

# Chapter 5

## Linear Transformations

### 5.1 Definition and examples

Mathematics features all kind of “categories”, which are sets endowed with a structure. This course is concerned with the category of vector spaces over a field  $\mathbb{F}$ , which are sets endowed with a notion of linear combinations. A major reason for defining vector spaces is that they are abundant—there are many vector spaces of interest in mathematics and its applications; indeed, it wouldn’t make sense to define a class of objects if there was only one such object in this class. Thus, we often encounter situations in which there are multiple vector spaces (over the same field). In such cases, we might be interested in looking at functions between two such objects.

Let  $(V, +, \mathbb{F}, \cdot)$  and  $(W, +, \mathbb{F}, \cdot)$  be two vector spaces over the same field. The set

$$\text{Func}(V, W) = \{f : V \rightarrow W\}$$

is the space of functions with domain (תחום)  $V$  and codomain (טווח)  $W$ . But just as with the linear forms on  $V$ , which are functions  $\mathbb{F}^V$ , we delineate a subset of all functions that “respect” the vector space structure:

**Definition 5.1** Let  $(V, +, \mathbb{F}, \cdot)$  and  $(W, +, \mathbb{F}, \cdot)$  be vector spaces. A **linear transformation** (העתקה ליניארית) from  $V$  to  $W$  is a function  $f : V \rightarrow W$ , satisfying

$$f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$$

and

$$f(a \mathbf{v}) = a f(\mathbf{v})$$

for all  $\mathbf{u}, \mathbf{v} \in V$  and  $a \in \mathbb{F}$ . The set of all linear transformations from  $V$  to  $W$  is denoted by  $\text{Hom}_{\mathbb{F}}(V, W)$ .

### Comments:

- (a) Note once again how addition and scalar multiplication on both sides of an equations are operations on different spaces.
- (b) Setting  $W = \mathbb{F}$ ,  $\text{Hom}_{\mathbb{F}}(V, \mathbb{F}) = V^{\vee}$ .

The following properties of linear transformation are easy to prove (cf. with their analogs for linear forms):

**Proposition 5.2** *Let  $(V, +, \mathbb{F}, \cdot)$  and  $(W, +, \mathbb{F}, \cdot)$  be vector spaces and let  $f \in \text{Hom}_{\mathbb{F}}(V, W)$ . Then,*

- (a)  $f(0_V) = 0_W$ .
- (b) For every  $\mathbf{v} \in V$ ,  $f(-\mathbf{v}) = -f(\mathbf{v})$ .
- (c) For every  $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$  and  $a^1, \dots, a^n \in \mathbb{F}$ ,

$$f(a^1 \mathbf{v}_1 + \dots + a^n \mathbf{v}_n) = a^1 f(\mathbf{v}_1) + \dots + a^n f(\mathbf{v}_n).$$

*Proof:* For the first item, for every  $\mathbf{v} \in V$ ,

$$f(0_V) = f(0_{\mathbb{F}} \mathbf{v}) = 0_{\mathbb{F}} f(\mathbf{v}) = 0_W.$$

For the second item,

$$f(-\mathbf{v}) = f((-1_{\mathbb{F}}) \mathbf{v}) = (-1_{\mathbb{F}}) f(\mathbf{v}) = -f(\mathbf{v}).$$

The third item follows by induction. Note that we can write it in matrix form,

$$f \left( \begin{pmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{pmatrix} \begin{bmatrix} a^1 \\ \vdots \\ a^n \end{bmatrix} \right) = \begin{pmatrix} f(\mathbf{v}_1) & \dots & f(\mathbf{v}_n) \end{pmatrix} \begin{bmatrix} a^1 \\ \vdots \\ a^n \end{bmatrix}.$$

■

**Example:** The zero transformation  $f : V \rightarrow W$  defined by

$$f(\mathbf{v}) = 0_W \quad \text{for all } \mathbf{v} \in V$$

is a linear transformation. ▲ ▲ ▲

**Example:** The **identity map** (העתקה הזהות)  $f : V \rightarrow V$  defined by

$$f(\mathbf{v}) = \mathbf{v} \quad \text{for all } \mathbf{v} \in V$$

is a linear transformation. ▲ ▲ ▲

**Example:** The inverse map  $f : V \rightarrow V$  defined by

$$f(\mathbf{v}) = -\mathbf{v} \quad \text{for all } \mathbf{v} \in V$$

is a linear transformation. ▲ ▲ ▲

**Example:** Linear forms are linear transformations  $\text{Hom}_{\mathbb{F}}(V, \mathbb{F})$ . ▲ ▲ ▲

**Example:** Maps  $f : V \rightarrow V$  defined by

$$f(\mathbf{v}) = a\mathbf{v} \quad \text{for all } \mathbf{v} \in V$$

for some  $a \in \mathbb{F}$  are linear transformations. They are called **homotheties** (הומוטתיות). ▲ ▲ ▲

**Example:** Let  $V$  be a finitely-generated vector space and let  $\mathfrak{B} = (\mathbf{v}_1 \ \dots \ \mathbf{v}_n)$  be an ordered basis. The **coordinate map**

$$f : V \rightarrow \mathbb{F}_{\text{col}}^n$$

defined by

$$f(\mathbf{v}) = [\mathbf{v}]_{\mathfrak{B}}$$

is a linear transformation. This was in fact proved in Proposition 3.46. ▲ ▲ ▲

**Example:** Let  $A \in M_{m \times n}(\mathbb{F})$ . Consider the transformations

$$f : \mathbb{F}_{\text{col}}^n \rightarrow \mathbb{F}_{\text{col}}^m \quad \text{and} \quad g : \mathbb{F}_{\text{row}}^m \rightarrow \mathbb{F}_{\text{row}}^n$$

defined by

$$f(\mathbf{v}) = A\mathbf{v} \quad \text{and} \quad g(\mathbf{w}) = \mathbf{w}A.$$

Both are linear transformations. ▲ ▲ ▲

## 5.2 Properties of linear transformations

Like for linear forms, we consider the case where  $V$  is finitely-generated. The following two propositions are analogous to Proposition 4.4 and Proposition 4.5

**Proposition 5.3** *Let  $V$  be a finitely-generated vector space and let  $W$  be a vector space over the same field. Let*

$$\mathfrak{B} = (\mathbf{v}_1 \quad \dots \quad \mathbf{v}_n)$$

*be an ordered basis for  $V$ . Then, for every sequence  $\mathbf{w}_1, \dots, \mathbf{w}_n \in W$  there exists a linear transformation  $f \in \text{Hom}_{\mathbb{F}}(V, W)$ , such that*

$$f(\mathbf{v}_i) = \mathbf{w}_i \quad \text{for every } i = 1, \dots, n.$$

*Proof:* There really is only one way to define such a transformation. Since every  $\mathbf{v} \in V$  has a unique representation as

$$\mathbf{v} = a^1 \mathbf{v}_1 + \dots + a^n \mathbf{v}_n,$$

then  $f(\mathbf{v})$  must be given by

$$f(\mathbf{v}) = a^1 f(\mathbf{v}_1) + \dots + a^n f(\mathbf{v}_n) = a^1 \mathbf{w}_1 + \dots + a^n \mathbf{w}_n.$$

To complete the proof, we have to verify that  $f$  is a linear transformation. Let  $\mathbf{u}, \mathbf{v} \in V$  be given by

$$\mathbf{u} = a^1 \mathbf{v}_1 + \dots + a^n \mathbf{v}_n$$

$$\mathbf{v} = b^1 \mathbf{v}_1 + \dots + b^n \mathbf{v}_n.$$

Then,

$$\mathbf{u} + \mathbf{v} = (a^1 + b^1) \mathbf{v}_1 + \dots + (a^n + b^n) \mathbf{v}_n.$$

By the way we defined  $f$ ,

$$f(\mathbf{u}) = a^1 \mathbf{w}_1 + \dots + a^n \mathbf{w}_n$$

$$f(\mathbf{v}) = b^1 \mathbf{w}_1 + \dots + b^n \mathbf{w}_n,$$

and

$$f(\mathbf{u} + \mathbf{v}) = (a^1 + b^1) \mathbf{w}_1 + \cdots + (a^n + b^n) \mathbf{w}_n,$$

so that indeed  $f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$ . We proceed similarly to show that  $f(k\mathbf{v}) = k f(\mathbf{v})$  for  $k \in \mathbb{F}$ . ■

The following complementing proposition asserts that there really was no other way to define  $f$ :

**Proposition 5.4** *Let  $V$  be a finitely-generated vector space and let  $W$  be a vector space over the same field. Let*

$$\mathfrak{B} = (\mathbf{v}_1 \quad \cdots \quad \mathbf{v}_n)$$

*be an ordered basis for  $V$ . If two linear transformations  $g, f \in \text{Hom}_{\mathbb{F}}(V, W)$  satisfy*

$$g(\mathbf{v}_i) = f(\mathbf{v}_i) \quad \text{for all } i = 1, \dots, n,$$

*then  $g = f$ .*

*Proof:* By the property of a basis in a finitely-generated vector space, every  $\mathbf{v} \in V$  can be represented uniquely as

$$\mathbf{v} = a^1 \mathbf{v}_1 + \cdots + a^n \mathbf{v}_n$$

for some scalars  $a^1, \dots, a^n$ . Then, by the linearity of  $g, f$ ,

$$g(\mathbf{v}) = a^1 g(\mathbf{v}_1) + \cdots + a^n g(\mathbf{v}_n) = a^1 f(\mathbf{v}_1) + \cdots + a^n f(\mathbf{v}_n) = f(\mathbf{v}).$$

■

**Example:** Let  $V$  be a finitely-generated vector space. Let

$$\mathfrak{B} = (\mathbf{v}_1 \quad \cdots \quad \mathbf{v}_n)$$

be an ordered basis for  $V$ . Let  $\lambda_1, \dots, \lambda_n \in \mathbb{F}$ . Then the linear transformation  $f \in \text{Hom}_{\mathbb{F}}(V, V)$  satisfying

$$f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$$

is defined uniquely. ▲ ▲ ▲

**Example:** Let  $V = \mathbb{F}_{\text{col}}^n$  and  $W = \mathbb{F}_{\text{col}}^m$ . We will show that to every  $f \in \text{Hom}_{\mathbb{F}}(V, W)$  corresponds a  $A \in M_{m \times n}(\mathbb{F})$  such that

$$f(\mathbf{v}) = A\mathbf{v}.$$

Take the standard basis  $\mathfrak{E} = (\mathbf{e}_1 \ \dots \ \mathbf{e}_n)$ . Every  $\mathbf{v} \in \mathbb{F}_{\text{col}}^n$  has a unique representation

$$\mathbf{v} = v^1 \mathbf{e}_1 + \dots + v^n \mathbf{e}_n,$$

hence

$$f(\mathbf{v}) = v^1 f(\mathbf{e}_1) + \dots + v^n f(\mathbf{e}_n) = A\mathbf{v},$$

where for every  $i = 1, \dots, n$ ,

$$\text{Col}_i(A) = f(\mathbf{e}_i).$$

▲ ▲ ▲

**Example:** Let  $V = \mathbb{R}^2$  and  $W = \mathbb{R}^3$ . Consider the basis for  $\mathbb{R}^2$ ,  $\mathfrak{B} = (\mathbf{v}_1, \mathbf{v}_2)$ , where

$$\mathbf{v}_1 = (1, 2) \quad \text{and} \quad \mathbf{v}_2 = (3, 4).$$

By the above propositions, there exists a unique linear transformations  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  satisfying

$$f(\mathbf{v}_1) = (3, 2, 1) \quad \text{and} \quad f(\mathbf{v}_2) = (6, 5, 4).$$

How do we find it. A direct calculation shows that

$$[(x, y)]_{\mathfrak{B}} = \begin{bmatrix} \frac{1}{2}(3y - 4x) \\ \frac{1}{2}(2x - y) \end{bmatrix}.$$

Hence,

$$\begin{aligned} f(x, y) &= \frac{1}{2}(3y - 4x) f(\mathbf{v}_1) + \frac{1}{2}(2x - y) f(\mathbf{v}_2) \\ &= \frac{1}{2}(3y - 4x)(3, 2, 1) + \frac{1}{2}(2x - y)(6, 5, 4). \end{aligned}$$

For example,

$$f(1, 0) = -2(3, 2, 1) + (6, 5, 4) = (0, 1, 2).$$

▲ ▲ ▲

**Exercises**

**(easy) 5.1** Which of the following functions  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is a linear transformation?

- (a)  $f(x, y) = (1 + x, y)$ .
- (b)  $f(x, y) = (y, x)$ .
- (c)  $f(x, y) = (x^2, y)$ .
- (d)  $f(x, y) = (\sin x, y)$ .
- (e)  $f(x, y) = (y - x, 0)$ .

**(easy) 5.2** Let  $V = \mathbb{R}^2$  and  $W = \mathbb{R}^3$ . Write in explicit form the linear transformation  $f \in \text{Hom}_{\mathbb{R}}(V, W)$  satisfying

$$f(1, 2) = (3, 2, 1) \quad \text{and} \quad f(3, 4) = (6, 5, 4).$$

**(easy) 5.3** Let

$$V = \mathbb{R}_{<2}[X] = \{p \in \mathbb{R}[X] : \deg p < 2\},$$

and let  $W = M_3(\mathbb{R})$ . Define the function  $f : V \rightarrow W$ ,

$$f(a + bX) = \begin{bmatrix} a & & \\ & a + b & \\ & & b \end{bmatrix}.$$

- (a) Show that  $f$  is a linear transformation.
- (b) Does there exist a  $p \in V$  such that

$$f(p) = \begin{bmatrix} 1 & & \\ & 0 & \\ & & 1 \end{bmatrix}?$$

- (c) Does there exist a non-zero  $p \in V$  such that  $f(p) = 0_W$ ?

**(intermediate) 5.4** Let  $f : \mathbb{C}^2 \rightarrow \mathbb{C}$  be defined by

$$f(z, w) = z + \bar{w},$$

where  $\bar{w}$  is the complex-conjugate of  $w$ . Is  $f$  a linear transformation when

- (a)  $\mathbb{C}^2$  and  $\mathbb{C}$  are vector spaces over  $\mathbb{C}$ ?  
 (b)  $\mathbb{C}^2$  and  $\mathbb{C}$  are vector spaces over  $\mathbb{R}$ ?

**(intermediate) 5.5** Does there exist a linear transformation  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ , which is not the zero transformation satisfying

$$f(\mathbf{v}_1) = f(\mathbf{v}_2) = f(\mathbf{v}_3) = f(\mathbf{v}_4),$$

where

$$\mathbf{v}_1 = (1, 0, 1) \quad \mathbf{v}_2 = (1, 2, 1) \quad \mathbf{v}_3 = (0, 1, 1) \quad \mathbf{v}_4 = (2, 3, 3)?$$

If it does, write it explicitly; otherwise explain why not.

**(intermediate) 5.6** Consider a linear transformation  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  satisfying

$$f(0, 1, 2) = (1, 0) \quad \text{and} \quad f(0, 0, 1) = (1, 1).$$

Based on this, it is possible to find

$$f(0, 2, 3) \quad \text{and} \quad f(1, 2, 3)?$$

**(intermediate) 5.7** Let  $V, W$  be vector spaces and let  $U \leq W$ . Let  $f : V \rightarrow W$  be a linear transformation. Show that

$$S = \{\mathbf{v} \in V : f(\mathbf{v}) \in U\} \leq V.$$

**(easy) 5.8** Let  $V$  be a vector space over  $\mathbb{F}$  and let  $\ell^1, \dots, \ell^n \in V^\vee$ . Define  $f : V \rightarrow \mathbb{F}_{\text{col}}^n$  by

$$f(\mathbf{v}) = \begin{bmatrix} \ell^1(\mathbf{v}) \\ \vdots \\ \ell^n(\mathbf{v}) \end{bmatrix}.$$

Show that  $f$  is a linear transformation.

