

# Linear Algebra 2: Lecture 23

Irena Penev

Summer 2023

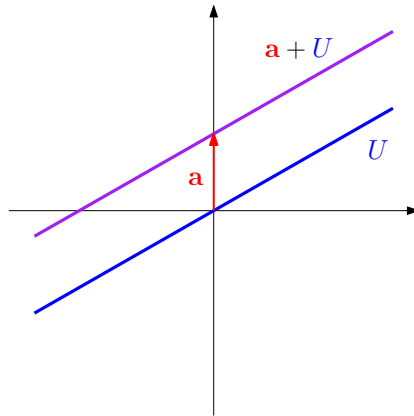
**Terminology:** So far, we have typically referred to vector/linear subspaces of a vector space simply as “subspaces.” In this lecture, we will study a generalization of linear subspaces, called “affine subspaces.” To avoid any confusion, in this lecture, we will not use the term “subspace” and will instead always write either “linear subspace” or “affine subspace.”

## 1 Affine subspaces

An *affine subspace* of a vector space  $V$  over a field  $\mathbb{F}$  is any set of the form

$$\mathbf{a} + U := \{\mathbf{a} + \mathbf{u} \mid \mathbf{u} \in U\},$$

where  $\mathbf{a}$  is a vector in  $V$  and  $U$  is a linear subspace of  $V$ .



Thus, an affine subspace of  $V$  is obtained by shifting a linear subspace  $U$  of  $V$  by some vector  $\mathbf{a}$ .

### Remarks:

1. For a vector space  $V$  over a field  $\mathbb{F}$ :

- every linear subspace  $U$  of  $V$  is also an affine subspace of  $V$ , since  $U = \mathbf{0} + U$ ;<sup>1</sup>

---

<sup>1</sup>Moreover, as we shall see, linear subspaces of  $V$  are precisely those affine subspaces of  $V$  that contain  $\mathbf{0}$  (see Corollary 1.2).

- $V$  is an affine subspace of itself (because  $V$  is a linear subspace of itself);
  - for every vector  $\mathbf{a} \in V$ ,  $\{\mathbf{a}\}$  is an affine subspace of  $V$ , since  $\{\mathbf{a}\} = \mathbf{a} + \{\mathbf{0}\}$  and  $\{\mathbf{0}\}$  is a linear subspace of  $V$ .
2. As we know, linear subspaces of  $\mathbb{R}^n$  are  $\{\mathbf{0}\}$ , lines through the origin, planes through the origin, and higher dimensional generalizations. So, affine subspaces of  $\mathbb{R}^n$  are  $\{\mathbf{a}\}$  (for any vector  $\mathbf{a} \in \mathbb{R}^n$ ), lines, planes, and higher dimensional generalizations (these lines, planes, and higher dimensional generalizations may, but need not, pass through the origin).

As Theorem 1.1 (below) shows, for an affine subspace  $M = \mathbf{a} + U$  (where  $\mathbf{a}$  and  $U$  are as in the definition above), the vector  $\mathbf{a}$  need not be unique (indeed, it can be any vector in  $M$ ),<sup>2</sup> but the linear subspace  $U$  is unique (it depends only on  $M$ , and not on the vector  $\mathbf{a}$ ).<sup>3</sup>

**Theorem 1.1.** *Let  $V$  be a vector space over a field  $\mathbb{F}$ , and let  $M = \mathbf{a} + U$  be an affine subspace of  $V$ , where  $\mathbf{a}$  is a vector and  $U$  a linear subspace of  $V$ . Then all the following hold:*

- (a)  $\mathbf{a} \in M$  (and in particular,  $M \neq \emptyset$ );
- (b) for all  $\mathbf{a}' \in M$ , we have that  $M = \mathbf{a}' + U$ ;
- (c) for all  $\mathbf{b} \in V \setminus M$ , we have that  $M \cap (\mathbf{b} + U) = \emptyset$ ;
- (d) for all vectors  $\mathbf{a}'$  and linear subspaces  $U'$  of  $V$  such that  $M = \mathbf{a}' + U'$ , we have that  $U' = U$ .

*Proof.* (a) Since  $U$  is a linear subspace of  $V$ , we have that  $\mathbf{0} \in U$ , and consequently,  $\mathbf{a} = \mathbf{a} + \mathbf{0} \in \mathbf{a} + U = M$ .

(b) Fix  $\mathbf{a}' \in M$ . Since  $\mathbf{a}' \in M = \mathbf{a} + U$ , there exists some  $\mathbf{u}' \in U$  such that  $\mathbf{a}' = \mathbf{a} + \mathbf{u}'$ . Now, we must show that  $M = \mathbf{a}' + U$ .

Let us first show that  $M \subseteq \mathbf{a}' + U$ . Fix  $\mathbf{x} \in M$ . Since  $M = \mathbf{a} + U$ , there exists some  $\mathbf{u} \in U$  such that  $\mathbf{x} = \mathbf{a} + \mathbf{u}$ . Then  $\mathbf{x} = \mathbf{a} + \mathbf{u} = (\mathbf{a}' - \mathbf{u}') + \mathbf{u} = \mathbf{a}' + (\mathbf{u} - \mathbf{u}')$ . Since  $\mathbf{u}, \mathbf{u}' \in U$ , and  $U$  is a linear subspace of  $V$ , we have that  $\mathbf{u} - \mathbf{u}' \in U$ ; so,  $\mathbf{x} = \mathbf{a}' + (\mathbf{u} - \mathbf{u}') \in \mathbf{a}' + U$ . This proves that  $M \subseteq \mathbf{a}' + U$ .

Let us now show that  $\mathbf{a}' + U \subseteq M$ . Fix  $\mathbf{u} \in U$ ; we must show that  $\mathbf{a}' + \mathbf{u} \in M$ . But note that  $\mathbf{a}' + \mathbf{u} = \mathbf{a} + \mathbf{u}' + \mathbf{u}$ . Since  $\mathbf{u}', \mathbf{u} \in U$ , and  $U$  is a linear subspace of  $V$ , we have that  $\mathbf{u}' + \mathbf{u} \in U$ ; consequently,  $\mathbf{a}' + \mathbf{u} = \mathbf{a} + \mathbf{u}' + \mathbf{u} \in \mathbf{a} + U = M$ . This proves that  $\mathbf{a}' + U \subseteq M$ .

(c) Fix  $\mathbf{b} \in V \setminus M$ . We must show that  $M \cap (\mathbf{b} + U) = \emptyset$ . Suppose otherwise, and fix  $\mathbf{x} \in M \cap (\mathbf{b} + U)$ . Since  $\mathbf{x} \in M = \mathbf{a} + U$ , there exists some

<sup>2</sup>This follows that Theorem 1.1(b)

<sup>3</sup>This follows from Theorem 1.1(d).

$\mathbf{u}_1 \in U$  such that  $\mathbf{x} = \mathbf{a} + \mathbf{u}_1$ ; on the other hand, since  $\mathbf{x} \in \mathbf{b} + U$ , there exists some  $\mathbf{u}_2 \in U$  such that  $\mathbf{x} = \mathbf{b} + \mathbf{u}_2$ . So,  $\mathbf{a} + \mathbf{u}_1 = \mathbf{b} + \mathbf{u}_2$ , and it follows that  $\mathbf{b} = \mathbf{a} + (\mathbf{u}_1 - \mathbf{u}_2)$ . Since  $\mathbf{u}_1, \mathbf{u}_2 \in U$ , and since  $U$  is a linear subspace of  $V$ , we have that  $\mathbf{u}_1 - \mathbf{u}_2 \in U$ ; consequently,  $\mathbf{b} = \mathbf{a} + (\mathbf{u}_1 - \mathbf{u}_2) \in \mathbf{a} + U = M$ , contrary to the fact that  $\mathbf{b} \in V \setminus M$ .

