Chapter 2
Basics of Affine Geometry

L'algèbre n'est qu'une géométrie écrite; la géométrie n'est qu'une algèbre figurée.
—Sophie Germain

2.1 Affine Spaces

Geometrically, curves and surfaces are usually considered to be sets of points with some special properties, living in a space consisting of "points." Typically, one is also interested in geometric properties invariant under certain transformations, for example, translations, rotations, projections, etc. One could model the space of points as a vector space, but this is not very satisfactory for a number of reasons. One reason is that the point corresponding to the zero vector (0), called the origin, plays a special role, when there is really no reason to have a privileged origin. Another reason is that certain notions, such as parallelism, are handled in an awkward manner. But the deeper reason is that vector spaces and affine spaces really have different geometries. The geometric properties of a vector space are invariant under the group of bijective linear maps, whereas the geometric properties of an affine space are invariant under the group of bijective affine maps, and these two groups are not isomorphic. Roughly speaking, there are more affine maps than linear maps.

Affine spaces provide a better framework for doing geometry. In particular, it is possible to deal with points, curves, surfaces, etc., in an **intrinsic manner**, that is, independently of any specific choice of a coordinate system. As in physics, this is highly desirable to really understand what is going on. Of course, coordinate systems have to be chosen to finally carry out computations, but one should learn to resist the temptation to resort to coordinate systems until it is really necessary.

Affine spaces are the right framework for dealing with motions, trajectories, and physical forces, among other things. Thus, affine geometry is crucial to a clean presentation of kinematics, dynamics, and other parts of physics (for example, elasticity). After all, a rigid motion is an affine map, but not a linear map in general.

J. Gallier, *Geometric Methods and Applications, For Computer Science and Engineering*,
Texts in Applied Mathematics 38, DOI 10.1007/978-1-4419-9961-0_2,

Also, given an $m \times n$ matrix A and a vector $b \in \mathbb{R}^m$, the set $U = \{x \in \mathbb{R}^n \mid Ax = b\}$ of solutions of the system $Ax = b$ is an affine space, but not a vector space (linear space) in general.

Use coordinate systems only when needed!

This chapter proceeds as follows. We take advantage of the fact that almost every affine concept is the counterpart of some concept in linear algebra. We begin by defining affine spaces, stressing the physical interpretation of the definition in terms of points (particles) and vectors (forces). Corresponding to linear combinations of vectors, we define affine combinations of points (barycenters), realizing that we are forced to restrict our attention to families of scalars adding up to 1. Corresponding to linear subspaces, we introduce affine subspaces as subsets closed under affine combinations. Then, we characterize affine subspaces in terms of certain vector spaces called their directions. This allows us to define a clean notion of parallelism. Next, corresponding to linear independence and bases, we define affine independence and affine frames. We also define convexity. Corresponding to linear maps, we define affine maps as maps preserving affine combinations. We show that every affine map is completely defined by the image of one point and a linear map. Then, we investigate briefly some simple affine maps, the translations and the central dilatations. At this point, we give a glimpse of affine geometry. We prove the theorems of Thales, Pappus, and Desargues. After this, the definition of affine hyperplanes in terms of affine forms is reviewed. The section ends with a closer look at the intersection of affine subspaces.

Our presentation of affine geometry is far from being comprehensive, and it is biased toward the algorithmic geometry of curves and surfaces. For more details, the reader is referred to Pedoe [9], Snapper and Troyer [11], Berger [2, 3], Coxeter [4], Samuel [10], Tisseron [13], and Hilbert and Cohn-Vossen [7].

Suppose we have a particle moving in 3D space and that we want to describe the trajectory of this particle. If one looks up a good textbook on dynamics, such as Greenwood [6], one finds out that the particle is modeled as a point, and that the position of this point x is determined with respect to a "frame" in $\mathbb{R}^3$ by a vector. Curiously, the notion of a frame is rarely defined precisely, but it is easy to infer that a frame is a pair $(O, (e_1, e_2, e_3))$ consisting of an origin O (which is a point) together with a basis of three vectors (e_1, e_2, e_3). For example, the standard frame in $\mathbb{R}^3$ has origin $O = (0,0,0)$ and the basis of three vectors $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, and $e_3 = (0,0,1)$. The position of a point x is then defined by the "unique vector" from O to x.

But wait a minute, this definition seems to be defining frames and the position of a point without defining what a point is! Well, let us identify points with elements of $\mathbb{R}^3$. If so, given any two points $a = (a_1, a_2, a_3)$ and $b = (b_1, b_2, b_3)$, there is a unique *free vector*, denoted by $\overrightarrow{ab}$, from a to b, the vector $\overrightarrow{ab} = (b_1 - a_1, b_2 - a_2, b_3 - a_3)$. Note that

$$b = a + \overrightarrow{ab},$$

addition being understood as addition in $\mathbb{R}^3$. Then, in the standard frame, given a point $x = (x_1, x_2, x_3)$, the position of x is the vector $\overrightarrow{Ox} = (x_1, x_2, x_3)$, which coincides with the point itself. In the standard frame, points and vectors are identified. Points and free vectors are illustrated in Figure 2.1.

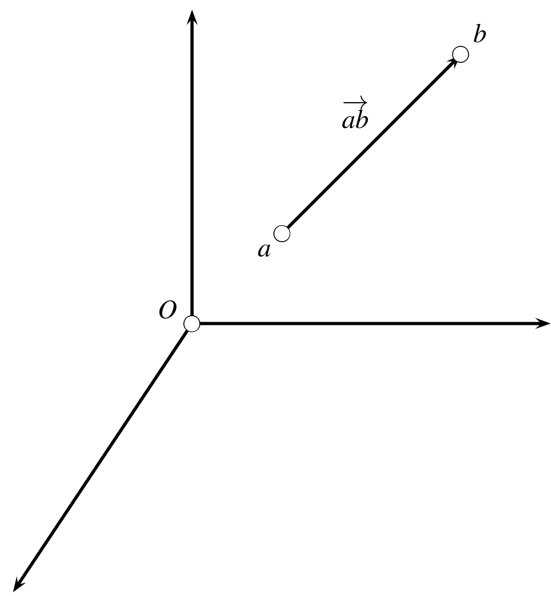

Fig. 2.1 Points and free vectors.

What if we pick a frame with a different origin, say $\Omega = (\omega_1, \omega_2, \omega_3)$, but the same basis vectors (e_1, e_2, e_3)? This time, the point $x = (x_1, x_2, x_3)$ is defined by two position vectors:

$$\overrightarrow{Ox} = (x_1, x_2, x_3)$$

in the frame $(O, (e_1, e_2, e_3))$ and

$$\overrightarrow{\Omega x} = (x_1 - \omega_1, x_2 - \omega_2, x_3 - \omega_3)$$

in the frame $(\Omega, (e_1, e_2, e_3))$.

This is because

$$\overrightarrow{Ox} = \overrightarrow{O\Omega} + \overrightarrow{\Omega x} \quad \text{and} \quad \overrightarrow{O\Omega} = (\omega_1, \omega_2, \omega_3).$$

We note that in the second frame $(\Omega, (e_1, e_2, e_3))$, points and position vectors are no longer identified. This gives us evidence that points are not vectors. It may be computationally convenient to deal with points using position vectors, but such a treatment is not frame invariant, which has undesirable effets.

Inspired by physics, we deem it important to define points and properties of points that are frame invariant. An undesirable side effect of the present approach shows up if we attempt to define linear combinations of points. First, let us review the notion of linear combination of vectors. Given two vectors u and v of coordinates (u_1,u_2,u_3) and (v_1,v_2,v_3) with respect to the basis (e_1,e_2,e_3), for any two scalars λ,μ, we can define the linear combination $\lambda u+\mu v$ as the vector of coordinates

$$(\lambda u_1+\mu v_1,\lambda u_2+\mu v_2,\lambda u_3+\mu v_3).$$

If we choose a different basis (e'_1,e'_2,e'_3) and if the matrix P expressing the vectors (e'_1,e'_2,e'_3) over the basis (e_1,e_2,e_3) is

$$P=\begin{pmatrix} a_1 & b_1 & c_1\\ a_2 & b_2 & c_2\\ a_3 & b_3 & c_3\end{pmatrix},$$

which means that the columns of P are the coordinates of the e'_j over the basis (e_1,e_2,e_3), since

$$u_1e_1+u_2e_2+u_3e_3=u'_1e'_1+u'_2e'_2+u'_3e'_3$$

and

$$v_1e_1+v_2e_2+v_3e_3=v'_1e'_1+v'_2e'_2+v'_3e'_3,$$

it is easy to see that the coordinates (u_1,u_2,u_3) and (v_1,v_2,v_3) of u and v with respect to the basis (e_1,e_2,e_3) are given in terms of the coordinates (u'_1,u'_2,u'_3) and (v'_1,v'_2,v'_3) of u and v with respect to the basis (e'_1,e'_2,e'_3) by the matrix equations

$$\begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix}=P\begin{pmatrix}u'_1\\u'_2\\u'_3\end{pmatrix}\quad\text{and}\quad\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}=P\begin{pmatrix}v'_1\\v'_2\\v'_3\end{pmatrix}.$$

From the above, we get

$$\begin{pmatrix}u'_1\\u'_2\\u'_3\end{pmatrix}=P^{-1}\begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix}\quad\text{and}\quad\begin{pmatrix}v'_1\\v'_2\\v'_3\end{pmatrix}=P^{-1}\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix},$$

and by linearity, the coordinates

$$(\lambda u'_1+\mu v'_1,\lambda u'_2+\mu v'_2,\lambda u'_3+\mu v'_3)$$

of $\lambda u+\mu v$ with respect to the basis (e'_1,e'_2,e'_3) are given by

$$\begin{pmatrix}\lambda u'_1+\mu v'_1\\ \lambda u'_2+\mu v'_2\\ \lambda u'_3+\mu v'_3\end{pmatrix}=\lambda P^{-1}\begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix}+\mu P^{-1}\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}=P^{-1}\begin{pmatrix}\lambda u_1+\mu v_1\\ \lambda u_2+\mu v_2\\ \lambda u_3+\mu v_3\end{pmatrix}.$$

Everything worked out because the change of basis does not involve a change of origin. On the other hand, if we consider the change of frame from the frame $(O,(e_1,e_2,e_3))$ to the frame $(\Omega,(e_1,e_2,e_3))$, where $\overrightarrow{O\Omega}=(\omega_1,\omega_2,\omega_3)$, given two points a, b of coordinates (a_1,a_2,a_3) and (b_1,b_2,b_3) with respect to the frame

$(O,(e_1,e_2,e_3))$ and of coordinates (a'_1,a'_2,a'_3) and (b'_1,b'_2,b'_3) with respect to the frame $(\Omega,(e_1,e_2,e_3))$, since

$$(a'_1,a'_2,a'_3) = (a_1-\omega_1, a_2-\omega_2, a_3-\omega_3)$$

and

$$(b'_1,b'_2,b'_3) = (b_1-\omega_1, b_2-\omega_2, b_3-\omega_3),$$

the coordinates of $\lambda a+\mu b$ with respect to the frame $(O,(e_1,e_2,e_3))$ are

$$(\lambda a_1+\mu b_1, \lambda a_2+\mu b_2, \lambda a_3+\mu b_3),$$

but the coordinates

$$(\lambda a'_1+\mu b'_1, \lambda a'_2+\mu b'_2, \lambda a'_3+\mu b'_3)$$

of $\lambda a+\mu b$ with respect to the frame $(\Omega,(e_1,e_2,e_3))$ are

$$(\lambda a_1+\mu b_1-(\lambda+\mu)\omega_1, \lambda a_2+\mu b_2-(\lambda+\mu)\omega_2, \lambda a_3+\mu b_3-(\lambda+\mu)\omega_3),$$

which are different from

$$(\lambda a_1+\mu b_1-\omega_1, \lambda a_2+\mu b_2-\omega_2, \lambda a_3+\mu b_3-\omega_3),$$

unless $\lambda+\mu=1$.

Thus, we have discovered a major difference between vectors and points: The notion of linear combination of vectors is basis independent, but the notion of linear combination of points is frame dependent. In order to salvage the notion of linear combination of points, some restriction is needed: The scalar coefficients must add up to 1.

A clean way to handle the problem of frame invariance and to deal with points in a more intrinsic manner is to make a clearer distinction between points and vectors. We duplicate $\mathbb{R}^3$ into two copies, the first copy corresponding to points, where we forget the vector space structure, and the second copy corresponding to free vectors, where the vector space structure is important. Furthermore, we make explicit the important fact that the vector space $\mathbb{R}^3$ acts on the set of points $\mathbb{R}^3$: Given any **point** $a=(a_1,a_2,a_3)$ and any **vector** $v=(v_1,v_2,v_3)$, we obtain the **point**

$$a+v=(a_1+v_1, a_2+v_2, a_3+v_3),$$

which can be thought of as the result of translating a to b using the vector v. We can imagine that v is placed such that its origin coincides with a and that its tip coincides with b. This action $+\colon \mathbb{R}^3\times\mathbb{R}^3\to\mathbb{R}^3$ satisfies some crucial properties. For example,

$$a+0=a,$$
$$(a+u)+v=a+(u+v),$$

and for any two points a, b, there is a unique free vector $\overrightarrow{ab}$ such that

$$b = a + \overrightarrow{ab}.$$

It turns out that the above properties, although trivial in the case of $\mathbb{R}^3$, are all that is needed to define the abstract notion of affine space (or affine structure). The basic idea is to consider two (distinct) sets E and $\overrightarrow{E}$, where E is a set of points (with no structure) and $\overrightarrow{E}$ is a vector space (of free vectors) acting on the set E.

Did you say "A fine space"?

Intuitively, we can think of the elements of $\overrightarrow{E}$ as forces moving the points in E, considered as physical particles. The effect of applying a force (free vector) $u \in \overrightarrow{E}$ to a point $a \in E$ is a translation. By this, we mean that for every force $u \in \overrightarrow{E}$, the action of the force u is to "move" every point $a \in E$ to the point $a + u \in E$ obtained by the translation corresponding to u viewed as a vector. Since translations can be composed, it is natural that $\overrightarrow{E}$ is a vector space.

For simplicity, it is assumed that all vector spaces under consideration are defined over the field $\mathbb{R}$ of real numbers. Most of the definitions and results also hold for an arbitrary field K, although some care is needed when dealing with fields of characteristic different from zero (see the problems). It is also assumed that all families $(\lambda_i)_{i \in I}$ of scalars have finite support. Recall that a family $(\lambda_i)_{i \in I}$ of scalars has *finite support* if $\lambda_i = 0$ for all $i \in I - J$, where J is a finite subset of I. Obviously, finite families of scalars have finite support, and for simplicity, the reader may assume that all families of scalars are finite. The formal definition of an affine space is as follows.

Definition 2.1. An *affine space* is either the degenerate space reduced to the empty set, or a triple $\langle E, \overrightarrow{E}, + \rangle$ consisting of a nonempty set E (of *points*), a vector space $\overrightarrow{E}$ (of *translations*, or *free vectors*), and an action $+\colon E \times \overrightarrow{E} \to E$, satisfying the following conditions.