**(harder) 5.9** Consider  $V = \text{Func}(\mathbb{R}, \mathbb{R})$  as a vector space over  $\mathbb{R}$ . Show that a function  $h : V \rightarrow V$  (it is a function mapping functions to functions!) defined for every  $f \in V$  by

$$(h(f))(x) = f(x + 1)$$

is a linear transformation.

### 5.3 The space $\text{Hom}_{\mathbb{F}}(V, W)$

Given what we learned about linear forms, you will not be surprised to know that the set of linear transformations  $\text{Hom}_{\mathbb{F}}(V, W)$  can be given a structure of a vector space. After all, it is a subset of the space of functions  $\text{Func}(V, W)$ , which are a vector space with respect to the addition of functions,

$$(f + g)(\mathbf{v}) = f(\mathbf{v}) + g(\mathbf{v}),$$

and scalar multiplication,

$$(a f)(\mathbf{v}) = a f(\mathbf{v}).$$

**Proposition 5.5** *Let  $V$  and  $W$  be vector spaces over a field  $\mathbb{F}$ . The set  $\text{Hom}_{\mathbb{F}}(V, W)$  is a linear subspace of  $\text{Func}(V, W)$ .*

*Proof:* The set  $\text{Hom}_{\mathbb{F}}(V, W)$  is non-empty because it contains the zero map. Let  $f, g \in \text{Hom}_{\mathbb{F}}(V, W)$  and  $b \in \mathbb{F}$ ; we need to show that  $f + g \in \text{Hom}_{\mathbb{F}}(V, W)$  and that  $b f \in \text{Hom}_{\mathbb{F}}(V, W)$ . For all  $\mathbf{u}, \mathbf{v} \in V$ ,

$$\begin{aligned} (f + g)(\mathbf{u} + \mathbf{v}) &= f(\mathbf{u} + \mathbf{v}) + g(\mathbf{u} + \mathbf{v}) = (f(\mathbf{u}) + f(\mathbf{v})) + (g(\mathbf{u}) + g(\mathbf{v})) \\ &= (f(\mathbf{u}) + g(\mathbf{u})) + (f(\mathbf{v}) + g(\mathbf{v})) = (f + g)(\mathbf{u}) + (f + g)(\mathbf{v}), \end{aligned}$$

and for every  $\mathbf{v} \in V$  and  $a \in \mathbb{F}$ ,

$$\begin{aligned} (f + g)(a \mathbf{v}) &= f(a \mathbf{v}) + g(a \mathbf{v}) = a f(\mathbf{v}) + a g(\mathbf{v}) \\ &= a(f(\mathbf{v}) + g(\mathbf{v})) = a((f + g)(\mathbf{v})), \end{aligned}$$

proving that  $f + g \in \text{Hom}_{\mathbb{F}}(V, W)$ .

Likewise, for all  $\mathbf{u}, \mathbf{v} \in V$ ,

$$\begin{aligned} (b f)(\mathbf{u} + \mathbf{v}) &= b(f(\mathbf{u} + \mathbf{v})) = b(f(\mathbf{u}) + f(\mathbf{v})) \\ &= b(f(\mathbf{u})) + b(f(\mathbf{v})) = (b f)(\mathbf{u}) + (b f)(\mathbf{v}), \end{aligned}$$

and for every  $\mathbf{v} \in V$  and  $a \in \mathbb{F}$ ,

$$(b f)(a \mathbf{v}) = b(f(a \mathbf{v})) = b(a f(\mathbf{v})) = a(b f(\mathbf{v})) = a((b f)(\mathbf{v})),$$

proving that  $b f \in \text{Hom}_{\mathbb{F}}(V, W)$ . ■

**Example:** Let  $V = \mathbb{F}_{\text{col}}^n$  and  $W = \mathbb{F}_{\text{col}}^m$ . For  $A, B \in M_{m \times n}(\mathbb{F})$ , we define  $f_A, f_B \in \text{Hom}_{\mathbb{F}}(V, W)$  by

$$f_A(\mathbf{v}) = A\mathbf{v} \quad \text{and} \quad f_B(\mathbf{v}) = B\mathbf{v}.$$

Then,  $f_A + f_B \in \text{Hom}_{\mathbb{F}}(V, W)$  is given by

$$(f_A + f_B)(\mathbf{v}) = f_A(\mathbf{v}) + f_B(\mathbf{v}) = A\mathbf{v} + B\mathbf{v} = (A + B)\mathbf{v},$$

where in the last equality we used the distributivity of matrix multiplication. Thus,  $f_A + f_B = f_{A+B}$ . We conclude that the addition of matrices of the same dimensions is really the addition of two linear transformations.  $\blacktriangle \blacktriangle \blacktriangle$

## 5.4 Projections and reflections

In this section we will see two interesting examples of linear transformations.

**Definition 5.6** Let  $V$  be a vector space over  $\mathbb{F}$  and let  $U, W \leq V$ . We say that  $U$  and  $W$  are **complementary** (משלימים) if

- (a)  $U + W = V$ .
- (b) To every  $\mathbf{v} \in V$  correspond unique  $\mathbf{u} \in U$  and  $\mathbf{w} \in W$ , such that

$$\mathbf{v} = \mathbf{u} + \mathbf{w}.$$

In such case we write

$$V = U \oplus W,$$

and such a sum is called a **direct sum** (סכום ישיר).

We have already seen earlier in this course that these two conditions are equivalent to the conditions that  $U + W = V$  and  $U \cap W = \{0_V\}$ .

**Example:** Let  $V = \mathbb{R}^3$ , then

$$U = \{(v^1, v^2, v^3) \in \mathbb{R}^3 : v^3 = 0\} \quad \text{and} \quad W = \{(v^1, v^2, v^3) \in \mathbb{R}^3 : v^1 = v^2 = 0\}$$

are complementary, because every  $\mathbf{v} \in \mathbb{R}^3$  is a sum of a vector in  $U$  and a vector in  $W$ , and this decomposition,

$$(v^1, v^2, v^3) = (v^1, v^2, 0) + (0, 0, v^3),$$

is unique.  $\blacktriangle \blacktriangle \blacktriangle$

**Example:** Let  $V = \mathbb{R}[X]$  be the space of polynomials in  $X$  with real-valued coefficients. Then,  $V = U \oplus W$ , where

$$U = \{p \in \mathbb{R}[X] : p = \sum_{i=0} p_i X^{2i}\}$$

$$W = \{p \in \mathbb{R}[X] : p = \sum_{i=0} p_i X^{2i+1}\}.$$

I.e., the polynomials of odd and even powers are complementary in the space of all polynomials. ▲ ▲ ▲

**Definition 5.7** Let  $V$  be a vector space over  $\mathbb{F}$  such that  $V = U_1 \oplus U_2$ . We define two **projection operators** (אופרטורי הטלה)

$$p_1 : V \rightarrow V \quad \text{and} \quad p_2 : V \rightarrow V,$$

by

$$p_1(\mathbf{v}) = \mathbf{u}_1 \quad \text{and} \quad p_2(\mathbf{v}) = \mathbf{u}_2,$$

where  $\mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2$  is the unique decomposition of  $\mathbf{v}$  as a sum of elements in  $U_1, U_2$ . The operator  $p_1$  is called the **projection on  $U_1$  parallel to  $U_2$** ; the operator  $p_2$  is called the **projection on  $U_2$  parallel to  $U_1$** .

### Comments:

- (a) We could have defined  $p_1 : V \rightarrow U_1$  and  $p_2 : V \rightarrow U_2$ .
- (b) For every  $\mathbf{v} \in V$ ,

$$(p_1 + p_2)(\mathbf{v}) = p_1(\mathbf{v}) + p_2(\mathbf{v}) = \mathbf{u}_1 + \mathbf{u}_2 = \mathbf{v},$$

so that  $p_1 + p_2$  is the identity  $V \rightarrow V$ .

**Example:** Let  $V = \mathbb{R}^3$ ,

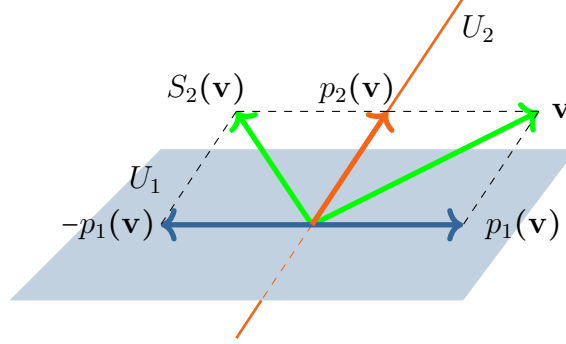
$$U_1 = \text{Span}\{(1, 0, 0), (0, 1, 0)\} \quad \text{and} \quad U_2 = \text{Span}\{(1, 1, 1)\}.$$

Then, every  $(x, y, z) \in \mathbb{R}^3$  has a unique decomposition

$$(x, y, z) = (x - z, y - z, 0) + (z, z, z),$$

so that

$$p_1(x, y, z) = (x - z, y - z, 0) \quad \text{and} \quad p_2(x, y, z) = (z, z, z).$$



▲ ▲ ▲

**Definition 5.8** Let  $V$  be a vector space over  $\mathbb{F}$  such that  $V = U_1 \oplus U_2$ . We define two **reflection operators** (אופרטורי שיקוף)

$$S_1 : V \rightarrow V \quad \text{and} \quad S_2 : V \rightarrow V,$$

by

$$S_1(\mathbf{v}) = \mathbf{u}_1 - \mathbf{u}_2 \quad \text{and} \quad S_2(\mathbf{v}) = \mathbf{u}_2 - \mathbf{u}_1,$$

where  $\mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2$  is the unique decomposition of  $\mathbf{v}$  as a sum of elements in  $U_1, U_2$ .

**Proposition 5.9** Let  $V$  be a vector space over  $\mathbb{F}$  such that  $V = U_1 \oplus U_2$ . Then the projection and the reflection operators are linear transformations.

*Proof:* The key is to observe that if  $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$  and  $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$ , where  $\mathbf{u}_1, \mathbf{v}_1 \in U_1$  and  $\mathbf{u}_2, \mathbf{v}_2 \in U_2$ , then

$$\mathbf{u} + \mathbf{v} = \underbrace{\mathbf{u}_1 + \mathbf{v}_1}_{\in U_1} + \underbrace{\mathbf{u}_2 + \mathbf{v}_2}_{\in U_2} \quad \text{and} \quad a\mathbf{v} = \underbrace{a\mathbf{v}_1}_{\in U_1} + \underbrace{a\mathbf{v}_2}_{\in U_2},$$

hence by definition

$$p_1(\mathbf{u} + \mathbf{v}) = \mathbf{u}_1 + \mathbf{v}_1 = p_1(\mathbf{u}) + p_1(\mathbf{v}) \quad \text{and} \quad p_1(a\mathbf{v}) = a\mathbf{v}_1 = a p_1(\mathbf{v}),$$

and similarly for the three other operators. ■

**Exercises**

**(easy) 5.10** Prove (possibly for the second time) that  $V = U \oplus W$  if and only if  $V = U + W$  and  $U \cap W = \{0_V\}$ .

**(easy) 5.11** Let  $V = \mathbb{R}^2$  and consider the linear subspaces

$$U = \text{Span}\{(1, 0)\} \quad \text{and} \quad W = \text{Span}\{(0, 1)\}.$$

- (a) Show that  $V = U \oplus W$ .
- (b) Write explicitly the linear transformations  $p_i$  and  $S_i$ .

**(easy) 5.12** Let  $V = \mathbb{R}^2$  and consider the linear subspaces

$$U = \text{Span}\{(1, 2)\} \quad \text{and} \quad W = \text{Span}\{(1, 1)\}.$$

- (a) Show that  $V = U \oplus W$ .
- (b) Write explicitly the linear transformations  $p_i$  and  $S_i$ .

**(intermediate) 5.13** Let  $V = \mathbb{R}^3$  and consider the linear subspaces

$$U = \text{Span}\{(1, 0, 0), (1, 1, 0)\} \quad \text{and} \quad W = \text{Span}\{(1, 1, 1)\}.$$

- (a) Show that  $V = U \oplus W$ .
- (b) Write explicitly the linear transformations  $p_i$  and  $S_i$ .

**(intermediate) 5.14** Let  $f \in \text{Hom}_{\mathbb{R}}(\mathbb{R}^3, \mathbb{R}^3)$  be the linear transformation defined by

$$f(x, y, z) = (x, y, -z).$$

Show that

- (a)  $f(\mathbf{u}) = \mathbf{u}$  if and only if  $\mathbf{u} \in \text{Span}(\mathbf{e}_1, \mathbf{e}_2) = U$ .
- (b)  $f(\mathbf{w}) = -\mathbf{w}$  if and only if  $\mathbf{w} \in \text{Span}(\mathbf{e}_3) = W$ .
- (c)  $f$  is the reflection through  $U$  parallel to  $W$ .

**(intermediate) 5.15** Let  $V = M_2(\mathbb{R})$ ,

$$U = \left\{ \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} : a, b \in \mathbb{R} \right\} \quad \text{and} \quad W = \left\{ \begin{bmatrix} -c & 0 \\ c & d \end{bmatrix} : c, d \in \mathbb{R} \right\}.$$

- (a) Show that  $U, W \leq V$  and  $V = U \oplus W$ .  
 (b) Write explicitly the projection and reflection operators.

**(harder) 5.16** Let  $V = \text{Func}(\mathbb{R}, \mathbb{R})$  and consider the linear subspaces

$$U = \{f \in \text{Func}(\mathbb{R}, \mathbb{R}) : f(x) = f(-x) \text{ for all } x \in \mathbb{R}\}$$

and

$$W = \{f \in \text{Func}(\mathbb{R}, \mathbb{R}) : f(x) = -f(-x) \text{ for all } x \in \mathbb{R}\}.$$

- (a) Show that  $V = U \oplus W$ .  
 (b) Write explicitly the linear transformations  $p_i$  and  $S_i$ .

**(harder) 5.17** Let  $V$  be a vector space over  $\mathbb{Q}$ . Let  $\mathfrak{B} = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$  be an ordered basis for  $V$ . Let  $f \in \text{Hom}_{\mathbb{Q}}(V, V)$  be the linear transformation satisfying

$$\begin{aligned} f(\mathbf{v}_1) &= \frac{5}{6}\mathbf{v}_1 - \frac{1}{3}\mathbf{v}_2 - \frac{1}{2}\mathbf{v}_3 \\ f(\mathbf{v}_2) &= -\frac{1}{6}\mathbf{v}_1 + \frac{2}{3}\mathbf{v}_2 - \frac{1}{2}\mathbf{v}_3 \\ f(\mathbf{v}_3) &= -\frac{1}{6}\mathbf{v}_1 - \frac{1}{3}\mathbf{v}_2 + \frac{1}{2}\mathbf{v}_3. \end{aligned}$$

Find subspaces  $U, W \leq V$  such that  $f$  is the projection on  $U$  parallel to  $W$ .

## 5.5 Kernel and image

**Definition 5.10** Let  $V$  and  $W$  be vector spaces over a field  $\mathbb{F}$ . Let  $f \in \text{Hom}_{\mathbb{F}}(V, W)$ . The **kernel** (גרעין) of  $f$  is the set of vectors in  $V$  that are mapped by  $f$  to the zero vector in  $W$ ,

$$\ker f = \{\mathbf{v} \in V : f(\mathbf{v}) = 0_W\}.$$

The **image** (תמונה) of  $f$  is those vectors in  $\mathbf{w} \in W$  for which there exists a vectors in  $\mathbf{v} \in V$ , such that  $\mathbf{w} = f(\mathbf{v})$ ,

$$\text{Image } f = \{f(\mathbf{v}) : \mathbf{v} \in V\} = \{\mathbf{w} \in W : \exists \mathbf{v} \in V, \mathbf{w} = f(\mathbf{v})\}.$$

Note that  $\ker f$  is a subset of  $V$  whereas  $\text{Image } f$  is a subset of  $W$ . The following proposition asserts that they are more than just subsets—they are linear subspaces. Furthermore, for the case where  $W = \mathbb{F}$  and  $\text{Hom}_{\mathbb{F}}(V, W) = V^\vee$ , then  $\ker f = \{f\}_0$  (the null space of  $f$ ).