(d) Fix a vector  $\mathbf{a}'$  and a linear subspace  $U'$  of  $V$  such that  $M = \mathbf{a}' + U'$ . By (a), we have that  $\mathbf{a}' \in M$ , and so by (b), we have that  $M = \mathbf{a}' + U$ . So,  $\mathbf{a}' + U' = \mathbf{a}' + U$ , and we deduce that  $U' = U$ .<sup>4</sup>  $\square$

Given a vector space  $V$  over a field  $\mathbb{F}$ , we define the *dimension* of an affine subspace  $M = \mathbf{a} + U$  of  $V$  (where  $\mathbf{a}$  is a vector and  $U$  a linear subspace of  $V$ ) to be  $\dim(U)$ . By Theorem 1.1(d), this is well defined.

**Corollary 1.2.** *Let  $V$  be a vector space over a field  $\mathbb{F}$ . Then linear subspaces of  $V$  are precisely those affine spaces of  $V$  that contain  $\mathbf{0}$ . In other words, for all  $U \subseteq V$ , the following are equivalent:*

- (i)  $U$  is a linear subspace of  $V$ ;
- (ii)  $U$  is an affine subspace of  $V$  and  $\mathbf{0} \in U$ .

*Proof.* Fix  $U \subseteq V$ . Suppose first that (i) holds. Then clearly,  $\mathbf{0} \in U$ , and moreover,  $U = \mathbf{0} + U$ . So, (ii) holds.

Suppose now that (ii) holds. Since  $U$  is an affine subspace of  $V$ , we know that there exists a vector  $\mathbf{a} \in V$  and a linear subspace  $U'$  of  $V$  such that  $U = \mathbf{a} + U'$ . Moreover, by (ii), we have that  $\mathbf{0} \in U$ , and so by Theorem 1.1(b), we have that  $U = \mathbf{0} + U'$ . So,  $U = U'$ , and so since  $U'$  is a linear subspace of  $V$ , we see that (i) holds.  $\square$

Recall that the intersection of two linear subspaces is a linear subspace.<sup>5</sup> In the case of affine subspaces, we have the following corollary.

**Corollary 1.3.** *Let  $V$  be a vector space over a field  $\mathbb{F}$ , and let  $M_1$  and  $M_2$  be affine subspaces of  $V$ . Then either  $M_1 \cap M_2 = \emptyset$ , or  $M_1 \cap M_2$  is an affine subspace of  $V$ .*

*Proof.* We may assume that  $M_1 \cap M_2 \neq \emptyset$ , for otherwise we are done. Fix any  $\mathbf{a} \in M_1 \cap M_2$ . By Theorem 1.1,  $M_1$  and  $M_2$  can be written as  $M_1 = \mathbf{a} + U_1$  and  $M_2 = \mathbf{a} + U_2$ , for some linear subspaces  $U_1$  and  $U_2$  of  $V$ . Then  $U := U_1 \cap U_2$  is a linear subspace of  $V$ .<sup>6</sup> Moreover, it is clear that  $M_1 \cap M_2 = \mathbf{a} + U$ ,<sup>7</sup> and so  $M_1 \cap M_2$  is an affine subspace.  $\square$

<sup>4</sup>This is “obvious,” but here is a formal proof. By symmetry, it suffices to show that  $U' \subseteq U$ . Fix  $\mathbf{u}' \in U'$ . Then  $\mathbf{a}' + \mathbf{u}' \in \mathbf{a}' + U' = \mathbf{a}' + U$ , and it follows that there exists some  $\mathbf{u} \in U$  such that  $\mathbf{a}' + \mathbf{u}' = \mathbf{a}' + \mathbf{u}$ . By subtracting  $\mathbf{a}'$  from both sides, we get  $\mathbf{u}' = \mathbf{u}$ ; since  $\mathbf{u} \in U$ , we deduce that  $\mathbf{u}' \in U$ . So,  $U' \subseteq U$ .

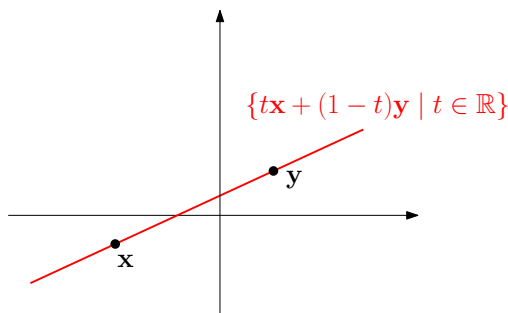
<sup>5</sup>This was Problem 4 of HW#6 from Linear Algebra 1, winter 2022.

<sup>6</sup>This follows from Problem 4 of HW#6 from Linear Algebra 1, winter 2022.

<sup>7</sup>This is “obvious,” but here is a full proof. It is clear that  $\mathbf{a} + U \subseteq M_1 \cap M_2$ . For

## 2 Affine combinations and affine hulls

Recall from analytic geometry that if  $\mathbf{x}$  and  $\mathbf{y}$  are distinct points (vectors) in  $\mathbb{R}^2$ , then the line in  $\mathbb{R}^2$  that passes through  $\mathbf{x}$  and  $\mathbf{y}$  is  $\{t\mathbf{x} + (1-t)\mathbf{y} \mid t \in \mathbb{R}\}$ . This in fact holds for all distinct points  $\mathbf{x}$  and  $\mathbf{y}$  in  $\mathbb{R}^n$  (not just  $\mathbb{R}^2$ ). Affine combinations are a generalization of this concept.



Suppose that  $\mathbf{x}_1, \dots, \mathbf{x}_n$  ( $n \geq 1$ ) are vectors in a vector space  $V$  over a field  $\mathbb{F}$ . An *affine combination* of  $\mathbf{x}_1, \dots, \mathbf{x}_n$  is any sum of the form  $\alpha_1\mathbf{x}_1 + \dots + \alpha_n\mathbf{x}_n$ , where  $\alpha_1, \dots, \alpha_n \in \mathbb{F}$  satisfy  $\alpha_1 + \dots + \alpha_n = 1$ . The set of all affine combinations of  $\mathbf{x}_1, \dots, \mathbf{x}_n$ , denoted  $\text{Aff}(\mathbf{x}_1, \dots, \mathbf{x}_n)$ , is called the *affine hull* (or *affine span*) of  $\mathbf{x}_1, \dots, \mathbf{x}_n$ . So, we have that

$$\text{Aff}(\mathbf{x}_1, \dots, \mathbf{x}_n) := \left\{ \sum_{i=1}^n \alpha_i \mathbf{x}_i \mid \alpha_1, \dots, \alpha_n \in \mathbb{F}, \sum_{i=1}^n \alpha_i = 1 \right\}.$$

Since  $\mathbf{x}_i = 0\mathbf{x}_1 + \dots + 0\mathbf{x}_{i-1} + 1\mathbf{x}_i + 0\mathbf{x}_{i+1} + \dots + 0\mathbf{x}_n$  for all  $i \in \{1, \dots, n\}$ , we see that  $\mathbf{x}_1, \dots, \mathbf{x}_n \in \text{Aff}(\mathbf{x}_1, \dots, \mathbf{x}_n)$ . As Theorem 2.1 (below) shows, affine subspaces of  $V$  are precisely those non-empty subsets of  $V$  that are closed under affine combinations. As a corollary (see Corollary 2.2), we deduce that all affine hulls are affine subspaces of  $V$ .