(A1) $a + 0 = a$, for every $a \in E$.

(A2) $(a + u) + v = a + (u + v)$, for every $a \in E$, and every $u, v \in \overrightarrow{E}$.

(A3) For any two points $a, b \in E$, there is a unique $u \in \overrightarrow{E}$ such that $a + u = b$.

The unique vector $u \in \overrightarrow{E}$ such that $a + u = b$ is denoted by $\overrightarrow{ab}$, or sometimes by **ab**, or even by $b - a$. Thus, we also write

$$b = a + \overrightarrow{ab}$$

(or $b = a + \mathbf{ab}$, or even $b = a + (b - a)$).

The *dimension of the affine space* $\langle E, \overrightarrow{E}, + \rangle$ is the dimension $\dim(\overrightarrow{E})$ of the vector space $\overrightarrow{E}$. For simplicity, it is denoted by $\dim(E)$.

Conditions (A1) and (A2) say that the (abelian) group $\overrightarrow{E}$ acts on E, and condition (A3) says that $\overrightarrow{E}$ acts transitively and faithfully on E. Note that

$$\overrightarrow{a(a+v)} = v$$

for all $a \in E$ and all $v \in \overrightarrow{E}$, since $\overrightarrow{a(a+v)}$ is the unique vector such that $a+v = a + \overrightarrow{a(a+v)}$. Thus, $b = a+v$ is equivalent to $\overrightarrow{ab} = v$. Figure 2.2 gives an intuitive picture of an affine space. It is natural to think of all vectors as having the same origin, the null vector.

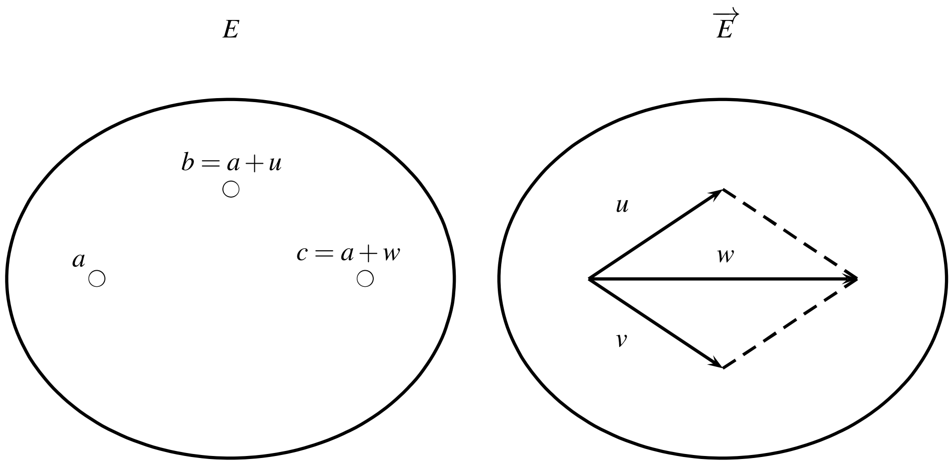

Fig. 2.2 Intuitive picture of an affine space.

The axioms defining an affine space $\langle E, \overrightarrow{E}, +\rangle$ can be interpreted intuitively as saying that E and $\overrightarrow{E}$ are two different ways of looking at the same object, but wearing different sets of glasses, the second set of glasses depending on the choice of an "origin" in E. Indeed, we can choose to look at the points in E, forgetting that every pair (a,b) of points defines a unique vector $\overrightarrow{ab}$ in $\overrightarrow{E}$, or we can choose to look at the vectors u in $\overrightarrow{E}$, forgetting the points in E. Furthermore, if we also pick any point a in E, a point that can be viewed as an *origin* in E, then we can recover all the points in E as the translated points $a+u$ for all $u \in \overrightarrow{E}$. This can be formalized by defining two maps between E and $\overrightarrow{E}$.

For every $a \in E$, consider the mapping from $\overrightarrow{E}$ to E given by

$$u \mapsto a+u,$$

where $u \in \overrightarrow{E}$, and consider the mapping from E to $\overrightarrow{E}$ given by

$$b \mapsto \overrightarrow{ab},$$

where $b \in E$. The composition of the first mapping with the second is

$$u \mapsto a + u \mapsto \overrightarrow{a(a+u)},$$

which, in view of (A3), yields u. The composition of the second with the first mapping is

$$b \mapsto \overrightarrow{ab} \mapsto a + \overrightarrow{ab},$$

which, in view of (A3), yields b. Thus, these compositions are the identity from $\overrightarrow{E}$ to $\overrightarrow{E}$ and the identity from E to E, and the mappings are both bijections.

When we identify E with $\overrightarrow{E}$ via the mapping $b \mapsto \overrightarrow{ab}$, we say that we consider E as the vector space obtained *by taking a as the origin in E*, and we denote it by E_a. Because E_a is a vector space, to be consistent with our notational conventions we should use the notation $\overrightarrow{E_a}$ (using an arrow), instead of E_a. However, for simplicity, we stick to the notation E_a.

Thus, an affine space $\langle E, \overrightarrow{E}, + \rangle$ is a way of defining a vector space structure on a set of points E, without making a commitment to a **fixed** origin in E. Nevertheless, as soon as we commit to an origin a in E, we can view E as the vector space E_a. However, we urge the reader to think of E as a physical set of points and of $\overrightarrow{E}$ as a set of forces acting on E, rather than reducing E to some isomorphic copy of $\mathbb{R}^n$. After all, points are points, and not vectors! For notational simplicity, we will often denote an affine space $\langle E, \overrightarrow{E}, + \rangle$ by $(E, \overrightarrow{E})$, or even by E. The vector space $\overrightarrow{E}$ is called the *vector space associated with E*.

One should be careful about the overloading of the addition symbol $+$. Addition is well-defined on vectors, as in $u + v$; the translate $a + u$ of a point $a \in E$ by a vector $u \in \overrightarrow{E}$ is also well-defined, but addition of points $a + b$ **does not make sense**. In this respect, the notation $b - a$ for the unique vector u such that $b = a + u$ is somewhat confusing, since it suggests that points can be subtracted (but not added!). Yet, we will see in Section 4.1 that it is possible to make sense of linear combinations of points, and even mixed linear combinations of points and vectors.

Any vector space $\overrightarrow{E}$ has an affine space structure specified by choosing $E = \overrightarrow{E}$, and letting $+$ be addition in the vector space $\overrightarrow{E}$. We will refer to the affine structure $\langle \overrightarrow{E}, \overrightarrow{E}, + \rangle$ on a vector space $\overrightarrow{E}$ as the *canonical (or natural) affine structure on $\overrightarrow{E}$*. In particular, the vector space $\mathbb{R}^n$ can be viewed as the affine space $\langle \mathbb{R}^n, \mathbb{R}^n, + \rangle$, denoted by $\mathbb{A}^n$. In general, if K is any field, the affine space $\langle K^n, K^n, + \rangle$ is denoted by $\mathbb{A}^n_K$. In order to distinguish between the double role played by members of $\mathbb{R}^n$, points and vectors, we will denote points by row vectors, and vectors by column vectors. Thus, the action of the vector space $\mathbb{R}^n$ over the set $\mathbb{R}^n$ simply viewed as a set of points is given by

$$(a_1,\ldots,a_n)+\begin{pmatrix}u_1\\ \vdots\\ u_n\end{pmatrix}=(a_1+u_1,\ldots,a_n+u_n).$$

We will also use the convention that if $x=(x_1,\ldots,x_n)\in\mathbb{R}^n$, then the column vector associated with x is denoted by $\mathbf{x}$ (in boldface notation). Abusing the notation slightly, if $a\in\mathbb{R}^n$ is a point, we also write $a\in\mathbb{A}^n$. The affine space $\mathbb{A}^n$ is called the *real affine space of dimension n*. In most cases, we will consider $n=1,2,3$.

2.2 Examples of Affine Spaces

Let us now give an example of an affine space that is not given as a vector space (at least, not in an obvious fashion). Consider the subset L of $\mathbb{A}^2$ consisting of all points (x,y) satisfying the equation

$$x+y-1=0.$$

The set L is the line of slope -1 passing through the points $(1,0)$ and $(0,1)$ shown in Figure 2.3.

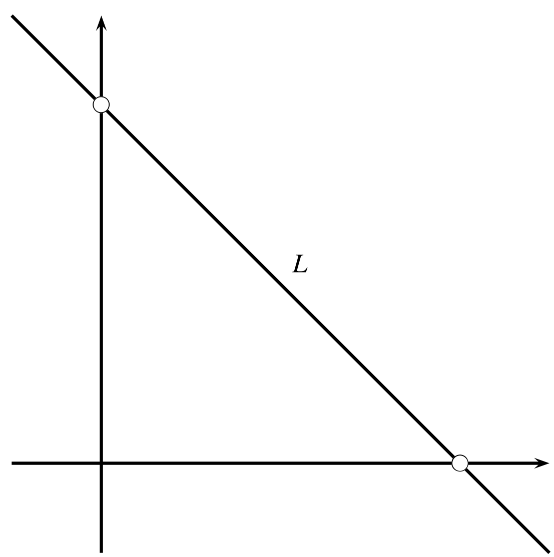

Fig. 2.3 An affine space: the line of equation $x+y-1=0$.

The line L can be made into an official affine space by defining the action $+\colon L\times\mathbb{R}\to L$ of $\mathbb{R}$ on L defined such that for every point $(x,1-x)$ on L and any $u\in\mathbb{R}$,

$$(x,1-x)+u=(x+u,1-x-u).$$

It is immediately verified that this action makes L into an affine space. For example, for any two points $a = (a_1, 1 - a_1)$ and $b = (b_1, 1 - b_1)$ on L, the unique (vector) $u \in \mathbb{R}$ such that $b = a + u$ is $u = b_1 - a_1$. Note that the vector space $\mathbb{R}$ is isomorphic to the line of equation $x + y = 0$ passing through the origin.

Similarly, consider the subset H of $\mathbb{A}^3$ consisting of all points (x, y, z) satisfying the equation

$$x + y + z - 1 = 0.$$

The set H is the plane passing through the points $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$. The plane H can be made into an official affine space by defining the action $+\colon H \times \mathbb{R}^2 \to H$ of $\mathbb{R}^2$ on H defined such that for every point $(x, y, 1 - x - y)$ on H and any $\begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{R}^2$,

$$(x, y, 1 - x - y) + \begin{pmatrix} u \\ v \end{pmatrix} = (x + u, y + v, 1 - x - u - y - v).$$

For a slightly wilder example, consider the subset P of $\mathbb{A}^3$ consisting of all points (x, y, z) satisfying the equation

$$x^2 + y^2 - z = 0.$$

The set P is a paraboloid of revolution, with axis Oz. The surface P can be made into an official affine space by defining the action $+\colon P \times \mathbb{R}^2 \to P$ of $\mathbb{R}^2$ on P defined such that for every point $(x, y, x^2 + y^2)$ on P and any $\begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{R}^2$,

$$(x, y, x^2 + y^2) + \begin{pmatrix} u \\ v \end{pmatrix} = (x + u, y + v, (x + u)^2 + (y + v)^2).$$

This should dispell any idea that affine spaces are dull. Affine spaces not already equipped with an obvious vector space structure arise in projective geometry. Indeed, we will see in Section 5.1 that the complement of a hyperplane in a projective space has an affine structure.

2.3 Chasles's Identity

Given any three points $a, b, c \in E$, since $c = a + \overrightarrow{ac}$, $b = a + \overrightarrow{ab}$, and $c = b + \overrightarrow{bc}$, we get

$$c = b + \overrightarrow{bc} = (a + \overrightarrow{ab}) + \overrightarrow{bc} = a + (\overrightarrow{ab} + \overrightarrow{bc})$$

by (A2), and thus, by (A3),

$$\overrightarrow{ab} + \overrightarrow{bc} = \overrightarrow{ac},$$

which is known as *Chasles's identity*, and illustrated in Figure 2.4.

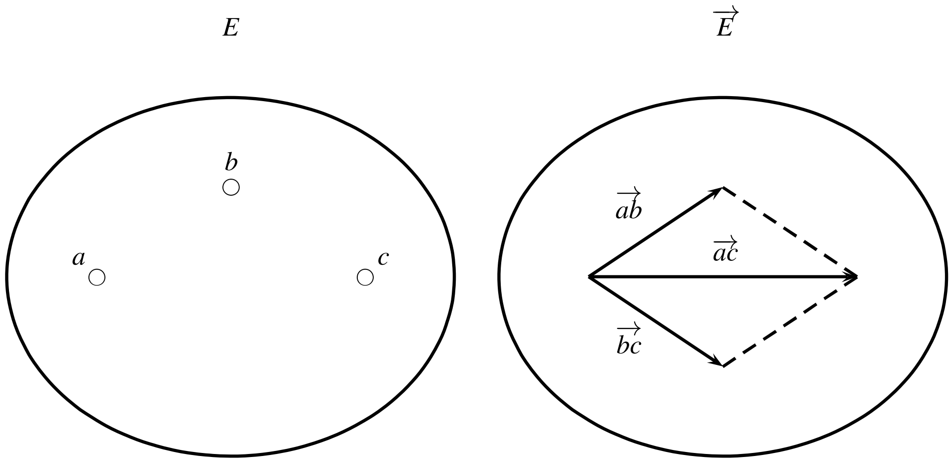

Fig. 2.4 Points and corresponding vectors in affine geometry.

Since $a = a + \overrightarrow{aa}$ and by (A1) $a = a + 0$, by (A3) we get

$$\overrightarrow{aa} = 0.$$

Thus, letting $a = c$ in Chasles's identity, we get

$$\overrightarrow{ba} = -\overrightarrow{ab}.$$

Given any four points $a, b, c, d \in E$, since by Chasles's identity

$$\overrightarrow{ab} + \overrightarrow{bc} = \overrightarrow{ad} + \overrightarrow{dc} = \overrightarrow{ac},$$

we have the *parallelogram law*

$$\overrightarrow{ab} = \overrightarrow{dc} \quad \text{iff} \quad \overrightarrow{bc} = \overrightarrow{ad}.$$

2.4 Affine Combinations, Barycenters

A fundamental concept in linear algebra is that of a linear combination. The corresponding concept in affine geometry is that of an *affine combination*, also called a *barycenter*. However, there is a problem with the naive approach involving a coordinate system, as we saw in Section 2.1. Since this problem is the reason for introducing affine combinations, at the risk of boring certain readers, we give another example showing what goes wrong if we are not careful in defining linear combinations of points.

Consider $\mathbb{R}^2$ as an affine space, under its natural coordinate system with origin $O = (0,0)$ and basis vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Given any two points $a = (a_1, a_2)$ and

$b = (b_1, b_2)$, it is natural to define the affine combination $\lambda a + \mu b$ as the point of coordinates

$$(\lambda a_1 + \mu b_1, \lambda a_2 + \mu b_2).$$

Thus, when $a = (-1,-1)$ and $b = (2,2)$, the point $a+b$ is the point $c = (1,1)$.