**Proposition 5.11** Let  $V$  and  $W$  be vector spaces over a field  $\mathbb{F}$ . Let  $f \in \text{Hom}_{\mathbb{F}}(V, W)$ . Then,

$$\ker f \leq V \quad \text{and} \quad \text{Image } f \leq W.$$

*Proof:* The set  $\ker f$  is not empty, because it contains  $0_V$ . Let  $\mathbf{u}, \mathbf{v} \in \ker f$  and  $a \in \mathbb{F}$ , i.e.,

$$f(\mathbf{u}) = f(\mathbf{v}) = 0_W.$$

Then,

$$f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v}) = 0_W \quad \text{and} \quad f(a\mathbf{v}) = a f(\mathbf{v}) = 0_W,$$

i.e.,  $\mathbf{u} + \mathbf{v} \in \ker f$  and  $a\mathbf{v} \in \ker f$ , proving that  $\ker f$  is a linear subspace of  $V$ . Likewise,  $\text{Image } f$  is not empty because it contains  $0_W$ . Let  $\mathbf{w}_1, \mathbf{w}_2 \in \text{Image } f$  and let  $a \in \mathbb{F}$ . By definition of the image, there exist  $\mathbf{v}_1, \mathbf{v}_2 \in V$  such that

$$\mathbf{w}_1 = f(\mathbf{v}_1) \quad \text{and} \quad \mathbf{w}_2 = f(\mathbf{v}_2).$$

By the linearity of  $f$ ,

$$f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2) = \mathbf{w}_1 + \mathbf{w}_2,$$

i.e.,  $\mathbf{w}_1 + \mathbf{w}_2 \in \text{Image } f$ , and

$$f(a\mathbf{v}_1) = a f(\mathbf{v}_1) = a\mathbf{w}_1,$$

i.e.,  $a\mathbf{w}_1 \in \text{Image } f$ , thus proving that  $\text{Image } f$  is a linear subspace of  $W$ . ■

**Example:** Let  $f \in \text{Hom}_{\mathbb{F}}(V, W)$  be the zero transformation. Then

$$\ker f = V \quad \text{and} \quad \text{Image } f = \{0_W\}.$$

▲ ▲ ▲

**Example:** Let  $V = U_1 \oplus U_2$  and let  $p_1, p_2 : V \rightarrow V$  be the projections on the components of the direct sum. Then,

$$\text{Image } p_1 = U_1,$$

as, by definition, for every  $\mathbf{v} \in V$ ,  $p_1(\mathbf{v}) = \mathbf{u}_1$  where  $\mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2$  for  $\mathbf{u}_1 \in U_1$  and  $\mathbf{u}_2 \in U_2$ . This shows that

$$\text{Image } p_1 \leq U_1.$$

On the other hand, for every  $\mathbf{u}_1 \in U_1$ ,  $p_1(\mathbf{u}_1) = \mathbf{u}_1$ , proving that

$$U_1 \leq \text{Image } p_1.$$

Likewise,

$$\ker p_1 = U_2,$$

because if  $\mathbf{u}_2 \in U_2$ , then  $p_1(\mathbf{u}_2) = 0_V$ , proving that

$$U_2 \leq \ker p_1.$$

Conversely, if  $\mathbf{u} \in \ker p_1$ , then  $p_1(\mathbf{u}) = 0_V$ , proving that  $\mathbf{u} \in U_2$ , i.e.,

$$\ker p_1 \leq U_2.$$

▲ ▲ ▲

Recall that a function  $f : V \rightarrow W$  is called one-to-one (חדר ערכית) (or **injective**), if  $f(\mathbf{u}) = f(\mathbf{v})$  implies that  $\mathbf{u} = \mathbf{v}$ . The following proposition relates the kernel of a linear transformation to its injectivity.

**Proposition 5.12** *Let  $f \in \text{Hom}_{\mathbb{F}}(V, W)$ . Then,  $f$  is one-to-one if and only if  $\ker f = \{0_V\}$ .*

*Proof:* Let  $f$  be one-to-one. Since  $f(0_V) = 0_W$ , it follows that  $f(\mathbf{v}) = 0_W$  only if  $\mathbf{v} = 0_V$ , proving that  $\ker f = \{0_V\}$ . Conversely, suppose that  $\ker f = \{0_V\}$ . Let  $\mathbf{u}, \mathbf{v} \in V$  satisfy  $f(\mathbf{u}) = f(\mathbf{v})$ . Then,

$$f(\mathbf{u} - \mathbf{v}) = f(\mathbf{u}) - f(\mathbf{v}) = 0_W,$$

i.e.,  $\mathbf{u} - \mathbf{v} \in \ker f$ , and by assumption  $\mathbf{u} - \mathbf{v} = 0_V$ , namely,  $\mathbf{u} = \mathbf{v}$ . ■

**Exercises**

(easy) **5.18** Let  $A \in M_{2 \times 2}(\mathbb{R})$  be given by

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix},$$

and consider the linear transformation  $f : \mathbb{R}_{\text{col}}^2 \rightarrow \mathbb{R}_{\text{col}}^2$ ,

$$f(\mathbf{v}) = A\mathbf{v}.$$

Find  $\ker f$  and  $\text{Image } f$ .

(easy) **5.19** Let

$$V = \mathbb{R}_{<3}[X] = \{p \in \mathbb{R}[X] : \deg p < 3\},$$

and let  $W = M_{2 \times 2}(\mathbb{R})$ . Consider the linear transformation  $f \in \text{Hom}_{\mathbb{R}}(V, W)$ ,

$$f(a + bX + cX^2) = \begin{bmatrix} a + b & 0 \\ b + c & c - a \end{bmatrix}.$$

Find  $\ker f$  and  $\text{Image } f$ .

## 5.6 linear transformations and subspaces

The content of the previous section is in fact particular cases to the more general interaction between linear transformations and subspaces. Since linear transformations “communicate” with the linear structure of vector spaces and so do linear subspaces, it turns out that linear transformations map subspaces of  $V$  into subspaces of  $W$ , and conversely, the set of vectors whose image under  $f$  lie in a subspace of  $W$  constitute a subspace of  $V$ .

We give here two useful definitions pertinent to functions between any pair of sets:

**Definition 5.13** Let  $f \in \text{Func}(D, C)$  be a function with **domain** (תחום)  $D$  and **codomain** (טווח)  $C$  (note that  $D$  and  $C$  need not have any algebraic structure). Let  $S \subseteq D$ . The **image** (תמונה) of  $S$  under  $f$  is the subset of  $C$ ,

$$f(S) = \{f(x) : x \in S\}.$$

That is,  $y \in f(S)$  if and only if there exists an  $x \in S$  such that  $y = f(x)$ .

In the particular case where  $S = D$ ,  $f(D)$  is the image of  $f$ ,

$$f(D) = \text{Image } f.$$

**Definition 5.14** Let  $f \in \text{Func}(D, C)$ . Let  $T \subseteq C$ . The **pre-image** (תמונה הפוכה) of  $T$  under  $f$  is the subset of  $D$ ,

$$f^{-1}(T) = \{x \in D : f(x) \in T\}.$$

That is,  $x \in f^{-1}(T)$  if and only if  $f(x) \in T$ .

It is important to emphasize that the pre-image is always well-defined, regardless of whether  $f$  is invertible! For invertible functions, the pre-image of every singleton in  $C$  is a singleton in  $D$ . In the particular case where  $T = C$ ,

$$f^{-1}(C) = \{x \in D : f(x) \in C\} = D.$$

Indeed, by definition, every element in  $D$  is mapped by  $f$  into an element in  $C$ .

Thus far, we dealt with general functions between sets. Next, these sets will be vector spaces, or linear subspaces (which by definitions are vector spaces on their own).

**Proposition 5.15** Let  $V$  and  $W$  be vector spaces over a field  $\mathbb{F}$ . Let  $f \in \text{Hom}_{\mathbb{F}}(V, W)$ . For every  $U \leq V$ ,

$$f(U) \leq W,$$

and for every  $Z \leq W$ ,

$$f^{-1}(Z) \leq V.$$

**Comment:** For  $U = V$  this proposition asserts that the image of  $f$ ,

$$\text{Image } f = f(V)$$

is a linear subspace of  $W$ ; for  $Z = \{0_W\}$ , this proposition asserts that  $\ker f$ ,

$$\ker f = f^{-1}(\{0_W\})$$

is a linear subspace of  $V$ .

*Proof:* The proof follows the same lines as the proof of Proposition 5.11. The set  $f(U)$  is not empty because  $0_V \in U$ , hence  $0_W \in f(U)$ . Let  $\mathbf{w}_1, \mathbf{w}_2 \in f(U)$  and let  $a \in \mathbb{F}$ . By definition, there exist  $\mathbf{u}_1, \mathbf{u}_2 \in U$  such that

$$\mathbf{w}_1 = f(\mathbf{u}_1) \quad \text{and} \quad \mathbf{w}_2 = f(\mathbf{u}_2).$$

By the linearity of  $f$ ,

$$f(\mathbf{u}_1 + \mathbf{u}_2) = f(\mathbf{u}_1) + f(\mathbf{u}_2) = \mathbf{w}_1 + \mathbf{w}_2,$$

and since  $\mathbf{u}_1 + \mathbf{u}_2 \in U$ , it follows that  $\mathbf{w}_1 + \mathbf{w}_2 \in f(U)$ . Similarly,

$$f(a\mathbf{u}_1) = a f(\mathbf{u}_1) = a\mathbf{w}_1,$$

and since  $a\mathbf{u}_1 \in U$ , it follows that  $a\mathbf{w}_1 \in f(U)$ , thus proving that  $f(U)$  is a linear subspace of  $W$ .

Conversely, The set  $f^{-1}(Z)$  is not empty, because  $0_W \in Z$  hence  $0_V \in f^{-1}(Z)$ . Let  $\mathbf{u}, \mathbf{v} \in f^{-1}(Z)$  and  $a \in \mathbb{F}$ . By definition,

$$f(\mathbf{u}) \in Z \quad \text{and} \quad f(\mathbf{v}) \in Z.$$

Since  $Z$  is a linear subspace of  $W$ ,

$$f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v}) \in Z \quad \text{and} \quad f(a\mathbf{v}) = a f(\mathbf{v}) \in Z,$$

i.e.,

$$\mathbf{u} + \mathbf{v} \in f^{-1}(Z) \quad \text{and} \quad a\mathbf{v} \in f^{-1}(Z).$$

proving that  $f^{-1}(Z)$  is a linear subspace of  $V$ . ■

### Exercises

**(easy) 5.20** Let  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  be the linear transformation given by

$$f(x, y, z) = (x + 2y, y - z, x + 2z).$$

Let

$$U = \text{Span}\{(1, 1, 1)\} \quad \text{and} \quad W = \text{Span}\{(1, 0, 1), (0, 1, 0)\}.$$

Find (a)  $\ker f$ , (b)  $\text{Image } f$ , (c)  $f(U)$ , (d)  $f(W)$ , (e)  $f^{-1}(U)$ , (f)  $f^{-1}(W)$ .

## 5.7 Nullity and Rank

The kernel and the image of a linear transformation are defined for transformations between any pair of vector spaces. We now examine the case where  $V$  is finitely-generated. First a lemma:

**Lemma 5.16** *Let  $V$  and  $W$  be vector spaces over a field  $\mathbb{F}$ , with  $V$  finitely-generated. Let  $f \in \text{Hom}_{\mathbb{F}}(V, W)$ . Then, both  $\ker f$  and  $\text{Image } f$  are finitely-generated.*

*Proof:* Since  $\ker f \leq V$  and  $V$  is finitely-generated, then  $\ker f$  is also finitely-generated. More surprising perhaps is the fact that  $\text{Image } f$  is finitely-generated, even though  $W$  may not be. Let

$$\mathfrak{B} = (\mathbf{v}_1 \quad \dots \quad \mathbf{v}_n)$$

be a generating set for  $V$ . We will show that

$$\mathfrak{C} = (f(\mathbf{v}_1) \quad \dots \quad f(\mathbf{v}_n))$$

is a generating set for  $\text{Image } f$ , hence  $\text{Image } f$  is of dimension at most  $n$ .

Let  $\mathbf{w} \in \text{Image } f$ . By definition, there exists a  $\mathbf{v} \in V$ , such that  $f(\mathbf{v}) = \mathbf{w}$ . Since  $\mathfrak{B}$  is a generating set for  $V$ , there exist  $n$  scalars  $a^1, \dots, a^n \in \mathbb{F}$ , such that

$$\mathbf{v} = a^1 \mathbf{v}_1 + \dots + a^n \mathbf{v}_n.$$

By the linearity of  $f$ ,

$$\mathbf{w} = f(\mathbf{v}) = a^1 f(\mathbf{v}_1) + \dots + a^n f(\mathbf{v}_n),$$

proving that  $\mathbf{w} \in \text{Span } \mathfrak{C}$ , i.e.,

$$\text{Image } f \leq \text{Span } \mathfrak{C}.$$

■

**Definition 5.17** Let  $V$  and  $W$  be vector spaces over a field  $\mathbb{F}$ , with  $V$  finitely-generated. Let  $f \in \text{Hom}_{\mathbb{F}}(V, W)$ . The **nullity** (אפסיות) of  $f$  is

$$\nu(f) = \dim_{\mathbb{F}} \ker f.$$

The **rank** (דרגה) of  $f$  is

$$\rho(f) = \dim_{\mathbb{F}} \text{Image } f.$$

Intuitively, the larger the nullity of a linear transformation, the more vectors in  $V$  are mapped to the zero vector in  $W$ . The larger the range of  $f$ , the more vectors in  $W$  are obtained by applying  $f$  on vectors in  $V$ .

**Example:** Let  $f \in \text{Hom}_{\mathbb{F}}(V, W)$  be the zero transformation. Then

$$\nu(f) = \dim_{\mathbb{F}} V \quad \text{and} \quad \rho(f) = 0.$$

▲ ▲ ▲

**Theorem 5.18 (Rank-nullity theorem (משפט הממדים))** Let  $V$  and  $W$  be vector spaces over a field  $\mathbb{F}$ , with  $V$  finitely-generated. Let  $f \in \text{Hom}_{\mathbb{F}}(V, W)$ . Then,

$$\nu(f) + \rho(f) = \dim_{\mathbb{F}} V.$$

(In other words, there is a “tradeoff” between “how many” vectors in  $V$  are mapped to the zero vector and “how many” vector in  $W$  can be obtained as the output of  $f$ .)

*Proof:* The idea of the proof is quite similar in essence to the proof of the theorems relating the annihilators of subspaces for linear forms. Denote by  $n$  the dimension of  $V$ . Let

$$(\mathbf{u}_1 \quad \dots \quad \mathbf{u}_k)$$

be a basis for  $\ker f$  (which is of dimension at most  $n$ ) and let

$$\mathfrak{B} = (\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}_1, \dots, \mathbf{v}_{n-k})$$

be its completion to a basis for  $V$  (recall that a basis for any subspace can be completed into a basis for the entire space). Since, by definition,  $\nu(f) = k$ , it remains to prove that  $\varrho(f) = n - k$ . Consider the set

$$\mathfrak{C} = (f(\mathbf{v}_1) \quad \dots \quad f(\mathbf{v}_{n-k})).$$

If we show that  $\mathfrak{C}$  is a basis for  $\text{Image } f$ , then we are done.