**Theorem 2.1.** *Let  $V$  be a vector space over a field  $\mathbb{F}$ , and let  $M$  be a non-empty subset of  $V$ . Then the following are equivalent:*

- (i)  $M$  is an affine subspace of  $V$ ;
- (ii)  $M$  is closed under affine combinations, that is, for all  $\mathbf{x}_1, \dots, \mathbf{x}_n \in M$  and  $\alpha_1, \dots, \alpha_n \in \mathbb{F}$  such that  $\alpha_1 + \dots + \alpha_n = 1$ , we have that  $\alpha_1\mathbf{x}_1 + \dots + \alpha_n\mathbf{x}_n \in M$ .

the reverse inclusion, we fix some  $\mathbf{x} \in M_1 \cap M_2$ , and we show that  $\mathbf{x} \in \mathbf{a} + U$ . Since  $\mathbf{x} \in M_1 = \mathbf{a} + U_1$ , we know that there exists some  $\mathbf{u}_1 \in U_1$  such that  $\mathbf{x} = \mathbf{a} + \mathbf{u}_1$ . Similarly, since  $\mathbf{x} \in M_2 = \mathbf{a} + U_2$ , there exists some  $\mathbf{u}_2 \in U_2$  such that  $\mathbf{x} = \mathbf{a} + \mathbf{u}_2$ . So,  $\mathbf{a} + \mathbf{u}_1 = \mathbf{a} + \mathbf{u}_2$ , and consequently,  $\mathbf{u}_1 = \mathbf{u}_2$ . Since  $\mathbf{u}_1 \in U_1$  and  $\mathbf{u}_2 \in U_2$ , we deduce that  $\mathbf{u}_1 = \mathbf{u}_2$  belongs to  $U_1 \cap U_2 = U$ . But now  $\mathbf{x} = \mathbf{a} + \mathbf{u}_1 \in \mathbf{a} + U$ , and we are done.

*Proof.* Assume first that (i) holds. Let us prove (ii). Set  $M = \mathbf{a} + U$ , where  $\mathbf{a}$  is a vector and  $U$  a linear subspace of  $V$ , as in the definition of an affine subspace. Fix  $\mathbf{x}_1, \dots, \mathbf{x}_n \in M$ , and fix  $\alpha_1, \dots, \alpha_n \in \mathbb{F}$  such that  $\alpha_1 + \dots + \alpha_n = 1$ ; we must show that  $\alpha_1 \mathbf{x}_1 + \dots + \alpha_n \mathbf{x}_n$  belongs to  $M$ . Since  $\mathbf{x}_1, \dots, \mathbf{x}_n \in M = \mathbf{a} + U$ , there exist vectors  $\mathbf{u}_1, \dots, \mathbf{u}_n \in U$  such that  $\mathbf{x}_1 = \mathbf{a} + \mathbf{u}_1, \dots, \mathbf{x}_n = \mathbf{a} + \mathbf{u}_n$ . We now have that

$$\begin{aligned} \alpha_1 \mathbf{x}_1 + \dots + \alpha_n \mathbf{x}_n &= \alpha_1(\mathbf{a} + \mathbf{u}_1) + \dots + \alpha_n(\mathbf{a} + \mathbf{u}_n) \\ &= \underbrace{(\alpha_1 + \dots + \alpha_n)}_{=1} \mathbf{a} + (\alpha_1 \mathbf{u}_1 + \dots + \alpha_n \mathbf{u}_n) \\ &= \mathbf{a} + \underbrace{(\alpha_1 \mathbf{u}_1 + \dots + \alpha_n \mathbf{u}_n)}_{:=\mathbf{u}}. \end{aligned}$$

Since  $\mathbf{u}_1, \dots, \mathbf{u}_n \in U$ , and  $U$  is a linear subspace of  $V$ , we have that  $\mathbf{u} \in U$ . So,  $\alpha_1 \mathbf{x}_1 + \dots + \alpha_n \mathbf{x}_n = \mathbf{a} + \mathbf{u} \in \mathbf{a} + U = M$ . This proves (ii).

Conversely, suppose that (ii) holds. We must prove (i). Using the fact that  $M \neq \emptyset$ , we fix some  $\mathbf{a} \in M$ . Set  $U := \{\mathbf{x} - \mathbf{a} \mid \mathbf{x} \in M\}$ . Clearly,  $M = \mathbf{a} + U$ . It remains to show that  $U$  is a linear subspace of  $V$ . By Theorem 2.7 of Lecture Notes 6, it suffices to show that  $\mathbf{0} \in U$ , and that  $U$  is closed under vector addition and scalar multiplication.

First, since  $\mathbf{a} \in M$ , we have that  $\mathbf{0} = \mathbf{a} - \mathbf{a} \in U$ .

Next, fix  $\mathbf{u}_1, \mathbf{u}_2 \in U$ . We must show that  $\mathbf{u}_1 + \mathbf{u}_2 \in U$ . Since  $\mathbf{u}_1, \mathbf{u}_2 \in U$ , there exist  $\mathbf{x}_1, \mathbf{x}_2 \in M$  such that  $\mathbf{u}_1 = \mathbf{x}_1 - \mathbf{a}$  and  $\mathbf{u}_2 = \mathbf{x}_2 - \mathbf{a}$ . Then

$$\mathbf{u}_1 + \mathbf{u}_2 = (\mathbf{x}_1 - \mathbf{a}) + (\mathbf{x}_2 - \mathbf{a}) = \left( \underbrace{1\mathbf{x}_1 + 1\mathbf{x}_2 + (-1)\mathbf{a}}_{:=\mathbf{y}} \right) - \mathbf{a}.$$

Since  $\mathbf{x}_1, \mathbf{x}_2, \mathbf{a} \in M$  and  $1 + 1 + (-1) = 1$ , and since (ii) holds, we see that  $\mathbf{y} \in M$ . But now  $\mathbf{u}_1 + \mathbf{u}_2 = \mathbf{y} - \mathbf{a} \in U$ .

Finally, fix  $\mathbf{u} \in U$  and  $\alpha \in \mathbb{F}$ ; we must show that  $\alpha \mathbf{u} \in U$ . Since  $\mathbf{u} \in U$ , we know that there exists some  $\mathbf{x} \in M$  such that  $\mathbf{u} = \mathbf{x} - \mathbf{a}$ . But now

$$\alpha \mathbf{u} = \alpha(\mathbf{x} - \mathbf{a}) = \left( \underbrace{\alpha \mathbf{x} + (1 - \alpha)\mathbf{a}}_{:=\mathbf{y}} \right) - \mathbf{a}.$$

Since  $\mathbf{x}, \mathbf{a} \in M$ , and since (ii) holds, we have that  $\mathbf{y} = \alpha \mathbf{x} + (1 - \alpha)\mathbf{a} \in M$ . But now  $\alpha \mathbf{u} = \mathbf{y} - \mathbf{a} \in U$ .