Let us now consider the new coordinate system with respect to the origin $c = (1,1)$ (and the same basis vectors). This time, the coordinates of a are $(-2,-2)$, the coordinates of b are $(1,1)$, and the point $a+b$ is the point d of coordinates $(-1,-1)$. However, it is clear that the point d is identical to the origin $O = (0,0)$ of the first coordinate system.

Thus, $a+b$ corresponds to two different points depending on which coordinate system is used for its computation!

This shows that some extra condition is needed in order for affine combinations to make sense. It turns out that if the scalars sum up to 1, the definition is intrinsic, as the following lemma shows.

Lemma 2.1. *Given an affine space E, let $(a_i)_{i\in I}$ be a family of points in E, and let $(\lambda_i)_{i\in I}$ be a family of scalars. For any two points $a, b \in E$, the following properties hold:*

(1) If $\sum_{i\in I} \lambda_i = 1$, then

$$a + \sum_{i\in I} \lambda_i \overrightarrow{aa_i} = b + \sum_{i\in I} \lambda_i \overrightarrow{ba_i}.$$

(2) If $\sum_{i\in I} \lambda_i = 0$, then

$$\sum_{i\in I} \lambda_i \overrightarrow{aa_i} = \sum_{i\in I} \lambda_i \overrightarrow{ba_i}.$$

Proof. (1) By Chasles's identity (see Section 2.3), we have

$$\begin{aligned}
a + \sum_{i\in I} \lambda_i \overrightarrow{aa_i} &= a + \sum_{i\in I} \lambda_i (\overrightarrow{ab} + \overrightarrow{ba_i}) \\
&= a + \left(\sum_{i\in I} \lambda_i\right) \overrightarrow{ab} + \sum_{i\in I} \lambda_i \overrightarrow{ba_i} \\
&= a + \overrightarrow{ab} + \sum_{i\in I} \lambda_i \overrightarrow{ba_i} && \text{since } \textstyle\sum_{i\in I} \lambda_i = 1 \\
&= b + \sum_{i\in I} \lambda_i \overrightarrow{ba_i} && \text{since } b = a + \overrightarrow{ab}.
\end{aligned}$$

(2) We also have

$$\begin{aligned}
\sum_{i\in I} \lambda_i \overrightarrow{aa_i} &= \sum_{i\in I} \lambda_i (\overrightarrow{ab} + \overrightarrow{ba_i}) \\
&= \left(\sum_{i\in I} \lambda_i\right) \overrightarrow{ab} + \sum_{i\in I} \lambda_i \overrightarrow{ba_i} \\
&= \sum_{i\in I} \lambda_i \overrightarrow{ba_i},
\end{aligned}$$

since $\sum_{i\in I}\lambda_i = 0$. $\square$

Thus, by Lemma 2.1, for any family of points $(a_i)_{i\in I}$ in E, for any family $(\lambda_i)_{i\in I}$ of scalars such that $\sum_{i\in I}\lambda_i = 1$, the point

$$x = a + \sum_{i\in I}\lambda_i\overrightarrow{aa_i}$$

is independent of the choice of the origin $a \in E$. This property motivates the following definition.

Definition 2.2. For any family of points $(a_i)_{i\in I}$ in E, for any family $(\lambda_i)_{i\in I}$ of scalars such that $\sum_{i\in I}\lambda_i = 1$, and for any $a \in E$, the point

$$a + \sum_{i\in I}\lambda_i\overrightarrow{aa_i}$$

(which is independent of $a \in E$, by Lemma 2.1) is called the *barycenter (or barycentric combination, or affine combination) of the points a_i assigned the weights λ_i*, and it is denoted by

$$\sum_{i\in I}\lambda_i a_i.$$

In dealing with barycenters, it is convenient to introduce the notion of a *weighted point*, which is just a pair (a,λ), where $a \in E$ is a point, and $\lambda \in \mathbb{R}$ is a scalar. Then, given a family of weighted points $((a_i,\lambda_i))_{i\in I}$, where $\sum_{i\in I}\lambda_i = 1$, we also say that the point $\sum_{i\in I}\lambda_i a_i$ is the *barycenter of the family of weighted points* $((a_i,\lambda_i))_{i\in I}$.

Note that the barycenter x of the family of weighted points $((a_i,\lambda_i))_{i\in I}$ is the unique point such that

$$\overrightarrow{ax} = \sum_{i\in I}\lambda_i\overrightarrow{aa_i} \quad \text{for every } a \in E,$$

and setting $a = x$, the point x is the unique point such that

$$\sum_{i\in I}\lambda_i\overrightarrow{xa_i} = 0.$$

In physical terms, the barycenter is the *center of mass* of the family of weighted points $((a_i,\lambda_i))_{i\in I}$ (where the masses have been normalized, so that $\sum_{i\in I}\lambda_i = 1$, and negative masses are allowed).

Remarks:

(1) Since the barycenter of a family $((a_i,\lambda_i))_{i\in I}$ of weighted points is defined for families $(\lambda_i)_{i\in I}$ of scalars with finite support (and such that $\sum_{i\in I}\lambda_i = 1$), we might as well assume that I is finite. Then, for all $m \geq 2$, it is easy to prove that the barycenter of m weighted points can be obtained by repeated computations of barycenters of two weighted points.

(2) This result still holds, provided that the field K has at least three distinct elements, but the proof is trickier!
(3) When $\sum_{i\in I}\lambda_i = 0$, the vector $\sum_{i\in I}\lambda_i\overrightarrow{aa_i}$ does not depend on the point a, and we may denote it by $\sum_{i\in I}\lambda_i a_i$. This observation will be used in Section 4.1 to define a vector space in which linear combinations of both points and vectors make sense, regardless of the value of $\sum_{i\in I}\lambda_i$.

Figure 2.5 illustrates the geometric construction of the barycenters g_1 and g_2 of the weighted points $\left(a,\frac{1}{4}\right)$, $\left(b,\frac{1}{4}\right)$, and $\left(c,\frac{1}{2}\right)$, and $(a,-1)$, $(b,1)$, and $(c,1)$.

The point g_1 can be constructed geometrically as the middle of the segment joining c to the middle $\frac{1}{2}a+\frac{1}{2}b$ of the segment (a,b), since

$$g_1 = \frac{1}{2}\left(\frac{1}{2}a+\frac{1}{2}b\right)+\frac{1}{2}c.$$

The point g_2 can be constructed geometrically as the point such that the middle $\frac{1}{2}b+\frac{1}{2}c$ of the segment (b,c) is the middle of the segment (a,g_2), since

$$g_2 = -a+2\left(\frac{1}{2}b+\frac{1}{2}c\right).$$

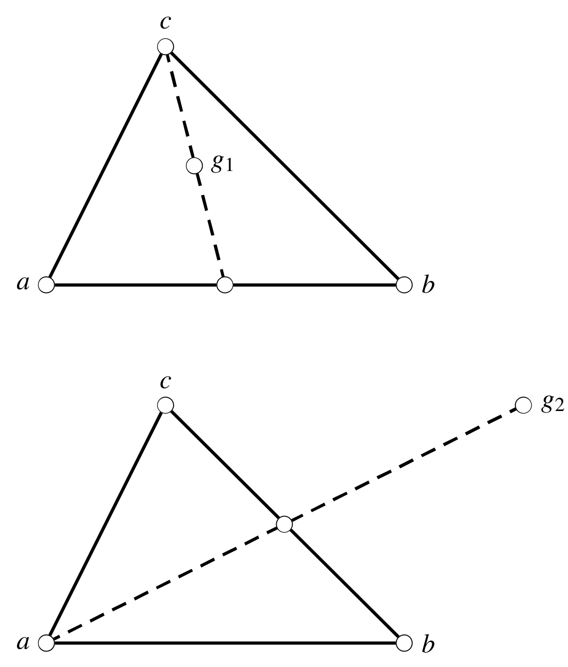

Fig. 2.5 Barycenters, $g_1 = \frac{1}{4}a+\frac{1}{4}b+\frac{1}{2}c$, $\quad g_2 = -a+b+c$.

Later on, we will see that a polynomial curve can be defined as a set of barycenters of a fixed number of points. For example, let (a,b,c,d) be a sequence of points in $\mathbb{A}^2$. Observe that

$$(1-t)^3 + 3t(1-t)^2 + 3t^2(1-t) + t^3 = 1,$$

since the sum on the left-hand side is obtained by expanding $(t+(1-t))^3 = 1$ using the binomial formula. Thus,

$$(1-t)^3\, a + 3t(1-t)^2\, b + 3t^2(1-t)\, c + t^3\, d$$

is a well-defined affine combination. Then, we can define the curve $F\colon \mathbb{A} \to \mathbb{A}^2$ such that

$$F(t) = (1-t)^3\, a + 3t(1-t)^2\, b + 3t^2(1-t)\, c + t^3\, d.$$

Such a curve is called a *Bézier curve*, and (a,b,c,d) are called its *control points*. Note that the curve passes through a and d, but generally not through b and c. We show in Chapter 18 (on the web site) how any point $F(t)$ on the curve can be constructed using an algorithm performing affine interpolation steps (the *de Casteljau algorithm*).

2.5 Affine Subspaces

In linear algebra, a (linear) subspace can be characterized as a nonempty subset of a vector space closed under linear combinations. In affine spaces, the notion corresponding to the notion of (linear) subspace is the notion of affine subspace. It is natural to define an affine subspace as a subset of an affine space closed under affine combinations.

Definition 2.3. Given an affine space $\langle E, \overrightarrow{E}, +\rangle$, a subset V of E is an *affine subspace (of* $\langle E, \overrightarrow{E}, +\rangle$) if for every family of weighted points $((a_i, \lambda_i))_{i\in I}$ in V such that $\sum_{i\in I} \lambda_i = 1$, the barycenter $\sum_{i\in I} \lambda_i a_i$ belongs to V.

An affine subspace is also called a *flat* by some authors. According to Definition 2.3, the empty set is trivially an affine subspace, and every intersection of affine subspaces is an affine subspace.

As an example, consider the subset U of $\mathbb{R}^2$ defined by

$$U = \left\{(x,y) \in \mathbb{R}^2 \mid ax + by = c\right\},$$

i.e., the set of solutions of the equation

$$ax + by = c,$$

where it is assumed that $a \neq 0$ or $b \neq 0$. Given any m points $(x_i, y_i) \in U$ and any m scalars λ_i such that $\lambda_1 + \cdots + \lambda_m = 1$, we claim that

$$\sum_{i=1}^{m} \lambda_i (x_i, y_i) \in U.$$

Indeed, $(x_i, y_i) \in U$ means that

$$ax_i + by_i = c,$$

and if we multiply both sides of this equation by λ_i and add up the resulting m equations, we get

$$\sum_{i=1}^{m} (\lambda_i a x_i + \lambda_i b y_i) = \sum_{i=1}^{m} \lambda_i c,$$

and since $\lambda_1 + \cdots + \lambda_m = 1$, we get

$$a\left(\sum_{i=1}^{m} \lambda_i x_i\right) + b\left(\sum_{i=1}^{m} \lambda_i y_i\right) = \left(\sum_{i=1}^{m} \lambda_i\right) c = c,$$

which shows that

$$\left(\sum_{i=1}^{m} \lambda_i x_i, \sum_{i=1}^{m} \lambda_i y_i\right) = \sum_{i=1}^{m} \lambda_i (x_i, y_i) \in U.$$

Thus, U is an affine subspace of $\mathbb{A}^2$. In fact, it is just a usual line in $\mathbb{A}^2$.

It turns out that U is closely related to the subset of $\mathbb{R}^2$ defined by

$$\overrightarrow{U} = \left\{(x, y) \in \mathbb{R}^2 \mid ax + by = 0\right\},$$

i.e., the set of solutions of the homogeneous equation

$$ax + by = 0$$

obtained by setting the right-hand side of $ax + by = c$ to zero. Indeed, for any m scalars λ_i, the same calculation as above yields that

$$\sum_{i=1}^{m} \lambda_i (x_i, y_i) \in \overrightarrow{U},$$

this time **without any restriction on the** λ_i, since the right-hand side of the equation is null. Thus, $\overrightarrow{U}$ is a subspace of $\mathbb{R}^2$. In fact, $\overrightarrow{U}$ is one-dimensional, and it is just a usual line in $\mathbb{R}^2$. This line can be identified with a line passing through the origin of $\mathbb{A}^2$, a line that is parallel to the line U of equation $ax + by = c$, as illustrated in Figure 2.6.

Now, if (x_0, y_0) is any point in U, we claim that

$$U = (x_0, y_0) + \overrightarrow{U},$$

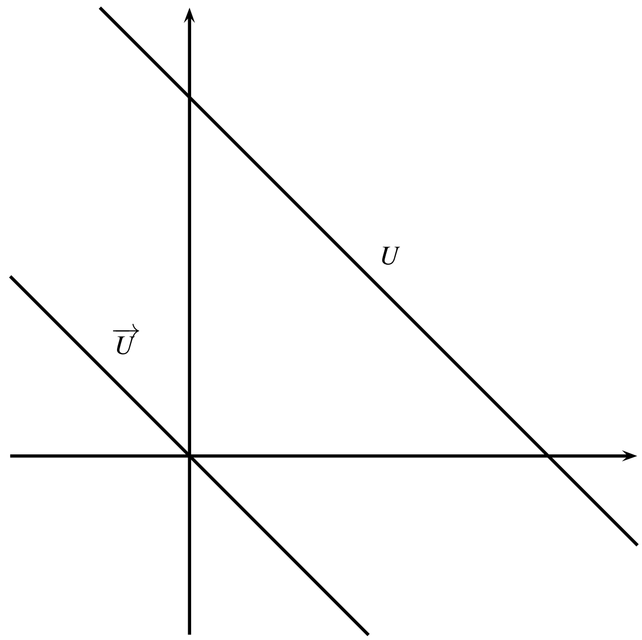

Fig. 2.6 An affine line U and its direction.

where

$$(x_0, y_0) + \overrightarrow{U} = \left\{ (x_0 + u_1, y_0 + u_2) \mid (u_1, u_2) \in \overrightarrow{U} \right\}.$$

First, $(x_0, y_0) + \overrightarrow{U} \subseteq U$, since $ax_0 + by_0 = c$ and $au_1 + bu_2 = 0$ for all $(u_1, u_2) \in \overrightarrow{U}$. Second, if $(x, y) \in U$, then $ax + by = c$, and since we also have $ax_0 + by_0 = c$, by subtraction, we get

$$a(x - x_0) + b(y - y_0) = 0,$$

which shows that $(x - x_0, y - y_0) \in \overrightarrow{U}$, and thus $(x, y) \in (x_0, y_0) + \overrightarrow{U}$. Hence, we also have $U \subseteq (x_0, y_0) + \overrightarrow{U}$, and $U = (x_0, y_0) + \overrightarrow{U}$.

