpha_{w^*}^{v^*} = \beta_{v^*}^{w^*}$, noting again that $\alpha_{w^*}^{v^*}$ need not have finite support on $\mathscr{B}_{W*}$. Thus, we then have column vectors $\beta_{v^*}^{w^*} \in V^*$ and row vectors $\alpha_{w^*}^{v^*} \in W^\bigstar$.

Again, we have $\mathrm{C}^{L^*} \colon \mathscr{B}_{W^*} \to V^*$ and a unique linear extension $L_{C^{L^*}} \colon W^* \to V^*$.

Exercise 73. Prove that $L^* = L_{C^{L^*}}$.

Chapter 4

Bilinear and Quadratic Functions

We will discussion bilinear forms on real vector spaces. We will require $\dim(V) \leq |\mathbf{N}|$ in some of the main topics of this chapter. Bilinear spaces are assumed to be non-degenerate mostly.

4.1 Definition and Concepts

4.1.1 General

Definition 4.1 (Bilinear Function)**.** Given vector spaces V, W, U, we say that a function $B\colon V \times W \to U$ is **bilinear** if for all $v_0, v_1, v_2 \in V$, $w_0, w_1, w_2 \in W$ and $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbf{R}$, we have

$$B(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2, w) = \alpha_1 \cdot B(v_1, w) + \alpha_2 \cdot B(v_2, w)$$
$$B(v, \beta_1 \cdot w_1 + \beta_2 \cdot w_2) = \beta_1 \cdot B(v, w_1) + \beta_2 \cdot B(v, w_2).$$

Definition 4.2 (Symmetric Bilinear Form)**.** Given a vector space V and a bilinear function $B\colon V \times V \to U$, we say that B is **symmetric** if $B(v, w) = B(w, v)$ for all $v, w \in V$.

Remark 29. We will now assume that **all bilinear functions are symmetric** when $V = W$ useless explicitly stated otherwise.

Given a bilinear function $B\colon V \times W \to U$, each $v_0 \in V$ defines a linear function $B_{v_0,\star}\colon W \to U$ defined by

$$B_{v_0,\star}(w) = B(v_0, w).$$

Similarly, each $w_0 \in W$ defines a linear function $B_{\star,w_0} \colon V \to U$ defined by

$$B_{\star,w_0}(v) = B(v,w_0).$$

These facts motivate the name "bilinear function".

We define functions $B_W \colon V \to \operatorname{Hom}(W,U)$ and $B_V \colon W \to \operatorname{Hom}(V,U)$ by

$$B_W(v) \stackrel{\text{def}}{=} B_{v,\star}, \quad B_V(w) \stackrel{\text{def}}{=} B_{\star,w}.$$

When $U = \mathbf{R}$, we see that $B_W \colon V \to W^*$ and $B_V \colon W \to V^*$. When $V = W$ and $U = \mathbf{R}$, we have $B \colon V \to V^*$. In this very special case, we see that a bilinear function is the same as an element of $\operatorname{Hom}(V,V^*)$. Indeed, given $\psi \in \operatorname{Hom}(V,V^*)$, we can define a bilinear function $B_\psi \colon V \times V^* \to \mathbf{R}$ by $B_\psi(v,\mathscr{L}) = (\psi(v))(\mathscr{L})$.

Define

$$\operatorname{BiL}(V,W;U) = \{B \colon V \times W \to U \ : \ B \text{ is bilinear}\}.$$

Lemma 4.3. $\operatorname{BiL}(V,W;U)$ *is a vector subspace of* $\operatorname{Fun}(V \times W, U)$.

Proof. This is straightforward. ♠

Exercise 74. Prove Lemma 4.3.

Lemma 4.4. $\operatorname{BiL}(V,V;\mathbf{R}) \cong \operatorname{Hom}(V,V^*)$.

Proof. This is straightforward (given the above discussion). ♠

Exercise 75. Prove Lemma 4.4.

Fixing $v \in V$, we have two maps $B_{v,\star}B_{\star,v} \colon V \to \mathbf{R}$. These functions form part of a "diamond" of vector spaces

$$\begin{tikzcd}
& V \times V \arrow[dl, "\Pi_L"'] \arrow[dr, "\Pi_R"] \arrow[dd, "B"] & \\
V \arrow[dr, "B_{\star,v}"'] & & V \arrow[dl, "B_{v,\star}"] \\
& \mathbf{R} &
\end{tikzcd}$$

where

$$\Pi_L, \Pi_R \colon V \times V \longrightarrow V$$

are defined by

$$\Pi_L(v_1 v_2) \stackrel{\text{def}}{=} v_1, \quad \Pi_R(v_1, v_2) \stackrel{\text{def}}{=} v_2.$$

In the above, notice that B gives two linear functions $B_{V,R}, B_{V,L} \colon V \to V^*$ where

$$B_{V,L}(v) \stackrel{\text{def}}{=} B_{\star,v}, \quad B_{V,R}(v) \stackrel{\text{def}}{=} B_{v,\star}.$$

We have

$$\begin{tikzcd}
V \arrow[r, bend left, "\mathrm{Eval}_V"] \arrow[rr, bend right, "B_{V,L}"'] & V^{\bigstar} \arrow[r, bend left, "(B_{V,R})^*"] & V^* & V^{\bigstar} \arrow[l, bend left, "(B_{V,L})^*"] & V \arrow[l, bend left, "\mathrm{Eval}_V"] \arrow[ll, bend right, "B_{V,R}"']
\end{tikzcd}$$

Proposition 4.5. *If V is finite dimensional, then $B_{V,R} = (B_{V,L})^* \circ \mathrm{Eval}_V$ and $B_{V,L} = (B_{V,R})^* \circ \mathrm{Eval}_V$.*

Proof. We will leave this for the reader. ♠

Exercise 76. Prove Proposition 4.5.

4.1.2 Main Interest

Our interest in this chapter will be in the case when $U = \mathbf{R}$ and $V = W$.

Definition 4.6 (Real Bilinear Form)**.** Given a vector space V, we call a bilinear function $B \colon V \times V \to \mathbf{R}$ a **real bilinear form**.

Example 22 (Euclidean Inner Product)**.** Let $V = \mathbf{R}^n$. Given $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, we define $\langle \cdot, \cdot \rangle : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ by

$$\langle x, y \rangle \stackrel{\text{def}}{=} \sum_{j=1}^{n} x_j y_j = x_1 y_1 + \cdots + x_n y_n.$$

We must show that

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

and

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

for all $x, y, z \in \mathbf{R}^n$ and all $\alpha, \beta \in \mathbf{R}$. Taking $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$, and $z = (z_1, \ldots, z_n)$, we see that

$$\begin{aligned}
\langle \alpha \cdot x + \beta \cdot z, y\rangle &= \sum_{j=1}^{n} (\alpha x_j + \beta z_j) y_j \\
&= \sum_{j=1}^{n} (\alpha x_j y_j + \beta z_j y_j) \\
&= \sum_{j=1}^{n} \alpha x_j y_j + \sum_{j=1}^{n} \beta z_j y_j \\
&= \alpha \left(\sum_{j=1}^{n} x_j y_j \right) + \beta \left(\sum_{j=1}^{n} z_j y_j \right) = \alpha \langle x, y\rangle + \beta \langle z, y\rangle .
\end{aligned}$$

Since $\langle x, y\rangle = \langle y, x\rangle$, we see that the other equality also holds.

Example 23 (Other Classical Examples)**.** Let $V = \mathbf{R}^n$ and let $p, q \geq 0$ with $p + q = n$. Define

$$\langle \cdot, \cdot \rangle_{p,q} : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$$

by

$$\langle x, y\rangle_{p,q} \stackrel{\text{def}}{=} \sum_{j=1}^{p} x_j y_j - \sum_{j=p+1}^{p+q=n} x_j y_j$$

where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. If $p, q > 0$, notice that for $x = (x_1, \ldots, x_n)$ has $x_j = 0$ for $j \neq p, p+1$ and $x_p = x_{p+1} = 1$, we have

$$\langle x, x\rangle_{p,q} = x_p^2 - x_{p+1}^2 = 1 - 1 = 0.$$

This cannot happen for $\langle \cdot, \cdot \rangle$. In fact, we see that $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_{n,0}$. One sometimes write $\mathbf{R}^{p,q}$ when viewing $\mathbf{R}^n$ with the bilinear function $\langle \cdot, \cdot \rangle_{p,q}$. The special family $\mathbf{R}^{n,1}$ is of particular interest in considerations beyond mathematics; the case $\mathbf{R}^{3,1}$ is sometimes referred to as **Minkowski space-time** as it arises in general relativity.

Exercise 77. Prove $\langle \cdot, \cdot \rangle_{p,q}$ is a bilinear form. Prove that each $v \neq 0_{\mathbf{R}^n}$, there exists $w \in \mathbf{R}^n$ such that $\langle v, w\rangle_{p,q} \neq 0$.

Example 24. Let $V = \mathrm{C}^0([0,1])$ be the vector subspace of $\mathrm{Fun}([0,1],\mathbf{R})$ of continuous functions. These functions are Riemann integrable as are products of continuous functions. Given $f,g \in \mathrm{C}^0([0,1])$, we define

$$B(g,f) \stackrel{\text{def}}{=} \int_0^1 f(x)g(x)dx.$$

Note that

$$B(f,f) = \int_0^1 (f(x))^2 dx > 0$$

if f is not the zero function (this requires continuity).

Example 25. Let $\ell^2(\mathbf{N})$ be the real sequences $\{x_n\}$ such that the series

$$\sum_{n=1}^{\infty} x_n$$

is absolutely convergent. This is a vector subspace of $\mathrm{Fun}(\mathbf{N},\mathbf{R})$, the vector space of real sequences. Define

$$B(\{x_n\},\{y_n\}) \stackrel{\text{def}}{=} \sum_{n=1}^{\infty} x_n y_n.$$

This series converges absolutely and so gives a well-defined value for B. Note that

$$B(\{x_n\},\{x_n\}) = \sum_{n=1}^{\infty} x_n^2 > 0$$

if $\{x_n\}$ is not the zero sequence. Checking that B is a bilinear form is straightforward.

Exercise 78. Prove B is a bilinear form.

Lemma 4.7. *If V is a vector space and $B\colon V \times V \to \mathbf{R}$ is a real bilinear form, then*

$$B(0_V, v) = B(v, 0_V) = 0$$

for all $v \in V$.

Proof. For $v \in V$, we have

$$B(0_V, v) = B(v - v, v) = B(v,v) - B(v,v) = 0$$

and

$$B(v, 0_V) = B(v, v - v) = B(v,v) - B(v,v) = 0.$$

♠

Definition 4.8 (Degenerate Bilinear Form)**.** Given a vector space V and a real bilinear form $B\colon V\times V\to\mathbf{R}$, we say that B is **degenerate** if there exists $v_0\in V$ with $v_0\neq 0$ such that $B(v_0,w)=0$ for all $w\in V$.

Definition 4.9 (Non-Degenerate Bilinear Form)**.** Given a vector space V and a real bilinear form $B\colon V\times V\to\mathbf{R}$, we say that B is **non-degenerate** if for each $v\in V$ with $v\neq 0$, there exists $w\in V$ such that $B(v,w)\neq 0$.