We have now shown that  $U$  is a linear subspace of  $V$ , and it follows that (i) holds.  $\square$

**Corollary 2.2.** *Let  $\mathbf{x}_1, \dots, \mathbf{x}_n$  ( $n \geq 1$ ) be vectors in a vector space  $V$  over a field  $\mathbb{F}$ . Then  $M := \text{Aff}(\mathbf{x}_1, \dots, \mathbf{x}_n)$  is an affine subspace of  $V$ .*

*Proof.* Since  $\mathbf{x}_1, \dots, \mathbf{x}_n \in M$ , we see that  $M \neq \emptyset$ . In view of Theorem 2.1, it now suffices to show that  $M$  is closed under affine combinations. Fix  $\mathbf{y}_1, \dots, \mathbf{y}_m \in M$  and  $\alpha_1, \dots, \alpha_m \in \mathbb{F}$  such that  $\alpha_1 + \dots + \alpha_m = 1$ . We must show that  $\mathbf{y} := \alpha_1 \mathbf{y}_1 + \dots + \alpha_m \mathbf{y}_m$  belongs to  $M$ . Since  $\mathbf{y}_1, \dots, \mathbf{y}_m \in M$ , we see that for all  $i \in \{1, \dots, m\}$ ,  $\mathbf{y}_i$  is an affine combination of vectors  $\mathbf{x}_1, \dots, \mathbf{x}_n$ , that is, there exist scalars  $\beta_{i,1}, \dots, \beta_{i,n} \in \mathbb{F}$  such that  $\mathbf{y}_i = \sum_{j=1}^n \beta_{i,j} \mathbf{x}_j$  and  $\sum_{j=1}^n \beta_{i,j} = 1$ . But now

$$\mathbf{y} = \sum_{i=1}^m \alpha_i \mathbf{y}_i = \sum_{i=1}^m \alpha_i \left( \sum_{j=1}^n \beta_{i,j} \mathbf{x}_j \right) = \sum_{j=1}^n \sum_{i=1}^m \alpha_i \beta_{i,j} \mathbf{x}_j.$$

For each  $j \in \{1, \dots, n\}$ , we set  $\gamma_j := \sum_{i=1}^m \alpha_i \beta_{i,j}$ . Then  $\mathbf{y} = \sum_{j=1}^n \gamma_j \mathbf{x}_j$ . It now remains to show that  $\sum_{j=1}^n \gamma_j = 1$ , for this will imply that  $\mathbf{y}$  is an affine combination of  $\mathbf{x}_1, \dots, \mathbf{x}_n$ , that is, that  $\mathbf{y} \in M$ , which is what we need to show. We compute:

$$\sum_{j=1}^n \gamma_j = \sum_{j=1}^n \sum_{i=1}^m \alpha_i \beta_{i,j} = \sum_{i=1}^m \alpha_i \left( \sum_{j=1}^n \beta_{i,j} \right) \stackrel{(*)}{=} \sum_{i=1}^m \alpha_i = 1,$$

where (\*) follows from the fact that  $\sum_{j=1}^n \beta_{i,j} = 1$ . This completes the argument.  $\square$

**Corollary 2.3.** *Let  $V$  be a vector space over a field  $\mathbb{F}$ , let  $M$  be an affine subspace of  $V$ , and let  $\mathbf{x}_1, \dots, \mathbf{x}_n$  ( $n \geq 1$ ) be vectors in  $V$ . Then the following are equivalent:*

- (i)  $M = \text{Aff}(\mathbf{x}_1, \dots, \mathbf{x}_n)$ ;
- (ii)  $\mathbf{x}_1, \dots, \mathbf{x}_n \in M$ , and every vector in  $M$  is an affine combination of  $\mathbf{x}_1, \dots, \mathbf{x}_n$ .

*Proof.* Obviously, (i) implies (ii). For the reverse implication, we assume that (ii) holds, and we prove (i). Since every vector in  $M$  is an affine combination of  $\mathbf{x}_1, \dots, \mathbf{x}_n$ , we have that  $M \subseteq \text{Aff}(\mathbf{x}_1, \dots, \mathbf{x}_n)$ . Let us prove the reverse inclusion. Fix  $\mathbf{x} \in \text{Aff}(\mathbf{x}_1, \dots, \mathbf{x}_n)$ . By (ii), we have that  $\mathbf{x}_1, \dots, \mathbf{x}_n \in M$ , and by Theorem 2.1, we know that  $M$  is closed under affine combinations. Since  $\mathbf{x}$  is an affine combination of  $\mathbf{x}_1, \dots, \mathbf{x}_n$ , we deduce that  $\mathbf{x} \in M$ . This proves that  $\text{Aff}(\mathbf{x}_1, \dots, \mathbf{x}_n) \subseteq M$ . Thus, (i) holds.  $\square$

### 3 Affine frames and affine bases

We have extensively studied bases of (finite-dimensional) vector spaces. For affine subspaces, we have two analogues of bases: “affine frames” and “affine bases.”

### 3.1 Affine frames

Let  $n$  be a non-negative integer, and let  $M$  be an  $n$ -dimensional affine subspace of a vector space  $V$  over a field  $\mathbb{F}$ . An *affine frame* of  $M$  is an ordered  $(n + 1)$ -tuple  $(\mathbf{a}, \mathbf{u}_1, \dots, \mathbf{u}_n)$  of vectors of  $V$  such that  $M$  can be written in the form  $M = \mathbf{a} + U$ , where  $U$  is a linear subspace of  $V$ , and  $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$  is a basis of  $U$ .

**Remark:** Infinite-dimensional affine subspaces do not have affine frames.

By Theorem 1.3 of Lecture Notes 7, if  $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  is a basis of a vector space  $V$  over a field  $\mathbb{F}$ , then every vector in  $V$  can be written as a linear combination of the vectors  $\mathbf{v}_1, \dots, \mathbf{v}_n$  in a unique way. Our next theorem is an analogue of this result for affine subspaces and affine frames.

**Theorem 3.1.** *Let  $M$  be an affine subspace of a vector space  $V$  over a field  $\mathbb{F}$ , and let  $(\mathbf{a}, \mathbf{u}_1, \dots, \mathbf{u}_n)$  be an affine frame of  $M$ . Then for all  $\mathbf{x} \in M$ , there exist unique scalars  $\alpha_1, \dots, \alpha_n \in \mathbb{F}$  such that  $\mathbf{x} = \mathbf{a} + \alpha_1 \mathbf{u}_1 + \dots + \alpha_n \mathbf{u}_n$ .*

*Proof.* Set  $U := \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_n)$ , so that  $M = \mathbf{a} + U$ . Fix  $\mathbf{x} \in M$ . We must show that there exist unique scalars  $\alpha_1, \dots, \alpha_n \in \mathbb{F}$  such that  $\mathbf{x} = \mathbf{a} + \alpha_1 \mathbf{u}_1 + \dots + \alpha_n \mathbf{u}_n$ .

We first prove existence. Since  $\mathbf{x} \in M = \mathbf{a} + U$ , there exists some  $\mathbf{u} \in U$  such that  $\mathbf{x} = \mathbf{a} + \mathbf{u}$ . Since  $\mathbf{u} \in U = \text{span}(\mathbf{u}_1, \dots, \mathbf{u}_n)$ , we know that there exist scalars  $\alpha_1, \dots, \alpha_n$  such that  $\mathbf{u} = \alpha_1 \mathbf{u}_1 + \dots + \alpha_n \mathbf{u}_n$ . So,  $\mathbf{x} = \mathbf{a} + \alpha_1 \mathbf{u}_1 + \dots + \alpha_n \mathbf{u}_n$ . This proves existence.