Let  $\mathbf{w} \in \text{Image } f$ . By definition there exists a  $\mathbf{v} \in V$  such that  $\mathbf{w} = f(\mathbf{v})$ . Since  $\mathfrak{B}$  is a basis for  $V$ ,

$$\mathbf{v} = a^1 \mathbf{u}_1 + \dots + a^k \mathbf{u}_k + b^1 \mathbf{v}_1 + \dots + b^{n-k} \mathbf{v}_{n-k}$$

for some scalars  $a^1, \dots, a^k, b^1, \dots, b^{n-k} \in \mathbb{F}$ . Applying  $f$  on both sides, using its linearity

$$\mathbf{w} = f(\mathbf{v}) = a^1 f(\mathbf{u}_1) + \dots + a^k f(\mathbf{u}_k) + b^1 f(\mathbf{v}_1) + \dots + b^{n-k} f(\mathbf{v}_{n-k}).$$

However,  $\mathbf{u}_i \in \ker f$ , namely,  $f(\mathbf{u}_i) = 0$ , from which we obtain that

$$\mathbf{w} = b^1 f(\mathbf{v}_1) + \dots + b^{n-k} f(\mathbf{v}_{n-k}),$$

i.e.,  $\mathbf{w} \in \text{Span } \mathfrak{C}$ , hence the latter is a generating set for  $\text{Image } f$ .

It remains to show that the sequence  $\mathfrak{C}$  is independent. Suppose that

$$b^1 f(\mathbf{v}_1) + \dots + b^{n-k} f(\mathbf{v}_{n-k}) = 0_W$$

for some scalars  $b^1, \dots, b^{n-k} \in \mathbb{F}$ . We need to show that  $b^i = 0$  for all  $i = 1, \dots, n - k$ .

Using the linearity of  $f$  in the “reverse direction”,

$$f(b^1 \mathbf{v}_1 + \dots + b^{n-k} \mathbf{v}_{n-k}) = 0_W.$$

This implies that

$$b^1 \mathbf{v}_1 + \dots + b^{n-k} \mathbf{v}_{n-k} \in \ker f.$$

Since the  $\mathbf{u}_i$ 's form a basis for  $\ker f$ , there exist scalars  $a^1, \dots, a^k \in \mathbb{F}$ , such that

$$b^1 \mathbf{v}_1 + \dots + b^{n-k} \mathbf{v}_{n-k} = a^1 \mathbf{u}_1 + \dots + a^k \mathbf{u}_k.$$

However,  $\mathfrak{B}$  is a basis for  $V$  (i.e., the set comprising both  $\mathbf{u}_i$ 's and  $\mathbf{v}_i$ 's is linearly-independent). It follows that  $a^i = 0$  and  $b^j = 0$  for all  $i$  and  $j$ , proving that the sequence  $\mathfrak{C}$  is linearly-independent. This completes the proof. ■

**Comment:** An implication of the rank-nullity theorem is that the dimension of the image of a linear transformation cannot exceed the dimension of its domain. For example, if  $V = \mathbb{R}$  and  $W = \mathbb{R}^{27}$ , then the image of a linear transformation  $f \in \text{Hom}_{\mathbb{F}}(V, W)$  is at most one-dimensional. In a sense, a linear transformation cannot “create a space from nothing”.

**Example:** Let  $A \in M_{m \times n}(\mathbb{F})$  and consider the linear map  $f \in \text{Hom}_{\mathbb{F}}(\mathbb{F}_{\text{col}}^n, \mathbb{F}_{\text{col}}^m)$  given by

$$f(\mathbf{v}) = A\mathbf{v}.$$

In this case, the image of  $f$  is the column space of  $A$ ,

$$\text{Image } A = \mathcal{C}(A),$$

whereas the kernel of  $f$  is the set of zeros of the rows of  $A$ , viewed as linear forms, i.e.,

$$\ker A = (\mathcal{R}(A))_0.$$

Then, by the rank-nullity theorem,

$$\dim_{\mathbb{F}} \mathcal{C}(A) + \dim_{\mathbb{F}} (\mathcal{R}(A))_0 = \dim_{\mathbb{F}} \mathbb{F}_{\text{col}}^n = n.$$

Recall that  $\dim_{\mathbb{F}} \mathcal{C}(A)$  is the column-rank of  $A$ , whereas  $n - \dim_{\mathbb{F}} (\mathcal{R}(A))_0$  is the row-rank of  $A$  (the number of non-zero rows in the reduced form). Thus, we have discovered once again that the row-rank of a matrix equals to its column-rank. ▲ ▲ ▲

## Exercises

**(intermediate) 5.21** Let  $V$  be a vector space and let

$$\mathfrak{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$$

be a basis for  $V$ . Let  $f \in \text{Hom}_{\mathbb{F}}(V, V)$  such that  $\{\mathbf{v}_1, \mathbf{v}_2\}$  is a basis for  $\ker f$ . Show that the set  $\{f(\mathbf{v}_3), f(\mathbf{v}_4)\}$  is linearly-independent.

**(harder) 5.22** Find a linear transformation  $f : \mathbb{R}_{<4}[X] \rightarrow M_{2 \times 3}(\mathbb{R})$ , such that

$$\ker f = \text{Span}\{X^3 - 2X + 1, X^3 + X^2 - X + 3\}$$

and

$$\text{Span}\left\{\begin{bmatrix} -1 & 2 & 1 \\ 3 & -1 & 0 \end{bmatrix}\right\} \subseteq \text{Image } f.$$

**(intermediate) 5.23** Consider a linear transformation  $g \in \text{Hom}_{\mathbb{R}}(\mathbb{R}^4, \mathbb{R}^3)$  satisfying

$$g(1, 3, -1, 0) = (1, 0, -4) \quad \text{and} \quad g(2, 1, 2, 1) = (2, 0, -8).$$

- (a) Can  $g$  be one-to-one? Find an example or argue why not.
- (b) Can  $g$  be onto? Find an example or argue why not.

**(intermediate) 5.24** Let  $f \in \text{Hom}_{\mathbb{R}}(\mathbb{R}^4, \mathbb{R}^3)$ . Let  $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^4$  be independent vectors satisfying  $f(\mathbf{v}_1) = f(\mathbf{v}_2) = 0_{\mathbb{R}^3}$ . Show that  $f$  is not onto.

**(intermediate) 5.25** Which of the following assertions is true? Provide an example or disprove:

- (a) There exists a linear transformation  $f \in \text{Hom}_{\mathbb{R}}(\mathbb{R}^2, \mathbb{R}^2)$  satisfying  $\ker f = \text{Image } f$ .
- (b) There exists a linear transformation  $f \in \text{Hom}_{\mathbb{R}}(\mathbb{R}^3, \mathbb{R}^3)$  satisfying  $\ker f = \text{Image } f$ .

**(intermediate) 5.26** Let  $V$  and  $W$  be vector spaces over  $\mathbb{F}$ . Let  $f \in \text{Hom}_{\mathbb{F}}(V, W)$  and let

$$(\mathbf{v}_1, \dots, \mathbf{v}_n)$$

be a sequence of vectors in  $V$ .

- (a) Suppose that  $f$  is one-to-one. Show that  $(\mathbf{v}_1, \dots, \mathbf{v}_n)$  are linearly-independent if and only if  $(f(\mathbf{v}_1), \dots, f(\mathbf{v}_n))$  are linearly-independent.
- (b) Suppose that  $f$  is onto. Show that if  $(\mathbf{v}_1, \dots, \mathbf{v}_n)$  is a generating set for  $V$ , then  $(f(\mathbf{v}_1), \dots, f(\mathbf{v}_n))$  is a generating set for  $W$ .
- (c) Show that it is not generally true that if  $(f(\mathbf{v}_1), \dots, f(\mathbf{v}_n))$  is a generating set for  $W$ , then  $(\mathbf{v}_1, \dots, \mathbf{v}_n)$  is a generating set for  $V$ .
- (d) Suppose that  $f$  is one-to-one and onto. Show that  $(\mathbf{v}_1, \dots, \mathbf{v}_n)$  is a basis for  $V$  if and only if  $(f(\mathbf{v}_1), \dots, f(\mathbf{v}_n))$  is a basis for  $W$ .

## 5.8 Composition of linear transformations

Let  $U, V, W$  be vector spaces over a field  $\mathbb{F}$ . For  $f \in \text{Hom}_{\mathbb{F}}(U, V)$  and  $g \in \text{Hom}_{\mathbb{F}}(V, W)$ , we can **compose** (להרכיב) the two functions, yielding a function

$$g \circ f : U \rightarrow W,$$

given for all  $\mathbf{u} \in U$  by

$$(g \circ f)(\mathbf{u}) = g(f(\mathbf{u})).$$

Note that the composition of functions is a notion pertinent to sets; there is nothing “linear” about it. The following proposition asserts that the composition of linear transformations is a linear transformation:

**Proposition 5.19** *Let  $U, V, W$  be vector spaces over a field  $\mathbb{F}$ . If  $f \in \text{Hom}_{\mathbb{F}}(U, V)$  and  $g \in \text{Hom}_{\mathbb{F}}(V, W)$ , then  $g \circ f \in \text{Hom}_{\mathbb{F}}(U, W)$ .*

*Proof:* For every  $\mathbf{u}_1, \mathbf{u}_2 \in U$ , since  $f$  and  $g$  are both linear transformations,

$$\begin{aligned} (g \circ f)(\mathbf{u}_1 + \mathbf{u}_2) &= g(f(\mathbf{u}_1 + \mathbf{u}_2)) \\ &= g(f(\mathbf{u}_1) + f(\mathbf{u}_2)) \\ &= g(f(\mathbf{u}_1)) + g(f(\mathbf{u}_2)) \\ &= (g \circ f)(\mathbf{u}_1) + (g \circ f)(\mathbf{u}_2), \end{aligned}$$

and for every  $\mathbf{u} \in U$  and  $a \in \mathbb{F}$ ,

$$(g \circ f)(a\mathbf{u}) = g(f(a\mathbf{u})) = g(a f(\mathbf{u})) = a g(f(\mathbf{u})) = a (g \circ f)(\mathbf{u}),$$

proving that  $g \circ f$  is indeed a linear transformation. ■

**Example:** Consider the case of  $U = \mathbb{F}_{\text{col}}^n$ ,  $V = \mathbb{F}_{\text{col}}^m$  and  $W = \mathbb{F}_{\text{col}}^k$ . Let  $A \in M_{m \times n}(\mathbb{F})$  and  $B \in M_{k \times m}(\mathbb{F})$ . Define the linear transformations  $f_A \in \text{Hom}_{\mathbb{F}}(U, V)$  and  $f_B \in \text{Hom}_{\mathbb{F}}(V, W)$  by

$$f_A(\mathbf{u}) = A\mathbf{u} \quad \text{and} \quad f_B(\mathbf{v}) = B\mathbf{v}.$$

Then,  $f_B \circ f_A \in \text{Hom}_{\mathbb{F}}(U, W)$  is given by

$$(f_B \circ f_A)(\mathbf{u}) = f_B(f_A(\mathbf{u})) = B(f_A(\mathbf{u})) = B(A\mathbf{u}) = (BA)\mathbf{u},$$

where in the last step we used the associativity of matrix multiplication. Thus,  $f_B \circ f_A = f_{BA}$ , showing that matrix multiplication is in fact a composition of linear transformations.  $\blacktriangle \blacktriangle \blacktriangle$

The following properties of composition are easy to verify:

**Proposition 5.20** *Let  $f, f_1, f_2 \in \text{Hom}_{\mathbb{F}}(U, V)$  and  $g, g_1, g_2 \in \text{Hom}_{\mathbb{F}}(V, W)$ . Then,*

$$(g_1 + g_2) \circ f = g_1 \circ f + g_2 \circ f,$$

and

$$g \circ (f_1 + f_2) = g \circ f_1 + g \circ f_2.$$

We next relate the composition of linear transformations to the notions of kernel and image:

**Proposition 5.21** *Let  $U, V, W$  be vector spaces over the same field  $\mathbb{F}$ . Let  $f \in \text{Hom}_{\mathbb{F}}(U, V)$  and  $g \in \text{Hom}_{\mathbb{F}}(V, W)$ . Then,*

$$\ker f \leq \ker(g \circ f),$$

and

$$\text{Image}(g \circ f) \leq \text{Image } g.$$

In other words, if  $f$  maps a vector to zero then the further application of  $g$  cannot yield a non-zero vector. Also,  $g \circ f$  cannot return a vector that  $g$  cannot return.

*Proof:* Let  $\mathbf{u} \in \ker f$ , i.e.,  $f(\mathbf{u}) = 0_V$ , then

$$(g \circ f)(\mathbf{u}) = g(f(\mathbf{u})) = g(0_V) = 0_W,$$

which means that  $\mathbf{u} \in \ker g \circ f$ , i.e.,  $\ker f \subseteq \ker g \circ f$ .

Let  $\mathbf{w} \in \text{Image}(g \circ f)$ . By definition, there exists a  $\mathbf{u} \in U$  such that

$$\mathbf{w} = (g \circ f)(\mathbf{u}),$$

but this means that

$$\mathbf{w} = g(f(\mathbf{u})),$$

proving that  $\mathbf{w} \in \text{Image } g$ , i.e.,  $\text{Image}(g \circ f) \subseteq \text{Image } g$ . ■

### Exercises

**(easy) 5.27** Prove Proposition 5.20.

**(easy) 5.28** Let  $f, g, h \in \text{Hom}_{\mathbb{R}}(\mathbb{R}^3, \mathbb{R}^3)$  be defined by

$$\begin{aligned} f(x, y, z) &= (x, y, -z) & g(x, y, z) &= (y + z, x, x + z) \\ h(x, y, z) &= (x + 2y, 2x + y, 0). \end{aligned}$$

Write explicitly the linear transformations

$$2f - g \quad f + 2h \quad f \circ g \quad g \circ f \quad h \circ f + 2g.$$

**(easy) 5.29** Let  $f \in \text{Hom}_{\mathbb{F}}(V, V)$ . Show that

$$\ker f \leq \ker(f \circ f) \quad \text{and} \quad \text{Image}(f \circ f) \leq \text{Image } f.$$

**(intermediate) 5.30** Let  $f \in \text{Hom}_{\mathbb{F}}(V, V)$ . Show that

$$\ker f = \ker(f \circ f)$$

if and only if  $\ker f \cap \text{Image } f = \{0_V\}$ .