Definition 4.10 (Degenerate Subspace)**.** Given a vector space V and a real bilinear form $B\colon V\times V\to\mathbf{R}$, we define the **degenerate subspace of** B to be the subset $\mathrm{Dead}(B)\subset V$ given by

$$\mathrm{Dead}(B)\overset{\mathrm{def}}{=}\{v\in V\ :\ B(v,w)=0\text{ for all }w\in V\}.$$

Lemma 4.11. *If V is a vector space and B is a real bilinear form, then* $\mathrm{Dead}(B)$ *is a vector subspace.*

Proof. Given $v_1,v_2\in\mathrm{Dead}(B)$, $\alpha_1,\alpha_2\in\mathbf{R}$, and $v\in V$, we have

$$B(\alpha_1\cdot v_1+\alpha_2\cdot v_2,v)=\alpha_1 B(v_1,v)+\alpha_2 B(v_2,v)=0+0=0.$$

Ss $0_V\in\mathrm{Dead}(B)$ by Lemma 4.7, we see that $\mathrm{Dead}(B)$ is a vector subspace. ♠

Definition 4.12 (Bilinear Space)**.** Given a vector space V and a real, symmetric, non-degenerate bilinear form B, we will call the pair (V,B) a **bilinear space**.

Definition 4.13 (Quadratic Form)**.** Given a vector space V and a function $q\colon V\to\mathbf{R}$, we say that q is a **quadratic form** if

(1) $q(\alpha\cdot v)=\alpha^2 q(v)$ for all $v\in V$.

(2) The function $B_q\colon V\times V\to\mathbf{R}$ defined by

$$B_q(v,w)\overset{\mathrm{def}}{=}q(v+w)-q(v)-q(w)$$

is a (symmetric) bilinear form.

We call B_q the **associated bilinear form**.

Definition 4.14. Given a vector space V and a quadratic form q, we call the pair (V,q) a **quadratic space**.

Definition 4.15 (Quadratic Form Associated to a Bilinear Form)**.** Given a bilinear space (V,B), the function $q_B\colon V \to \mathbf{R}$ defined by

$$q_B(v) \stackrel{\text{def}}{=} B(v,v)$$

is called the **associated quadratic form**.

Lemma 4.16. *If (V,B) is a bilinear space, then q_B is a quadratic form.*

The proof of this lemma below is not how one should prove this lemma. Instead, one should use Lemma 4.17. This proof is included as it is direct (i.e. via the definition of what should be shown).

Proof. For this we need to prove that q_B is a quadratic form. First, we have

$$q_B(\alpha \cdot v) \stackrel{\text{def}}{=} B(\alpha \cdot v, \alpha \cdot v) = \alpha^2 B(v,v) = \alpha^2 q_B(v).$$

Next, we must prove that

$$B_{q_B}(v,w) \stackrel{\text{def}}{=} q_B(v+w) - q_B(v) - q_B(w)$$

is a bilinear form. Since B_{q_B} is visibly symmetric, we need only check that

$$B_{q_B}(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2, w) = \alpha_1 B_{q_B}(v_1,w) + \alpha_2 B_{q_B}(v_2,w).$$

By definition, we have

$$B_{q_B}(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2, w) = q_B(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + w) - q_B(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2) - q_B(w).$$

We see that

$$\begin{aligned}
q_B(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + w) &= B(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + w, \alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + w) \\
q_B(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2) &= B(\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2, \alpha_1 \cdot v_1 + \alpha_2 \cdot v_2) \\
q_B(w) &= B(w,w) \\
\alpha_1 B_{q_B}(v_1,w) &= \alpha_1 \left(q_B(v_1+w) - q_B(v_1) - q_B(w)\right) = \alpha_1 \left(B(v_1+w, v_1+w) - B(v_1,v_1) - B(w,w)\right) \\
\alpha_2 B_{q_B}(v_2,w) &= \alpha_2 \left(q_B(v_2+w) - q_B(v_2) - q_B(w)\right) = \alpha_2 \left(B(v_2+w, v_2+w) - B(v_2,v_2) - B(w,w)\right).
\end{aligned}$$

Now

$$\begin{aligned}
\alpha_1 \left(B(v_1+w, v_1+w) - B(v_1,v_1) - B(w,w)\right) &= \alpha_1 B(v_1,v_1) 2\alpha_1 B(v_1,w) + \alpha_1 B(w,w) - \alpha_1 B(v_1,v_1) - \alpha_1 B(w,w) \\
&= 2\alpha_1 B(v_1,w)
\end{aligned}$$

and

$$\begin{aligned}\alpha_2\left(B(v_2+w,v_2+w)-B(v_2,v_2)-B(w,w)\right) &= \alpha_2 B(v_2,v_2)2\alpha_2 B(v_2,w)+\alpha_2 B(w,w)-\alpha_2 B(v_2,v_2)-\alpha_2 B(w,w)\\ &= 2\alpha_2 B(v_2,w).\end{aligned}$$

Hence

$$\alpha_1 B_{q_B}(v_1,w)+\alpha_2 B_{q_B}(v_2,w)=2\alpha_1 B(v_1,w)+2\alpha_2 B(v_2,w).$$

Now

$$\begin{aligned}B(\alpha_1\cdot v_1+\alpha_2\cdot v_2+w,\alpha_1\cdot v_1+\alpha_2\cdot v_2+w) &= \alpha_1^2 B(v_1,v_1)+\alpha_1\alpha_2 B(v_1,v_2)+\alpha_1 B(v_1,w)\\ &\quad+\alpha_1\alpha_2 B(v_2,v_1)+\alpha_2^2 B(v_2,v_2)+\alpha_2 B(v_2,w)\\ &\quad+\alpha_1 B(w,v_1)+\alpha_2 B(w,v_2)+B(w,w)\\ &= \alpha_1^2 B(v_1,v_1)+\alpha_2^2 B(v_2,v_2)+B(w,w)\\ &\quad+2\alpha_1\alpha_2 B(v_1,v_2)+2\alpha_1 B(v_1,w)+2\alpha_2 B(v_2,w).\end{aligned}$$

Also

$$B(\alpha_1\cdot v_1+\alpha_2\cdot v_2,\alpha_1\cdot v_1+\alpha_2\cdot v_2)=\alpha_1^2 B(v_1,v_1)+2\alpha_1\alpha_2 B(v_1,v_2)+\alpha_2^2 B(v_2,v_2).$$

Hence

$$q_B(\alpha_1\cdot v_1+\alpha_2\cdot v_2+w)-q_B(\alpha_1\cdot v_1+\alpha_2\cdot v_2)-q_B(w)=2\alpha_1 B(v_1,w)+2\alpha_2 B(v_2,w).$$

Hence B_{q_B} is bilinear. ♠

Lemma 4.17. *If (V,B) is a bilinear space, then $B_{q_B}=2B$.*

Proof. By definition of B_{q_B}, we have

$$B_{q_B}(v,w)=q(v+w)-q(v)-q(w)=B(v+w,v+w)-B(v,v)-B(w,w)=2B(v,w).$$

♠

Remark 30. Recall that $\mathrm{BiL}(V,V)$ is a vector space by Lemma 4.3. As $B\in\mathrm{BiL}(V,V)$, and $B_{q_B}=2B$, we see that $B_{q_B}\in\mathrm{BiL}(V,V)$. Hence Lemma 4.16 follows from Lemma 4.17.

Lemma 4.18. *If (V,q) is a quadratic space, then $q_{B_q}=2q$.*

Proof. We know that

$$B_q(v,w)=q(v+w)-q(v)-q(w).$$

Hence,

$$q_{B_q}(v)=B_q(v,v)=q(2\cdot v)-2q(v)=B(2\cdot v,2\cdot v)-2B(v,v)=2B(v,v)=2q(v).$$

♠

4.1.3 Skew-Symmetric and Alternating Bilinear Forms

Definition 4.19 (Skew-Symmetric)**.** If V is a vector space and $B\colon V \times V \to \mathbf{R}$ is a bilinear form, we say B is **skew-symmetric** if $B(v,w) = -B(w,v)$ for all $v,w \in V$.

Definition 4.20 (Alternating)**.** If V is a vector space and $B\colon V \times V \to \mathbf{R}$ is a bilinear form, we say B is **alternating** if $B(v,v) = 0$ for all $v \in V$.

Example 26. Consider $B\colon \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}$ given by

$$B(x,y) = x_1y_2 - x_2y_1$$

where $x = (x_1,x_2)$ and $y = (y_1,y_2)$. This is called the **determinant** and is denotes by det or more carefully by $\det_2$. More generally, we have $B\colon \mathbf{R}^{2n} \times \mathbf{R}^{2n} \to \mathbf{R}$ given by

$$\sum_{j=1}^{n} (x_{2j-1}y_{2j} - x_{2j}y_{2j-1}) = x_1y_2 - x_2y_1 + x_3y_4 - x_4y_3 + \cdots + x_{2n-1}y_{2n} - x_{2n}y_{2n-1}.$$

These are both alternating and skew-symmetric. These examples on $\mathbf{R}^{2n}$ of one of the standard examples of what are also called symplectic forms.

Lemma 4.21. *If V is a vector space and $B\colon V \times V \to \mathbf{R}$ is a bilinear function, then the following are equivalent:*

(i) B is skew-symmetric.

(ii) B is alternating.

Proof. If B is skew-symmetric, then $B(v,w) = -B(w,v)$ for all $v,w \in V$. Taking $v = w$, we see that $B(v,v) = -B(v,v)$ and so $B(v,v) = 0$. If B is alternating, then

$$B(v+w,v+w) = 0$$

for all $v,w \in V$. However,

$$B(v+w,v+w) = B(v,v) + B(v,w) + B(w,v) + B(w,w) = B(v,w) + B(w,v) = 0.$$

(Hence $B(v,w) = -B(w,v)$. ♠

Exercise 79. Prove that the subset $\Lambda^2(\mathbf{R}^2)$ of $\mathrm{BiL}(\mathbf{R}^2,\mathbf{R}^2)$ of alternating bilinear forms satisfies $\dim(\Lambda^2(\mathbf{R}^2) = 1$. In particular, every $B \in \Lambda^2(\mathbf{R}^2)$ is given by $\alpha \cdot \det_2$.

Alternating forms/Skew-Symmetric forms are central to the theory of integrating differential forms over subsets of $\mathbf{R}^n$ (e.g. line/surface integrals and Stokes' Theorem). A reader familiar with more computational linear algebra likely knows the definition of the determinant of a $A \in \mathrm{M}(n, \mathbf{R})$. If we view the matrix A has n (column) vectors in $\mathbf{R}^n$, then

$$\mathrm{M}(n,\mathbf{R}) = \underbrace{\mathbf{R}^n \times \mathbf{R}^n \times \cdots \times \mathbf{R}^n}_{n \text{ times}} \xrightarrow{\det_n} \mathbf{R}.$$

This function is **multi-linear** in the sense that if you fix all but one of the n variables, then the function is a linear function from $\mathbf{R}^n \to \mathbf{R}$. One remembers that interchanging the columns of an n by n matrix changes the sign of $\det_n$.

4.1.4 Geometric Concepts

Definition 4.22 (Orthogonal Vectors)**.** Given a bilinear space space (V, B) and $v, w \in V$, we say that v, w are ***B*–orthogonal** if $B(v, w) = 0$. We will write $v \perp w$ to denote that v, w are B–orthogonal.

Definition 4.23 (Orthogonal Vector and Subspace)**.** Given a bilinear space space (V, B), $v \in V$, and $W \leq V$, we say that v **is *B*–orthogonal to** W if $B(v, w) = 0$ for all $w \in W$. We will write $v \perp W$ to denote this.