Let us prove uniqueness. Suppose that  $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in \mathbb{F}$  are such that  $\mathbf{x} = \mathbf{a} + \alpha_1 \mathbf{u}_1 + \dots + \alpha_n \mathbf{u}_n$  and  $\mathbf{x} = \mathbf{a} + \beta_1 \mathbf{u}_1 + \dots + \beta_n \mathbf{u}_n$ . Then  $\mathbf{a} + \alpha_1 \mathbf{u}_1 + \dots + \alpha_n \mathbf{u}_n = \mathbf{a} + \beta_1 \mathbf{u}_1 + \dots + \beta_n \mathbf{u}_n$ , and consequently,  $(\alpha_1 - \beta_1) \mathbf{u}_1 + \dots + (\alpha_n - \beta_n) \mathbf{u}_n = \mathbf{0}$ . Since the set  $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$  is linearly independent,<sup>8</sup> we have that  $\alpha_1 - \beta_1 = \dots = \alpha_n - \beta_n = 0$ . It follows that  $\alpha_i = \beta_i$  for all  $i \in \{1, \dots, n\}$ . This proves uniqueness.  $\square$

### 3.2 Affine independence

Given vectors  $\mathbf{x}_1, \dots, \mathbf{x}_n$  in a vector space  $V$  over a field  $\mathbb{F}$ , we say that vectors  $\mathbf{x}_1, \dots, \mathbf{x}_n \in V$  are *affinely independent*, or that the set  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  is *affinely independent*, if for all  $\alpha_1, \dots, \alpha_n \in \mathbb{F}$  such that

$$\alpha_1 \mathbf{x}_1 + \dots + \alpha_n \mathbf{x}_n = \mathbf{0} \quad \text{and} \quad \alpha_1 + \dots + \alpha_n = 0,$$

we have that  $\alpha_1 = \dots = \alpha_n = 0$ .

**Proposition 3.2.** *Let  $V$  be a vector space over a field  $\mathbb{F}$ , and let  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n$  ( $n \geq 0$ ) be vectors in  $V$ . Then the following are equivalent:*

---

<sup>8</sup>This follows from the fact that  $(\mathbf{a}, \mathbf{u}_1, \dots, \mathbf{u}_n)$  is an affine frame of  $M$ .

(i)  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n$  are affinely independent;

(ii) there exists some  $i \in \{0, 1, \dots, n\}$  such that vectors

$$\mathbf{x}_0 - \mathbf{x}_i, \dots, \mathbf{x}_{i-1} - \mathbf{x}_i, \mathbf{x}_{i+1} - \mathbf{x}_i, \dots, \mathbf{x}_n - \mathbf{x}_i$$

are linearly independent;

(iii) for all  $i \in \{0, 1, \dots, n\}$ , vectors

$$\mathbf{x}_0 - \mathbf{x}_i, \dots, \mathbf{x}_{i-1} - \mathbf{x}_i, \mathbf{x}_{i+1} - \mathbf{x}_i, \dots, \mathbf{x}_n - \mathbf{x}_i$$

are linearly independent.

*Proof.* Obviously, (iii) implies (ii). We will show that (ii) implies (i), and that (i) implies (iii).

Suppose that (ii) holds. Let us prove (i). By (ii) and by symmetry, we may assume that  $\mathbf{x}_1 - \mathbf{x}_0, \dots, \mathbf{x}_n - \mathbf{x}_0$  are linearly independent. Now, fix scalars  $\alpha_0, \alpha_1, \dots, \alpha_n \in \mathbb{F}$  such that  $\alpha_0 \mathbf{x}_0 + \alpha_1 \mathbf{x}_1 + \dots + \alpha_n \mathbf{x}_n = \mathbf{0}$  and  $\alpha_0 + \alpha_1 + \dots + \alpha_n = 0$ . We must show that  $\alpha_0 = \alpha_1 = \dots = \alpha_n = 0$ . Since  $\alpha_0 + \alpha_1 + \dots + \alpha_n = 0$ , we have that  $\alpha_0 = -\alpha_1 - \dots - \alpha_n$ , and so

$$\begin{aligned} \mathbf{0} &= \alpha_0 \mathbf{x}_0 + \alpha_1 \mathbf{x}_1 + \dots + \alpha_n \mathbf{x}_n \\ &= (-\alpha_1 - \dots - \alpha_n) \mathbf{x}_0 + \alpha_1 \mathbf{x}_1 + \dots + \alpha_n \mathbf{x}_n \\ &= \alpha_1 (\mathbf{x}_1 - \mathbf{x}_0) + \dots + \alpha_n (\mathbf{x}_n - \mathbf{x}_0). \end{aligned}$$

Since vectors  $\mathbf{x}_1 - \mathbf{x}_0, \dots, \mathbf{x}_n - \mathbf{x}_0$  are linearly independent, we see that  $\alpha_1 = \dots = \alpha_n = 0$ . Since  $\alpha_0 = -\alpha_1 - \dots - \alpha_n$ , it follows that  $\alpha_0 = 0$ . This proves (i).

Suppose now that (i) holds. Let us prove (iii). By symmetry, it suffices to show that  $\mathbf{x}_1 - \mathbf{x}_0, \dots, \mathbf{x}_n - \mathbf{x}_0$  are linearly independent. Fix scalars  $\alpha_1, \dots, \alpha_n \in \mathbb{F}$  such that  $\alpha_1 (\mathbf{x}_1 - \mathbf{x}_0) + \dots + \alpha_n (\mathbf{x}_n - \mathbf{x}_0) = \mathbf{0}$ . Then

$$\underbrace{(-\alpha_1 - \dots - \alpha_n)}_{:=\alpha_0} \mathbf{x}_0 + \alpha_1 \mathbf{x}_1 + \dots + \alpha_n \mathbf{x}_n = \mathbf{0}.$$

Since  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n$  are affinely independent, we now get that  $\alpha_0 = \alpha_1 = \dots = \alpha_n = 0$ , and we deduce that (iii) holds.  $\square$

### 3.3 Affine bases

Let  $M$  be an affine subspace of a vector space  $V$  over a field  $\mathbb{F}$ . An *affine basis* (also called a *barycentric frame*) of  $M$  is a non-empty ordered set  $\{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n\}$  of vectors in  $M$  such that

- vectors  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n$  are affinely independent;
- $M = \text{Aff}(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n)$ .

As we shall see (see Theorem 3.4 and the remark following it), for any non-negative integer  $n$ , every affine basis of an  $n$ -dimensional affine subspace contains exactly  $n + 1$  vectors.

**Remark:** Suppose that  $M$  is an affine subspace of a vector space  $V$  over a field  $\mathbb{F}$ , and let  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n$  be vectors in  $V$ . In view of Corollary 2.3, we have that  $\{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n\}$  is an affine basis of  $M$  if and only if all the following hold:

1.  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n \in M$ ;
2. vectors  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n$  are affinely independent;
3. every vector in  $M$  can be expressed as an affine combination of  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n$ .

Recall that by Theorem 1.3 of Lecture Notes 7, if  $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  is a basis of a vector space  $V$  over a field  $\mathbb{F}$ , then every vector in  $V$  can be written as a linear combination of the vectors  $\mathbf{v}_1, \dots, \mathbf{v}_n$  in a unique way. Theorem 3.1 was an analogue of this result for affine frames. Theorem 3.3 (below) is an analogue of that same result for affine bases.