**(intermediate) 5.31** Let  $V$  be a finitely-generated vector space over  $\mathbb{R}$ . Let  $f \in \text{Hom}_{\mathbb{R}}(V, V)$  satisfy

$$f \circ f = 2f.$$

Show that  $\ker f \cap \text{Image } f = \{0_V\}$  and that

$$V = \text{Image } f \oplus \ker f.$$

**(intermediate) 5.32** Let  $V$  be a finitely-generated vector space over  $\mathbb{R}$ . Let  $f \in \text{Hom}_{\mathbb{R}}(V, V)$ . Show that

$$\ker f = \ker(f \circ f) \quad \text{implies} \quad \text{Image } f = \text{Image}(f \circ f).$$

**(intermediate) 5.33** Find vector spaces  $U, V, W$  and linear transformations  $f \in \text{Hom}_{\mathbb{F}}(U, V)$  and  $g \in \text{Hom}_{\mathbb{F}}(V, W)$ , such that

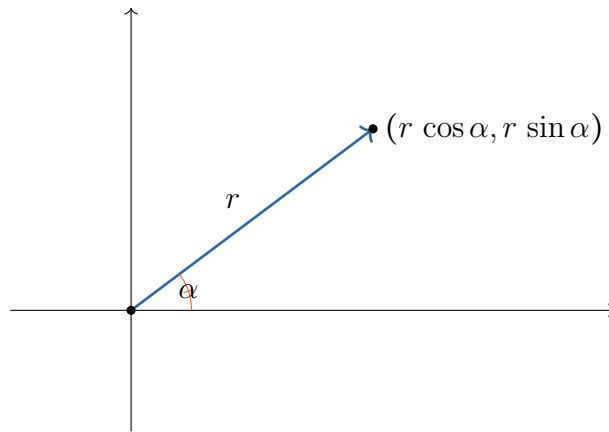
$$\ker f < \ker g \circ f,$$

and

$$\text{Image } g \circ f < \text{Image } g.$$

## 5.9 Rotations of the plane

In this section we introduce yet another family of linear transformations—this time transformations from the plane to itself,  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ . Recall that a point  $(x, y) \in \mathbb{R}^2$  can be represented by its distance from the origin,  $r$ , and the angle  $\alpha$  formed between the arrow pointing to it from the origin and the  $x$  axis:



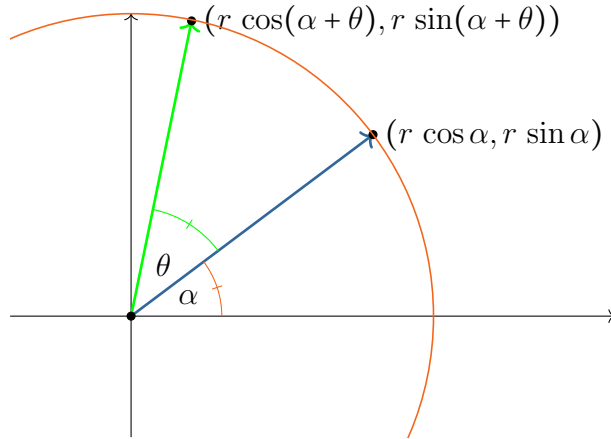
We define a family of linear transformations,

$$\text{Rot}_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2,$$

where  $\theta \in \mathbb{R}$ , where

$$\text{Rot}_\theta(r \cos \alpha, r \sin \alpha) = (r \cos(\alpha + \theta), r \sin(\alpha + \theta)).$$

That is, this transformation rotates vector about the origin by an angle  $\theta$ .



On the face of it, this transformation doesn't seem linear; the trigonometric functions are nonlinear. However, using the trigonometric identities,

$$\begin{aligned}\cos(\alpha + \theta) &= \cos \alpha \cos \theta - \sin \alpha \sin \theta \\ \sin(\alpha + \theta) &= \sin \alpha \cos \theta + \cos \alpha \sin \theta,\end{aligned}$$

setting  $x = r \cos \alpha$  and  $y = r \sin \alpha$ , we find that

$$\text{Rot}_\theta(x, y) = (\cos \theta x - \sin \theta y, \sin \theta x + \cos \theta y),$$

that is  $\text{Rot}_\theta \in \text{Hom}_{\mathbb{R}}(\mathbb{R}^2, \mathbb{R}^2)$ .

If we rather write the components of vectors relative to the standard basis,

$$\text{Rot}_\theta \left( \begin{bmatrix} x \\ y \end{bmatrix} \right) = \begin{bmatrix} \cos \theta x - \sin \theta y \\ \sin \theta x + \cos \theta y \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Thus, a rotation in the plane by an angle  $\theta$  is represented (in standard coordinates) by a multiplication by a **rotation matrix** (מטריצה סיבוב)

$$R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.$$

What happens if we compose two rotations? What happens if we rotate a vector by an angle  $\theta$  and then rotate the result by an addition angle of  $\varphi$ . Clearly, we expect

$$\text{Rot}_\varphi \circ \text{Rot}_\theta = \text{Rot}_{\varphi+\theta}.$$

A straightforward calculation shows that indeed

$$\begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} \cos(\varphi + \theta) & -\sin(\varphi + \theta) \\ \sin(\varphi + \theta) & \cos(\varphi + \theta) \end{bmatrix},$$

i.e.,

$$R_\varphi R_\theta = R_{\varphi+\theta}.$$

Note that

$$R_{2\pi} = R_0 = I,$$

and

$$R_\theta R_{-\theta} = I.$$

## 5.10 The dimension of $\text{Hom}_{\mathbb{F}}(V, W)$

Since the set of linear transformations  $\text{Hom}_{\mathbb{F}}(V, W)$  is a vector space in its own right, a number of questions arise right away: under what conditions is it finitely-generated? What would be a natural basis for it?

The following lemma is the key to answering these questions:

**Lemma 5.22** *Let  $V$  and  $W$  be finitely-generated vector spaces. Let*

$$\mathfrak{B} = (\mathbf{v}_1 \ \dots \ \mathbf{v}_n) \quad \text{and} \quad \mathfrak{C} = (\mathbf{w}_1 \ \dots \ \mathbf{w}_m)$$

*be ordered bases for  $V$  and  $W$ . Then, there exists for every  $i = 1, \dots, n$  and  $j = 1, \dots, m$  a unique linear transformation  $f_j^i \in \text{Hom}_{\mathbb{F}}(V, W)$ , such that*

$$f_j^i(\mathbf{v}_k) = \begin{cases} \mathbf{w}_j & k = i \\ 0_W & k \neq i. \end{cases} \quad (5.1)$$

*Proof:* This is an immediate consequence of Proposition 5.3 and Proposition 5.4, whereby a linear transformation is uniquely determined by its action on basis vectors. It is worth though to examine this in more detail.

Let  $f \in \text{Hom}_{\mathbb{F}}(V, W)$  and consider the vector  $f(\mathbf{v}_1)$ : it has a unique representation as a linear combination of the basis vectors  $\mathbf{w}_i$ , which we may write as

$$f(\mathbf{v}_1) = (\mathbf{w}_1 \quad \dots \quad \mathbf{w}_m) \begin{bmatrix} a_1^1 \\ \vdots \\ a_m^1 \end{bmatrix}.$$

Repeating this for each of the  $n$  vectors  $f(\mathbf{v}_j)$ , we obtain that  $f$  is uniquely determined by an  $m \times n$  matrix

$$(f(\mathbf{v}_1) \quad \dots \quad f(\mathbf{v}_n)) = (\mathbf{w}_1 \quad \dots \quad \mathbf{w}_m) \begin{bmatrix} a_1^1 & \dots & a_1^n \\ \vdots & \vdots & \vdots \\ a_m^1 & \dots & a_m^n \end{bmatrix}.$$

The function  $f_j^i$  corresponds to the matrix  $A$  which is zero everywhere, except for the element on the  $i$ -th column and  $j$ -th row, which is equal to one. ■

**Example:** Let  $n = 3$  and  $m = 5$ , then, for example,

$$f_4^2(\mathbf{v}_1) = 0_W \quad f_4^2(\mathbf{v}_2) = \mathbf{w}_4 \quad \text{and} \quad f_4^2(\mathbf{v}_3) = 0_W,$$

namely,

$$(f_4^2(\mathbf{v}_1) \quad f_4^2(\mathbf{v}_2) \quad f_4^2(\mathbf{v}_3)) = (\mathbf{w}_1 \quad \mathbf{w}_2 \quad \mathbf{w}_3 \quad \mathbf{w}_4 \quad \mathbf{w}_5) \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

▲ ▲ ▲

**Theorem 5.23** Let  $V$  and  $W$  be finitely-generated vector spaces. Let

$$\mathfrak{B} = (\mathbf{v}_1 \quad \dots \quad \mathbf{v}_n) \quad \text{and} \quad \mathfrak{C} = (\mathbf{w}_1 \quad \dots \quad \mathbf{w}_m)$$

be ordered bases for  $V$  and  $W$ . Then, the linear transformations  $f_j^i$  defined by (5.1) are a basis for  $\text{Hom}_{\mathbb{F}}(V, W)$ . In particular,

$$\dim_{\mathbb{F}} \text{Hom}_{\mathbb{F}}(V, W) = \dim_{\mathbb{F}} V \times \dim_{\mathbb{F}} W.$$

**Corollary 5.24** *In the particular case where  $W = \mathbb{F}$ . Theorem 5.23 asserts that*

$$\dim_{\mathbb{F}} \underbrace{\text{Hom}_{\mathbb{F}}(V, \mathbb{F})}_{V^{\vee}} = \dim_{\mathbb{F}} V \times \underbrace{\dim_{\mathbb{F}} \mathbb{F}}_1 = \dim_{\mathbb{F}} V,$$

*which we already know.*

*Proof:* We need to show that the set of linear transformations

$$\{f_j^i : i = 1, \dots, n, j = 1, \dots, m\}$$

is generating  $\text{Hom}_{\mathbb{F}}(V, W)$  and independent.

Let  $f \in \text{Hom}_{\mathbb{F}}(V, W)$ . We want to show that it can be represented as

$$f = \sum_{i=1}^n \sum_{j=1}^m a_i^j f_j^i,$$

where  $a_i^j \in \mathbb{F}$  are the coefficients. As we know, a linear transformation is uniquely determined by its action on basis vectors: substituting  $\mathbf{v}_k$ ,  $k = 1, \dots, n$  on both sides,

$$f(\mathbf{v}_k) = \sum_{i=1}^n \sum_{j=1}^m a_i^j f_j^i(\mathbf{v}_k) = \sum_{i=1}^n \sum_{j=1}^m a_i^j \delta_k^i \mathbf{w}_j = \sum_{j=1}^m a_k^j \mathbf{w}_j.$$

For every fixed  $k = 1, \dots, n$ , the coefficients  $a_k^j$  are the coordinates of  $f(\mathbf{v}_k) \in W$  relative the basis  $\mathfrak{C}$ ,

$$a_k^j = ([f(\mathbf{v}_k)]_{\mathfrak{C}})^j.$$

In other words, every  $f \in \text{Hom}_{\mathbb{F}}(V, W)$  can be represented as

$$f = \sum_{i=1}^n \sum_{j=1}^m ([f(\mathbf{v}_i)]_{\mathfrak{C}})^j f_j^i,$$

thus proving that the linear transformations  $f_j^i$  is a generating set for  $\text{Hom}_{\mathbb{F}}(V, W)$ .

It remains to show that they are also independent. Let  $a_j^i$  be scalars and suppose that

$$\sum_{i=1}^n \sum_{j=1}^m a_i^j f_j^i = 0_{\text{Hom}_{\mathbb{F}}(V, W)}.$$

Substituting  $\mathbf{v}_k$  on both sides,

$$\sum_{i=1}^n \sum_{j=1}^m a_i^j f_j^i(\mathbf{v}_k) = \sum_{j=1}^m a_k^j \mathbf{w}_j = 0_W.$$

Since  $\mathfrak{C}$  is a basis for  $W$ , it follows that  $a_k^j = 0_{\mathbb{F}}$  for all  $j = 1, \dots, m$  (and for all  $k = 1, \dots, n$ ), which completes the proof. ■

## 5.11 Isomorphisms

The notion of isomorphism (איזומורפיזם) is fundamental in mathematics: loosely speaking, two sets are said to be isomorphic if they are “the same” up to a renaming of their elements. The most basic notion of isomorphism is between plain sets: two sets  $S$  and  $T$  are isomorphic if there exists a function  $f : S \rightarrow T$  that is **one-to-one** (חד-חד ערכית) and **onto** (על); then, the function  $f$  induces a relation where every element in  $S$  can be identified with an element in  $T$  and vice-versa, so that we could say that  $f(x) \in T$  is a “renaming” of  $x \in S$ . We then say that  $f$  is an isomorphism and that  $S$  and  $T$  are **isomorphic** (איזומורפיים). An alternative way of stating that two sets are isomorphic is that there exist two functions  $f : S \rightarrow T$  and  $g : T \rightarrow S$ , such that

$$g(f(x)) = x \quad \text{for every } x \in S,$$

and

$$f(g(y)) = y \quad \text{for every } y \in T.$$

In other words,  $g \circ f = \text{Id}_S$  and  $f \circ g = \text{Id}_T$ .

Vector spaces are not just plain sets; they are endowed with a linear structure, so for two vector spaces to be considered isomorphic, we require more than being equivalent as sets. The function identifying an element in one space with an element in the other space has to “respect” linear operations. This leads us to the following definition:

**Definition 5.25** Let  $V$  and  $W$  be vector spaces over a field  $\mathbb{F}$ . The spaces are called *isomorphic* if there exists a linear transformation  $f \in \text{Hom}_{\mathbb{F}}(V, W)$  and a linear transformation  $g \in \text{Hom}_{\mathbb{F}}(W, V)$  such that  $g \circ f = \text{Id}_V$  and  $f \circ g = \text{Id}_W$ . The function  $f$  is called an *isomorphism from  $V$  to  $W$*  and  $g$  is called an *isomorphism from  $W$  to  $V$* .

Note that  $f$  and  $g$  both necessarily one-to-one and onto. Take  $f$  for example: for every  $\mathbf{w} \in W$ ,

$$\mathbf{w} = f(g(\mathbf{w})),$$

showing that  $f$  is onto. Likewise, if

$$f(\mathbf{v}_1) = f(\mathbf{v}_2),$$

then

$$\mathbf{v}_1 = g(f(\mathbf{v}_1)) = g(f(\mathbf{v}_2)) = \mathbf{v}_2,$$

showing that  $f$  is one-to-one.

The next proposition provides a sufficient condition for two vector spaces to be isomorphic.

**Proposition 5.26** *Let  $V$  and  $W$  be vector spaces over a field  $\mathbb{F}$ . Let  $f \in \text{Hom}_{\mathbb{F}}(V, W)$  be one-to-one and onto. Then,  $f$  is an isomorphism from  $V$  to  $W$  (implying that  $V$  and  $W$  are isomorphic).*

**Comment:** This proposition states that if a linear transformation is invertible, then its inverse is necessarily also a linear transformation.

*Proof:* Since  $f$  is one-to-one and onto, it has an inverse, which we denote by  $g$ . It remains to prove that  $g$  is a linear transformation. Let  $\mathbf{w}_1, \mathbf{w}_2 \in W$ . By definition, there exist unique  $\mathbf{v}_1, \mathbf{v}_2 \in V$ , such that

$$\mathbf{w}_1 = f(\mathbf{v}_1) \quad \text{and} \quad \mathbf{w}_2 = f(\mathbf{v}_2).$$

Reciprocally,

$$\mathbf{v}_1 = g(\mathbf{w}_1) \quad \text{and} \quad \mathbf{v}_2 = g(\mathbf{w}_2).$$

By the linearity of  $f$ ,

$$\mathbf{w}_1 + \mathbf{w}_2 = f(\mathbf{v}_1) + f(\mathbf{v}_2) = f(\mathbf{v}_1 + \mathbf{v}_2),$$

and reciprocally,

$$g(\mathbf{w}_1 + \mathbf{w}_2) = \mathbf{v}_1 + \mathbf{v}_2 = g(\mathbf{w}_1) + g(\mathbf{w}_2),$$

thus proving the first condition of linearity for  $g$ .