The above example shows that the affine line U defined by the equation

$$ax + by = c$$

is obtained by "translating" the parallel line $\overrightarrow{U}$ of equation

$$ax + by = 0$$

passing through the origin. In fact, given any point $(x_0, y_0) \in U$,

$$U = (x_0, y_0) + \overrightarrow{U}.$$

More generally, it is easy to prove the following fact. Given any $m \times n$ matrix A and any vector $b \in \mathbb{R}^m$, the subset U of $\mathbb{R}^n$ defined by

$$U = \{x \in \mathbb{R}^n \mid Ax = b\}$$

is an affine subspace of $\mathbb{A}^n$.

Actually, observe that $Ax = b$ should really be written as $Ax^\top = b$, to be consistent with our convention that points are represented by row vectors. We can also use the boldface notation for column vectors, in which case the equation is written as $A\mathbf{x} = b$. For the sake of minimizing the amount of notation, we stick to the simpler (yet incorrect) notation $Ax = b$. If we consider the corresponding homogeneous equation $Ax = 0$, the set

$$\overrightarrow{U} = \{x \in \mathbb{R}^n \mid Ax = 0\}$$

is a subspace of $\mathbb{R}^n$, and for any $x_0 \in U$, we have

$$U = x_0 + \overrightarrow{U}.$$

This is a general situation. Affine subspaces can be characterized in terms of subspaces of $\overrightarrow{E}$. Let V be a nonempty subset of E. For every family $(a_1, \ldots, a_n)$ in V, for any family $(\lambda_1, \ldots, \lambda_n)$ of scalars, and for every point $a \in V$, observe that for every $x \in E$,

$$x = a + \sum_{i=1}^{n} \lambda_i \overrightarrow{aa_i}$$

is the barycenter of the family of weighted points

$$\Big((a_1, \lambda_1), \ldots, (a_n, \lambda_n), \Big(a, 1 - \sum_{i=1}^{n} \lambda_i\Big)\Big),$$

since

$$\sum_{i=1}^{n} \lambda_i + \Big(1 - \sum_{i=1}^{n} \lambda_i\Big) = 1.$$

Given any point $a \in E$ and any subset $\overrightarrow{V}$ of $\overrightarrow{E}$, let $a + \overrightarrow{V}$ denote the following subset of E:

$$a + \overrightarrow{V} = \Big\{a + v \mid v \in \overrightarrow{V}\Big\}.$$

Lemma 2.2. *Let $\langle E, \overrightarrow{E}, + \rangle$ be an affine space.*

(1) A nonempty subset V of E is an affine subspace iff for every point $a \in V$, the set

$$\overrightarrow{V_a} = \{\overrightarrow{ax} \mid x \in V\}$$

is a subspace of $\overrightarrow{E}$. Consequently, $V = a + \overrightarrow{V_a}$. Furthermore,

$$\overrightarrow{V} = \{\overrightarrow{xy} \mid x, y \in V\}$$

is a subspace of $\overrightarrow{E}$ *and* $\overrightarrow{V_a} = \overrightarrow{V}$ *for all* $a \in E$. *Thus,* $V = a + \overrightarrow{V}$.

(2) For any subspace $\overrightarrow{V}$ *of* $\overrightarrow{E}$ *and for any* $a \in E$*, the set* $V = a + \overrightarrow{V}$ *is an affine subspace.*

Proof. The proof is straightforward, and is omitted. It is also given in Gallier [5]. □

In particular, when E is the natural affine space associated with a vector space $\overrightarrow{E}$, Lemma 2.2 shows that every affine subspace of E is of the form $u + \overrightarrow{U}$, for a subspace $\overrightarrow{U}$ of $\overrightarrow{E}$. The subspaces of $\overrightarrow{E}$ are the affine subspaces of E that contain 0.

The subspace $\overrightarrow{V}$ associated with an affine subspace V is called the *direction of* V. It is also clear that the map $+ \colon V \times \overrightarrow{V} \to V$ induced by $+ \colon E \times \overrightarrow{E} \to E$ confers to $\langle V, \overrightarrow{V}, + \rangle$ an affine structure. Figure 2.7 illustrates the notion of affine subspace.

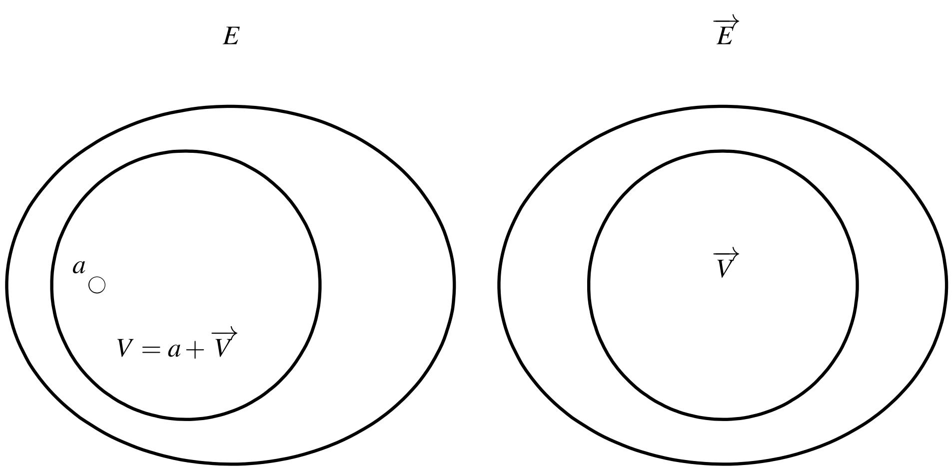

Fig. 2.7 An affine subspace V and its direction $\overrightarrow{V}$.

By the dimension of the subspace V, we mean the dimension of $\overrightarrow{V}$.

An affine subspace of dimension 1 is called a *line*, and an affine subspace of dimension 2 is called a *plane*.

An affine subspace of codimension 1 is called a *hyperplane* (recall that a subspace F of a vector space E has codimension 1 iff there is some subspace G of dimension 1 such that $E = F \oplus G$, the direct sum of F and G, see Strang [12] or Lang [8]).

We say that two affine subspaces U and V are *parallel* if their directions are identical. Equivalently, since $\overrightarrow{U} = \overrightarrow{V}$, we have $U = a + \overrightarrow{U}$ and $V = b + \overrightarrow{U}$ for any $a \in U$ and any $b \in V$, and thus V is obtained from U by the translation $\overrightarrow{ab}$.

In general, when we talk about n points $a_1, \ldots, a_n$, we mean the sequence $(a_1, \ldots, a_n)$, and not the set $\{a_1, \ldots, a_n\}$ (the a_i's need not be distinct).

By Lemma 2.2, a line is specified by a point $a \in E$ and a nonzero vector $v \in \overrightarrow{E}$, i.e., a line is the set of all points of the form $a + \lambda v$, for $\lambda \in \mathbb{R}$.

We say that three points a, b, c are *collinear* if the vectors $\overrightarrow{ab}$ and $\overrightarrow{ac}$ are linearly dependent. If two of the points a, b, c are distinct, say $a \neq b$, then there is a unique $\lambda \in \mathbb{R}$ such that $\overrightarrow{ac} = \lambda \overrightarrow{ab}$, and we define the ratio $\frac{\overrightarrow{ac}}{\overrightarrow{ab}} = \lambda$.

A plane is specified by a point $a \in E$ and two linearly independent vectors $u, v \in \overrightarrow{E}$, i.e., a plane is the set of all points of the form $a + \lambda u + \mu v$, for $\lambda, \mu \in \mathbb{R}$.

We say that four points a, b, c, d are *coplanar* if the vectors $\overrightarrow{ab}, \overrightarrow{ac}$, and $\overrightarrow{ad}$ are linearly dependent. Hyperplanes will be characterized a little later.

Lemma 2.3. *Given an affine space $\langle E, \overrightarrow{E}, + \rangle$, for any family $(a_i)_{i \in I}$ of points in E, the set V of barycenters $\sum_{i \in I} \lambda_i a_i$ (where $\sum_{i \in I} \lambda_i = 1$) is the smallest affine subspace containing $(a_i)_{i \in I}$.*

Proof. If $(a_i)_{i \in I}$ is empty, then $V = \emptyset$, because of the condition $\sum_{i \in I} \lambda_i = 1$. If $(a_i)_{i \in I}$ is nonempty, then the smallest affine subspace containing $(a_i)_{i \in I}$ must contain the set V of barycenters $\sum_{i \in I} \lambda_i a_i$, and thus, it is enough to show that V is closed under affine combinations, which is immediately verified. $\square$

Given a nonempty subset S of E, the smallest affine subspace of E generated by S is often denoted by $\langle S \rangle$. For example, a line specified by two distinct points a and b is denoted by $\langle a, b \rangle$, or even (a, b), and similarly for planes, etc.

Remarks:

(1) Since it can be shown that the barycenter of n weighted points can be obtained by repeated computations of barycenters of two weighted points, a nonempty subset V of E is an affine subspace iff for every two points $a, b \in V$, the set V contains all barycentric combinations of a and b. If V contains at least two points, then V is an affine subspace iff for any two distinct points $a, b \in V$, the set V contains the line determined by a and b, that is, the set of all points $(1 - \lambda)a + \lambda b$, $\lambda \in \mathbb{R}$.
(2) This result still holds if the field K has at least three distinct elements, but the proof is trickier!

2.6 Affine Independence and Affine Frames

Corresponding to the notion of linear independence in vector spaces, we have the notion of affine independence. Given a family $(a_i)_{i \in I}$ of points in an affine space E, we will reduce the notion of (affine) independence of these points to the (linear) independence of the families $(\overrightarrow{a_i a_j})_{j \in (I - \{i\})}$ of vectors obtained by choosing any a_i

as an origin. First, the following lemma shows that it is sufficient to consider only one of these families.

Lemma 2.4. *Given an affine space* $\langle E, \overrightarrow{E}, +\rangle$, *let* $(a_i)_{i\in I}$ *be a family of points in* E. *If the family* $(\overrightarrow{a_ia_j})_{j\in(I-\{i\})}$ *is linearly independent for some* $i \in I$, *then* $(\overrightarrow{a_ia_j})_{j\in(I-\{i\})}$ *is linearly independent for every* $i \in I$.

Proof. Assume that the family $(\overrightarrow{a_ia_j})_{j\in(I-\{i\})}$ is linearly independent for some specific $i \in I$. Let $k \in I$ with $k \neq i$, and assume that there are some scalars $(\lambda_j)_{j\in(I-\{k\})}$ such that

$$\sum_{j\in(I-\{k\})} \lambda_j \overrightarrow{a_ka_j} = 0.$$

Since

$$\overrightarrow{a_ka_j} = \overrightarrow{a_ka_i} + \overrightarrow{a_ia_j},$$

we have

$$\begin{aligned}
\sum_{j\in(I-\{k\})} \lambda_j \overrightarrow{a_ka_j} &= \sum_{j\in(I-\{k\})} \lambda_j \overrightarrow{a_ka_i} + \sum_{j\in(I-\{k\})} \lambda_j \overrightarrow{a_ia_j}, \\
&= \sum_{j\in(I-\{k\})} \lambda_j \overrightarrow{a_ka_i} + \sum_{j\in(I-\{i,k\})} \lambda_j \overrightarrow{a_ia_j}, \\
&= \sum_{j\in(I-\{i,k\})} \lambda_j \overrightarrow{a_ia_j} - \Bigg(\sum_{j\in(I-\{k\})} \lambda_j\Bigg) \overrightarrow{a_ia_k},
\end{aligned}$$

and thus

$$\sum_{j\in(I-\{i,k\})} \lambda_j \overrightarrow{a_ia_j} - \Bigg(\sum_{j\in(I-\{k\})} \lambda_j\Bigg) \overrightarrow{a_ia_k} = 0.$$

Since the family $(\overrightarrow{a_ia_j})_{j\in(I-\{i\})}$ is linearly independent, we must have $\lambda_j = 0$ for all $j \in (I - \{i,k\})$ and $\sum_{j\in(I-\{k\})} \lambda_j = 0$, which implies that $\lambda_j = 0$ for all $j \in (I - \{k\})$. $\square$

We define affine independence as follows.

Definition 2.4. Given an affine space $\langle E, \overrightarrow{E}, +\rangle$, a family $(a_i)_{i\in I}$ of points in E is *affinely independent* if the family $(\overrightarrow{a_ia_j})_{j\in(I-\{i\})}$ is linearly independent for some $i \in I$.

Definition 2.4 is reasonable, since by Lemma 2.4, the independence of the family $(\overrightarrow{a_ia_j})_{j\in(I-\{i\})}$ does not depend on the choice of a_i. A crucial property of linearly independent vectors $(u_1, \ldots, u_m)$ is that if a vector v is a linear combination

$$v = \sum_{i=1}^{m} \lambda_i u_i$$

of the u_i, then the λ_i are unique. A similar result holds for affinely independent points.

Lemma 2.5. *Given an affine space $\langle E, \overrightarrow{E}, +\rangle$, let $(a_0, \ldots, a_m)$ be a family of $m+1$ points in E. Let $x \in E$, and assume that $x = \sum_{i=0}^{m} \lambda_i a_i$, where $\sum_{i=0}^{m} \lambda_i = 1$. Then, the family $(\lambda_0, \ldots, \lambda_m)$ such that $x = \sum_{i=0}^{m} \lambda_i a_i$ is unique iff the family $(\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m})$ is linearly independent.*

Proof. The proof is straightforward and is omitted. It is also given in Gallier [5]. □

Lemma 2.5 suggests the notion of affine frame. Affine frames are the affine analogues of bases in vector spaces. Let $\langle E, \overrightarrow{E}, +\rangle$ be a nonempty affine space, and let $(a_0, \ldots, a_m)$ be a family of $m+1$ points in E. The family $(a_0, \ldots, a_m)$ determines the family of m vectors $(\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m})$ in $\overrightarrow{E}$. Conversely, given a point a_0 in E and a family of m vectors $(u_1, \ldots, u_m)$ in $\overrightarrow{E}$, we obtain the family of $m+1$ points $(a_0, \ldots, a_m)$ in E, where $a_i = a_0 + u_i$, $1 \leq i \leq m$.

Thus, for any $m \geq 1$, it is equivalent to consider a family of $m+1$ points $(a_0, \ldots, a_m)$ in E, and a pair $(a_0, (u_1, \ldots, u_m))$, where the u_i are vectors in $\overrightarrow{E}$. Figure 2.8 illustrates the notion of affine independence.

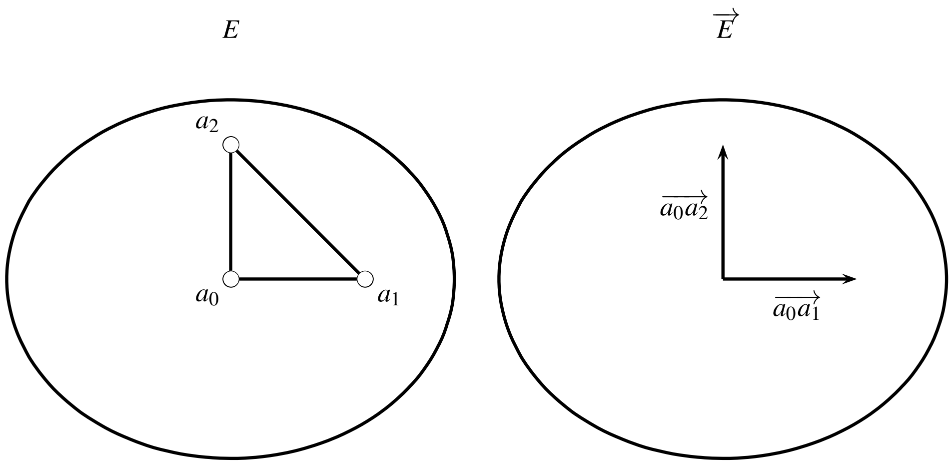

Fig. 2.8 Affine independence and linear independence.