**Theorem 3.3.** *Let  $M$  be an affine subspace of a vector space  $V$  over a field  $\mathbb{F}$ , and let  $\{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n\}$  be an affine basis of  $M$ . Then for all  $\mathbf{x} \in M$ , there exist unique scalars  $\alpha_0, \alpha_1, \dots, \alpha_n \in \mathbb{F}$ , called the barycentric coordinates of  $\mathbf{x}$  with respect to the affine basis  $\{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n\}$ , such that  $\mathbf{x} = \sum_{i=0}^n \alpha_i \mathbf{x}_i$  and  $\sum_{i=0}^n \alpha_i = 1$ .*

*Proof.* Fix  $\mathbf{x} \in M$ . The existence of scalars  $\alpha_0, \alpha_1, \dots, \alpha_n \in \mathbb{F}$  such that  $\mathbf{x} = \sum_{i=0}^n \alpha_i \mathbf{x}_i$  and  $\sum_{i=0}^n \alpha_i = 1$  follows from the fact that  $M = \text{Aff}(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n)$ .

It remains to prove uniqueness. So, fix  $\alpha_0, \alpha_1, \dots, \alpha_n, \beta_0, \beta_1, \dots, \beta_n \in \mathbb{F}$  such that

- $\mathbf{x} = \sum_{i=0}^n \alpha_i \mathbf{x}_i$  and  $\sum_{i=0}^n \alpha_i = 1$ ;
- $\mathbf{x} = \sum_{i=0}^n \beta_i \mathbf{x}_i$  and  $\sum_{i=0}^n \beta_i = 1$ .

Then  $\sum_{i=0}^n \alpha_i \mathbf{x}_i = \sum_{i=0}^n \beta_i \mathbf{x}_i$ , and we deduce that  $\sum_{i=0}^n (\alpha_i - \beta_i) \mathbf{x}_i = \mathbf{0}$ . On the other hand,

$$\sum_{i=0}^n (\alpha_i - \beta_i) = \left( \sum_{i=0}^n \alpha_i \right) - \left( \sum_{i=0}^n \beta_i \right) = 1 - 1 = 0.$$

Since vectors  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n$  are affinely independent, we now deduce that  $\alpha_0 - \beta_0 = \alpha_1 - \beta_1 = \dots = \alpha_n - \beta_n = 0$ . Therefore,  $\alpha_i = \beta_i$  for all  $i \in \{0, 1, \dots, n\}$ . This proves uniqueness.  $\square$

### 3.4 A relationship between affine bases and affine frames

**Theorem 3.4.** *Let  $V$  be a vector space over a field  $\mathbb{F}$ , let  $M$  be an affine subspace of  $V$ , and let  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n \in V$ . Then the following are equivalent:*

- (i)  $\{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n\}$  is an affine basis of  $M$ ;
- (ii)  $(\mathbf{x}_0, \mathbf{x}_1 - \mathbf{x}_0, \dots, \mathbf{x}_n - \mathbf{x}_0)$  is an affine frame of  $M$ .

**Remark:** Since every affine frame of an  $n$ -dimensional affine subspace contains  $n + 1$  vectors, Theorem 3.4 implies that every affine basis of an  $n$ -dimensional affine subspace contains exactly  $n + 1$  vectors.

*Proof.* Suppose first that (i) holds. Let us prove (ii). Since (i) holds, we know that  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n \in M$ . So, by Theorem 1.1(b),  $M$  can be written in the form  $M = \mathbf{x}_0 + U$  for some linear subspace  $U$  of  $V$ . It remains to show that  $\{\mathbf{x}_1 - \mathbf{x}_0, \dots, \mathbf{x}_n - \mathbf{x}_0\}$  is a basis of  $U$ . Since  $\mathbf{x}_1, \dots, \mathbf{x}_n \in M = \mathbf{x}_0 + U$ , it is clear that  $\mathbf{x}_1 - \mathbf{x}_0, \dots, \mathbf{x}_n - \mathbf{x}_0 \in U$ . Moreover, since the set  $\{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n\}$  is affinely independent (by (i)), Proposition 3.2 guarantees that  $\{\mathbf{x}_1 - \mathbf{x}_0, \dots, \mathbf{x}_n - \mathbf{x}_0\}$  is linearly independent.

It remains to show that any vector in  $U$  can be written as a linear combination of  $\mathbf{x}_1 - \mathbf{x}_0, \dots, \mathbf{x}_n - \mathbf{x}_0$ . Fix any  $\mathbf{u} \in U$ . Then  $\mathbf{x} := \mathbf{x}_0 + \mathbf{u} \in \mathbf{x}_0 + U = M$ , and so since (i) holds, we know that there exist scalars  $\alpha_0, \alpha_1, \dots, \alpha_n \in \mathbb{F}$  such that  $\mathbf{x} = \alpha_0 \mathbf{x}_0 + \alpha_1 \mathbf{x}_1 + \dots + \alpha_n \mathbf{x}_n$  and  $\alpha_0 + \alpha_1 + \dots + \alpha_n = 1$ . We now compute:

$$\begin{aligned} \mathbf{u} &= \mathbf{x} - \mathbf{x}_0 \\ &= \underbrace{(\alpha_0 \mathbf{x}_0 + \alpha_1 \mathbf{x}_1 + \dots + \alpha_n \mathbf{x}_n)}_{=\mathbf{x}} - \underbrace{(\alpha_0 + \alpha_1 + \dots + \alpha_n)}_{=1} \mathbf{x}_0 \\ &= \alpha_1 (\mathbf{x}_1 - \mathbf{x}_0) + \dots + \alpha_n (\mathbf{x}_n - \mathbf{x}_0). \end{aligned}$$

So,  $\mathbf{u}$  is indeed a linear combination of vectors  $\mathbf{x}_1 - \mathbf{x}_0, \dots, \mathbf{x}_n - \mathbf{x}_0$ . We now deduce that (ii) holds.

Suppose conversely that (ii) holds. Let us prove (i). Since (ii) holds, we have that  $M = \mathbf{x}_0 + U$ , where  $U = \text{Span}(\mathbf{x}_1 - \mathbf{x}_0, \dots, \mathbf{x}_n - \mathbf{x}_0)$ . Clearly,  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n \in M$ . Moreover, since the set  $\{\mathbf{x}_1 - \mathbf{x}_0, \dots, \mathbf{x}_n - \mathbf{x}_0\}$  is linearly independent (by (ii)), Proposition 3.2 guarantees that the set  $\{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n\}$  is affinely independent.

It remains to show that every vector in  $M$  is an affine combination of vectors  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n$ . Fix any  $\mathbf{x} \in M$ . Since  $M = \mathbf{x}_0 + U$ , there exists some  $\mathbf{u} \in U$  such that  $\mathbf{x} = \mathbf{x}_0 + \mathbf{u}$ . Since  $U = \text{Span}(\mathbf{x}_1 - \mathbf{x}_0, \dots, \mathbf{x}_n - \mathbf{x}_0)$ , there exist scalars  $\alpha_1, \dots, \alpha_n \in \mathbb{F}$  such that  $\mathbf{u} = \alpha_1 (\mathbf{x}_1 - \mathbf{x}_0) + \dots + \alpha_n (\mathbf{x}_n - \mathbf{x}_0)$ . Now, set  $\alpha_0 := 1 - \alpha_1 - \dots - \alpha_n$ ; then  $\alpha_0 + \alpha_1 + \dots + \alpha_n = 1$ , and we have the following:

$$\begin{aligned}
\mathbf{x} &= \mathbf{x}_0 + \mathbf{u} \\
&= \mathbf{x}_0 + \left( \alpha_1(\mathbf{x}_1 - \mathbf{x}_0) + \cdots + \alpha_n(\mathbf{x}_n - \mathbf{x}_0) \right) \\
&= (1 - \alpha_1 - \cdots - \alpha_n)\mathbf{x}_0 + \alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n \\
&= \alpha_0\mathbf{x}_0 + \alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n.
\end{aligned}$$

So,  $\mathbf{x}$  is indeed an affine combination of  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n$ . This proves that (i) holds.  $\square$

## 4 Affine transformations

Suppose that  $V_1$  and  $V_2$  are vector spaces over a field  $\mathbb{F}$ . A function  $f : V_1 \rightarrow V_2$  is called an *affine transformation* (or an *affine function*) if there exists a linear transformation  $g : V_1 \rightarrow V_2$  and a vector  $\mathbf{b} \in V_2$  such that for all  $\mathbf{x} \in V_1$ , we have that  $f(\mathbf{x}) = g(\mathbf{x}) + \mathbf{b}$ .