Definition 4.24 (Orthogonal Complement of a Subset)**.** Given a bilinear space space (V, B) and $S \subset V$, we define the ***B*–orthogonal complement of** S to be the subset $S^\perp$ defined by

$$S^\perp \stackrel{\text{def}}{=} \{ w \in V \ : \ v \perp w \text{ for all } v \in S \}.$$

Remark 31. It can happen that $S \subset S^\perp$. Take $B = \langle \cdot, \cdot \rangle_{2,1}$ and $v = (0,1,1)$. Recall that

$$\langle (x,y,z), (x,y,z) \rangle_{2,1} = x^2 + y^2 - z^2.$$

In particular, $\langle v, v \rangle_{2,1} = 0$ and so $v \in v^\perp$. When B is positive definite, then $v \notin v^\perp$ unless $v = 0_V$.

Lemma 4.25. *If (V, B) is a bilinear space and $v \in V$ then $v^\perp$ is a vector subspace.*

Proof. $v^\perp = \ker(B_{v,\star})$ and so is a subspace by Lemma 1.67. ♠

Lemma 4.26. *If (V,B) is a bilinear space and $W \leq V$, then $W^{\perp} \leq V$.*

Proof. We leave this for the reader. ♠

Exercise 80. Prove Lemma 4.26.

Lemma 4.27. *If (V,B) is a bilinear space and $S \subset V$, then $S \subset (S^{\perp})^{\perp}$.*

Proof. We leave this for the reader. ♠

Exercise 81. Prove Lemma 4.27.

Lemma 4.28. *If V is a vector space and B is a bilinear form, then B is non-degenerate if and only if $v^{\perp} \neq V$ for all $v \in V - \{0_V\}$.*

Proof. If B is non-degenerate and $v \in V - \{0_V\}$, then by definition, there exists $w \in V$ such that $B(v,w) \neq 0$. Hence $w \notin v^{\perp}$. If $v^{\perp} \neq V$ for all $v \neq 0_V$, then there must exist $w \in V - v^{\perp}$. Hence $B(v,w) \neq 0$ and so B is non-degenerate. ♠

Definition 4.29 (Positive Definite)**.** Given a vector space V and a real bilinear form B, we say that B is **positive definite** if $B(v,v) > 0$ for all $v \in V - \{0_V\}$.

Definition 4.30 (Negative Definite)**.** Given a vector space V and a real bilinear form B, we say that B is **negative definite** if $B(v,v) < 0$ for all $v \in V - \{0_V\}$.

Example 27. $\langle \cdot, \cdot \rangle$ on $\mathbf{R}^n$ is positive definite. For this, given $x = (x_1, \dots, x_n) \in \mathbf{R}^n$, we have

$$\langle x, x \rangle = \sum_{j=1}^{n} x_j^2 > 0$$

provided $x \neq 0_{\mathbf{R}^n}$. In fact, $\langle \cdot, \cdot \rangle_{p,q}$ is positive definite if and only if $q = 0$. Additionally, $\langle \cdot, \cdot \rangle_{p,q}$ is negative definite if and only if $p = 0$ (i.e. $\langle \cdot, \cdot \rangle_{0,n}$).

Definition 4.31 (Inner Product Space)**.** If (V,B) is a bilinear space and B is positive definite, then we will call (V,B) an **inner product space**.

Definition 4.32 (Norm of a Vector)**.** If (V,B) is an inner product space and $v \in V$, we define the **B–norm of** v to be defined by

$$||v||_B \stackrel{\text{def}}{=} \sqrt{B(v,v)}.$$

4.2 Isometries

Definition 4.33 (Isometry)**.** Given bilinear spaces (V, B_V), (W, B_W) and a linear function $L\colon V \to W$, we say that L is an **isometry** if L is an isomorphism and

$$B_V(v_1, v_2) = B_W(L(v_1), L(v_2))$$

for all $v_1, v_2 \in V$.

Definition 4.34 (Isometric Bilinear Spaces)**.** We say two bilinear spaces (V, B_V) and (W, B_W) are isometric if there exists an isometry $L\colon V \to W$. If (V, B_V) and (W, B_W) are isometric, we write $(V, B_V) \cong (W, B_W)$.

Lemma 4.35. *If (V, B_V) and (W, B_W) are bilinear spaces then the following are equivalent:*

(i) $(V, B_V) \cong (W, B_W)$.

(ii) There exists a isomorphism $L\colon V \to W$ such that

$$\begin{tikzcd} V \arrow[r, "L"] \arrow[d, "B_{V,R}"'] & W \arrow[d, "B_{W,R}"] \\ V^* & W^* \arrow[l, "L^*"] \end{tikzcd}$$

commutes. That is

$$B_{V,R} = L^* \circ B_{W,R} \circ L.$$

Proof. This is a (follow your nose) exercise for the reader. ♠

Exercise 82. Prove Lemma 4.35.

Definition 4.36. Given quadratic spaces (V, q_V), (W, q_W) and a linear function $L\colon V \to W$, we say that L is an **isometry** if L is an isomorphism and

$$q_V(v) = q_W(L(v))$$

for all $v \in V$.

Definition 4.37. Given quadratic spaces (V, q_V) and (W, q_W), we say (V, q_V) and (W, q_W) are **isometric** if there exists an isometry $L\colon V \to W$, If (V, q_V) and (W, q_W) are isometric, we write $(V, q_V) \cong (W, q_W)$.

Theorem 4.38. *Let (V, B_V) and (W, B_W) be bilinear spaces. Then the following are equivalent:*

(i) $(V, B_V) \cong (W, B_W)$.

(ii) $(V, q_{B_V}) \cong (W, q_{B_W})$.

Proof. This follows from Lemma 4.17. ♠

Corollary 4.39. *Let (V, q_V) and (W, q_W) be quadratic spaces. Then the following are equivalent:*

(i) $(V, q_V) \cong (W, q_W)$.

(ii) $(V, B_{q_V}) \cong (W, B_{q_W})$.

Proof. This follows from Theorem 4.38. ♠

Exercise 83. If $(V < B_V), (W, B_W), (U, B_U)$ are bilinear spaces such that $L_1 : V \to W$ and $L_2 : W \to U$ are isometries, then $L_2 \circ L_1$ is an isometry.

Exercise 84. Prove that if $L\colon (V, B_V) \to (W, B_W)$ is an isometry of bilinear spaces, then $L^{-1}\colon (W, B_W) \to (V, B_V)$ is an isometry.

4.3 Orthogonal Bases and the Gram–Schmidt Process

In this section we will present an algorithm for producing orthogonal basis with respect to a fixed bilinear form B.

Definition 4.40 (Orthogonal Sets)**.** Given a bilinear space (V, B) and a subset $S \subset V$, we say that S is **orthogonal** if for each $v, w \in S$ with $v \neq w$, then $B(v, w) = 0$.

Exercise 85. Let $g, h\colon [0, 1] \to \mathbf{R}$ by given by

$$g(x) = ax + b, \quad h(x) = cx^2 + dx + e$$

where $a, b, c, d, e \in \mathbf{R}$. Find values for a, b, c, d, e such that

$$\int_0^1 g(x) = \int_0^1 h(x) = \int_0^1 g(x)h(x) = 0, \quad \int_0^1 (g(x))^2 = \int_0^1 (h(x))^2 = 1.$$

Exercise 86. Let $g_1, \ldots, g_n \colon [0,1] \to \mathbf{R}$ be given by

$$g_j(x) \stackrel{\text{def}}{=} \sum_{k=0}^{j} \alpha_{j,k} x^k.$$

Find values for the $\alpha_{j,k}$ such that

$$\int_0^1 g_j(x) = \int_0^1 g_j(x) g_k(x) = 0$$

for all j and all $j \neq k$ and

$$\int_0^1 (g_j(x))^2 = 1$$

for all j.

Definition 4.41 (Orthogonal Basis)**.** Given a bilinear space (V,B), we say a basis $\mathscr{B}$ is an **orthogonal basis** if $\mathscr{B}$ is orthogonal.

Definition 4.42 (Normal Sets)**.** Given a bilinear space (V,B) and a subset $S \subset V$, we say that S is **normal** if $B(v,v) = 1$ for all $v \in S$.

Definition 4.43 (Normal Basis)**.** Given a bilinear space (V,B), we say a basis $\mathscr{B}$ is a **normal basis** if $\mathscr{B}$ is normal.

Definition 4.44 (Orthonormal Basis)**.** Given a bilinear space (V,B), we say a basis $\mathscr{B}$ is an **orthonormal basis** if $\mathscr{B}$ is orthogonal and normal.

Definition 4.45 (Projection)**.** Given a bilinear space (V,B) and $v, w \in V$ with $B(v,v) \neq 0$, we define the **projection of** w **to** v to be given by

$$\operatorname{Proj}_v(w) \stackrel{\text{def}}{=} \left(\frac{B(v,w)}{B(v,v)} \right) \cdot v.$$

Exercise 87. Let $V = \mathbf{R}^2$ and $B = \langle \cdot, \cdot \rangle$. Compute $\operatorname{Proj}_v(w)$ for $v = e_1$ and $w = (a,b)$.

Exercise 88. Let V be the vector space of continuous functions $f \colon [0,1] \to \mathbf{R}$. Compute $\operatorname{Proj}_v(w)$ for $v = x^2$ and

$$w = \sum_{k=0}^{n} \alpha_k x^k.$$

Exercise 89. Let $V = \mathbf{R}^4$ and $B = \langle \cdot, \cdot \rangle_{3,1}$. Compute $\operatorname{Proj}_v(w)$ for $v = (1,0,1,1)$ and $w = (a,b,c,d)$.

Lemma 4.46. *If (V,B) is a bilinear space and $v,w \in V$ with $B(v,v) \neq 0$, then v and $w - \mathrm{Proj}_v(w)$ are orthogonal.*

Proof. Define

$$u = w - \left(\frac{B(v,w)}{B(v,v)}\right) \cdot v.$$

We see that

$$\begin{aligned} B(u,v) &= B\left(w - \left(\frac{B(v,w)}{B(v,v)}\right) \cdot v, v\right) \\ &= B(w,v) - \left(\frac{B(v,w)}{B(v,v)}\right) B(v,v) \\ &= B(v,w) - B(v,w) = 0. \end{aligned}$$

♠

Lemma 4.47. *If (V,B) is a bilinear space and $v \perp w$, then $v \perp \alpha \cdot w$ for every $\alpha \in \mathbf{R}$.*

Proof. Since $B(v,w) = 0$, we see that

$$B(v, \alpha \cdot w) = \alpha B(v,w) = 0.$$

♠

Lemma 4.48. *If V is a vector space with a basis $\mathscr{B}$ and $\mathscr{B}' = (\mathscr{B} - \{v_0\}) \cup \{w\}$ where $v_0 \in \mathscr{B}$, $\alpha_0 \neq 0$, and*

$$w = \alpha_0 \cdot v_0 + \sum_{v \in \mathscr{B} - \{v_0\}} \alpha_v \cdot v$$

for some α_v with finite support, then $\mathscr{B}'$ is a basis.