Likewise, let  $\mathbf{w} \in W$  and  $a \in \mathbb{F}$ . There exists a unique  $\mathbf{v} \in V$ , such that

$$\mathbf{w} = f(\mathbf{v}) \quad \text{and} \quad \mathbf{v} = g(\mathbf{w}).$$

By the linearity of  $f$ ,

$$a\mathbf{w} = a f(\mathbf{v}) = f(a\mathbf{v}),$$

and reciprocally,

$$g(a\mathbf{w}) = a\mathbf{v} = a g(\mathbf{w}),$$

thus proving the second condition of linearity for  $g$ . ■

**Proposition 5.27** *An isomorphism between vector spaces is an equivalence relation.*

*Proof:* Recall that an equivalence relation has three criteria: reflexivity, symmetry and transitivity. Every vector space is isomorphic to itself. Why? Take the identity  $f : V \rightarrow V$ , defined by

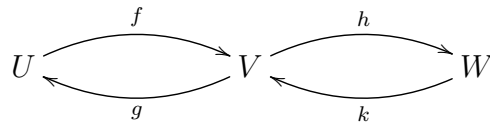
$$f(\mathbf{v}) = \mathbf{v} \quad \text{for all } \mathbf{v} \in V.$$

It is invertible (its inverse being also the identity) and linear, proving that  $V$  is isomorphic to itself. Next, if  $V$  is isomorphic to  $W$  then  $W$  is isomorphic to  $V$ , because an isomorphism is symmetric by construction. Remains transitivity: suppose that  $U$  and  $V$  are isomorphic and  $V$  and  $W$  are isomorphic. By definition, there exist

$$\begin{aligned} f &\in \text{Hom}_{\mathbb{F}}(U, V) & g &\in \text{Hom}_{\mathbb{F}}(V, U) \\ h &\in \text{Hom}_{\mathbb{F}}(V, W) & k &\in \text{Hom}_{\mathbb{F}}(W, V), \end{aligned}$$

such that

$$g \circ f = \text{Id}_U \quad f \circ g = \text{Id}_V \quad k \circ h = \text{Id}_V \quad \text{and} \quad h \circ k = \text{Id}_W.$$



Consider the functions

$$h \circ f : U \rightarrow W \quad \text{and} \quad g \circ k : W \rightarrow U.$$

Since they are compositions of linear transformations, they are linear transformations, i.e.,

$$h \circ f \in \text{Hom}_{\mathbb{F}}(U, W) \quad \text{and} \quad g \circ k \in \text{Hom}_{\mathbb{F}}(W, U).$$

Finally, for every  $\mathbf{u} \in U$ ,

$$(g \circ k) \circ (h \circ f)(\mathbf{u}) = g(k(h(f(\mathbf{u}))) = g(f(\mathbf{u})) = \mathbf{u},$$

and for every  $\mathbf{w} \in W$ ,

$$(h \circ f) \circ (g \circ k)(\mathbf{w}) = h(f(g(k(\mathbf{w}))) = h(k(\mathbf{w})) = \mathbf{w},$$

proving that  $U$  and  $W$  are isomorphic. ■

Now that we know what are isomorphic vector spaces, we will see examples, and in particular, discover that we have already encountered isomorphisms without being aware of it...

**Example:** Let  $V$  be finitely-generated vector space over  $\mathbb{F}$  and let

$$\mathfrak{B} = (\mathbf{v}_1 \quad \dots \quad \mathbf{v}_n)$$

be an ordered basis. The mapping  $f : V \rightarrow \mathbb{F}_{\text{col}}^n$ ,

$$f : \mathbf{v} \mapsto [\mathbf{v}]_{\mathfrak{B}},$$

mapping every vector to its coordinate matrix relative to  $\mathfrak{B}$  is an isomorphism. We know that this is a linear transformation; it is also one-to-one—every vector has a unique coordinate representation—and onto—every column of  $n$  scalars is the coordinate matrix of some vector. ▲ ▲ ▲

**Lemma 5.28** *Let  $V$  and  $W$  be finitely-generated vector spaces over  $\mathbb{F}$  having the same dimension. Let  $f \in \text{Hom}_{\mathbb{F}}(V, W)$ . Then,  $f$  is one-to-one if and only if  $f$  is onto.*

*Proof:* This is an immediate consequence of the rank-nullity theorem (Theorem 5.18), whereby

$$\dim_{\mathbb{F}} \ker f + \dim_{\mathbb{F}} \text{Image } f = \dim_{\mathbb{F}} V = \dim_{\mathbb{F}} W.$$

Recall that  $f$  is one-to-one if and only if  $\dim_{\mathbb{F}} \ker f = 0$  (Proposition 5.12);  $f$  is onto if and only if  $\text{Image } f = W$ . Thus, if  $f$  is one-to-one, then

$$\dim_{\mathbb{F}} \text{Image } f = \dim_{\mathbb{F}} W,$$

which implies that  $\text{Image } f = W$ , proving that  $f$  is onto. Conversely, if  $f$  is onto, namely,  $\text{Image } f = W$ , then

$$\dim_{\mathbb{F}} \ker f + \dim_{\mathbb{F}} W = \dim_{\mathbb{F}} W,$$

i.e.,  $\ker f = \{0_V\}$ , hence  $f$  is one-to-one. ■

**Proposition 5.29** *Every two finitely-generated vector spaces over the same field having the same dimension are isomorphic.*

*Proof:* Let  $V$  and  $W$  be vector spaces of dimension  $n$  over  $\mathbb{F}$ . Let

$$\mathfrak{B} = (\mathbf{v}_1 \quad \dots \quad \mathbf{v}_n) \quad \text{and} \quad \mathfrak{C} = (\mathbf{w}_1 \quad \dots \quad \mathbf{w}_n)$$

be ordered bases for  $V$  and  $W$ . Define  $f \in \text{Hom}_{\mathbb{F}}(V, W)$  as the unique linear transformation satisfying

$$f(\mathbf{v}_i) = \mathbf{w}_i \quad \text{for all } i = 1, \dots, n.$$

If we show that  $f$  is one-to-one and onto, then we are done, but from Lemma 5.28 it suffices to show just one of them. Let  $\mathbf{w} \in W$ ; by the definition of a basis, there exist scalars  $a^1, \dots, a^n \in \mathbb{F}$ , such that

$$\mathbf{w} = a^1 \mathbf{w}_1 + \dots + a^n \mathbf{w}_n.$$

Then,

$$\mathbf{w} = a^1 f(\mathbf{v}_1) + \dots + a^n f(\mathbf{v}_n) = f(a^1 \mathbf{v}_1 + \dots + a^n \mathbf{v}_n),$$

i.e.,  $\mathbf{w} \in \text{Image } f$ , which proves that  $f$  is onto. ■

**Corollary 5.30** *Every finitely-generated vector space is isomorphic to its dual.*

**Comment:** Isomorphisms are commonly split into two categories: natural isomorphisms and “unnatural” ones. An isomorphism is called natural if its definition does not rely on arbitrary choices. When we say that two vector spaces of the same dimension are isomorphic, the isomorphism depends on a choice of bases, therefore it is not considered natural.

**Example:** Let  $V$  be a finitely-generated vector space over  $\mathbb{F}$ . It is isomorphic to its dual, and its dual is isomorphic to its own dual (the so called **double-dual**). By transitivity,  $V$  is isomorphic to  $(V^\vee)^\vee$ . In this case, there exists a natural isomorphism. Consider the map

$$f : V \rightarrow (V^\vee)^\vee$$

assigning to every  $\mathbf{v} \in V$  a linear form (on linear forms...)  $f(\mathbf{v})$  defined by

$$(f(\mathbf{v}))(\ell) = \ell(\mathbf{v}) \quad \text{for every } \ell \in V^\vee.$$

We claim that  $f$  is an isomorphism, i.e., it is a linear transformation, one-to-one and onto.

To show that it is linear, we note that for every  $\mathbf{u}, \mathbf{v} \in V$  and every  $\ell \in V^\vee$ ,

$$(f(\mathbf{u} + \mathbf{v}))(\ell) = \ell(\mathbf{u} + \mathbf{v}) = \ell(\mathbf{u}) + \ell(\mathbf{v}) = (f(\mathbf{u}))(\ell) + (f(\mathbf{v}))(\ell),$$

and since this hold for every  $\ell \in V^\vee$ ,

$$f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v}).$$

Similarly, for  $\mathbf{v} \in V$ ,  $a \in \mathbb{F}$  and  $\ell \in V^\vee$ ,

$$(f(a\mathbf{v}))(\ell) = \ell(a\mathbf{v}) = a\ell(\mathbf{v}) = a(f(\mathbf{v}))(\ell),$$

and since this hold for every  $\ell \in V^\vee$ ,

$$f(a\mathbf{v}) = a f(\mathbf{v}).$$

This completes the proof that  $f$  is a linear transformation.

Since  $V$  and  $(V^\vee)^\vee$  are of the same dimension, it suffices to show that  $f$  is one-to-one, and equivalently, that its kernel is trivial. Let  $\mathbf{v} \in \ker f$ . This means that

$$(f(\mathbf{v}))(\ell) = \ell(\mathbf{v}) = 0$$

for all  $\ell \in V^\vee$ . We have seen that if  $\mathbf{v} \neq 0$  then there exists a linear form  $\ell \in V^\vee$  such that  $\ell(\mathbf{v}) \neq 0$ . We conclude that  $\mathbf{v} = 0_V$ , i.e.,

$$\ker f = \{0_V\},$$

completing the proof that  $f$  is an isomorphism. This isomorphism is considered natural because it does not hinge on any arbitrary construct.  $\blacktriangle \blacktriangle \blacktriangle$

We end this section with one more manifestation of isomorphisms respecting the linear structure of vector spaces:

**Proposition 5.31** *Let  $V, W$  be finitely-generated vector spaces over  $\mathbb{F}$  and let  $f \in \text{Hom}_{\mathbb{F}}(V, W)$  be an isomorphism. If*

$$\mathfrak{B} = (\mathbf{v}_1 \quad \dots \quad \mathbf{v}_n)$$

*is a basis for  $V$ , then*

$$\mathfrak{C} = (f(\mathbf{v}_1) \quad \dots \quad f(\mathbf{v}_n))$$

*is a basis for  $W$ . In particular, two finitely-generated vector spaces are isomorphism if and only if they are of the same dimension.*

*Proof:* We need to prove that  $\mathfrak{C}$  is generating  $W$  and independent. Denote by  $g \in \text{Hom}_{\mathbb{F}}(W, V)$  the map inverse to  $f$ . Let  $\mathbf{w} \in W$  and let  $\mathbf{v} = g(\mathbf{w})$ . Since  $\mathfrak{B}$  is a basis for  $V$ , we can write

$$\mathbf{v} = a^1 \mathbf{v}_1 + \dots + a^n \mathbf{v}_n$$

for some scalars  $a^1, \dots, a^n \in \mathbb{F}$ . Then,

$$\mathbf{w} = f(\mathbf{v}) = a^1 f(\mathbf{v}_1) + \dots + a^n f(\mathbf{v}_n),$$

proving that  $\mathfrak{C}$  is a generating set for  $W$ .

Let

$$a^1 f(\mathbf{v}_1) + \dots + a^n f(\mathbf{v}_n) = 0_W.$$

Applying  $g$  on both sides, using its linearity and the fact that  $g \circ f = \text{Id}$ , we obtain that

$$a^1 \mathbf{v}_1 + \dots + a^n \mathbf{v}_n = 0_V.$$

Since  $\mathfrak{B}$  is a basis,  $a^i = 0$  for all  $i = 1, \dots, n$ , proving that  $\mathfrak{C}$  is an independent set. This completes the proof.  $\blacksquare$

### Exercises

**(intermediate) 5.34** Complete the proof of Proposition 5.31: deduce that if  $\dim_{\mathbb{F}} V = m$  and  $\dim_{\mathbb{F}} W = n$ , where  $m \neq n$ , then  $V$  and  $W$  are not isomorphic.

**(intermediate) 5.35** Let  $A \in M_n(\mathbb{F})$ . Prove that the linear transformation  $f : \mathbb{F}_{\text{col}}^n \rightarrow \mathbb{F}_{\text{col}}^n$ ,

$$f(\mathbf{v}) = A\mathbf{v}$$

is an isomorphism if and only if  $A \in \text{GL}_n(\mathbb{F})$ .

## 5.12 Matrix representation

Recall that in finitely-generated vector spaces, the introduction of ordered bases enables us to encode vectors as coordinate matrices. In a similar way, if  $V$  and  $W$  are finitely-generated vector spaces, we can encode linear transformations in  $\text{Hom}_{\mathbb{F}}(V, W)$  as matrices acting on the coordinate representation of  $\mathbf{v} \in V$ , returning the coordinate representation of  $f(\mathbf{v}) \in W$ .

Consider the following diagram:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \mathfrak{B} \downarrow & & \downarrow \mathfrak{C} \\ \mathbb{F}_{\text{col}}^n & \xrightarrow{A_f} & \mathbb{F}_{\text{col}}^m \end{array}$$

In this diagram there are four vector spaces:  $V$ ,  $W$ ,  $\mathbb{F}_{\text{col}}^n$  and  $\mathbb{F}_{\text{col}}^m$ . The arrows represent linear transformations between the tail of the arrow and the head of the arrow. Thus,  $f$  is a linear transformation from  $V$  to  $W$ . Assume that  $\dim_{\mathbb{F}} V = n$  and  $\dim_{\mathbb{F}} W = m$ . The introduction of ordered bases,  $\mathfrak{B}$  for  $V$  and  $\mathfrak{C}$  for  $W$ , induces two linear transformations, one from  $V$  to the space of its coordinate matrices  $\mathbb{F}_{\text{col}}^n$ , and one from  $W$  to the space of its coordinate matrices  $\mathbb{F}_{\text{col}}^m$ . In this section, we show that to every  $f \in \text{Hom}_{\mathbb{F}}(V, W)$  corresponds a unique matrix  $A_f \in M_{m \times n}(\mathbb{F})$ , which we view as a linear transformation in  $\text{Hom}_{\mathbb{F}}(\mathbb{F}_{\text{col}}^n, \mathbb{F}_{\text{col}}^m)$ , such that this diagram “commutes”. To explain what this means, consider the same diagram through its action on a vector  $\mathbf{v} \in V$ :

$$\begin{array}{ccc}
\mathbf{v} & \xrightarrow{f} & f(\mathbf{v}) \\
\mathfrak{B} \downarrow & & \downarrow \mathfrak{C} \\
[\mathbf{v}]_{\mathfrak{B}} & \xrightarrow{A_f} & [f(\mathbf{v})]_{\mathfrak{C}}
\end{array}$$

Take the vector  $\mathbf{v} \in V$ ; if we apply  $f$  on it we obtain a vector  $f(\mathbf{v}) \in W$ . If we apply on the latter the linear transformation returning its coordinate matrix relative to  $\mathfrak{C}$ , we obtain  $[f(\mathbf{v})]_{\mathfrak{C}} \in \mathbb{F}_{\text{col}}^m$ . Alternatively, apply on  $\mathbf{v}$  first the linear transformation returning its coordinate matrix  $[\mathbf{v}]_{\mathfrak{B}} \in \mathbb{F}_{\text{col}}^n$  relative to  $\mathfrak{B}$ . Multiply it then by the matrix  $A_f$ , yielding a matrix  $A_f[\mathbf{v}]_{\mathfrak{B}}$ . When we say that the diagram commutes, we mean that either path yields the same outcome, namely,

$$[f(\mathbf{v})]_{\mathfrak{C}} = A_f [\mathbf{v}]_{\mathfrak{B}}.$$

The matrix  $A_f$  is called the **matrix representing** (המטריצה המיצגת) the linear transformation  $f$  relative to the ordered bases  $\mathfrak{B}$  and  $\mathfrak{C}$ ; we denote it by  $[f]_{\mathfrak{C}}^{\mathfrak{B}}$ , i.e.,

$$[f(\mathbf{v})]_{\mathfrak{C}} = [f]_{\mathfrak{C}}^{\mathfrak{B}} [\mathbf{v}]_{\mathfrak{B}}.$$

**Theorem 5.32** Let  $V, W$  be finitely-generated vector spaces over  $\mathbb{F}$ . Let

$$\mathfrak{B} = (\mathbf{v}_1 \ \dots \ \mathbf{v}_n) \quad \text{and} \quad \mathfrak{C} = (\mathbf{w}_1 \ \dots \ \mathbf{w}_m)$$

be ordered bases for  $V$  and  $W$ . Every  $f \in \text{Hom}_{\mathbb{F}}(V, W)$  has a unique  $A \in M_{m \times n}(\mathbb{F})$ , such that

$$[f(\mathbf{v})]_{\mathfrak{C}} = A [\mathbf{v}]_{\mathfrak{B}} \quad \text{for every } \mathbf{v} \in V.$$

*Proof:* Consider the transformation taking  $\mathbf{v} \in V$  and returning  $[f(\mathbf{v})]_{\mathfrak{C}}$ . This is a mapping  $V \rightarrow \mathbb{F}_{\text{col}}^m$ , which is a composition of two linear transformations, hence it is a linear transformation. For every  $i = 1, \dots, n$ , substituting  $\mathbf{v}_i$ , we obtain a coordinate matrix

$$[f(\mathbf{v}_i)]_{\mathfrak{C}} \in \mathbb{F}_{\text{col}}^m.$$

We define then the matrix  $A$  to be the unique linear transformation in  $\text{Hom}_{\mathbb{F}}(\mathbb{F}_{\text{col}}^n, \mathbb{F}_{\text{col}}^m)$  satisfying

$$A[\mathbf{v}_i]_{\mathfrak{B}} = [f(\mathbf{v}_i)]_{\mathfrak{C}}.$$

Here we used several facts: first, by Proposition 5.3 and Proposition 5.4, a linear transformation is uniquely determined by its action on basis vectors. But second, we used the fact that  $[\mathbf{v}_i]_{\mathfrak{B}}$  is a basis for  $\mathbb{F}_{\text{col}}^n$ ; this follows from the fact that the mapping from a vector to its coordinate matrix is an isomorphism (Proposition 5.31).