Obviously, every linear transformation  $f$  is affine (we simply take  $g := f$  and  $\mathbf{b} := \mathbf{0}$ ). Moreover, we have the following proposition.

**Proposition 4.1.** *Let  $V_1$  and  $V_2$  be vector spaces over a field  $\mathbb{F}$ , and let  $f : V_1 \rightarrow V_2$  be a linear transformation. Then  $f$  is linear if and only if  $f(\mathbf{0}) = \mathbf{0}$ .*

*Proof.* If  $f$  is linear, then Proposition 1.4 of Lecture Notes 8 guarantees that  $f(\mathbf{0}) = \mathbf{0}$ . For the reverse implication, we assume that  $f(\mathbf{0}) = \mathbf{0}$ , and we show that  $f$  is linear. Since  $f$  is an affine transformation, we know that there exists a linear transformation  $g : V_1 \rightarrow V_2$  and a vector  $\mathbf{b} \in V_2$  such that for all  $\mathbf{x} \in V_1$ , we have that  $f(\mathbf{x}) = g(\mathbf{x}) + \mathbf{b}$ . But now

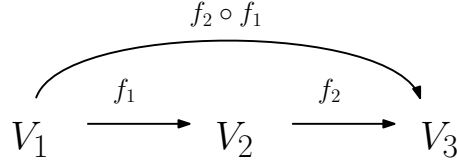
$$\mathbf{0} = f(\mathbf{0}) = g(\mathbf{0}) + \mathbf{b} \stackrel{(*)}{=} \mathbf{0} + \mathbf{b} = \mathbf{b}$$

where (\*) follows from the fact that  $g$  is linear, and so  $g(\mathbf{0}) = \mathbf{0}$  (by Proposition 1.4 of Lecture Notes 8). So,  $f(\mathbf{x}) = g(\mathbf{x})$  for all  $\mathbf{x} \in V_1$ , that is,  $f = g$ . Since the function  $g$  is linear, so is  $f$ .  $\square$

**Theorem 4.2.** *Let  $V_1, V_2, V_3$  be vector spaces over a field  $\mathbb{F}$ . Then all the following hold:*

- (a) *for all affine transformations  $f_1, f_2 : V_1 \rightarrow V_2$ , we have that  $f_1 + f_2$  is an affine transformation;*
- (b) *for all affine transformations  $f : V_1 \rightarrow V_2$  and scalars  $\alpha$ , we have that  $\alpha f$  is an affine transformation;*

(c) for all affine transformations  $f_1 : V_1 \rightarrow V_2$  and  $f_2 : V_2 \rightarrow V_3$ , we have that  $f_2 \circ f_1$  is an affine transformation.



*Proof.* We prove (c). The proofs of (a) and (b) are left as an exercise. Fix affine transformations  $f_1 : V_1 \rightarrow V_2$  and  $f_2 : V_2 \rightarrow V_3$ . Then there exists a linear transformation  $g_1 : V_1 \rightarrow V_2$  and a vector  $\mathbf{b}_2 \in V_2$  such that for all  $\mathbf{x} \in V_1$ , we have that  $f_1(\mathbf{x}) = g_1(\mathbf{x}) + \mathbf{b}_2$ . Similarly, since  $f_2 : V_2 \rightarrow V_3$  is an affine transformation, there exists a linear transformation  $g_2 : V_2 \rightarrow V_3$  and a vector  $\mathbf{b}_3 \in V_3$  such that for all  $\mathbf{x} \in V_2$ , we have that  $f_2(\mathbf{x}) = g_2(\mathbf{x}) + \mathbf{b}_3$ . But now for all  $\mathbf{x} \in V_1$ , we have that

$$\begin{aligned}
 (f_2 \circ f_1)(\mathbf{x}) &= f_2(f_1(\mathbf{x})) \\
 &= f_2(g_1(\mathbf{x}) + \mathbf{b}_2) \\
 &= g_2(g_1(\mathbf{x}) + \mathbf{b}_2) + \mathbf{b}_3 \\
 &= g_2(g_1(\mathbf{x})) + g_2(\mathbf{b}_2) + \mathbf{b}_3 && \text{because } g_2 \text{ is linear} \\
 &= (g_2 \circ g_1)(\mathbf{x}) + (g_2(\mathbf{b}_2) + \mathbf{b}_3).
 \end{aligned}$$

Since  $g_1$  and  $g_2$  are linear, Proposition 1.3(c) of Lecture Notes 8 implies that  $g_2 \circ g_1$  is linear. On the other hand,  $(g_2(\mathbf{b}_2) + \mathbf{b}_3)$  is a vector in  $V_3$ . So,  $f_2 \circ f_1$  is an affine transformation.  $\square$

We now need some terminology and notation. For any function  $f : A \rightarrow B$ , we define the following:

- for all  $A_1 \subseteq A$ , the *image* of  $A_1$  under  $f$  is the set

$$f[A_1] := \{f(a) \mid a \in A_1\};$$

- the *image* (or *range*) of  $f$  is the set  $\text{Im}(f) := f[A]$ ;

- for all  $B_1 \subseteq B$ , the *preimage* of  $B_1$  under  $f$  is the set

$$f^{-1}[B_1] := \{a \in A \mid f(a) \in B_1\}.$$

**Theorem 4.3.** Let  $V_1$  and  $V_2$  be vector spaces over a field  $\mathbb{F}$ , let  $g : V_1 \rightarrow V_2$  be a linear transformation, and let  $U_2$  be a linear subspace of  $V_2$ . Then  $g^{-1}[U_2]$  is a linear subspace of  $V_1$ .

*Proof.* By Theorem 2.7 of Lecture Notes 6, it suffices to show that  $\mathbf{0} \in g^{-1}[U]$ , and that  $g^{-1}[U]$  is closed under vector addition and scalar multiplication.

Since  $g$  is linear, we know that  $g(\mathbf{0}) = \mathbf{0}$ , and since  $U_2$  is a linear subspace of  $V_2$ , we know that  $\mathbf{0} \in U_2$ . So,  $g(\mathbf{0}) = \mathbf{0} \in U_2$ , and consequently,  $\mathbf{0} \in g^{-1}[U_2]$ .

Next, fix  $\mathbf{x}, \mathbf{y} \in g^{-1}[U_2]$ . Then  $g(\mathbf{x}), g(\mathbf{y}) \in U_2$ , and since  $U_2$  is a linear subspace of  $V_2$ , we have that  $g(\mathbf{x}) + g(\mathbf{y}) \in U_2$ . But since  $g$  is linear, we have that  $g(\mathbf{x}) + g(\mathbf{y}) = g(\mathbf{x} + \mathbf{y})$ , and so  $g(\mathbf{x} + \mathbf{y}) \in U_2$ . Consequently,  $\mathbf{x} + \mathbf{y} \in g^{-1}[U_2]$ .