Proof. As $\mathscr{B} \subset \mathrm{Span}(\mathscr{B}')$, we see that $\mathscr{B}'$ spans by Lemma 1.40. If

$$\beta_w \cdot w + \sum_{v \neq v_0} \beta_v \cdot v = 0_V$$

for some $\beta_w \in \mathbf{R}$ and some β_v with finite support, then

$$\begin{aligned} 0_V &= \beta_w \cdot \left(\alpha_0 \cdot v_0 + \sum_{v \in \mathscr{B} - \{v_0\}} \alpha_v \cdot v\right) + \sum_{v \neq v_0} \beta_v \cdot v \\ &= \beta_w \alpha_0 \cdot v_0 + \sum_{v \neq v_0} (\beta_w \alpha_v + \beta_v) \cdot . \end{aligned}$$

Hence

$$\beta_w \alpha_0 = 0, \quad \beta_w \alpha_v + \beta_v = 0$$

for all $v \neq v_0$. Since $\alpha_0 \neq 0$, we see that $\beta_w = 0$. Thus, $\beta_v = 0$ for all $v \neq v_0$. Hence $\mathscr{B}'$ is linearly independent. ♠

4.3.1 Gram–Schmidt Process

Theorem 4.49 (Gram–Schmidt: Positive Definite Version)**.** *If (V,B) is an inner product space with* $\dim(V) \leq |\mathbf{N}|$*, then there exists an orthonormal basis.*

Proof of Theorem 4.49. We will assume that $\dim(V) = |\mathbf{N}|$ as the finite case will be handled in the proof of this case. We will build the basis recursively using a process called the **Gram–Schmidt process**. By assumption, we have a basis

$$\mathscr{B} = \{v_1, v_2, v_3, \dots\}.$$

We will replace the vectors in $\mathscr{B}$ one at a time.

In the first stage, we replace v_1 with

$$u_1 \stackrel{\text{def}}{=} \left(\frac{1}{||v_1||}\right) \cdot v_1.$$

We obtain $\mathscr{B}_1 = \{u_1, v_2, v_3, \dots\}$. This is a basis by Lemma 4.48. Before proceeding to the second stage, we verify that $||u_1|| = 1$. For that, we have

$$B\left(\left(\frac{1}{||v_1||}\right) \cdot v_1, \left(\frac{1}{||v_1||}\right) \cdot v_1\right) = \frac{B(v_1,v_1)}{||v_1||^2} = \frac{B(v_1,v_1)}{\left(\sqrt{B(v_1,v_1)}\right)^2} = 1.$$

In the second stage, we take two steps: first we replace v_2 with

$$w_2 \stackrel{\text{def}}{=} v_2 - \mathrm{Proj}_{u_1}(v_2)$$

and then replace w_2 with u_2 given by

$$u_2 \stackrel{\text{def}}{=} \left(\frac{1}{||w_2||}\right) \cdot w_2.$$

We define $\mathscr{B}_2 = \{u_1, u_2, v_3, \dots\}$. This is a basis by Lemma 4.48. Before proceeding to the next stage, we check that $||u_2|| = 1$ and $u_1 \perp u_2$. For $||u_2|| = 1$, the argument is identical to the first stage. We see that $u_1 \perp u_2$ by Lemma 4.46 and Lemma 4.47.

At the third stage, we replace v_3 in two stages again. First, we define

$$w_3 \stackrel{\text{def}}{=} v_3 - \text{Proj}_{u_1}(v_3) - \text{Proj}_{u_2}(v_3)$$

and

$$u_3 \stackrel{\text{def}}{=} \left(\frac{1}{||w_3||}\right) \cdot w_3.$$

We define $\mathscr{B}_3 = \{u_1, u_2, u_3, v_4, \dots\}$. This is a basis by Lemma 4.48. We again check that $u_3 \perp u_1$ and $u_3 \perp u_2$. By Lemma 4.47, it is enough to check that $w_3 \perp u_1, u_2$. For that, we have

$$\begin{aligned}
B(w_3, u_1) &= B(v_3 - \text{Proj}_{u_1}(v_3) - \text{Proj}_{u_2}(v_3), u_1) \\
&= B(v_3 - B(v_3, u_1) \cdot u_1 - B(v_3, u_2) \cdot u_2, u_1) \\
&= B(v_3, u_1) - B(B(v_3, u_1) \cdot u_1, u_1) - B(B(v_3, u_2) \cdot u_2, u_1) \\
&= B(v_3, u_1) - B(v_3, u_1)B(u_1, u_1) - B(v_3, u_2)B(u_2, u_1) \\
&= B(v_3, u_1) - B(v_3, u_1) = 0
\end{aligned}$$

since $B(u_2, u_1) = 0$ and $B(u_1, u_1) = 1$. Likewise, we have

$$\begin{aligned}
B(w_3, u_2) &= B(v_3 - \text{Proj}_{u_1}(v_3) - \text{Proj}_{u_2}(v_3), u_2) \\
&= B(v_3 - B(v_3, u_2) \cdot u_1 - B(v_3, u_2) \cdot u_2, u_2) \\
&= B(v_3, u_2) - B(B(v_3, u_2) \cdot u_1, u_2) - B(B(v_3, u_2) \cdot u_2, u_2) \\
&= B(v_3, u_2) - B(v_3, u_2)B(u_1, u_2) - B(v_3, u_2)B(u_2, u_2) \\
&= B(v_3, u_2) - B(v_3, u_2) = 0
\end{aligned}$$

since $B(u_2, u_1) = 0$ and $B(u_2, u_2) = 1$.

Continuing to the jth stage, we have $\mathscr{B}_{j-1} = \left\{u_1, u_2, \dots, u_{j-1}, v_j, v_{j+1}, \dots\right\}$ with $||u_k|| = 1$ and $u_k \perp u_\ell$ for every $k, \ell \in \{1, \dots, j-1\}$ with $k \neq \ell$. We define

$$w_j \stackrel{\text{def}}{=} v_j - \sum_{k=1}^{j-1} \text{Proj}_{u_k}(v_j)$$

and

$$u_j \stackrel{\text{def}}{=} \left(\frac{1}{||w_j||}\right) \cdot w_j.$$

We define $\mathscr{B}_j = \{u_1, \dots, u_j, v_{j+1}, \dots\}$. This is a basis by Lemma 4.48. It remains to prove that $u_j \perp u_k$ for $k \in \{1, \dots, j-1\}$. For that, we have

$$\begin{aligned} B(w_j, u_k) &= B\left(v_k - \sum_{\ell=1}^{j-1} \mathrm{Proj}_{u_\ell}(v_j), u_k\right) \\ &= B(v_k, u_k) - \sum_{\ell=1}^{j-1} B(\mathrm{Proj}_{u_\ell}(v_j), u_k) \\ &= B(v, u_k) - \sum_{\ell=1}^{j-1} B(B(v_j, u_\ell) \cdot u_\ell, u_k) \\ &= B(v, u_k) - B(v, u_k) = 0 \end{aligned}$$

since $B(u_\ell, u_k) = 0$ for $\ell \neq k$. Continuing further, we see that the desired basis exists recursively. ♠

In the proof of Theorem 4.49, if we instead start with a linearly independent subset S of V, then we can run the Gram–Schmidt process of S, obtaining an orthonormal set S_B with the same span as S. We record this in the following scholium.

Scholium 4.50. *If (V, B) is an inner product space with $\dim(V) \leq \mathbf{N}$ and S is a linearly independent subset of V, then there exists an orthonormal set S_B such that $\mathrm{Span}(S_B) = \mathrm{Span}(S)$. In particular, if $W \leq V$ is a subspace, then there exists an orthonormal basis $\mathscr{B}_V$ for V such that $\mathscr{B}_W = W \cap \mathscr{B}_V$ is a orthonormal basis for W.*

Proof. For the readers' sake, we briefly sketch the proof. Starting with S, we can extend S to a basis $\mathscr{B} = \{v_1, v_2, \dots\}$ by Corollary 2.19. At the jth stage in the Gram–Schmidt process, we have the basis $\mathscr{B}_j = \{u_1, \dots, u_j, v_{j+1}, \dots\}$. The key point needed is that

$$\mathrm{Span}(v_1, \dots, v_j) = \mathrm{Span}(u_1, \dots, u_j)$$

by construction. Hence, if we write $\mathscr{B} = S \cup (\mathscr{B} - S)$ and order it so that the elements of S are listed first, we will produce the desired orthonormal basis. ♠

Exercise 90. Apply the Gram–Schmidt process to $\{v_1, v_2, v_3\} \in \mathbf{R}^3$ with $\langle \cdot, \cdot \rangle$ where

$$v_1 = (1, 0, 0), \quad v_2 = (1, 1, 0), \quad v_3 = (1, 1, 1).$$

We next extend Theorem 4.49 to general bilinear spaces (V, B) with $\dim(V) \leq |\mathbf{N}|$.

Theorem 4.51 (Gram–Schmidt: General Version)**.** *If (V,B) is a bilinear space with* $\dim(V) \leq |\mathbf{N}|$*, then there exists an orthogonal basis.*

To prove this, we require a proposition so that we can run the Gram–Schmidt process. For the proposition, we require the following lemma.

Lemma 4.52. *If (V,B) is a bilinear space, $\mathscr{B}$ is a basis for V, and $v_0 \in V$ is such that $B(v_0,v) = 0$ for all $v \in \mathscr{B}$, then $v_0 = 0_V$.*

Proof. Assume that $B(v_0,v) = 0$ for every $v \in \mathscr{B}$ for some basis $\mathscr{B}$ of V. By Theorem 2.22, given $u \in V$, there a unique α_v with finite support such that

$$u = \sum_{v \in \mathscr{B}} \alpha_v \cdot v.$$

We see that

$$\begin{aligned} B(v_0,u) &= B\left(v_0, \sum_{v \in \mathscr{B}} \alpha_v \cdot v\right) \\ &= \sum_{\alpha \in \mathscr{B}} \alpha_v B(v_0,v) = 0. \end{aligned}$$

Since B is non-degenerate by assumption, we see that $v_0 = 0_V$. ♠

Definition 4.53 (B–good Bases)**.** Given a bilinear space (V,B), we say that a basis $\mathscr{B}$ for V is ***B*–good** if $B(v,v) \neq 0$ for each $v \in \mathscr{B}$.

Proposition 4.54. *If (V,B) is a bilinear space with* $\dim(V) \leq |\mathbf{N}|$*, then there exists a B–good basis.*

Proof. We will construct the basis in stages. As such, the finite case will be taken care off in the infinite case. By assumption, we have a basis

$$\mathscr{B} = \{v_1, v_2, v_3, \dots\}.$$

In the first stage, if $B(v_1,v_1) \neq 0$, then we define $\mathscr{B}_1 = \mathscr{B}$. If $B(v_1,v_1) = 0$, then by Lemma 4.52, there exists $v_j \in \mathscr{B}$ such that $B(v_1,v_{k_1}) \neq 0$. If

$$B(v_{k_1},v_{k_1}) + 2B(v_1,v_{k_1}) \neq 0$$

then define

$$u_1 = v_1 + v_{k_1}$$

and $\mathscr{B}_1 = \{u_1, v_2, \dots\}$. If
$$B(v_{k_1}, v_{k_1}) + 2B(v_1, v_{k_1}) = 0,$$
then define
$$u_1 = v_1 - v_{k_1}$$
and $\mathscr{B}_1 = \{u_1, v_2, \dots\}$. In there case, this is a basis by Lemma 4.48. Before moving to stage two, we verify that $B(u_1, u_1) \neq 0$. We see that
$$B(v_1 \pm v_{k_1}, v_1 \pm v_{k_1}) = B(v_1, v_1) \pm 2B(v_1, v_{k_1}) + B(v_{k_1}, v_{k-1} = B(v_{k_1}, v_{k_1}) \pm 2B(v_1, v_{k_1}).$$
If $B(v_{k_1}, v_{k_1}) + 2B(v_1, v_{k_1}) \neq 0$, then $B(v_1 + v_{k_1}, v_1 + v_{k_1}) \neq 0$. Otherwise, $B(v_1 - v_{k_1}, v_1 - v_{k_1}) \neq 0$.