We claim that  $A$  has the desired property: for every  $\mathbf{v} \in V$ , which we write as

$$\mathbf{v} = a^1 \mathbf{v}_1 + \cdots + a^n \mathbf{v}_n,$$

we have

$$\begin{aligned} A[\mathbf{v}]_{\mathfrak{B}} &= A[a^1 \mathbf{v}_1 + \cdots + a^n \mathbf{v}_n]_{\mathfrak{B}} \\ &= A(a^1 [\mathbf{v}_1]_{\mathfrak{B}} + \cdots + a^n [\mathbf{v}_n]_{\mathfrak{B}}) \\ &= a^1 A[\mathbf{v}_1]_{\mathfrak{B}} + \cdots + a^n A[\mathbf{v}_n]_{\mathfrak{B}} \\ &= a^1 [f(\mathbf{v}_1)]_{\mathfrak{C}} + \cdots + a^n [f(\mathbf{v}_n)]_{\mathfrak{C}} \\ &= [a^1 f(\mathbf{v}_1) + \cdots + a^n f(\mathbf{v}_n)]_{\mathfrak{C}} \\ &= [f(a^1 \mathbf{v}_1 + \cdots + a^n \mathbf{v}_n)]_{\mathfrak{C}} \\ &= [f(\mathbf{v})]_{\mathfrak{C}}, \end{aligned}$$

which completes the proof. ■

We've already seen this matrix. Recall that there is a matrix  $A \in M_{m \times n}(\mathbb{F})$ , such that

$$(f(\mathbf{v}_1) \ \cdots \ f(\mathbf{v}_n)) = (\mathbf{w}_1 \ \cdots \ \mathbf{w}_m) A,$$

The entries of  $A$  are precisely  $([f(\mathbf{v}_i)]_{\mathfrak{C}})^j$ . That is,

$$(f(\mathbf{v}_1) \ \cdots \ f(\mathbf{v}_n)) = (\mathbf{w}_1 \ \cdots \ \mathbf{w}_m) [f]_{\mathfrak{C}}^{\mathfrak{B}}.$$

**Example:** The zero transformation is represented by the zero matrix. ▲▲▲

**Example:** Let  $\dim_{\mathbb{F}} V = n$  and consider the identity function  $\text{Id} \in \text{Hom}_{\mathbb{F}}(V, V)$ , i.e.,

$$\text{Id}(\mathbf{v}) = \mathbf{v} \quad \text{for all } \mathbf{v} \in V.$$

Let  $\mathfrak{B} = (\mathbf{v}_1 \ \dots \ \mathbf{v}_n)$  be a basis for  $V$ . Then,

$$(\text{Id}(\mathbf{v}_1) \ \dots \ \text{Id}(\mathbf{v}_n)) = (\mathbf{v}_1 \ \dots \ \mathbf{v}_n) \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix},$$

i.e., the matrix representing the identity map of a vector space is the identity matrix,

$$[\text{Id}]_{\mathfrak{B}}^{\mathfrak{B}} = I_n.$$

In this very special case it does not depend on the choice of basis, as long as we use the same basis both for the domain and the codomain.  $\blacktriangle \blacktriangle \blacktriangle$

**Example:** Let  $\dim_{\mathbb{F}} V = n$  and for  $a \in \mathbb{F}$  consider the homothety  $f \in \text{Hom}_{\mathbb{F}}(V, V)$ , given by

$$f(\mathbf{v}) = a\mathbf{v} \quad \text{for all } \mathbf{v} \in V.$$

Let  $\mathfrak{B} = (\mathbf{v}_1 \ \dots \ \mathbf{v}_n)$  be a basis for  $V$ . Then,

$$(f(\mathbf{v}_1) \ \dots \ f(\mathbf{v}_n)) = (\mathbf{v}_1 \ \dots \ \mathbf{v}_n) \begin{bmatrix} a & & \\ & \ddots & \\ & & a \end{bmatrix},$$

i.e., the matrix representing a homothety is a multiple of the identity matrix, and for every  $\mathbf{v} \in V$ ,

$$[f(\mathbf{v})]_{\mathfrak{B}} = \begin{bmatrix} a & & \\ & \ddots & \\ & & a \end{bmatrix} [\mathbf{v}]_{\mathfrak{B}}.$$

$\blacktriangle \blacktriangle \blacktriangle$

**Example:** Let  $\mathfrak{B}$  and  $\mathfrak{C}$  be two ordered bases for  $V$  (which is finitely-generated). Recall that the two bases are connected via transition matrices,  $P, Q \in \text{GL}_n(\mathbb{F})$ ,

$$\mathfrak{C} = \mathfrak{B}P \quad \text{and} \quad \mathfrak{B} = \mathfrak{C}Q,$$

where  $Q = P^{-1}$ . Furthermore, for every  $\mathbf{v} \in V$ ,

$$[\mathbf{v}]_{\mathfrak{B}} = P[\mathbf{v}]_{\mathfrak{C}} \quad \text{and} \quad [\mathbf{v}]_{\mathfrak{C}} = Q[\mathbf{v}]_{\mathfrak{B}}.$$

Since we can equivalently write

$$[\text{Id}(\mathbf{v})]_{\mathfrak{B}} = P[\mathbf{v}]_{\mathfrak{C}},$$

it follows that the transition matrix  $P$  is the matrix representing the identity map  $\text{Id} \in \text{Hom}_{\mathbb{F}}(V, V)$  relative to the bases  $\mathfrak{B}$  and  $\mathfrak{C}$ , namely,

$$P = [\text{Id}]_{\mathfrak{B}}^{\mathfrak{C}}.$$

For example, let  $V = \mathbb{R}^2$ , with

$$\mathfrak{B} = ((1, 2), (2, 1)) \quad \text{and} \quad \mathfrak{C} = ((1, 1), (1, -1)).$$

You may verify once again that

$$\underbrace{((1, 1) \ (1, -1))}_{\mathfrak{C}} = \underbrace{((1, 2) \ (2, 1))}_{\mathfrak{B}} \underbrace{\begin{bmatrix} 1/3 & -1 \\ 1/3 & 1 \end{bmatrix}}_P,$$

Then,

$$[\text{Id}]_{\mathfrak{B}}^{\mathfrak{C}} = \begin{bmatrix} 1/3 & -1 \\ 1/3 & 1 \end{bmatrix},$$

implying that for all  $\mathbf{v} \in \mathbb{R}^2$ ,

$$[\mathbf{v}]_{\mathfrak{B}} = \begin{bmatrix} 1/3 & -1 \\ 1/3 & 1 \end{bmatrix} [\mathbf{v}]_{\mathfrak{C}}.$$

▲ ▲ ▲

### Exercises

**(easy) 5.36** Let  $f \in \text{Hom}_{\mathbb{R}}(\mathbb{R}^2, \mathbb{R}^2)$  be given by

$$f(x, y) = (2x, x + y).$$

Calculate  $[f]_{\mathfrak{B}}^{\mathfrak{C}}$  and  $[f]_{\mathfrak{C}}^{\mathfrak{B}}$  for

$$\mathfrak{C} = ((1, 0) \ (0, 1)) \quad \text{and} \quad \mathfrak{B} = ((1, 1) \ (0, 1)).$$

**(easy) 5.37** Let  $A \in M_2(\mathbb{F})$  and let  $f : M_2(\mathbb{F}) \rightarrow M_2(\mathbb{F})$  be given by  $f(B) = AB$ .

- (a) Show that  $f$  is a linear transformation.  
 (b) Does there exist a basis  $\mathfrak{B}$  for  $M_2(\mathbb{F})$ , such that  $[f]_{\mathfrak{B}}^{\mathfrak{B}} = A$ ?

**(easy) 5.38** Let  $\mathfrak{E}_n$  denote the standard ordered basis for  $\mathbb{R}^n$ . Let  $f \in \text{Hom}_{\mathbb{R}}(\mathbb{R}^2, \mathbb{R}^3)$  be given by

$$f(x, y) = (2x - y, x + y, -x + 3y).$$

- (a) Write the matrix  $[f]_{\mathfrak{E}_3}^{\mathfrak{E}_2}$ .  
 (b) Find the linear transformation  $g \in \text{Hom}_{\mathbb{R}}(\mathbb{R}^2, \mathbb{R}^3)$  for which

$$[g]_{\mathfrak{E}_3}^{\mathfrak{E}_2} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ -1 & 2 \end{bmatrix}.$$

**(easy) 5.39** Repeat the previous exercise, this time using the ordered bases

$$\mathfrak{B} = ((1, 1) \ (1, -1)) \quad \text{and} \quad \mathfrak{C} = ((1, 0, 0) \ (1, 1, 0) \ (1, 1, 1)).$$

**(easy) 5.40** Find the linear transformation  $f \in \text{Hom}_{\mathbb{R}}(\mathbb{R}^2, \mathbb{R}^3)$  satisfying  $[f]_{\mathfrak{C}}^{\mathfrak{B}} = I_3$  relative to the ordered bases

$$\begin{aligned} \mathfrak{B} &= ((1, 0, -1) \ (1, -1, 0), (0, 1, 1)) \\ \mathfrak{C} &= ((1, 0, 0) \ (1, 1, 0) \ (1, 1, 1)). \end{aligned}$$

**(intermediate) 5.41** Let  $V = M_2(\mathbb{R})$ ,

$$U = \left\{ \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} : a, b \in \mathbb{R} \right\} \quad \text{and} \quad W = \left\{ \begin{bmatrix} -c & 0 \\ c & d \end{bmatrix} : c, d \in \mathbb{R} \right\}.$$

In Exercise 5.15 you showed that  $V = U \oplus W$  and wrote explicitly the projections  $p_i$  and reflections  $S_i$ .

- (a) Find an ordered basis  $\mathfrak{B} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{w}_1, \mathbf{w}_2\}$  for  $V$ , such that  $\{\mathbf{u}_1, \mathbf{u}_2\}$  is a basis for  $U$  and  $\{\mathbf{w}_1, \mathbf{w}_2\}$  is a basis for  $W$ .  
 (b) Find the matrices  $[p_1]_{\mathfrak{B}}^{\mathfrak{B}}$  and  $[S_1]_{\mathfrak{B}}^{\mathfrak{B}}$ .

**(intermediate) 5.42** Let  $V = (\mathbb{C}, +, \mathbb{R}, \cdot)$  and consider the linear transformation  $f : \mathbb{C} \rightarrow \mathbb{C}$  defined by

$$f(z) = \bar{z}.$$

Find  $[f]_{\mathfrak{B}}^{\mathfrak{B}}$  for  $\mathfrak{B} = (1, -i)$ .

**(intermediate) 5.43** Let

$$V = \mathbb{R}_{<4}[X] = \{p \in \mathbb{R}[X] : \deg p < 4\}.$$

Let  $f : V \rightarrow V$  be defined by

$$(f(p))(X) = X p'(X),$$

where  $p'$  is the derivative of  $p$ , viewed as a function, e.g.,  $f(3X - X^2) = X(3 - 2X) = 3X - 2X^2$ .

- (a) Show that  $f$  is a linear transformation.
- (b) Find  $[f]_{\mathfrak{B}}^{\mathfrak{B}}$  for  $\mathfrak{B} = (1, X, X^2, X^3)$ .
- (c) Find the kernel and the image of  $f$ .

**(intermediate) 5.44** Let  $V = M_2(\mathbb{R})$  and let  $W = \mathbb{R}_{<3}[X]$ . Let  $f : V \rightarrow W$  be the linear transformation defined by

$$f\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (a + 2b + c) + (3a - d)X + (a - 4b - 2c - d)X^2.$$

- (a) Find  $[f]_{\mathfrak{C}}^{\mathfrak{B}}$  for  $\mathfrak{C} = (1, X, X^2)$  and

$$\mathfrak{B} = \left( \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right).$$

- (b) Find the kernel and the image of  $f$ .

**(intermediate) 5.45** Let  $f : \mathbb{R}_{<3}[X] \rightarrow \mathbb{R}_{<3}[X]$  be the linear transformation represented by the matrix

$$[f]_{\mathfrak{B}}^{\mathfrak{B}} = \begin{bmatrix} 1 & 2 & 5 \\ -1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}$$

relative to the ordered basis  $\mathfrak{B} = (1, 1 + X, 1 - X + X^2)$ . Find Image  $f$ .

**(intermediate) 5.46** Let

$$\mathfrak{B} = ((1, 3) \quad (3, 0))$$

be an ordered basis for  $\mathbb{R}^2$ . Let  $f \in \text{Hom}_{\mathbb{R}}(\mathbb{R}^3, \mathbb{R}^2)$  satisfy

$$[f]_{\mathfrak{B}}^{\mathfrak{C}} = \begin{bmatrix} -2 & 5 & 0 \\ 1 & 0 & -1 \end{bmatrix},$$

where  $\mathfrak{C}$  is the standard basis. Find a matrix  $A \in M_{2 \times 3}(\mathbb{R})$  such that

$$f(x, y, z) = (x, y, z)A.$$

**(intermediate) 5.47** Let  $V$  be a vector space over  $\mathbb{R}$  and let  $\mathfrak{B} = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$  be an ordered basis for  $V$ . Let  $f \in \text{Hom}_{\mathbb{R}}(V, \mathbb{R}^2)$  be the linear transformation satisfying

$$[f]_{\mathfrak{C}}^{\mathfrak{B}} = \begin{bmatrix} 5 & -3 & 4 \\ -1 & 6 & 2 \end{bmatrix}$$

where

$$\mathfrak{C} = ((1, 2) \quad (0, -1)).$$

Let  $\mathbf{v} \in V$  satisfy

$$[\mathbf{v}]_{\mathfrak{B}} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

Find  $f(\mathbf{v})$ .