Remark: The above observation also applies to infinite families $(a_i)_{i\in I}$ of points in E and families $(\overrightarrow{u_i})_{i\in I-\{0\}}$ of vectors in $\overrightarrow{E}$, provided that the index set I contains 0.

When $(\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m})$ is a basis of $\overrightarrow{E}$ then, for every $x \in E$, since $x = a_0 + \overrightarrow{a_0x}$, there is a unique family $(x_1, \ldots, x_m)$ of scalars such that

$$x = a_0 + x_1\overrightarrow{a_0a_1} + \cdots + x_m\overrightarrow{a_0a_m}.$$

The scalars $(x_1, \ldots, x_m)$ may be considered as coordinates with respect to $(a_0, (\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m}))$. Since

$$x = a_0 + \sum_{i=1}^{m} x_i \overrightarrow{a_0a_i} \quad \text{iff} \quad x = \left(1 - \sum_{i=1}^{m} x_i\right) a_0 + \sum_{i=1}^{m} x_i a_i,$$

$x \in E$ can also be expressed uniquely as

$$x = \sum_{i=0}^{m} \lambda_i a_i$$

with $\sum_{i=0}^{m} \lambda_i = 1$, and where $\lambda_0 = 1 - \sum_{i=1}^{m} x_i$, and $\lambda_i = x_i$ for $1 \le i \le m$. The scalars $(\lambda_0, \ldots, \lambda_m)$ are also certain kinds of coordinates with respect to $(a_0, \ldots, a_m)$. All this is summarized in the following definition.

Definition 2.5. Given an affine space $\langle E, \overrightarrow{E}, + \rangle$, an *affine frame with origin* a_0 is a family $(a_0, \ldots, a_m)$ of $m+1$ points in E such that the list of vectors $(\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m})$ is a basis of $\overrightarrow{E}$. The pair $(a_0, (\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m}))$ is also called an *affine frame with origin* a_0. Then, every $x \in E$ can be expressed as

$$x = a_0 + x_1 \overrightarrow{a_0a_1} + \cdots + x_m \overrightarrow{a_0a_m}$$

for a unique family $(x_1, \ldots, x_m)$ of scalars, called the *coordinates of x w.r.t. the affine frame* $(a_0, (\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m}))$. Furthermore, every $x \in E$ can be written as

$$x = \lambda_0 a_0 + \cdots + \lambda_m a_m$$

for some unique family $(\lambda_0, \ldots, \lambda_m)$ of scalars such that $\lambda_0 + \cdots + \lambda_m = 1$ called the *barycentric coordinates of x with respect to the affine frame* $(a_0, \ldots, a_m)$.

The coordinates $(x_1, \ldots, x_m)$ and the barycentric coordinates $(\lambda_0, \ldots, \lambda_m)$ are related by the equations $\lambda_0 = 1 - \sum_{i=1}^{m} x_i$ and $\lambda_i = x_i$, for $1 \le i \le m$. An affine frame is called an *affine basis* by some authors. A family $(a_i)_{i \in I}$ of points in E is *affinely dependent* if it is not affinely independent. We can also characterize affinely dependent families as follows.

Lemma 2.6. *Given an affine space $\langle E, \overrightarrow{E}, + \rangle$, let $(a_i)_{i \in I}$ be a family of points in E. The family $(a_i)_{i \in I}$ is affinely dependent iff there is a family $(\lambda_i)_{i \in I}$ such that $\lambda_j \neq 0$ for some $j \in I$, $\sum_{i \in I} \lambda_i = 0$, and $\sum_{i \in I} \lambda_i \overrightarrow{xa_i} = 0$ for every $x \in E$.*

Proof. By Lemma 2.5, the family $(a_i)_{i \in I}$ is affinely dependent iff the family of vectors $(\overrightarrow{a_ia_j})_{j \in (I - \{i\})}$ is linearly dependent for some $i \in I$. For any $i \in I$, the family $(\overrightarrow{a_ia_j})_{j \in (I - \{i\})}$ is linearly dependent iff there is a family $(\lambda_j)_{j \in (I - \{i\})}$ such that $\lambda_j \neq 0$ for some j, and such that

$$\sum_{j \in (I - \{i\})} \lambda_j \overrightarrow{a_ia_j} = 0.$$

Then, for any $x \in E$, we have

$$\sum_{j\in(I-\{i\})}\lambda_j\overrightarrow{a_ia_j} = \sum_{j\in(I-\{i\})}\lambda_j(\overrightarrow{xa_j}-\overrightarrow{xa_i})$$
$$= \sum_{j\in(I-\{i\})}\lambda_j\overrightarrow{xa_j} - \Bigg(\sum_{j\in(I-\{i\})}\lambda_j\Bigg)\overrightarrow{xa_i},$$

and letting $\lambda_i = -\big(\sum_{j\in(I-\{i\})}\lambda_j\big)$, we get $\sum_{i\in I}\lambda_i\overrightarrow{xa_i} = 0$, with $\sum_{i\in I}\lambda_i = 0$ and $\lambda_j \neq 0$ for some $j \in I$. The converse is obvious by setting $x = a_i$ for some i such that $\lambda_i \neq 0$, since $\sum_{i\in I}\lambda_i = 0$ implies that $\lambda_j \neq 0$, for some $j \neq i$. $\square$

Even though Lemma 2.6 is rather dull, it is one of the key ingredients in the proof of beautiful and deep theorems about convex sets, such as Carathéodory's theorem, Radon's theorem, and Helly's theorem (see Section 3.1).

A family of two points (a,b) in E is affinely independent iff $\overrightarrow{ab} \neq 0$, iff $a \neq b$. If $a \neq b$, the affine subspace generated by a and b is the set of all points $(1-\lambda)a+\lambda b$, which is the unique line passing through a and b. A family of three points (a,b,c) in E is affinely independent iff $\overrightarrow{ab}$ and $\overrightarrow{ac}$ are linearly independent, which means that a, b, and c are not on the same line (they are not collinear). In this case, the affine subspace generated by (a,b,c) is the set of all points $(1-\lambda-\mu)a+\lambda b+\mu c$, which is the unique plane containing a, b, and c. A family of four points (a,b,c,d) in E is affinely independent iff $\overrightarrow{ab}$, $\overrightarrow{ac}$, and $\overrightarrow{ad}$ are linearly independent, which means that a, b, c, and d are not in the same plane (they are not coplanar). In this case, a, b, c, and d are the vertices of a tetrahedron. Figure 2.9 shows affine frames and their convex hulls for $|I| = 0,1,2,3$.

Given $n+1$ affinely independent points $(a_0,\ldots,a_n)$ in E, we can consider the set of points $\lambda_0a_0+\cdots+\lambda_na_n$, where $\lambda_0+\cdots+\lambda_n = 1$ and $\lambda_i \geq 0$ ($\lambda_i \in \mathbb{R}$). Such affine combinations are called *convex combinations*. This set is called the *convex hull* of $(a_0,\ldots,a_n)$ (or *n-simplex spanned by* $(a_0,\ldots,a_n)$). When $n = 1$, we get the segment between a_0 and a_1, including a_0 and a_1. When $n = 2$, we get the interior of the triangle whose vertices are a_0,a_1,a_2, including boundary points (the edges). When $n = 3$, we get the interior of the tetrahedron whose vertices are a_0,a_1,a_2,a_3, including boundary points (faces and edges). The set

$$\{a_0+\lambda_1\overrightarrow{a_0a_1}+\cdots+\lambda_n\overrightarrow{a_0a_n} \mid \text{where } 0 \leq \lambda_i \leq 1 \ (\lambda_i \in \mathbb{R})\}$$

is called the *parallelotope spanned by* $(a_0,\ldots,a_n)$. When E has dimension 2, a parallelotope is also called a *parallelogram*, and when E has dimension 3, a *parallelepiped*.

More generally, we say that a subset V of E is *convex* if for any two points $a,b \in V$, we have $c \in V$ for every point $c = (1-\lambda)a+\lambda b$, with $0 \leq \lambda \leq 1$ ($\lambda \in \mathbb{R}$).

Points are not vectors! The following example illustrates why treating points as vectors may cause problems. Let a,b,c be three affinely independent points in $\mathbb{A}^3$. Any point x in the plane (a,b,c) can be expressed as

$$x = \lambda_0a+\lambda_1b+\lambda_2c,$$

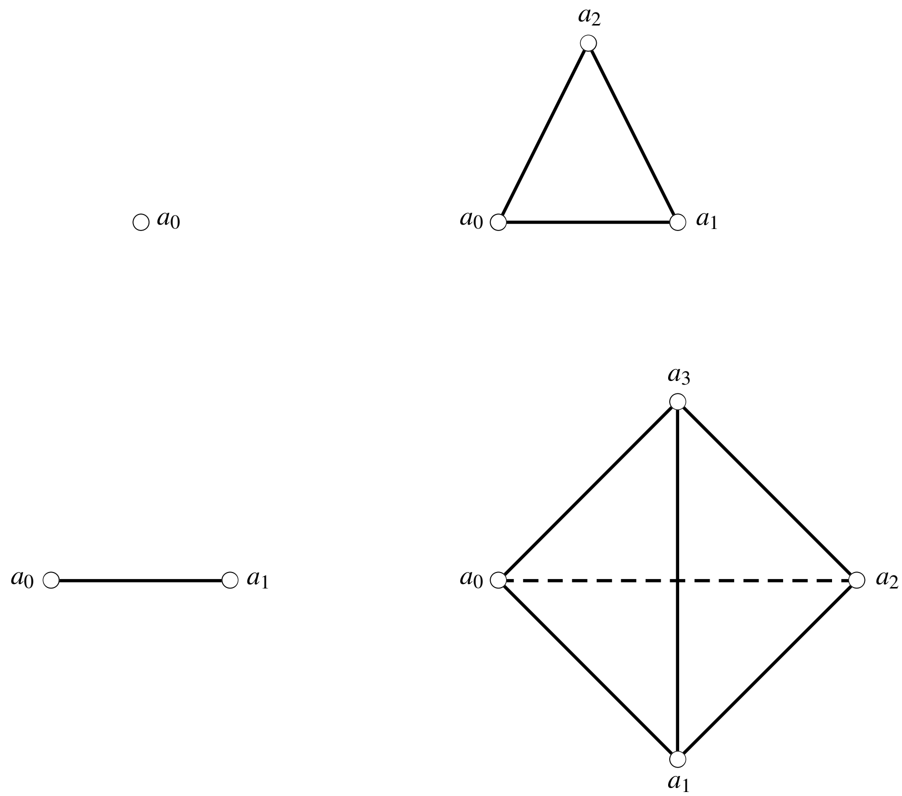

Fig. 2.9 Examples of affine frames and their convex hulls.

where $\lambda_0 + \lambda_1 + \lambda_2 = 1$. How can we compute $\lambda_0, \lambda_1, \lambda_2$? Letting $a = (a_1, a_2, a_3)$, $b = (b_1, b_2, b_3)$, $c = (c_1, c_2, c_3)$, and $x = (x_1, x_2, x_3)$ be the coordinates of a, b, c, x in the standard frame of $\mathbb{A}^3$, it is tempting to solve the system of equations

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} \lambda_0 \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

However, there is a problem when the origin of the coordinate system belongs to the plane (a, b, c), since in this case, the matrix is not invertible! What we should really be doing is to solve the system

$$\lambda_0 \overrightarrow{Oa} + \lambda_1 \overrightarrow{Ob} + \lambda_2 \overrightarrow{Oc} = \overrightarrow{Ox},$$

where O is any point **not** in the plane (a, b, c). An alternative is to use certain well-chosen cross products.

It can be shown that barycentric coordinates correspond to various ratios of areas and volumes; see the problems.

2.7 Affine Maps

Corresponding to linear maps we have the notion of an affine map. An affine map is defined as a map preserving affine combinations.

Definition 2.6. Given two affine spaces $\langle E, \overrightarrow{E}, +\rangle$ and $\langle E', \overrightarrow{E'}, +'\rangle$, a function $f\colon E \to E'$ is an *affine map* iff for every family $((a_i, \lambda_i))_{i\in I}$ of weighted points in E such that $\sum_{i\in I} \lambda_i = 1$, we have

$$f\left(\sum_{i\in I} \lambda_i a_i\right) = \sum_{i\in I} \lambda_i f(a_i).$$

In other words, f preserves barycenters.

Affine maps can be obtained from linear maps as follows. For simplicity of notation, the same symbol $+$ is used for both affine spaces (instead of using both $+$ and $+'$).

Given any point $a \in E$, any point $b \in E'$, and any linear map $h\colon \overrightarrow{E} \to \overrightarrow{E'}$, we claim that the map $f\colon E \to E'$ defined such that

$$f(a+v) = b + h(v)$$

is an affine map. Indeed, for any family $(\lambda_i)_{i\in I}$ of scalars with $\sum_{i\in I} \lambda_i = 1$ and any family $(\overrightarrow{v_i})_{i\in I}$, since

$$\sum_{i\in I} \lambda_i (a + v_i) = a + \sum_{i\in I} \lambda_i \overrightarrow{a(a+v_i)} = a + \sum_{i\in I} \lambda_i v_i$$

and

$$\sum_{i\in I} \lambda_i (b + h(v_i)) = b + \sum_{i\in I} \lambda_i \overrightarrow{b(b+h(v_i))} = b + \sum_{i\in I} \lambda_i h(v_i),$$

we have

$$\begin{aligned}
f\left(\sum_{i\in I} \lambda_i (a+v_i)\right) &= f\left(a + \sum_{i\in I} \lambda_i v_i\right) \\
&= b + h\left(\sum_{i\in I} \lambda_i v_i\right) \\
&= b + \sum_{i\in I} \lambda_i h(v_i) \\
&= \sum_{i\in I} \lambda_i (b + h(v_i)) \\
&= \sum_{i\in I} \lambda_i f(a+v_i).
\end{aligned}$$

□

Note that the condition $\sum_{i\in I}\lambda_i = 1$ was implicitly used (in a hidden call to Lemma 2.1) in deriving that

$$\sum_{i\in I}\lambda_i(a+v_i) = a+\sum_{i\in I}\lambda_i v_i$$

and

$$\sum_{i\in I}\lambda_i(b+h(v_i)) = b+\sum_{i\in I}\lambda_i h(v_i).$$

As a more concrete example, the map

$$\begin{pmatrix}x_1\\x_2\end{pmatrix} \mapsto \begin{pmatrix}1&2\\0&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}+\begin{pmatrix}3\\1\end{pmatrix}$$

defines an affine map in $\mathbb{A}^2$. It is a "shear" followed by a translation. The effect of this shear on the square (a,b,c,d) is shown in Figure 2.10. The image of the square (a,b,c,d) is the parallelogram (a',b',c',d').