At the jth stage, we have a basis $\mathscr{B}_{j-1} = \left\{u_1, \dots, u_{j-1}, v_j, v_{j+1}, \dots\right\}$ with $B(u_k, u_k) \neq 0$ for all $k \in \{1, \dots, j-1\}$. If $B(v_j, v_j) \neq 0$, we set $\mathscr{B}_j = \mathscr{B}_{j-1}$. Otherwise, there exists $v_{k_j} \in \mathscr{B}$ such that $B(v_j, v_{k_j}) \neq 0$. We define
$$u_j = v_j + v_{k_j}$$
if $B(v_{k_j}, v_{k_j}) + 2B(v_j, v_{k_j}) \neq 0$ and
$$u_j = v_j - v_{k_j}$$
if $B(v_{k_j}, v_{k_j}) + 2B(v_j, v_{k_j}) = 0$. We then define $\mathscr{B}_j = \left\{u_1, \dots, u_j, v_{j+1}, \dots\right\}$. This is a basis by Lemma 4.48. The desired basis is obtain now inductively. ♠

Proof of Theorem 4.51. By Proposition 4.54, there exists a B–good basis $\mathscr{B}$. The remainder of the proof is similar to the proof of Theorem 4.49. The only modification is that we do not normalize the vectors in this version of the Gram-Schmidt process. In particular, we define
$$\begin{aligned}
&\mathscr{B}_1 = \mathscr{B} = \{v_1, v_2, v_3, \dots\} = \{u_1, v_2, v_3, \dots\} \\
&\mathscr{B}_2 = \left\{u_1, v_2 - \mathrm{Proj}_{u_1}(v_2), v_3, \dots\right\} = \{u_1, u_2, v_3, \dots\} \\
&\quad \vdots \\
&\mathscr{B}_k = \left\{ u_1, \dots, u_{k-1}, v_k - \sum_{j=1}^{k-1} \mathrm{Proj}_{u_j}(v_k), v_{k+1}, \dots \right\} = \{u_1, \dots, u_k, v_{k+1}, \dots\} \\
&\quad \vdots \\
&\mathscr{B}_\infty = \mathscr{B}.
\end{aligned}$$

♠

Exercise 91. Let $(V,B) = (\mathbf{R}^2, \langle\cdot,\cdot\rangle_{1,1})$. Apply the Gram–Schmidt process to the set $\{v_1, v_2\}$ where

$$v_1 = (1,1), \quad v_2 = (1,-1).$$

Exercise 92. Let $(V,B) = (\mathbf{R}^3, \langle\cdot,\cdot\rangle_{2,1}$. Apply the Gram–Schmidt process to $\{v_1, v_2, v_3\}$ where

$$v_1 = (0,1,1), \quad v_2 = (0,1,-1), \quad v_3 = (1,1,1).$$

Corollary 4.55. *If (V,B) is a bilinear space with a B–orthogonal basis $\mathscr{B}$, then $\mathscr{B}$ is B–good.*

Proof. This follows immediately from the assumption that B is non-degenerate and Lemma 4.52. ♠

Corollary 4.56. *If (V,B) is a bilinear space with* $\dim(V) \leq |\mathbf{N}|$, *then there exists an orthogonal basis $\mathscr{B}$ for V such that $B(v,v) = \pm 1$ for all $v \in \mathscr{B}$.*

Proof. We are afforded an orthogonal basis $\mathscr{B}$ such that $B(v,v) \neq 0$ for all $v \in \mathscr{B}$ by Theorem 4.51. Replacing each $v \in \mathscr{B}$ by

$$u_v \overset{\text{def}}{=} \left(\frac{1}{\sqrt{|B(v,v)|}} \right) v$$

we obtain an orthogonal basis $\mathscr{B}'$ with $B(u,u) = \pm 1$ for all $u \in \mathscr{B}'$. ♠

We also have a version of Scholium 4.50 for general bilinear spaces.

Scholium 4.57. *If (V,B) is a bilinear space with* $\dim(V) \leq \mathbf{N}$ *and S is a linearly independent subset of V, then there exists an orthogonal set S_B such that* $\mathrm{Span}(S_B) = \mathrm{Span}(S)$. *In particular, if $W \leq V$ is a subspace, then there exists an orthogonal basis $\mathscr{B}_V$ for V such that $\mathscr{B}_W = W \cap \mathscr{B}_V$ is a orthogonal basis for W.*

4.3.2 Positive and Negative Definite Subspaces

Definition 4.58 (Positive Definite Subspaces)**.** Given a bilinear space (V,B) and a subspace $W \leq V$, we say that B **is positive definite on** W if the restriction of B to W given by

$$B_W : W \times W \longrightarrow \mathbf{R}$$

is positive definite.

Definition 4.59 (Negative Definite Subspaces)**.** Given a bilinear space (V,B) and a subspace $W \leq V$, we say that B **is negative definite on** W if the restriction of B to W given by

$$B_W : W \times W \longrightarrow \mathbf{R}$$

is negative definite.

Lemma 4.60. *If (V,B) is a bilinear space and $W \leq V$ is positive definite with respect to B, then $W^{\perp} \cap W = \{0_V\}$.*

Proof. Given $w \in W$ with $w \neq 0_V$, we know that $B(w,w) \neq 0$ and so $w \notin W^{\perp}$. ♠

Lemma 4.61. *If (V,B) is a bilinear space and $W \leq V$ is negative definite with respect to B, then $W^{\perp} \cap W = \{0_V\}$.*

Proof. We leave this as an exercise. ♠

Exercise 93. Prove Lemma 4.61.

Definition 4.62 (Maximal Positive Definite Subspaces)**.** Given a bilinear space (V,B) and a subspace $W \leq V$, we say that **a maximal positive definite subspace** if W is positive definite with respect to B and satisfies the following condition: If $W' \leq V$ is positive definite with respect to B and $W \subset W'$, then $W = W'$.

Definition 4.63 (Maximal Negative Definite Subspaces)**.** Given a bilinear space (V,B) and a subspace $W \leq V$, we say that **a maximal negative definite subspace** if W is negative definite with respect to B and satisfies the following condition: If $W' \leq V$ is negative definite with respect to B and $W \subset W'$, then $W = W'$.

Lemma 4.64. *If (V,B) is a bilinear space, $W \leq V$ is a positive definite subspace, and $v \in w^{\perp}$ with $B(v,v) > 0$, then $U = \mathrm{Span}(W \cup \{v\})$ is positive definite.*

Proof. Given $u \in U$, there exists a unique $w_0 \in W$ and $\alpha_u \in \mathbf{R}$ such that

$$u = w_0 + \alpha_u \cdot v.$$

In particular,

$$\begin{aligned} B(u,u) &= B(w_0 + \alpha_u \cdot u, w_0 + \alpha_u \cdot u) \\ &= B(w_0,w_0) + 2\alpha_u B(w_0,v) + \alpha_u^2 B(v,v) \\ &= B(w_0,w_0) + \alpha_u^2 B(v,v) > 0 \end{aligned}$$

since W is positive definite. ♠

Lemma 4.65. *If (V,B) is a bilinear space, $W \leq V$ is a negative definite subspace, and $v \in w^{\perp}$ with $B(v,v) < 0$, then $U = \mathrm{Span}(W \cup \{v\})$ is negative definite.*

Proof. The proof is left for the reader. ♠

Exercise 94. Prove Lemma 4.65.

Corollary 4.66. *If (V,B) is a bilinear space and W is a maximal positive definite subspace, then $W^{\perp}$ is negative definite subspace.*

Proof. If $W^{\perp}$ is not negative definite, then there exists $u \in W^{\perp}$ with $B(u,u) > 0$. By Lemma 4.60, we know that $u \notin W$ and by Lemma 4.64, we know that $U = \mathrm{Span}(W \cup \{u\})$ is positive definite. This contradicts that W is a maximal positive definite subspace. ♠

Corollary 4.67. *If (V,B) is a bilinear space and W is a maximal negative definite subspace, then $W^{\perp}$ is a positive definite subspace.*

Proof. We leave this as an exercise. ♠

Exercise 95. Prove Corollary 4.67.

Lemma 4.68. *If (V,B) is a bilinear space and $W \leq V$ is a maximal positive definite subspace such that W^T is a maximal negative definite subspace, then $(W^{\perp})^{\perp} = W$.*

Proof. It follows that $W \subset (W^{\perp})^{\perp}$ by Lemma 4.27 and $(W^{\perp})^{\perp}$ is positive definite by Lemma 4.69. Hence by maximality of W, we must have $W = (W^{\perp})^{\perp}$. ♠

Remark 32. If $W^{\perp}$ is not a maximal negative definite subspace, then there exists a maximal negative definite subspace U with $W^{\perp} \subset U$. We know that $U^{\perp}$ is positive definite by Corollary 4.67.

Lemma 4.69. *If (V,B) is a bilinear space with a positive definite subspace $W \leq V$ and a negative definite subspace $U \leq V$, then $U \cap W = \{0_V\}$.*

Lemma 4.70. *If (V,B) is a bilinear space with $\dim(V) \leq |\mathbf{N}|$ and $W_1, W_2 \leq V$ are maximal positive definite subspaces, then $\dim(W_1) = \dim(W_2)$.*

Proof. We need to show that if W_1, W_2 are maximal positive definite subspaces for B, then $\dim(W_1) = \dim(W_2)$. We have two cases to consider:

Case 1: $\dim(W_1)$ and $\dim(W_2)$ are both finite.

In this case, by Scholium 4.57, we have orthogonal bases $\mathscr{B}_1$ and $\mathscr{B}_2$ of V such that $\mathscr{B}_1 \cap W_1$ is a basis for W_1 and $\mathscr{B}_2 \cap W_2$ is a basis for W_2. We can write

$$\mathscr{B}_1 = \{v_1, \dots, v_r, v_{r+1}, \dots, \}, \quad \mathscr{B}_2 = \{w_1, \dots, w_s, w_{s+1}, \dots\}$$

where

$$W_1 \cap \mathscr{B}_1 = \{v_1, \dots, v_r\}, \quad W_2 \cap \mathscr{B}_2 = \{w_1, \dots, w_s\}.$$

Now we will define

$$U = W_1 + W_2.$$

By Corollary 2.64, we know that

$$\dim(U) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2).$$

Restricting B to U, we obtain the finite dimensional bilinear space (U, B_U). By Scholium 4.57, we can find orthogonal bases for U of the form

$$\mathscr{B}_1' = \left\{v_1, \dots, v_r, v_{r+1}', \dots, v_t'\right\}, \quad \mathscr{B}_2 = \left\{w_1, \dots, w_s, w_{s+1}', \dots, w_t'\right\}$$

where $t = \dim(U)$. Define

$$U' = \mathrm{Span}(w_{s+1}', \dots, w_t').$$

If $B(w_k', w_k') > 0$ for some $k \in \{r+1, \dots, t\}$, then $W_1 + \mathrm{Span}(w_k')$ is a positive definite subspace that properly contains W_1. As W_1 is maximal with respect to being positive definite, we see that U' must be negative definite. Since U' and W_1 is positive definite, we must have $U' \cap W_1 = \{0_V\}$. By Corollary 2.59, we have

$$\dim(W_1) + \dim(U') \leq \dim(U)$$

and so $r + (t - s) \leq t$. Thus, $r \leq s$. Taking

$$U'' = \mathrm{Span}(v_{r+1}', \dots, v_r')$$

we see that $W_2 \cap U'' = \{0_V\}$ and arguing as before, we obtain $s + (t - r) \leq t$. Thus $s \leq r$ and so $r = s$.