Finally, fix  $\mathbf{x} \in g^{-1}[U_2]$  and  $\alpha \in \mathbb{F}$ . Then  $g(\mathbf{x}) \in U_2$ , and since  $U_2$  is a linear subspace of  $V_2$ , we have that  $\alpha g(\mathbf{x}) \in U_2$ . But since  $g$  is linear, we have that  $\alpha g(\mathbf{x}) = g(\alpha \mathbf{x})$ , and so  $g(\alpha \mathbf{x}) \in U_2$ . Consequently,  $\alpha \mathbf{x} \in g^{-1}[U_2]$ .

It now follows that  $g^{-1}[U_2]$  is a linear subspace of  $V_1$ .  $\square$

**Theorem 4.4.** *Let  $V_1$  and  $V_2$  be vector spaces over a field  $\mathbb{F}$ , and let  $f : V_1 \rightarrow V_2$  be an affine transformation. Then all the following hold:*

- (a) *for every affine subspace  $M_1$  of  $V_1$ , we have that  $f[M_1]$  is an affine subspace of  $V_2$ ;*
- (b)  *$Im(f)$  is an affine subspace of  $V_2$ ;*
- (c) *for every affine subspace  $M_2$  of  $V_2$ ,  $f^{-1}[M_2]$  is either empty or an affine subspace of  $V_1$ ;*
- (d) *for every  $\mathbf{b} \in V$ , the set of solutions of the equation  $f(\mathbf{x}) = \mathbf{b}$  is either empty or an affine subspace of  $V_1$ .*

*Proof.* Fix a linear transformation  $g : V_1 \rightarrow V_2$  and a vector  $\mathbf{b} \in V_2$  such that for all  $\mathbf{x} \in V_1$ , we have that  $f(\mathbf{x}) = g(\mathbf{x}) + \mathbf{b}$ .

We first prove (a). Fix an affine subspace  $M_1$  of  $V_1$ , and set  $M_1 := \mathbf{a}_1 + U_1$ , where  $\mathbf{a}_1$  is a vector and  $U_1$  is a linear subspace of  $V_1$ , as in the definition of an affine subspace. Then

$$\begin{aligned}
 f[M_1] &= \{f(\mathbf{x}) \mid \mathbf{x} \in M_1\} \\
 &= \{f(\mathbf{a}_1 + \mathbf{u}) \mid \mathbf{u} \in U_1\} && \text{because } M_1 = \mathbf{a}_1 + U_1 \\
 &= \{g(\mathbf{a}_1 + \mathbf{u}) + \mathbf{b} \mid \mathbf{u} \in U_1\} && \text{because } f(\mathbf{x}) = g(\mathbf{x}) + \mathbf{b} \\
 & && \text{for all } \mathbf{x} \in V_1 \\
 &= \{g(\mathbf{a}_1) + g(\mathbf{u}) + \mathbf{b} \mid \mathbf{u} \in U_1\} && \text{because } g \text{ is linear} \\
 &= (g(\mathbf{a}_1) + \mathbf{b}) + \{g(\mathbf{u}) \mid \mathbf{u} \in U_1\} \\
 &= (g(\mathbf{a}_1) + \mathbf{b}) + g[U_1].
 \end{aligned}$$

Since  $g : V_1 \rightarrow V_2$  is a linear transformation and  $U_1$  is a linear subspace of  $V_1$ , Theorem 1.6(b) of Lecture Notes 8 guarantees that  $g[U_1]$  is a linear subspace of  $V_2$ . So,  $f[M_1] = (g(\mathbf{a}_1) + \mathbf{b}) + g[U_1]$  is an affine subspace of  $V_2$ . This proves (a). Since  $V_1$  is an affine subspace of itself, (b) follows immediately from (a).

We now prove (c). Fix an affine subspace  $M_2$  of  $V_2$ . We may assume that  $f^{-1}[M_2] \neq \emptyset$ , for otherwise we are done. So, fix  $\mathbf{a}_1 \in f^{-1}[M_2]$ ; then  $f(\mathbf{a}_1) \in M_2$ . By Theorem 1.1, we have that  $M_2 = f(\mathbf{a}_1) + U_2$  for some linear subspace  $U_2$  of  $V_2$ . We now have the following:

$$\begin{aligned}
f^{-1}[M_2] &= \{\mathbf{x} \in V_1 \mid f(\mathbf{x}) \in M_2\} \\
&= \{\mathbf{x} \in V_1 \mid f(\mathbf{x}) \in f(\mathbf{a}_1) + U_2\} \\
&= \{\mathbf{x} \in V_1 \mid g(\mathbf{x}) + \mathbf{b} \in (g(\mathbf{a}_1) + \mathbf{b}) + U_2\} \\
&= \{\mathbf{x} \in V_1 \mid g(\mathbf{x}) \in g(\mathbf{a}_1) + U_2\} \\
&= \{\mathbf{x} \in V_1 \mid g(\mathbf{x}) - g(\mathbf{a}_1) \in U_2\} \\
&\stackrel{(*)}{=} \{\mathbf{x} \in V_1 \mid g(\mathbf{x} - \mathbf{a}_1) \in U_2\} \\
&\stackrel{(**)}{=} \mathbf{a}_1 + \{\mathbf{y} \in V_1 \mid g(\mathbf{y}) \in U_2\} \\
&= \mathbf{a}_1 + g^{-1}[U_2],
\end{aligned}$$

where (\*) follows from the fact that  $g$  is linear, and in (\*\*) we set  $\mathbf{y} = \mathbf{x} - \mathbf{a}_1$ . By Theorem 4.3,  $g^{-1}[U_2]$  is a linear subspace of  $V_1$ , and we deduce that  $f^{-1}[M_2] = \mathbf{a}_1 + g^{-1}[U_2]$  is an affine subspace of  $V_1$ .

For (d), we fix  $\mathbf{b} \in V_2$ , and we observe that the solution set of  $f(\mathbf{x}) = \mathbf{b}$  is precisely the set  $f^{-1}(\{\mathbf{b}\})$ . Since  $\{\mathbf{b}\}$  is an affine subspace of  $V_2$ , (c) guarantees that  $f^{-1}(\{\mathbf{b}\})$  is either empty or an affine subspace of  $V_1$ . This proves (d).  $\square$

**Corollary 4.5.** *Let  $\mathbb{F}$  be a field, and let  $A \in \mathbb{F}^{n \times m}$  and  $\mathbf{b} \in \mathbb{F}^n$ . Then the solution set of the matrix-vector equation  $A\mathbf{x} = \mathbf{b}$  is either empty or an affine subspace of  $\mathbb{F}^m$ .*

*Proof.* Let  $f_A : \mathbb{F}^m \rightarrow \mathbb{F}^n$  be given by  $f_A(\mathbf{x}) = A\mathbf{x}$  for all  $\mathbf{x} \in \mathbb{F}^m$ . Then  $f_A$  is a linear (and therefore affine) transformation. Moreover, the solution set of the matrix-vector equation  $A\mathbf{x} = \mathbf{b}$  is precisely the solution set of the equation  $f_A(\mathbf{x}) = \mathbf{b}$ . The result now follows from Theorem 4.4(d).  $\square$