## 5.13 Algebra of transformations and matrix algebra

We now connect the composition of linear transformations to their matrix representation. Let  $U, V, W$  be vector spaces over  $\mathbb{F}$ , let  $f \in \text{Hom}_{\mathbb{F}}(U, V)$  and let  $g \in \text{Hom}_{\mathbb{F}}(V, W)$ . Let  $\mathfrak{B}$ ,  $\mathfrak{C}$  and  $\mathfrak{D}$  be ordered bases for  $U$ ,  $V$  and  $W$ . The following diagram is useful:

$$\begin{array}{ccccc}
 U & \xrightarrow{f} & V & \xrightarrow{g} & W \\
 \mathfrak{B} \downarrow & & \downarrow \mathfrak{C} & & \downarrow \mathfrak{D} \\
 \mathbb{F}_{\text{col}}^n & \xrightarrow{[f]_{\mathfrak{C}}^{\mathfrak{B}}} & \mathbb{F}_{\text{col}}^m & \xrightarrow{[g]_{\mathfrak{D}}^{\mathfrak{C}}} & \mathbb{F}_{\text{col}}^k
 \end{array}$$

**Proposition 5.33** *The above diagram commutes, namely,*

$$[g \circ f]_{\mathfrak{D}}^{\mathfrak{B}} = [g]_{\mathfrak{D}}^{\mathfrak{C}} [f]_{\mathfrak{C}}^{\mathfrak{B}}.$$

*In other words, the matrix representation of a composition is the product of the matrix representations.*

*Proof:* By definition, for every  $\mathbf{u} \in U$ ,

$$[f(\mathbf{u})]_{\mathfrak{C}} = [f]_{\mathfrak{C}}^{\mathfrak{B}} [\mathbf{u}]_{\mathfrak{B}}$$

and for every  $\mathbf{v} \in V$ ,

$$[g(\mathbf{v})]_{\mathfrak{D}} = [g]_{\mathfrak{D}}^{\mathfrak{C}} [\mathbf{v}]_{\mathfrak{C}}.$$

Combining the two,

$$[g(f(\mathbf{u}))]_{\mathfrak{D}} = [g]_{\mathfrak{D}}^{\mathfrak{C}} [f(\mathbf{u})]_{\mathfrak{C}} = [g]_{\mathfrak{D}}^{\mathfrak{C}} [f]_{\mathfrak{C}}^{\mathfrak{B}} [\mathbf{u}]_{\mathfrak{B}},$$

and by definition,

$$[g]_{\mathfrak{D}}^{\mathfrak{C}} [f]_{\mathfrak{C}}^{\mathfrak{B}} = [g \circ f]_{\mathfrak{D}}^{\mathfrak{B}}.$$

■

Linear transformation from  $V$  to  $W$  can also be added. The addition of linear transformations is represented by the addition of the corresponding transition matrices:

**Proposition 5.34** *Let  $V, W$  be vector spaces over  $\mathbb{F}$  and let  $f, g \in \text{Hom}_{\mathbb{F}}(V, W)$ . Let  $\mathfrak{B}$  and  $\mathfrak{C}$  be ordered bases for  $V$  and  $W$ . Then,*

$$[f + g]_{\mathfrak{C}}^{\mathfrak{B}} = [f]_{\mathfrak{C}}^{\mathfrak{B}} + [g]_{\mathfrak{C}}^{\mathfrak{B}}.$$

*Proof:* By definition, for every  $\mathbf{v} \in V$ ,

$$[f(\mathbf{v})]_{\mathfrak{C}} = [f]_{\mathfrak{C}}^{\mathfrak{B}}[\mathbf{v}]_{\mathfrak{B}} \quad \text{and} \quad [g(\mathbf{v})]_{\mathfrak{C}} = [g]_{\mathfrak{C}}^{\mathfrak{B}}[\mathbf{v}]_{\mathfrak{B}}.$$

Combining the two,

$$\begin{aligned} [(f+g)(\mathbf{v})]_{\mathfrak{C}} &= [f(\mathbf{v}) + g(\mathbf{v})]_{\mathfrak{C}} \\ &= [f(\mathbf{v})]_{\mathfrak{C}} + [g(\mathbf{v})]_{\mathfrak{C}} \\ &= [f]_{\mathfrak{C}}^{\mathfrak{B}}[\mathbf{v}]_{\mathfrak{B}} + [g]_{\mathfrak{C}}^{\mathfrak{B}}[\mathbf{v}]_{\mathfrak{B}} \\ &= ([f]_{\mathfrak{C}}^{\mathfrak{B}} + [g]_{\mathfrak{C}}^{\mathfrak{B}})[\mathbf{v}]_{\mathfrak{B}}, \end{aligned}$$

and by definition

$$[f]_{\mathfrak{C}}^{\mathfrak{B}} + [g]_{\mathfrak{C}}^{\mathfrak{B}} = [f+g]_{\mathfrak{C}}^{\mathfrak{B}}.$$

■

Similarly, we can prove:

**Proposition 5.35** *Let  $V, W$  be vector spaces over  $\mathbb{F}$ ; let  $f \in \text{Hom}_{\mathbb{F}}(V, W)$  and  $a \in \mathbb{F}$ . Let  $\mathfrak{B}$  and  $\mathfrak{C}$  be ordered bases for  $V$  and  $W$ . Then,*

$$[af]_{\mathfrak{C}}^{\mathfrak{B}} = a[f]_{\mathfrak{C}}^{\mathfrak{B}}.$$

*Proof:* We leave this as an exercise. ■

**Example:** Let  $V = U_1 \oplus U_2$ , where

$$\dim_{\mathbb{F}} U_1 = \dim_{\mathbb{F}} U_2 = 1.$$

Recall that every  $\mathbf{v} \in V$  has a unique representation as  $\mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2$ , and we defined the projection operators  $p_1, p_2 : V \rightarrow V$  by

$$p_1(\mathbf{v}) = \mathbf{u}_1 \quad \text{and} \quad p_2(\mathbf{v}) = \mathbf{u}_2,$$

and the reflection operators  $S_1, S_2 : V \rightarrow V$  by

$$S_1(\mathbf{v}) = \mathbf{u}_1 - \mathbf{u}_2 \quad \text{and} \quad S_2(\mathbf{v}) = \mathbf{u}_2 - \mathbf{u}_1.$$

These linear transformations satisfy the following additive relations,

$$\begin{aligned} p_1 + p_2 &= \text{Id}_V \\ p_1 - p_2 &= S_1 \\ p_2 - p_1 &= S_2 \\ S_1 + S_2 &= 0_{\text{Hom}_{\mathbb{F}}(V,V)}. \end{aligned}$$

In the present case there exist  $\mathbf{u}, \mathbf{w} \in V$ , such that

$$U_1 = \text{Span}\{\mathbf{u}\} \quad \text{and} \quad U_2 = \text{Span}\{\mathbf{w}\}.$$

Take  $\mathfrak{B} = (\mathbf{u}, \mathbf{w})$  as an ordered basis for  $V$ . We have

$$[\mathbf{u}]_{\mathfrak{B}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad [\mathbf{w}]_{\mathfrak{B}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Since  $p_1(\mathbf{u}) = \mathbf{u}$  and  $p_1(\mathbf{w}) = 0_V$ , the matrix representation of  $p_1$  relative to the basis  $\mathfrak{B}$  is

$$[p_1]_{\mathfrak{B}}^{\mathfrak{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Likewise,

$$[p_2]_{\mathfrak{B}}^{\mathfrak{B}} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Note that  $[p_1]_{\mathfrak{B}}^{\mathfrak{B}} + [p_2]_{\mathfrak{B}}^{\mathfrak{B}} = I_2 = [\text{Id}_V]_{\mathfrak{B}}^{\mathfrak{B}}$ . Further

$$[S_1]_{\mathfrak{B}}^{\mathfrak{B}} = [p_1]_{\mathfrak{B}}^{\mathfrak{B}} - [p_2]_{\mathfrak{B}}^{\mathfrak{B}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

and

$$[S_2]_{\mathfrak{B}}^{\mathfrak{B}} = [p_2]_{\mathfrak{B}}^{\mathfrak{B}} - [p_1]_{\mathfrak{B}}^{\mathfrak{B}} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix},$$

so that  $[S_1]_{\mathfrak{B}}^{\mathfrak{B}} + [S_2]_{\mathfrak{B}}^{\mathfrak{B}} = 0_{M_2(\mathbb{F})}$ , as expected.

Consider now compositions of these operators, for example,

$$p_1 \circ p_1 = p_1 \quad \text{and} \quad p_1 \circ p_2 = 0_V.$$

Indeed,

$$[p_1]_{\mathfrak{B}}^{\mathfrak{B}}[p_1]_{\mathfrak{B}}^{\mathfrak{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = [p_1 \circ p_1]_{\mathfrak{B}}^{\mathfrak{B}},$$

and

$$[p_1]_{\mathfrak{B}}[p_2]_{\mathfrak{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = [p_1 \circ p_2]_{\mathfrak{B}}.$$

▲ ▲ ▲

### Exercises

**(easy) 5.48** Show explicitly in the last example that  $p_1 \circ S_2 = -p_1$  and

$$[p_1]_{\mathfrak{B}}[S_1]_{\mathfrak{B}} = [p_1 \circ S_1]_{\mathfrak{B}}.$$

**(harder) 5.49** Let  $V$  be a three-dimensional vector space over a field  $\mathbb{F}$  and let  $f \in \text{Hom}_{\mathbb{F}}(V, V)$  be a linear transformation, which is not the zero transformation, satisfying

$$f \circ f = 0_{\text{Hom}_{\mathbb{F}}(V, V)}.$$

Show that there exists an ordered basis  $\mathfrak{B}$  for  $V$ , such that

$$[f]_{\mathfrak{B}} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hint: start by finding the dimensions of  $\ker f$  and  $\text{Image } f$ . Is one of those subspaces contained in the other?

**(harder) 5.50** Let  $V$  be a three-dimensional vector space over a field  $\mathbb{F}$  and let  $f \in \text{Hom}_{\mathbb{F}}(V, V)$  be a linear transformation satisfying

$$f \circ f \neq 0_{\text{Hom}_{\mathbb{F}}(V, V)} \quad \text{and} \quad f \circ f \circ f = 0_{\text{Hom}_{\mathbb{F}}(V, V)}.$$

Show that there exists an ordered basis  $\mathfrak{B}$  for  $V$ , such that

$$[f]_{\mathfrak{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hint: start by finding the dimensions of  $\ker f$  and  $\text{Image } f$ . Is one of those subspaces contained in the other? What is the implication of  $f \circ f$  not being the zero transformation?

## 5.14 Change of basis

The change of an ordered basis induces a change in the coordinate matrices of vectors. Likewise, it also induces a change in the matrix representation of linear transformations. The following theorem provides a formula for the change of the matrix representation.

**Theorem 5.36** *Let  $V$  be a finitely-generated vector space over  $\mathbb{F}$ ,  $\dim_{\mathbb{F}} V = n$ . Let  $\mathfrak{B}$  and  $\mathfrak{C}$  be ordered bases for  $V$ , such that*

$$\mathfrak{C} = \mathfrak{B}P,$$

*for some  $P \in \text{GL}_n(\mathbb{F})$ . Then, for  $f \in \text{Hom}_{\mathbb{F}}(V, V)$ ,*

$$[f]_{\mathfrak{C}}^{\mathfrak{C}} = P^{-1}[f]_{\mathfrak{B}}^{\mathfrak{B}}P.$$

*Proof:* By definition of the representing matrix, for every  $\mathbf{v} \in V$ ,

$$[f(\mathbf{v})]_{\mathfrak{B}} = [f]_{\mathfrak{B}}^{\mathfrak{B}}[\mathbf{v}]_{\mathfrak{B}} \quad \text{and} \quad [f(\mathbf{v})]_{\mathfrak{C}} = [f]_{\mathfrak{C}}^{\mathfrak{C}}[\mathbf{v}]_{\mathfrak{C}}.$$

Moreover,

$$P[\mathbf{v}]_{\mathfrak{C}} = [\mathbf{v}]_{\mathfrak{B}} \quad \text{and} \quad P[f(\mathbf{v})]_{\mathfrak{C}} = [f(\mathbf{v})]_{\mathfrak{B}},$$

from which we obtain,

$$[f]_{\mathfrak{C}}^{\mathfrak{C}}[\mathbf{v}]_{\mathfrak{C}} = [f(\mathbf{v})]_{\mathfrak{C}} = P^{-1}[f(\mathbf{v})]_{\mathfrak{B}} = P^{-1}[f]_{\mathfrak{B}}^{\mathfrak{B}}[\mathbf{v}]_{\mathfrak{B}} = P^{-1}[f]_{\mathfrak{B}}^{\mathfrak{B}}P[\mathbf{v}]_{\mathfrak{C}}.$$

This holds for every  $\mathbf{v} \in V$ , hence

$$[f]_{\mathfrak{C}}^{\mathfrak{C}} = P^{-1}[f]_{\mathfrak{B}}^{\mathfrak{B}}P.$$

■

**Example:** Let  $V = \mathbb{R}^2$  and let  $f \in \text{Hom}_{\mathbb{R}}(\mathbb{R}^2, \mathbb{R}^2)$  be given by

$$f(x, y) = (3x + 7y, 2x - 5y).$$

With respect to the standard basis  $\mathfrak{E}$ ,

$$[(x, y)]_{\mathfrak{E}} = \begin{bmatrix} x \\ y \end{bmatrix},$$

and

$$[f(x, y)]_{\mathfrak{E}} = \begin{bmatrix} 3x + 7y \\ 2x - 5y \end{bmatrix} = \begin{bmatrix} 3 & 7 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

namely

$$[f]_{\mathfrak{E}} = \begin{bmatrix} 3 & 7 \\ 2 & -5 \end{bmatrix}.$$

Let now

$$\mathfrak{B} = ((1, 2) \quad (2, 1))$$

be another ordered basis for  $\mathbb{R}^2$ . Then,

$$((1, 2) \quad (2, 1)) = ((1, 0) \quad (0, 1)) \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix},$$

and for  $(x, y) \in \mathbb{R}^2$ ,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} [(x, y)]_{\mathfrak{B}} \quad \text{and} \quad [(x, y)]_{\mathfrak{B}} = \begin{bmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Now,

$$\begin{aligned} [f(x, y)]_{\mathfrak{B}} &= [(3x + 7y, 2x - 5y)]_{\mathfrak{B}} \\ &= \begin{bmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{bmatrix} \begin{bmatrix} 3x + 7y \\ 2x - 5y \end{bmatrix} \\ &= \begin{bmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{bmatrix} \begin{bmatrix} 3 & 7 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{bmatrix} \begin{bmatrix} 3 & 7 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} [(x, y)]_{\mathfrak{B}}. \end{aligned}$$

We conclude that

$$[f]_{\mathfrak{B}} = \begin{bmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{bmatrix} \begin{bmatrix} 3 & 7 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}.$$

▲ ▲ ▲

**Definition 5.37** Square matrices  $A, B \in M_n(\mathbb{F})$  are called **similar** (דומות), if there exists an invertible matrix  $P \in \text{GL}_n(\mathbb{F})$  such that

$$B = P^{-1}AP.$$

Thus, we have proved that matrices representing the same linear transformation  $f \in \text{Hom}_{\mathbb{F}}(V, V)$  relative to different bases are similar. The opposite is also true: two matrices that are similar represent the same linear transformation relative to different bases.

**Proposition 5.38** Similarity between matrices is an equivalence relation.

*Proof:* This is left as an exercise. ■

## Exercises

**(intermediate) 5.51** Let

$$\mathfrak{B} = ((2, 1) \quad (3, 2)) \quad \text{and} \quad \mathfrak{C} = ((1, -1) \quad (-1, 2))$$

be ordered bases for  $\mathbb{R}^2$ . Let  $f \in \text{Hom}_{\mathbb{R}}(\mathbb{R}^2, \mathbb{R}^2)$  be the linear transformation satisfying

$$[f]_{\mathfrak{B}}^{\mathfrak{B}} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}.$$

Find  $[f]_{\mathfrak{C}}^{\mathfrak{C}}$ .

**(easy) 5.52** Prove that similarity between matrices is an equivalence relation (first, remind yourself what it takes to be a similarity relation).

**(easy) 5.53** Prove that similar matrices represent the same linear transformation relative to different bases.

**(easy) 5.54** Show that for any scalar  $a$ , the matrix  $aI_n$  is similar only to itself. Interpret this result in terms of the matrix representation of linear transformations.

**(intermediate) 5.55** Let  $A, B \in M_n(\mathbb{R})$ . Prove or disprove each of the following statements:

- (a) If  $A$  and  $B$  are row-equivalent, then they are similar.
- (b) If  $A$  and  $B$  are similar, then they are row-equivalent.
- (c) If  $A$  and  $B$  are similar and  $A$  is invertible, then  $B$  is invertible.
- (d) If  $A$  is not invertible, then it is similar to a matrix having a row of zeros.

**(intermediate) 5.56** Let  $A \in M_2(\mathbb{R})$  be similar to the matrix

$$D = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$$

Prove that

$$(A - 3I_2)(A - 2I_2) = 0.$$