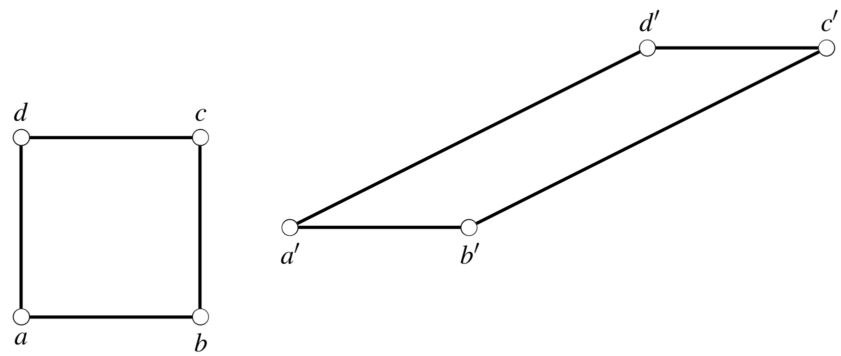

Fig. 2.10 The effect of a shear.

Let us consider one more example. The map

$$\begin{pmatrix}x_1\\x_2\end{pmatrix} \mapsto \begin{pmatrix}1&1\\1&3\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}+\begin{pmatrix}3\\0\end{pmatrix}$$

is an affine map. Since we can write

$$\begin{pmatrix}1&1\\1&3\end{pmatrix} = \sqrt{2}\begin{pmatrix}\sqrt{2}/2&-\sqrt{2}/2\\2/2&\sqrt{2}/2\end{pmatrix}\begin{pmatrix}1&2\\0&1\end{pmatrix},$$

this affine map is the composition of a shear, followed by a rotation of angle $\pi/4$, followed by a magnification of ratio $\sqrt{2}$, followed by a translation. The effect of this map on the square (a,b,c,d) is shown in Figure 2.11. The image of the square (a,b,c,d) is the parallelogram (a',b',c',d').

The following lemma shows the converse of what we just showed. Every affine map is determined by the image of any point and a linear map.

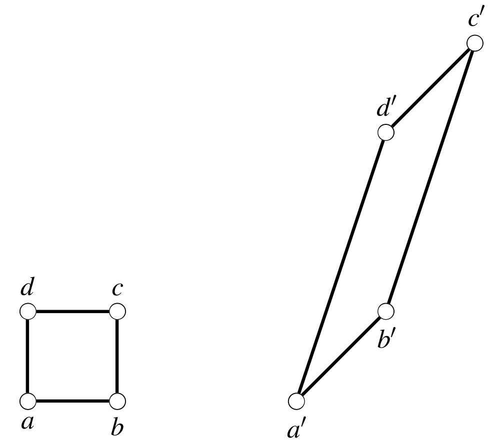

Fig. 2.11 The effect of an affine map.

Lemma 2.7. *Given an affine map $f\colon E \to E'$, there is a unique linear map $\overrightarrow{f}\colon \overrightarrow{E} \to \overrightarrow{E'}$ such that*

$$f(a+v) = f(a) + \overrightarrow{f}(v),$$

for every $a \in E$ and every $v \in \overrightarrow{E}$.

Proof. Let $a \in E$ be any point in E. We claim that the map defined such that

$$\overrightarrow{f}(v) = \overrightarrow{f(a)f(a+v)}$$

for every $v \in \overrightarrow{E}$ is a linear map $\overrightarrow{f}\colon \overrightarrow{E} \to \overrightarrow{E'}$. Indeed, we can write

$$a + \lambda v = \lambda(a+v) + (1-\lambda)a,$$

since $a + \lambda v = a + \lambda\overrightarrow{a(a+v)} + (1-\lambda)\overrightarrow{aa}$, and also

$$a + u + v = (a+u) + (a+v) - a,$$

since $a + u + v = a + \overrightarrow{a(a+u)} + \overrightarrow{a(a+v)} - \overrightarrow{aa}$. Since f preserves barycenters, we get

$$f(a + \lambda v) = \lambda f(a+v) + (1-\lambda) f(a).$$

If we recall that $x = \sum_{i\in I} \lambda_i a_i$ is the barycenter of a family $((a_i, \lambda_i))_{i\in I}$ of weighted points (with $\sum_{i\in I} \lambda_i = 1$) iff

$$\overrightarrow{bx} = \sum_{i\in I} \lambda_i \overrightarrow{ba_i} \quad \text{for every } b \in E,$$

we get

$$\overrightarrow{f(a)f(a+\lambda v)} = \lambda\overrightarrow{f(a)f(a+v)} + (1-\lambda)\overrightarrow{f(a)f(a)} = \lambda\overrightarrow{f(a)f(a+v)},$$

showing that $\overrightarrow{f}(\lambda v) = \lambda \overrightarrow{f}(v)$. We also have

$$f(a+u+v) = f(a+u) + f(a+v) - f(a),$$

from which we get

$$\overrightarrow{f(a)f(a+u+v)} = \overrightarrow{f(a)f(a+u)} + \overrightarrow{f(a)f(a+v)},$$

showing that $\overrightarrow{f}(u+v) = \overrightarrow{f}(u) + \overrightarrow{f}(v)$. Consequently, $\overrightarrow{f}$ is a linear map. For any other point $b \in E$, since

$$b+v = a + \overrightarrow{ab} + v = a + \overrightarrow{a(a+v)} - \overrightarrow{aa} + \overrightarrow{ab},$$

$b+v = (a+v) - a + b$, and since f preserves barycenters, we get

$$f(b+v) = f(a+v) - f(a) + f(b),$$

which implies that

$$\begin{aligned}
\overrightarrow{f(b)f(b+v)} &= \overrightarrow{f(b)f(a+v)} - \overrightarrow{f(b)f(a)} + \overrightarrow{f(b)f(b)}, \\
&= \overrightarrow{f(a)f(b)} + \overrightarrow{f(b)f(a+v)}, \\
&= \overrightarrow{f(a)f(a+v)}.
\end{aligned}$$

Thus, $\overrightarrow{f(b)f(b+v)} = \overrightarrow{f(a)f(a+v)}$, which shows that the definition of $\overrightarrow{f}$ does not depend on the choice of $a \in E$. The fact that $\overrightarrow{f}$ is unique is obvious: We must have $\overrightarrow{f}(v) = \overrightarrow{f(a)f(a+v)}$. $\square$

The unique linear map $\overrightarrow{f} : \overrightarrow{E} \to \overrightarrow{E'}$ given by Lemma 2.7 is called the *linear map associated with the affine map f*.

Note that the condition

$$f(a+v) = f(a) + \overrightarrow{f}(v),$$

for every $a \in E$ and every $v \in \overrightarrow{E}$, can be stated equivalently as

$$f(x) = f(a) + \overrightarrow{f}(\overrightarrow{ax}), \quad \text{or} \quad \overrightarrow{f(a)f(x)} = \overrightarrow{f}(\overrightarrow{ax}),$$

for all $a, x \in E$. Lemma 2.7 shows that for any affine map $f \colon E \to E'$, there are points $a \in E$, $b \in E'$, and a unique linear map $\overrightarrow{f} : \overrightarrow{E} \to \overrightarrow{E'}$, such that

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

for all $v \in \overrightarrow{E}$ (just let $b = f(a)$, for any $a \in E$). Affine maps for which $\overrightarrow{f}$ is the identity map are called *translations*. Indeed, if $\overrightarrow{f} = \mathrm{id}$,

$$\begin{aligned} f(x) &= f(a) + \overrightarrow{f}(\overrightarrow{ax}) = f(a) + \overrightarrow{ax} = x + \overrightarrow{xa} + \overrightarrow{af(a)} + \overrightarrow{ax} \\ &= x + \overrightarrow{xa} + \overrightarrow{af(a)} - \overrightarrow{xa} = x + \overrightarrow{af(a)}, \end{aligned}$$

and so

$$\overrightarrow{xf(x)} = \overrightarrow{af(a)},$$

which shows that f is the translation induced by the vector $\overrightarrow{af(a)}$ (which does not depend on a).

Since an affine map preserves barycenters, and since an affine subspace V is closed under barycentric combinations, the image $f(V)$ of V is an affine subspace in E'. So, for example, the image of a line is a point or a line, and the image of a plane is either a point, a line, or a plane.

It is easily verified that the composition of two affine maps is an affine map. Also, given affine maps $f\colon E \to E'$ and $g\colon E' \to E''$, we have

$$g(f(a+v)) = g\Big(f(a) + \overrightarrow{f}(v)\Big) = g(f(a)) + \overrightarrow{g}\Big(\overrightarrow{f}(v)\Big),$$

which shows that $\overrightarrow{g \circ f} = \overrightarrow{g} \circ \overrightarrow{f}$. It is easy to show that an affine map $f\colon E \to E'$ is injective iff $\overrightarrow{f}\colon \overrightarrow{E} \to \overrightarrow{E'}$ is injective, and that $f\colon E \to E'$ is surjective iff $\overrightarrow{f}\colon \overrightarrow{E} \to \overrightarrow{E'}$ is surjective. An affine map $f\colon E \to E'$ is constant iff $\overrightarrow{f}\colon \overrightarrow{E} \to \overrightarrow{E'}$ is the null (constant) linear map equal to 0 for all $v \in \overrightarrow{E}$.

If E is an affine space of dimension m and $(a_0, a_1, \ldots, a_m)$ is an affine frame for E, then for any other affine space F and for any sequence $(b_0, b_1, \ldots, b_m)$ of $m+1$ points in F, there is a unique affine map $f\colon E \to F$ such that $f(a_i) = b_i$, for $0 \le i \le m$. Indeed, f must be such that

$$f(\lambda_0 a_0 + \cdots + \lambda_m a_m) = \lambda_0 b_0 + \cdots + \lambda_m b_m,$$

where $\lambda_0 + \cdots + \lambda_m = 1$, and this defines a unique affine map on all of E, since $(a_0, a_1, \ldots, a_m)$ is an affine frame for E.

Using affine frames, affine maps can be represented in terms of matrices. We explain how an affine map $f\colon E \to E$ is represented with respect to a frame $(a_0, \ldots, a_n)$ in E, the more general case where an affine map $f\colon E \to F$ is represented with respect to two affine frames $(a_0, \ldots, a_n)$ in E and $(b_0, \ldots, b_m)$ in F being analogous. Since

$$f(a_0 + x) = f(a_0) + \overrightarrow{f}(x)$$

for all $x \in \overrightarrow{E}$, we have

$$\overrightarrow{a_0 f(a_0 + x)} = \overrightarrow{a_0 f(a_0)} + \overrightarrow{f}(x).$$

Since x, $\overrightarrow{a_0 f(a_0)}$, and $\overrightarrow{a_0 f(a_0+x)}$, can be expressed as

$$\begin{aligned} x &= x_1\overrightarrow{a_0a_1} + \cdots + x_n\overrightarrow{a_0a_n}, \\ \overrightarrow{a_0 f(a_0)} &= b_1\overrightarrow{a_0a_1} + \cdots + b_n\overrightarrow{a_0a_n}, \\ \overrightarrow{a_0 f(a_0+x)} &= y_1\overrightarrow{a_0a_1} + \cdots + y_n\overrightarrow{a_0a_n}, \end{aligned}$$

if $A = (a_{i\,j})$ is the $n \times n$ matrix of the linear map $\overrightarrow{f}$ over the basis $(\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_n})$, letting x, y, and b denote the column vectors of components $(x_1, \ldots, x_n)$, $(y_1, \ldots, y_n)$, and $(b_1, \ldots, b_n)$,

$$\overrightarrow{a_0 f(a_0+x)} = \overrightarrow{a_0 f(a_0)} + \overrightarrow{f}(x)$$

is equivalent to

$$y = Ax + b.$$

Note that $b \neq 0$ unless $f(a_0) = a_0$. Thus, f is generally not a linear transformation, unless it has a *fixed point*, i.e., there is a point a_0 such that $f(a_0) = a_0$. The vector b is the "translation part" of the affine map. Affine maps do not always have a fixed point. Obviously, nonnull translations have no fixed point. A less trivial example is given by the affine map

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

This map is a reflection about the x-axis followed by a translation along the x-axis. The affine map

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3}/4 & 1/4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

can also be written as

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

which shows that it is the composition of a rotation of angle $\pi/3$, followed by a stretch (by a factor of 2 along the x-axis, and by a factor of $\frac{1}{2}$ along the y-axis), followed by a translation. It is easy to show that this affine map has a unique fixed point. On the other hand, the affine map

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} 8/5 & -6/5 \\ 3/10 & 2/5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

has no fixed point, even though

$$\begin{pmatrix} 8/5 & -6/5 \\ 3/10 & 2/5 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} 4/5 & -3/5 \\ 3/5 & 4/5 \end{pmatrix},$$

and the second matrix is a rotation of angle θ such that $\cos\theta = \frac{4}{5}$ and $\sin\theta = \frac{3}{5}$. For more on fixed points of affine maps, see the problems.

There is a useful trick to convert the equation $y = Ax + b$ into what looks like a linear equation. The trick is to consider an $(n+1)\times(n+1)$ matrix. We add 1 as the $(n+1)$th component to the vectors x, y, and b, and form the $(n+1)\times(n+1)$ matrix

$$\begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}$$

so that $y = Ax + b$ is equivalent to

$$\begin{pmatrix} y \\ 1 \end{pmatrix} = \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix}.$$

This trick is very useful in kinematics and dynamics, where A is a rotation matrix. Such affine maps are called *rigid motions*.

If $f\colon E \to E'$ is a bijective affine map, given any three collinear points a, b, c in E, with $a \neq b$, where, say, $c = (1-\lambda)a + \lambda b$, since f preserves barycenters, we have $f(c) = (1-\lambda)f(a) + \lambda f(b)$, which shows that $f(a), f(b), f(c)$ are collinear in E'. There is a converse to this property, which is simpler to state when the ground field is $K = \mathbb{R}$. The converse states that given any bijective function $f\colon E \to E'$ between two real affine spaces of the same dimension $n \geq 2$, if f maps any three collinear points to collinear points, then f is affine. The proof is rather long (see Berger [2] or Samuel [10]).

Given three collinear points a, b, c, where $a \neq c$, we have $b = (1-\beta)a + \beta c$ for some unique β, and we define the *ratio of the sequence a, b, c*, as

$$\mathrm{ratio}(a,b,c) = \frac{\beta}{(1-\beta)} = \frac{\overrightarrow{ab}}{\overrightarrow{bc}},$$

provided that $\beta \neq 1$, i.e., $b \neq c$. When $b = c$, we agree that $\mathrm{ratio}(a,b,c) = \infty$. We warn our readers that other authors define the ratio of a, b, c as $-\mathrm{ratio}(a,b,c) = \frac{\overrightarrow{ba}}{\overrightarrow{bc}}$. Since affine maps preserve barycenters, it is clear that affine maps preserve the ratio of three points.