Case 2: $\dim(W_1)$ is finite and $\dim(W_2)$ is infinite.

In this case, by Scholium 4.57, we can find an orthogonal basis $\mathscr{B}$ of V such that $W_1 \cap \mathscr{B}$ is a basis for W_1. In particular,

$$\mathscr{B} = \{w_1, \dots, w_r, w_{r+1}, \dots\}, \quad W_1 \cap \mathscr{B} = \{w_1, \dots, w_r\}$$

where $r = \dim(W_1)$. Define

$$U = \mathrm{Span}(\mathscr{B} - \{w_1, \ldots, w_r\}).$$

As before, U must be negative definite since W_1 is maximal with respect to being positive definite for B. Since W_2 is positive definite, we see that $U \cap W_2 = \{0_V\}$. In particular, we see that

$$U + W_2 \leq U + W_1 = V.$$

Now, we have the quotient $L_U : V \to V/U \cong W_1$. Since $\ker(L_U) = U$ and $U \cap W_2 = \{0_V\}$, we see that $L_U(W_2) \cong W_2$. In particular, the restriction of L_U to W_2 is injective. However, this is impossible by Theorem 2.35 since $\dim(W_2) > \dim(W_1)$. ♠

Remark 33. When $\dim(V) < \infty$, we only need to consider Case 1 in the proof of Lemma 4.70. The proof we give in Case 1 is fairly standard. We note that Lemma 4.70 can be proven without cases and direct the reader to the proof of Lemma 4.72 we give below for more on this.

Remark 34. I do not know if there exists a bilinear space (V, B) with maximal positive definite spaces W_1, W_2 such that $\dim(W_1) \neq \dim(W_2)$.

Exercise 96. Prove or disprove: if is a bilinear space (V, B) with maximal positive definite spaces W_1, W_2, then $\dim(W_1) = \dim(W_2)$.

Definition 4.71 (Positive Definite Dimension)**.** Given a bilinear space (V, B) with $\dim(V) \leq |\mathbf{N}|$, we define the **positive definite dimension of** B to be $\dim_+(V, B) \stackrel{\text{def}}{=} \dim(W)$ for any $W \leq V$ that is a maximal positive definite subspace.

Remark 35. I do not know if there exists a bilinear space (V, B) with maximal positive definite spaces W_1, W_2 such that $\dim(W_1) \neq \dim(W_2)$.

Exercise 97. Prove or disprove: If is a bilinear space (V, B) with maximal positive definite spaces W_1, W_2, then $\dim(W_1) = \dim(W_2)$.

Lemma 4.72. *If (V, B) is a bilinear space with $\dim(V) \leq |\mathbf{N}|$ and $W_1, W_2 \leq V$ are maximal negative definite subspaces, then $\dim(W_1) = \dim(W_2)$.*

Proof. One can prove this using a logically identical proof as that given for Lemma 4.70. For the enjoyment of the reader, we will give a slightly different proof of Lemma 4.70 without cases. Let W_1, W_2 be maximal negative definite subspaces. By Scholium 4.57, we can find orthogonal bases $\mathscr{B}_1, \mathscr{B}_2$ such that $\mathscr{B}_{W_j} = W_j \cap \mathscr{B}_j$ is a basis for W_j. Define

$$U_j \stackrel{\text{def}}{=} \mathrm{Span}(\mathscr{B}_j - \mathscr{B}_{W_j})$$

for $j = 1,2$. It follows that $U_j \cap W_j = \{0_V\}$ and that U_j is positive definite for $j = 1,2$. Hence $W_j \cap U_k = \{0_V\}$ for $j,k = 1,2$. Thus the restriction of L_{U_j} to W_k is injective for $j,k = 1,2$. As $V/U_j \cong W_j$, we see that L_{U_1} provides a linear injection of W_2 into W_1 and L_{U_2} provides a linear injection of W_1 into W_2. Thus $\dim(W_1) = \dim(W_2)$ by Theorem 2.35. ♠

Definition 4.73 (Negative Definite Dimension)**.** Given a bilinear space (V,B) with $\dim(V) \leq |\mathbf{N}|$, we define the **negative definite dimension of** B to be $\dim_+(V,B) \stackrel{\text{def}}{=} \dim(W)$ for any $W \leq V$ that is a maximal negative definite subspace.

Remark 36. I do not know if there exists a bilinear space (V,B) with maximal negative definite spaces W_1, W_2 such that $\dim(W_1) \neq \dim(W_2)$.

Exercise 98. Prove or disprove: If is a bilinear space (V,B) with maximal negative definite spaces W_1, W_2, then $\dim(W_1) = \dim(W_2)$.

Theorem 4.74. *If (V,B) is a bilinear space with $\dim(V) \leq |\mathbf{N}|$, then the following is true:*

(a) If $W \leq V$ is a maximal positive definite subspace, then $W^\perp$ is a maximal negative definite subspace and $V = W + W^\perp$.

(b) If $W \leq V$ is a maximal negative definite subspace, then $W^\perp$ is a maximal positive definite subspace and $V = W + W^\perp$.

Proof. For (a), given $W \leq V$, we can find a orthogonal basis $\mathscr{B}$ for V such that $\mathscr{B}_W \stackrel{\text{def}}{=} W \cap \mathscr{B}$ is a basis for W. Define

$$U \stackrel{\text{def}}{=} \operatorname{Span}(\mathscr{B} - \mathscr{B}_W).$$

Since $W \perp v$ for every $v \in \mathscr{B} - \mathscr{B}_W$, we see that $U \subset W^\perp$. Since $V = W + U$, we see that $V = W + W^\perp$. If $W^\perp$ is not a maximal negative definite subspace, then $W^\perp \subset U'$ for some maximal negative definite subspace $U' \leq V$. It follows that $w \in U'$ for some $w \in W$ since $V = W + W^\perp$. But $B(w,w) > 0$ and so U' cannot be negative definite. Hence $W^\perp$ is a maximal negative definite subspace.

We leave (b) to the reader. ♠

Exercise 99. Prove (b) of Theorem 4.74.

Remark 37. Theorem 4.74 does not likely hold in general without the assumption $\dim(V) \leq |\mathbf{N}|$.

4.3.3 Signature and the Law of Inertia

Definition 4.75 (Signature)**.** Given a bilinear space (V,B) with $\dim(V) \leq |\mathbf{N}|$, we define the **signature of** B to be $\sigma(V,B) \overset{\text{def}}{=} (\dim_+(V,B), \dim_-(V,B))$.

Theorem 4.76 (Law of Inertia)**.** *If $(V,B_1),(V,B_2)$ are bilinear spaces with $\dim(V) \leq |\mathbf{N}|$, then the following are equivalent:*

(i) $\sigma(V,B_1) = \sigma(V,B_2)$.

(ii) $(V,B_1) \cong (V,B_2)$.

Proof. We leave this for the reader. ♠

Exercise 100. Prove Theorem 4.76.

Corollary 4.77. *If V is a vector space with $\dim(V) \leq |\mathbf{N}|$ with two positive definite bilinear forms B_1, B_2, then $(V,B_1) \cong (V,B_2)$.*

Proof. This follows from Theorem 4.76. ♠

Exercise 101. Let (V,B) be a bilinear space, $W \leq V$ be a maximal positive definite subspace, and $L\colon V \to V$ be an isometry.

(a) Prove that $L(W) \leq V$ is a maximal positive definite subspace.

(b) Prove that $L^{-1}(W) \leq V$ is a maximal positive definite subspace.

Exercise 102. Let (V,B) be a bilinear space, $W \leq V$ be a maximal negative definite subspace, and $L\colon V \to V$ be an isometry.

(a) Prove that $L(W) \leq V$ is a maximal negative definite subspace.

(b) Prove that $L^{-1}(W) \leq V$ is a maximal negative definite subspace.

Exercise 103. Let (V,B) be a finite dimensional bilinear space and let $W_1, W_2 \leq V$ be maximal positive definite subspaces. Prove that there exists an isometry $L\colon V \to V$ such that $L(W_1) = W_2$.

Exercise 104. Let (V,B) be a finite dimensional bilinear space and let $W_1, W_2 \leq V$ be negative positive definite subspaces. Prove that there exists an isometry $L\colon V \to V$ such that $L(W_1) = W_2$.

4.4 Classic Inner Product Space Results

We now establish some standard results for inner product spaces.

Lemma 4.78 (Pythagorean Theorem). *If (V,B) is a positive definite space and $v,w \in V$ with $v \perp w$, then*

$$||v+w||_B^2 = ||v||_B^2 + ||w||_B^2 .$$

Proof. For this, we have

$$\begin{aligned} ||v+w||_B^2 &= B(v+w, v+w) \\ &= B(v,v) + 2B(v,w) + B(w,w) \\ &= B(v,v) + B(w,w) = ||v||_B^2 + ||w||_B^2 . \end{aligned}$$

♠

Lemma 4.79 (Cauchy–Schwartz). *If (V,B) is an inner product space and $v,w \in V$, then*

$$|B(v,w)| \leq ||v||_B \, ||w||_B$$

with equality if and only if $w = \alpha \cdot v$ for some $\alpha \in \mathbf{R}$.

Proof. If $v = 0_V$, then this is true by Lemma 4.7. Otherwise, define

$$u = w - \left(\frac{B(v,w)}{B(v,v)} \right) \cdot v = w - \mathrm{Proj}_v(w).$$

Now

$$w = u + \left(\frac{B(v,w)}{B(v,v)} \right) \cdot v.$$

By Lemma 4.46, Lemma 4.47, and Lemma 4.78, we have

$$B(w,w) = ||w||_B^2 = ||u||_B^2 + \left|\left| \left(\frac{B(v,w)}{B(v,v)} \right) \cdot v \right|\right|_B^2 = ||u||_B^2 + \frac{(B(v,w))^2}{B(v,v)} \geq \frac{(B(v,w))^2}{B(v,v)}.$$

Hence

$$B(v,v)B(w,w) \geq (B(v,w))^2$$

and so

$$||v||_B \, ||w||_B = \sqrt{B(v,v)}\sqrt{B(w,w)} \geq |B(v,w)| .$$

We have equality only when $u = 0$ since $||u||_B^2 > 0$. Hence, $w = \alpha \cdot v$ where $\alpha = B(v,w)/B(v,v)$.

♠

Remark 38. Cauchy–Schwartz implies that

$$-1 \leq \frac{B(v,w)}{||v||_B\,||w||_B} \leq 1.$$

In particular,

$$\cos^{-1}\left(\frac{B(v,w)}{||v||_B\,||w||_B}\right)$$

makes sense since $\cos^{-1}\colon [-1,1] \to [0,\pi]$.