2.8 Affine Groups

We now take a quick look at the bijective affine maps. Given an affine space E, the set of affine bijections $f\colon E \to E$ is clearly a group, called the *affine group of E*, and denoted by $\mathbf{GA}(E)$. Recall that the group of bijective linear maps of the vector space $\overrightarrow{E}$ is denoted by $\mathbf{GL}(\overrightarrow{E})$. Then, the map $f \mapsto \overrightarrow{f}$ defines a group homomorphism $L\colon \mathbf{GA}(E) \to \mathbf{GL}(\overrightarrow{E})$. The kernel of this map is the set of translations on E.

The subset of all linear maps of the form $\lambda\,\mathrm{id}_{\overrightarrow{E}}$, where $\lambda \in \mathbb{R} - \{0\}$, is a subgroup of $\mathbf{GL}(\overrightarrow{E})$, and is denoted by $\mathbb{R}^*\mathrm{id}_{\overrightarrow{E}}$ (where $\lambda\,\mathrm{id}_{\overrightarrow{E}}(u) = \lambda u$, and $\mathbb{R}^* = \mathbb{R} - \{0\}$). The subgroup $\mathbf{DIL}(E) = L^{-1}(\mathbb{R}^*\mathrm{id}_{\overrightarrow{E}})$ of $\mathbf{GA}(E)$ is particularly interesting. It turns out that it is the disjoint union of the translations and of the dilatations of ratio $\lambda \neq 1$. The elements of $\mathbf{DIL}(E)$ are called *affine dilatations*.

Given any point $a \in E$, and any scalar $\lambda \in \mathbb{R}$, a *dilatation or central dilatation (or homothety) of center a and ratio* λ is a map $H_{a,\lambda}$ defined such that

$$H_{a,\lambda}(x) = a + \lambda\overrightarrow{ax},$$

for every $x \in E$.

Remark: The terminology does not seem to be universally agreed upon. The terms *affine dilatation* and *central dilatation* are used by Pedoe [9]. Snapper and Troyer use the term *dilation* for an affine dilatation and *magnification* for a central dilatation [11]. Samuel uses *homothety* for a central dilatation, a direct translation of the French “homothétie” [10]. Since dilation is shorter than dilatation and somewhat easier to pronounce, perhaps we should use that!

Observe that $H_{a,\lambda}(a) = a$, and when $\lambda \neq 0$ and $x \neq a$, $H_{a,\lambda}(x)$ is on the line defined by a and x, and is obtained by “scaling” $\overrightarrow{ax}$ by λ.

Figure 2.12 shows the effect of a central dilatation of center d. The triangle (a,b,c) is magnified to the triangle (a',b',c'). Note how every line is mapped to a parallel line.

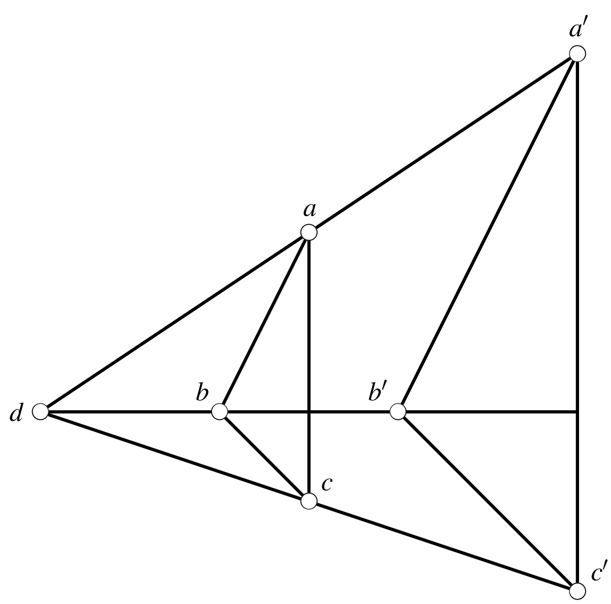

Fig. 2.12 The effect of a central dilatation.

When $\lambda = 1$, $H_{a,1}$ is the identity. Note that $\overrightarrow{H_{a,\lambda}} = \lambda\,\mathrm{id}_{\overrightarrow{E}}$. When $\lambda \neq 0$, it is clear that $H_{a,\lambda}$ is an affine bijection. It is immediately verified that

$$H_{a,\lambda} \circ H_{a,\mu} = H_{a,\lambda\mu}.$$

We have the following useful result.

Lemma 2.8. *Given any affine space E, for any affine bijection $f \in \mathbf{GA}(E)$, if $\overrightarrow{f} = \lambda\,\mathrm{id}_{\overrightarrow{E}}$, for some $\lambda \in \mathbb{R}^*$ with $\lambda \neq 1$, then there is a unique point $c \in E$ such that $f = H_{c,\lambda}$.*

Proof. The proof is straightforward, and is omitted. It is also given in Gallier [5]. $\square$

Clearly, if $\overrightarrow{f} = \mathrm{id}_{\overrightarrow{E}}$, the affine map f is a translation. Thus, the group of affine dilatations $\mathbf{DIL}(E)$ is the disjoint union of the translations and of the dilatations of ratio $\lambda \neq 0, 1$. Affine dilatations can be given a purely geometric characterization.

Another point worth mentioning is that affine bijections preserve the ratio of volumes of parallelotopes. Indeed, given any basis $B = (u_1, \ldots, u_m)$ of the vector space $\overrightarrow{E}$ associated with the affine space E, given any $m+1$ affinely independent points $(a_0, \ldots, a_m)$, we can compute the determinant $\det_B(\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m})$ w.r.t. the basis B. For any bijective affine map $f\colon E \to E$, since

$$\det_B\Big(\overrightarrow{f}(\overrightarrow{a_0a_1}), \ldots, \overrightarrow{f}(\overrightarrow{a_0a_m})\Big) = \det\big(\overrightarrow{f}\big)\det_B(\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m})$$

and the determinant of a linear map is intrinsic (i.e., depends only on $\overrightarrow{f}$, and not on the particular basis B), we conclude that the ratio

$$\frac{\det_B\Big(\overrightarrow{f}(\overrightarrow{a_0a_1}), \ldots, \overrightarrow{f}(\overrightarrow{a_0a_m})\Big)}{\det_B(\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m})} = \det\big(\overrightarrow{f}\big)$$

is independent of the basis B. Since $\det_B(\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m})$ is the volume of the parallelotope spanned by $(a_0, \ldots, a_m)$, where the parallelotope spanned by any point a and the vectors $(u_1, \ldots, u_m)$ has unit volume (see Berger [2], Section 9.12), we see that affine bijections preserve the ratio of volumes of parallelotopes. In fact, this ratio is independent of the choice of the parallelotopes of unit volume. In particular, the affine bijections $f \in \mathbf{GA}(E)$ such that $\det\big(\overrightarrow{f}\big) = 1$ preserve volumes. These affine maps form a subgroup $\mathbf{SA}(E)$ of $\mathbf{GA}(E)$ called the *special affine group of E*. We now take a glimpse at affine geometry.

2.9 Affine Geometry: A Glimpse

In this section we state and prove three fundamental results of affine geometry. Roughly speaking, affine geometry is the study of properties invariant under affine bijections. We now prove one of the oldest and most basic results of affine geometry, the theorem of Thales.

Lemma 2.9. *Given any affine space E, if H_1, H_2, H_3 are any three distinct parallel hyperplanes, and A and B are any two lines not parallel to H_i, letting $a_i = H_i \cap A$ and $b_i = H_i \cap B$, then the following ratios are equal:*

$$\frac{\overrightarrow{a_1a_3}}{\overrightarrow{a_1a_2}} = \frac{\overrightarrow{b_1b_3}}{\overrightarrow{b_1b_2}} = \rho.$$

Conversely, for any point d on the line A, if $\frac{\overrightarrow{a_1d}}{\overrightarrow{a_1a_2}} = \rho$, then $d = a_3$.

Proof. Figure 2.13 illustrates the theorem of Thales. We sketch a proof, leaving the details as an exercise. Since H_1, H_2, H_3 are parallel, they have the same direction $\overrightarrow{H}$, a hyperplane in $\overrightarrow{E}$. Let $u \in \overrightarrow{E} - \overrightarrow{H}$ be any nonnull vector such that $A = a_1 + \mathbb{R}u$. Since A is not parallel to H, we have $\overrightarrow{E} = \overrightarrow{H} \oplus \mathbb{R}u$, and thus we can define the linear map $p\colon \overrightarrow{E} \to \mathbb{R}u$, the projection on $\mathbb{R}u$ parallel to $\overrightarrow{H}$. This linear map induces an affine map $f\colon E \to A$, by defining f such that

$$f(b_1 + w) = a_1 + p(w),$$

for all $w \in \overrightarrow{E}$. Clearly, $f(b_1) = a_1$, and since H_1, H_2, H_3 all have direction $\overrightarrow{H}$, we also have $f(b_2) = a_2$ and $f(b_3) = a_3$. Since f is affine, it preserves ratios, and thus

$$\frac{\overrightarrow{a_1a_3}}{\overrightarrow{a_1a_2}} = \frac{\overrightarrow{b_1b_3}}{\overrightarrow{b_1b_2}}.$$

The converse is immediate. $\square$

We also have the following simple lemma, whose proof is left as an easy exercise.

Lemma 2.10. *Given any affine space E, given any two distinct points $a, b \in E$, and for any affine dilatation f different from the identity, if $a' = f(a)$, $D = \langle a, b\rangle$ is the line passing through a and b, and D' is the line parallel to D and passing through a', the following are equivalent:*

(i) $b' = f(b)$;
(ii) *If f is a translation, then b' is the intersection of D' with the line parallel to $\langle a, a'\rangle$ passing through b;*
If f is a dilatation of center c, then $b' = D' \cap \langle c, b\rangle$.

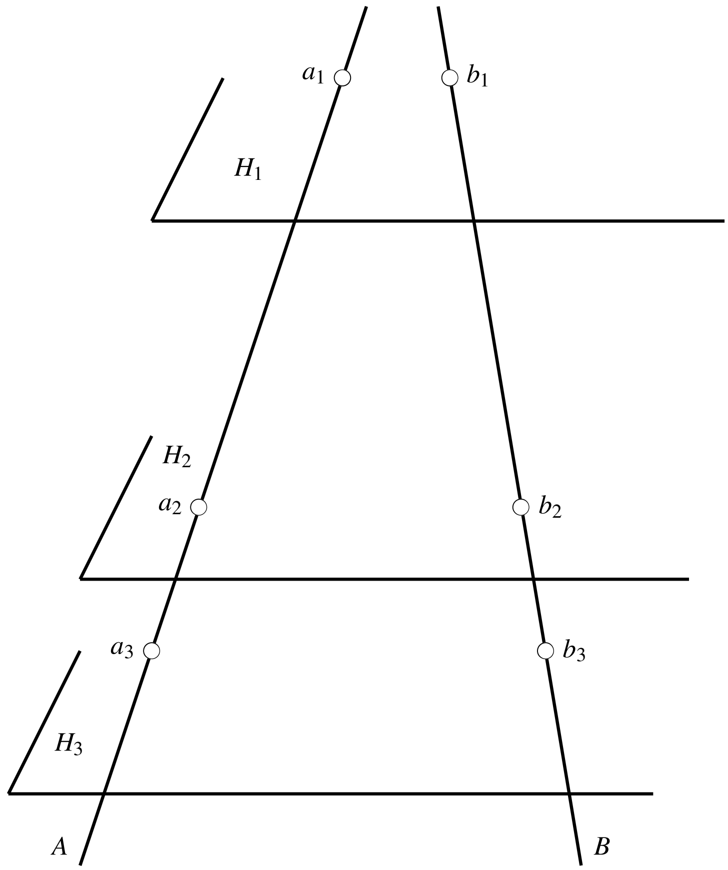

Fig. 2.13 The theorem of Thales.

The first case is the parallelogram law, and the second case follows easily from Thales' theorem.

We are now ready to prove two classical results of affine geometry, Pappus's theorem and Desargues's theorem. Actually, these results are theorems of projective geometry, and we are stating affine versions of these important results. There are stronger versions that are best proved using projective geometry.

Lemma 2.11. *Given any affine plane E, any two distinct lines D and D', then for any distinct points a,b,c on D and a',b',c' on D', if a,b,c,a', b', c' are distinct from the intersection of D and D' (if D and D' intersect) and if the lines $\langle a,b'\rangle$ and $\langle a',b\rangle$ are parallel, and the lines $\langle b,c'\rangle$ and $\langle b',c\rangle$ are parallel, then the lines $\langle a,c'\rangle$ and $\langle a',c\rangle$ are parallel.*

Proof. Pappus's theorem is illustrated in Figure 2.14. If D and D' are not parallel, let d be their intersection. Let f be the dilatation of center d such that $f(a)=b$, and let g be the dilatation of center d such that $g(b)=c$. Since the lines $\langle a,b'\rangle$ and $\langle a',b\rangle$ are parallel, and the lines $\langle b,c'\rangle$ and $\langle b',c\rangle$ are parallel, by Lemma 2.10 we

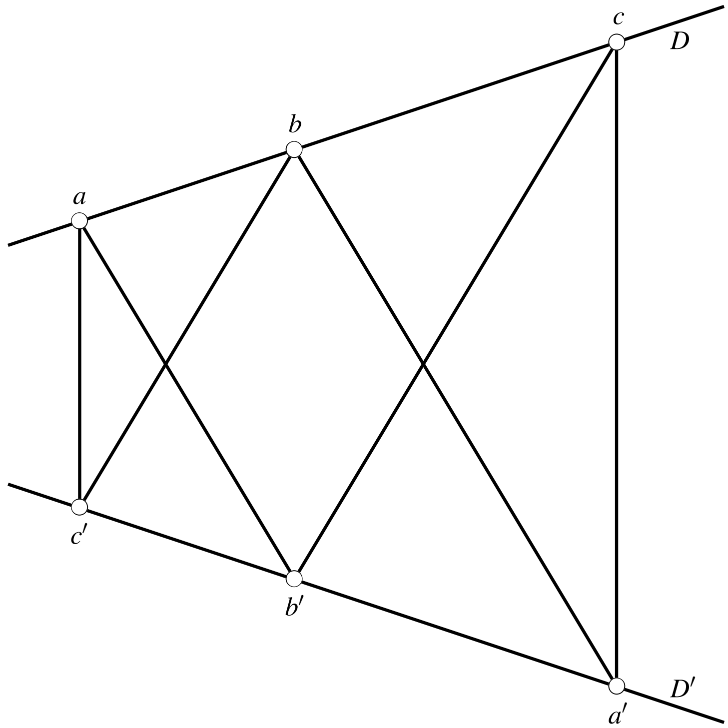

Fig. 2.14 Pappus's theorem (affine version).

have $a' = f(b')$ and $b' = g(c')$. However, we observed that dilatations with the same center commute, and thus $f \circ g = g \circ f$, and thus, letting $h = g \circ f$, we get $c = h(a)$ and $a' = h(c')$. Again, by Lemma 2.10, the lines $\langle a, c' \rangle$ and $\langle a', c \rangle$ are parallel. If D and D' are parallel, we use translations instead of dilatations. $\square$

There is a converse to Pappus's theorem, which yields a fancier version of Pappus's theorem, but it is easier to prove it using projective geometry. It should be noted that in axiomatic presentations of projective geometry, Pappus's theorem is equivalent to the commutativity of the ground field K (in the present case, $K = \mathbb{R}$). We now prove an affine version of Desargues's theorem.