Definition 4.80 (Angle Between Vectors)**.** If (V,B) is an inner product space and $v,w \in V - \{0_V\}$, we define the angle $\theta_{v,w} \in [0,\pi]$ between v and w to be defined by

$$\theta_{v,w} \stackrel{\text{def}}{=} \cos^{-1}\left(\frac{B(v,w)}{||v||_B\,||w||_B}\right).$$

Lemma 4.81. *If (V,B) is an inner product space and $v,w \in V$ are non-zero, then $v \perp w$ if and only if $\theta_{v,w} = \pi/2$.*

Proof. This follows from the definition of orthogonal and $\theta_{v,w}$. ♠

Lemma 4.82 (Triangle Inequality)**.** *If (V,B) is an inner product space and $v,w \in V$, then*

$$||v+w||_B \leq ||v||_B + ||w||_B\,.$$

Proof. For this, we have

$$\begin{aligned}
||v+w||_B^2 &= B(v+w,v+w) = B(v,v)+2B(v,w)+B(w,w) \\
&\leq B(v,v)+2\,|B(v,w)|+B(w,w) \leq B(v,v)+2\,||v||_B\,||w||_B + B(w,w) \\
&= ||v||_B^2 + 2\,||v||_B\,||w||_B + ||w||_B^2 = (||v||_B + ||w||_B)^2\,.
\end{aligned}$$

Hence

$$||v+w||_B^2 \leq ||v||_B^2 + ||w||_B^2$$

and so the result follows from taking the square root of both sides. ♠

Proposition 4.83 (Parseval's Identity)**.** *If (V,B) is a positive definite space and $v_1,\ldots,v_n$ are orthogonal vectors, then*

$$\left|\left|\sum_{j=1}^{n} v_j\right|\right|_B^2 = \sum_{j=1}^{n} ||v_j||_B^2\,.$$

Proof. This follows from induction and Lemma 4.78. ♠

Exercise 105. Prove that if (V,B) is an bilinear space and $\{v_1,\dots,v_n\} \in V$ such that

$$B(v_j,v_k) = 0, \quad B(v_j,v_j) \neq 0$$

for all $j,k \in \{1,\dots,n\}$ with $j \neq k$ and all $j \in \{1,\dots,n\}$, then $\{v_1,\dots,v_n\}$ is linearly independent.

Exercise 106. Prove that if (V,B) is an inner product space with vectors $v_1,\dots,v_n,w \in V$ such that

$$B(w,v_j) > 0, \quad B(v_j,v_k) \leq 0$$

for all $j \in \{1,\dots,n\}$ and all $j,k \in \{1,\dots,n\}$ with $j \neq k$, then $\{v_1,\dots,v_n\}$ is linearly independent.

Exercise 107. Let (V,B) be an inner product space with an orthogonal set $\{v_1,\dots,v_n\}$. Prove that if $u,w \in \mathrm{Span}(v_1,\dots,v_n)$, then

$$B(u,w) = \sum_{k=1}^{n} B(u,v_k)B(w,v_k).$$

Exercise 108. Let (V,B) be a bilinear space with an orthogonal set $\{v_1,\dots,v_n\}$. Prove that

$$B\left(\sum_{k=1}^{n} v_k, \sum_{k=1}^{n} v_k\right) = \sum_{k=1}^{n} B(v_k,v_k).$$

Exercise 109. If (V,B) is an inner product space and $v,w \in V$ such that $\theta_{v,w} = 0$ or π, then $v = \alpha \cdot w$ for some $\alpha \in \mathbf{R}$.

4.5 Finite Dimensional Bilinear Spaces

In this section, we will consider finite dimensional bilinear spaces.

Definition 4.84 (Finite Dimensional Bilinear Spaces)**.** Given a bilinear space (V,B), we say (V,B) is **finite dimensional** when $\dim(V) < \infty$.

If (V,B) is finite dimensional with $n = \dim(V)$, we know that there exists a B–orthogonal basis

$$\mathscr{B} = \{u_1,\dots,u_n\}$$

with $B(u_j,u_j) = \pm 1$. Define

$$\varepsilon_j \stackrel{\text{def}}{=} B(u_j,u_j) \in \{-1,1\},$$

and note that both $\varepsilon_j^{-1} = \varepsilon_j$ and $\varepsilon_j^2 = 1$ hold always. Given $v \in V$, there exist unique $\alpha_1, \ldots, \alpha_n \in \mathbf{R}$ such that

$$v = \sum_{j=1}^{n} \alpha_j \cdot u_j.$$

Lemma 4.85. *If (V,B) is a bilinear space with B–orthogonal basis $\mathscr{B}$ and $v,w \in V$ with*

$$B(v,u) = B(w,u)$$

for all $u \in \mathscr{B}$, then $v = w$.

Proof. There exist unique α_u, β_u with finite support such that

$$v = \sum_{u \in \mathscr{B}} \alpha_u \cdot u, \quad \sum_{u \in \mathscr{B}} \beta_u \cdot u.$$

For $u' \in \mathscr{B}$, We see that

$$\begin{aligned} B(v,u') &= B\left(\sum_{u \in \mathscr{B}} \alpha_u \cdot u, u'\right) \\ &= \sum_{u \in \mathscr{B}} \alpha_u B(u,u') = \alpha_u B(u',u'). \end{aligned}$$

Thus we have by assumption on v,w that

$$\alpha_{u'} B(u',u') = \beta_{u'} B(u',u')$$

for every $u' \in \mathscr{B}$. Hence $\alpha_{u'} = \beta_{u'}$ since $B(u',u') \neq 0$ and so $v = w$. ♠

Proposition 4.86. *If (V,B) is a finite dimensional bilinear space with B–orthogonal basis $\mathscr{B} = \{u_1, \ldots, u_n\}$ and $v \in V$, then*

$$v = \sum_{j=1}^{n} \varepsilon_j B(v,u_j) \cdot u_j.$$

Proof. We know that there exist unique $\alpha_1, \ldots, \alpha_n \in \mathbf{R}$ such that

$$v = \sum_{j=1}^{n} \alpha_j \cdot u_j.$$

Define

$$w = \sum_{j=1}^{n} \varepsilon_j B(v,u_j) \cdot u_j.$$

We see that

$$\begin{aligned} B(w,u_k) &= B\left(\sum_{j=1}^{n} \varepsilon_j B(v,u_j)\cdot u_j, u_k\right) \\ &= \sum_{j=1}^{n} \varepsilon_j^2 B(v,u_j) = B(v,u_k). \end{aligned}$$

As $B(v,u_k) = B(w,u_k)$ for all $k \in \{1,\dots,n\}$, we have $v = w$ by Lemma 4.85. ♠

Proposition 4.87. *If (V,B) is a finite dimensional bilinear space with B–orthogonal basis $\mathscr{B} = \{u_1,\dots,u_n\}$ and $v \in V$ with*

$$v = \sum_{j=1}^{n} \alpha_j \cdot u_j$$

then

$$B(v,v) = \sum_{j=1}^{n} \alpha_j^2 \varepsilon_j.$$

Proof. This is straightforward. ♠

Exercise 110. Prove Proposition 4.87.

Corollary 4.88. *If (V,B) is a finite dimensional positive definite space with a B–orthonormal basis $\{u_1,\dots,u_n\}$ and $v \in V$, then*

$$v = \sum_{j=1}^{n} B(v,u_j)\cdot u_j.$$

If

$$v = \sum_{j=1}^{n} \alpha_j \cdot u_j,$$

then

$$||v||_B = \sqrt{\sum_{j=1}^{n} \alpha_j^2} = \sqrt{\sum_{j=1}^{n} (B(v,u_j))^2}.$$

Exercise 111. Prove Corollary 4.88.

Corollary 4.89. *If (V,B) is a finite dimensional positive definite space with B–orthonormal basis $\{u_1,\dots,u_n\}$ and $v,w\in V$, then*

$$B(v,w)=\sum_{j=1}^{n}B(v,u_j)B(w,u_j).$$

Moreover, if

$$v=\sum_{j=1}^{n}\alpha_j\cdot u_j,\quad w=\sum_{j=1}^{n}\beta_j\cdot u_j,$$

then

$$B(v,w)=\sum_{j=1}^{n}\alpha_j\beta_j.$$

Exercise 112. Prove Corollary 4.89.

Corollary 4.90. *If (V,B) is a finite dimensional positive definite space with $\dim(V)=n$, then $(V,B)\cong(\mathbf{R}^n,\langle\cdot,\cdot\rangle)$*

Exercise 113. Prove Corollary 4.90.

If (V,B) is a finite dimensional bilinear space with a basis $\mathscr{B}=\{v_1,\dots,v_n\}$, then we can define

$$(M_B)_{j,k}\stackrel{\text{def}}{=}B(v_j,v_k).$$

This defines a matrix $M_B\in\mathrm{M}(n,\mathbf{R})$ by

$$M_B\stackrel{\text{def}}{=}\begin{pmatrix}B(v_1,v_1)&B(v_1,v_2)&B(v_1,v_3)&\dots&B(v_1,v_n)\\B(v_2,v_1)&B(v_2,v_2)&B(v_2,v_3)&\dots&B(v_2,v_n)\\B(v_3,v_1)&B(v_3,v_2)&B(v_3,v_3)&\dots&B(v_3,v_n)\\\vdots&\vdots&\vdots&\ddots&\vdots\\B(v_n,v_1)&B(v_n,v_2)&B(v_n,v_3)&\dots&B(v_n,v_n)\end{pmatrix}.$$

Since $B(v_j,v_k)=B(v_k,v_j)$, we see that M_B is symmetric. That is

$$(M_B)_{j,k}=(M_B)_{k,j}$$

or

$$M_B^T=M_B$$

where M_B^T is the transpose.

Definition 4.91. If (V,B) is a finite dimensional bilinear space with a basis $\mathscr{B} = \{v_1,\dots,v_n\}$, then we define the **symmetric matric associated to B and $\mathscr{B}$** to be $M_B \in \mathrm{M}(n,\mathbf{R})$ given by

$$M_B \overset{\text{def}}{=} \begin{pmatrix} B(v_1,v_1) & B(v_1,v_2) & B(v_1,v_3) & \dots & B(v_1,v_n) \\ B(v_2,v_1) & B(v_2,v_2) & B(v_2,v_3) & \dots & B(v_2,v_n) \\ B(v_3,v_1) & B(v_3,v_2) & B(v_3,v_3) & \dots & B(v_3,v_n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B(v_n,v_1) & B(v_n,v_2) & B(v_n,v_3) & \dots & B(v_n,v_n) \end{pmatrix}.$$

Remark 39. M_B depends on the basis $\mathscr{B}$. Indeed, if $\mathscr{B}$ happens to be an orthogonal basis for B, then

$$M_B = \begin{pmatrix} B(v_1,v_1) & 0 & 0 & \dots & 0 \\ 0 & B(v_2,v_2) & 0 & \dots & 0 \\ 0 & 0 & B(v_3,v_3) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & B(v_n,v_n) \end{pmatrix}.$$

Using Gram–Schmidt, we can find an orthogonal basis $\mathscr{B} = \{v_1,\dots,v_n\}$ with $B(v_j,v_j) = \pm 1$. Reordering the basis, we can list the positive ones first and the negative ones last. If there are r positive ones and s negative ones, then M_B in this basis will be

$$M_B = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & -1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & -1 \end{pmatrix}.$$

Lemma 4.92. *If (V,B) is a finite dimensional bilinear space with a basis $\mathscr{B} = \{v_1,\dots,v_n\}$, then*

$$B(v,w) = \langle M_B v, w \rangle .$$

Proof. Given $v,w \in V$, we need to show that

$$B(v,w) = \langle M_B v, w \rangle .$$

We know there exist unique $\alpha_1,\dots,\alpha_n,\beta_1,\dots,\beta_n \in \mathbf{R}$ such that

$$v = \sum_{j=1}^{n} \alpha_j \cdot v_j, \quad w = \sum_{j=1}^{n} \beta_j \cdot v_j.$$

We see then that

$$
\begin{aligned}
B(v,w) &= B\left(\sum_{j=1}^{n} \alpha_j \cdot v_j, \sum_{k=1}^{n} \beta_k \cdot v_k\right) \\
&= \sum_{j=1}^{n}\sum_{k=1}^{n} \alpha_j \beta_k B(v_j, v_k).
\end{aligned}
$$

Now, we have

$$
\begin{aligned}
M_B v &= \begin{pmatrix} B(v_1,v_1) & B(v_1,v_2) & B(v_1,v_3) & \dots & B(v_1,v_n) \\ B(v_2,v_1) & B(v_2,v_2) & B(v_2,v_3) & \dots & B(v_2,v_n) \\ B(v_3,v_1) & B(v_3,v_2) & B(v_3,v_3) & \dots & B(v_3,v_n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B(v_n,v_1) & B(v_n,v_2) & B(v_n,v_3) & \dots & B(v_n,v_n) \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_n \end{pmatrix} \\
&= \begin{pmatrix} \sum_{j=1}^{n} \alpha_j B(v_1, v_j) \\ \sum_{j=1}^{n} \alpha_j B(v_2, v_j) \\ \sum_{j=1}^{n} \alpha_j B(v_3, v_j) \\ \vdots \\ \sum_{j=1}^{n} \alpha_j B(v_n, v_j) \end{pmatrix}.
\end{aligned}
$$

Finally, we have

$$
\begin{aligned}
\langle M_B v, w\rangle &= \left\langle \begin{pmatrix} \sum_{j=1}^{n} \alpha_j B(v_1, v_j) \\ \sum_{j=1}^{n} \alpha_j B(v_2, v_j) \\ \sum_{j=1}^{n} \alpha_j B(v_3, v_j) \\ \vdots \\ \sum_{j=1}^{n} \alpha_j B(v_n, v_j) \end{pmatrix}, \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_n \end{pmatrix} \right\rangle \\
&= \sum_{k=1}^{n} \beta_k \left(\sum_{j=1}^{n} \alpha_j B(v_k, v_j) \right) = \sum_{j=1}^{n}\sum_{k=1}^{n} \alpha_j \beta_k B(v_j, v_k).
\end{aligned}
$$

as needed. ♠

Given a symmetric matrix $M \in \mathrm{M}(n, \mathbf{R})$, we can define a function $B_M \colon V \times V \to \mathbf{R}$ by

$$
B_M(v,w) \stackrel{\text{def}}{=} \langle Mv, w\rangle .
$$

Lemma 4.93. *If $M \in \mathrm{M}(n, \mathbf{R})$, then*

$$
\langle Mv, w\rangle = \left\langle v, w^T M^T \right\rangle .
$$

Proof. If $M = (\mu_{j,k})$, $v = (\alpha_1, \ldots, \alpha_n)$, and $w = (\beta_1, \ldots, \beta_n)$, we see that

$$\begin{aligned} w^T M^T &= \begin{pmatrix} \beta_1 & \beta_2 & \beta_3 & \ldots & \beta_n \end{pmatrix} \begin{pmatrix} \mu_{1,1} & \mu_{2,1} & \mu_{3,1} & \cdots & \mu_{n,1} \\ \mu_{1,2} & \mu_{2,2} & \mu_{3,2} & \cdots & \mu_{n,2} \\ \mu_{1,3} & \mu_{2,3} & \mu_{3,3} & \cdots & \mu_{n,3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{1,n} & \mu_{2,n} & \mu_{3,n} & \cdots & \mu_{n,n} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{j=1}^n \mu_{1,j}\beta_j & \sum_{j=1}^n \mu_{2,j}\beta_j & \sum_{j=1}^n \mu_{3,j}\beta_j & \ldots & \sum_{j=1}^n \mu_{n,j}\beta_j \end{pmatrix} \end{aligned}$$

Next, we have

$$\begin{aligned} \langle v, M^T w \rangle &= \left\langle \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_n \end{pmatrix}, \begin{pmatrix} \sum_{j=1}^n \mu_{1,j}\beta_j \\ \sum_{j=1}^n \mu_{2,j}\beta_j \\ \sum_{j=1}^n \mu_{3,j}\beta_j \\ \vdots \\ \sum_{j=1}^n \mu_{n,j}\beta_j \end{pmatrix} \right\rangle \\ &= \sum_{k=1}^n \alpha_k \left(\sum_{j=1}^n \beta_j \mu_{k,j} \right) = \sum_{j,k=1}^n \alpha_k \beta_j \mu_{k,j}. \end{aligned}$$

We saw in the proof of the previous lemma that

$$\langle Mv, w \rangle = \sum_{j,k=1}^n \alpha_j \beta_k \mu_{j,k} = \sum_{j,k=1}^n \alpha_k \beta_j \mu_{k,j}$$

and so $\langle Mv, w \rangle = \langle v, w^T M^T \rangle$. ♠

Corollary 4.94. *If $M \in \mathrm{M}(n, \mathbf{R})$ is symmetric, then B_M is a symmetric bilinear form where*

$$B_M(v, w) = \langle Mv, w \rangle.$$

Exercise 114. Prove Corollary 4.94.

4.6 Orthogonal Complements

Lemma 4.95. *If (V, B) is a bilinear space, $S \subset V$, and $v \in S^\perp$, then $v \in (\mathrm{Span}(S))^\perp$.*

Proof. This is straightforward. ♠

Exercise 115. Prove Lemma 4.95.

Lemma 4.96. *If (V,B) is an inner product space with $\dim(V) \leq |\mathbf{N}|$, $W \leq V$, and $\mathscr{B}_V$ is an orthonormal basis for V such that $\mathscr{B}_W \stackrel{def}{=} \mathscr{B}_V \cap W$ is a basis for W, then $\mathscr{B}_{W^\perp} \stackrel{def}{=} \mathscr{B}_V - \mathscr{B}_W$ is a basis for $W^\perp$.*

Proof. We need to show that $\mathrm{Span}(\mathscr{B}_{W^\perp}) = W^\perp$ since we know that $\mathscr{B}_{W^\perp}$ is linearly independent. Given $u_0 \in \mathrm{Span}(\mathscr{B}_{W^\perp})$, we know that

$$u_0 = \sum_{u \in \mathscr{B}_{W^\perp}} \beta_u \cdot u$$

for some β_u with finite support. Given $w_0 \in W$, we know that

$$w_0 = \sum_{w \in \mathscr{B}_W} \alpha_w \cdot w$$

for some α_w with finite support. Now, we have

$$\begin{aligned} B(w_0,u_0) &= B\left(\sum_{w \in \mathscr{B}_W} \alpha_w \cdot w, \sum_{u \in \mathscr{B}_{W^\perp}} \beta_u \cdot u\right) \\ &= \sum_{w \in \mathscr{B}_W} \sum_{u \in \mathscr{B}_{W^\perp}} \alpha_w \beta_u B(w,u) = 0 \end{aligned}$$

as $B(w,u) = 0$ for all $w \in \mathscr{B}_W$ and $u \in \mathscr{B}_{W^\perp}$ since $\mathscr{B}_V$ is an orthogonal basis. Hence $\mathrm{Span}(\mathscr{B}_{W^\perp}) \subset W^\perp$. Given $u_0 \in W^\perp$, we know that

$$u_0 = \sum_{w \in \mathscr{B}_W} \alpha_w \cdot w + \sum_{u \in \mathscr{B}_{W^\perp}} \beta_u \cdot u$$

for some α_w, β_u with finite support (and defined on $\mathscr{B}_W, \mathscr{B}_{W^\perp}$ respectively). Given $w_0 \in \mathscr{B}_W$, we see that

$$B(w_0,u_0) = \alpha_{w_0}.$$

Since $u_0 \in W^\perp$, we have $B(w_0,u_0) = 0$ for all $w_0 \in \mathscr{B}$. Thus,

$$u_0 = \sum_{u \in \mathscr{B}_{W^\perp}} \beta_u \cdot u \in \mathrm{Span}(\mathscr{B}_{W^\perp}).$$

Hence $W^\perp \subset \mathrm{Span}(\mathscr{B}_{W^\perp})$ and thus $\mathrm{Span}(\mathscr{B}_{W^\perp}) = W^\perp$. ♠

Corollary 4.97. *If (V,B) is an inner product space with $\dim(V) \leq |\mathbf{N}|$ and $W \leq V$, then $V \cong W \times W^{\perp}$.*

Proof. This follows from Lemma 4.96 and Corollary 2.59. ♠

Exercise 116. Given a vector space V and a linearly independent set $S \subset V$, prove there exists a bilinear form B such that S is B–orthonormal.

Exercise 117. Let $(V,B) = (\mathbf{R} \cdot \langle \cdot,\cdot \rangle)$, let

$$W = \left\{(x,y,z) \in \mathbf{R}^3 \ : \ x+y+z = 0\right\}, \quad U = \left\{(x,y,z) \in \mathbf{R}^3 \ : \ 2x-3y+9z = 0\right\}.$$

(a) Compute $W \cap U$.

(b) Compute $W^{\perp}$ and $U^{\perp}$.

(c) Compute $(W \cap U)^{\perp}$.

(d) Prove that $(W \cap U)^{\perp} = W^{\perp} \cap U^{\perp}$.

Exercise 118. Let $(V,B) = (\mathbf{R} \cdot \langle \cdot,\cdot \rangle_{2,1})$, let

$$W = \left\{(x,y,z) \in \mathbf{R}^3 \ : \ x+y+z = 0\right\}, \quad U = \left\{(x,y,z) \in \mathbf{R}^3 \ : \ 2x-3y+9z = 0\right\}.$$

(a) Compute $W \cap U$.

(b) Compute $W^{\perp}$ and $U^{\perp}$.

(c) Compute $(W \cap U)^{\perp}$.

(d) Prove that $(W \cap U)^{\perp} = W^{\perp} \cap U^{\perp}$.

Exercise 119. Let (V,B) be a bilinear space with $U,W \leq V$. Prove or disprove:

$$U^{\perp} \cap W^{\perp} \leq (U \cap W)^{\perp}.$$

Exercise 120. Let (V,B) be a bilinear space with $U,W \leq V$. Prove or disprove:

$$U^{\perp} \cap W^{\perp} \geq (U \cap W)^{\perp}.$$

Exercise 121. Let (V,B) be a bilinear space and $L\colon V \to V$ be a linear function. We say that L is **B–self-adjoint** if $B(L(v),w) = B(v,L(w))$. Prove that the subset of $\mathrm{Hom}(V,V)$ of B–self-adjoint linear functions is a vector subspace.

Exercise 122. Let (V,B) be a bilinear space and $L\colon V \to V$ be a linear function. Prove or disprove: There exists a linear function $L'\colon V \to V$ such that $B(L(v),w) = B(v,L'(w))$ for all $v,w \in V$.