Lemma 2.12. *Given any affine space E, and given any two triangles (a, b, c) and (a', b', c'), where a, b, c, a', b', c' are all distinct, if $\langle a, b \rangle$ and $\langle a', b' \rangle$ are parallel and $\langle b, c \rangle$ and $\langle b', c' \rangle$ are parallel, then $\langle a, c \rangle$ and $\langle a', c' \rangle$ are parallel iff the lines $\langle a, a' \rangle$, $\langle b, b' \rangle$, and $\langle c, c' \rangle$ are either parallel or concurrent (i.e., intersect in a common point).*

Proof. We prove half of the lemma, the direction in which it is assumed that $\langle a, c \rangle$ and $\langle a', c' \rangle$ are parallel, leaving the converse as an exercise. Since the lines $\langle a, b \rangle$ and $\langle a', b' \rangle$ are parallel, the points a, b, a', b' are coplanar. Thus, either $\langle a, a' \rangle$ and $\langle b, b' \rangle$ are parallel, or they have some intersection d. We consider the second case where they intersect, leaving the other case as an easy exercise. Let f be the dilatation of center d such that $f(a) = a'$. By Lemma 2.10, we get $f(b) = b'$. If $f(c) = c''$, again by Lemma 2.10 twice, the lines $\langle b, c\rangle$ and $\langle b', c''\rangle$ are parallel, and the lines $\langle a, c\rangle$ and $\langle a', c''\rangle$ are parallel. From this it follows that $c'' = c'$. Indeed, recall that $\langle b, c\rangle$ and $\langle b', c'\rangle$ are parallel, and similarly $\langle a, c\rangle$ and $\langle a', c'\rangle$ are parallel. Thus, the lines $\langle b', c''\rangle$ and $\langle b', c'\rangle$ are identical, and similarly the lines $\langle a', c''\rangle$ and $\langle a', c'\rangle$ are identical. Since $\overrightarrow{a'c'}$ and $\overrightarrow{b'c'}$ are linearly independent, these lines have a unique intersection, which must be $c'' = c'$.

The direction where it is assumed that the lines $\langle a, a'\rangle$, $\langle b, b'\rangle$ and $\langle c, c'\rangle$, are either parallel or concurrent is left as an exercise (in fact, the proof is quite similar). $\square$

Desargues's theorem is illustrated in Figure 2.15.

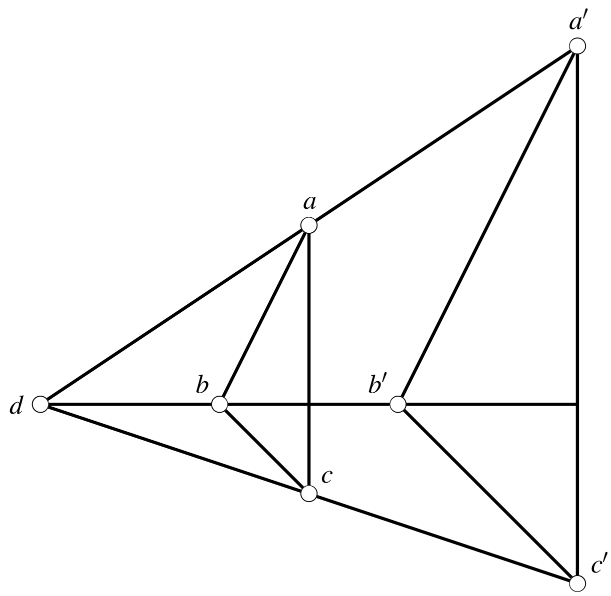

Fig. 2.15 Desargues's theorem (affine version).

There is a fancier version of Desargues's theorem, but it is easier to prove it using projective geometry. It should be noted that in axiomatic presentations of projective geometry, Desargues's theorem is related to the associativity of the ground field K (in the present case, $K = \mathbb{R}$). Also, Desargues's theorem yields a geometric characterization of the affine dilatations. An affine dilatation f on an affine space E is a bijection that maps every line D to a line $f(D)$ parallel to D. We leave the proof as an exercise.

2.10 Affine Hyperplanes

We now consider affine forms and affine hyperplanes. In Section 2.5 we observed that the set L of solutions of an equation

$$ax + by = c$$

is an affine subspace of $\mathbb{A}^2$ of dimension 1, in fact, a line (provided that a and b are not both null). It would be equally easy to show that the set P of solutions of an equation

$$ax + by + cz = d$$

is an affine subspace of $\mathbb{A}^3$ of dimension 2, in fact, a plane (provided that a, b, c are not all null). More generally, the set H of solutions of an equation

$$\lambda_1 x_1 + \cdots + \lambda_m x_m = \mu$$

is an affine subspace of $\mathbb{A}^m$, and if $\lambda_1, \ldots, \lambda_m$ are not all null, it turns out that it is a subspace of dimension $m-1$ called a *hyperplane*.

We can interpret the equation

$$\lambda_1 x_1 + \cdots + \lambda_m x_m = \mu$$

in terms of the map $f \colon \mathbb{R}^m \to \mathbb{R}$ defined such that

$$f(x_1, \ldots, x_m) = \lambda_1 x_1 + \cdots + \lambda_m x_m - \mu$$

for all $(x_1, \ldots, x_m) \in \mathbb{R}^m$. It is immediately verified that this map is affine, and the set H of solutions of the equation

$$\lambda_1 x_1 + \cdots + \lambda_m x_m = \mu$$

is the *null set*, or *kernel,* of the affine map $f \colon \mathbb{A}^m \to \mathbb{R}$, in the sense that

$$H = f^{-1}(0) = \{x \in \mathbb{A}^m \mid f(x) = 0\},$$

where $x = (x_1, \ldots, x_m)$.

Thus, it is interesting to consider *affine forms*, which are just affine maps $f \colon E \to \mathbb{R}$ from an affine space to $\mathbb{R}$. Unlike linear forms f^*, for which $\operatorname{Ker} f^*$ is never empty (since it always contains the vector 0), it is possible that $f^{-1}(0) = \emptyset$ for an affine form f. Given an affine map $f \colon E \to \mathbb{R}$, we also denote $f^{-1}(0)$ by $\operatorname{Ker} f$, and we call it the *kernel* of f. Recall that an (affine) hyperplane is an affine subspace of codimension 1. The relationship between affine hyperplanes and affine forms is given by the following lemma.

Lemma 2.13. *Let E be an affine space. The following properties hold:*

(a) Given any nonconstant affine form $f\colon E \to \mathbb{R}$, its kernel $H = \mathrm{Ker}\, f$ is a hyperplane.
(b) For any hyperplane H in E, there is a nonconstant affine form $f\colon E \to \mathbb{R}$ such that $H = \mathrm{Ker}\, f$. For any other affine form $g\colon E \to \mathbb{R}$ such that $H = \mathrm{Ker}\, g$, there is some $\lambda \in \mathbb{R}$ such that $g = \lambda f$ (with $\lambda \neq 0$).
(c) Given any hyperplane H in E and any (nonconstant) affine form $f\colon E \to \mathbb{R}$ such that $H = \mathrm{Ker}\, f$, every hyperplane H' parallel to H is defined by a nonconstant affine form g such that $g(a) = f(a) - \lambda$, for all $a \in E$ and some $\lambda \in \mathbb{R}$.

Proof. The proof is straightforward, and is omitted. It is also given in Gallier [5]. $\square$

When E is of dimension n, given an affine frame $(a_0, (u_1, \ldots, u_n))$ of E with origin a_0, recall from Definition 2.5 that every point of E can be expressed uniquely as $x = a_0 + x_1 u_1 + \cdots + x_n u_n$, where $(x_1, \ldots, x_n)$ are the *coordinates* of x with respect to the affine frame $(a_0, (u_1, \ldots, u_n))$.

Also recall that every linear form f^* is such that $f^*(x) = \lambda_1 x_1 + \cdots + \lambda_n x_n$, for every $x = x_1 u_1 + \cdots + x_n u_n$ and some $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Since an affine form $f\colon E \to \mathbb{R}$ satisfies the property $f(a_0 + x) = f(a_0) + \overrightarrow{f}(x)$, denoting $f(a_0 + x)$ by $f(x_1, \ldots, x_n)$, we see that we have

$$f(x_1, \ldots, x_n) = \lambda_1 x_1 + \cdots + \lambda_n x_n + \mu,$$

where $\mu = f(a_0) \in \mathbb{R}$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Thus, a hyperplane is the set of points whose coordinates $(x_1, \ldots, x_n)$ satisfy the (affine) equation

$$\lambda_1 x_1 + \cdots + \lambda_n x_n + \mu = 0.$$

2.11 Intersection of Affine Spaces

In this section we take a closer look at the intersection of affine subspaces. This subsection can be omitted at first reading.

First, we need a result of linear algebra. Given a vector space E and any two subspaces M and N, there are several interesting linear maps. We have the canonical injections $i\colon M \to M + N$ and $j\colon N \to M + N$, the canonical injections $in_1\colon M \to M \oplus N$ and $in_2\colon N \to M \oplus N$, and thus, injections $f\colon M \cap N \to M \oplus N$ and $g\colon M \cap N \to M \oplus N$, where f is the composition of the inclusion map from $M \cap N$ to M with in_1, and g is the composition of the inclusion map from $M \cap N$ to N with in_2. Then, we have the maps $f + g\colon M \cap N \to M \oplus N$, and $i - j\colon M \oplus N \to M + N$.

Lemma 2.14. *Given a vector space E and any two subspaces M and N, with the definitions above,*

$$0 \longrightarrow M \cap N \xrightarrow{f+g} M \oplus N \xrightarrow{i-j} M + N \longrightarrow 0$$

is a short exact sequence, which means that $f+g$ is injective, $i-j$ is surjective, and that $\mathrm{Im}\,(f+g)=\mathrm{Ker}\,(i-j)$*. As a consequence, we have the Grassmann relation*

$$\dim(M)+\dim(N)=\dim(M+N)+\dim(M\cap N).$$

Proof. It is obvious that $i-j$ is surjective and that $f+g$ is injective. Assume that $(i-j)(u+v)=0$, where $u\in M$, and $v\in N$. Then, $i(u)=j(v)$, and thus, by definition of i and j, there is some $w\in M\cap N$, such that $i(u)=j(v)=w\in M\cap N$. By definition of f and g, $u=f(w)$ and $v=g(w)$, and thus $\mathrm{Im}\,(f+g)=\mathrm{Ker}\,(i-j)$, as desired. The second part of the lemma follows from standard results of linear algebra (see Artin [1], Strang [12], or Lang [8]). $\square$

We now prove a simple lemma about the intersection of affine subspaces.

Lemma 2.15. *Given any affine space E, for any two nonempty affine subspaces M and N, the following facts hold:*

(1) $M\cap N\neq\emptyset$ iff $\overrightarrow{ab}\in\overrightarrow{M}+\overrightarrow{N}$ for some $a\in M$ and some $b\in N$.
(2) $M\cap N$ consists of a single point iff $\overrightarrow{ab}\in\overrightarrow{M}+\overrightarrow{N}$ for some $a\in M$ and some $b\in N$, and $\overrightarrow{M}\cap\overrightarrow{N}=\{0\}$.
(3) If S is the least affine subspace containing M and N, then $\overrightarrow{S}=\overrightarrow{M}+\overrightarrow{N}+K\overrightarrow{ab}$ (the vector space $\overrightarrow{E}$ is defined over the field K).

Proof. (1) Pick any $a\in M$ and any $b\in N$, which is possible, since M and N are nonempty. Since $\overrightarrow{M}=\{\overrightarrow{ax}\mid x\in M\}$ and $\overrightarrow{N}=\{\overrightarrow{by}\mid y\in N\}$, if $M\cap N\neq\emptyset$, for any $c\in M\cap N$ we have $\overrightarrow{ab}=\overrightarrow{ac}-\overrightarrow{bc}$, with $\overrightarrow{ac}\in\overrightarrow{M}$ and $\overrightarrow{bc}\in\overrightarrow{N}$, and thus, $\overrightarrow{ab}\in\overrightarrow{M}+\overrightarrow{N}$. Conversely, assume that $\overrightarrow{ab}\in\overrightarrow{M}+\overrightarrow{N}$ for some $a\in M$ and some $b\in N$. Then $\overrightarrow{ab}=\overrightarrow{ax}+\overrightarrow{by}$, for some $x\in M$ and some $y\in N$. But we also have

$$\overrightarrow{ab}=\overrightarrow{ax}+\overrightarrow{xy}+\overrightarrow{yb},$$

and thus we get $0=\overrightarrow{xy}+\overrightarrow{yb}-\overrightarrow{by}$, that is, $\overrightarrow{xy}=2\overrightarrow{by}$. Thus, b is the middle of the segment $[x,y]$, and since $\overrightarrow{yx}=2\overrightarrow{yb}$, $x=2b-y$ is the barycenter of the weighted points $(b,2)$ and $(y,-1)$. Thus x also belongs to N, since N being an affine subspace, it is closed under barycenters. Thus, $x\in M\cap N$, and $M\cap N\neq\emptyset$.

(2) Note that in general, if $M\cap N\neq\emptyset$, then

$$\overrightarrow{M\cap N}=\overrightarrow{M}\cap\overrightarrow{N},$$

because

$$\overrightarrow{M\cap N}=\{\overrightarrow{ab}\mid a,b\in M\cap N\}=\{\overrightarrow{ab}\mid a,b\in M\}\cap\{\overrightarrow{ab}\mid a,b\in N\}=\overrightarrow{M}\cap\overrightarrow{N}.$$

Since $M\cap N=c+\overrightarrow{M\cap N}$ for any $c\in M\cap N$, we have

$$M \cap N = c + \overrightarrow{M} \cap \overrightarrow{N} \quad \text{for any } c \in M \cap N.$$

From this it follows that if $M \cap N \neq \emptyset$, then $M \cap N$ consists of a single point iff $\overrightarrow{M} \cap \overrightarrow{N} = \{0\}$. This fact together with what we proved in (1) proves (2).

(3) This is left as an easy exercise. □

Remarks:

(1) The proof of Lemma 2.15 shows that if $M \cap N \neq \emptyset$, then $\overrightarrow{ab} \in \overrightarrow{M} + \overrightarrow{N}$ for all $a \in M$ and all $b \in N$.

(2) Lemma 2.15 implies that for any two nonempty affine subspaces M and N, if $\overrightarrow{E} = \overrightarrow{M} \oplus \overrightarrow{N}$, then $M \cap N$ consists of a single point. Indeed, if $\overrightarrow{E} = \overrightarrow{M} \oplus \overrightarrow{N}$, then $\overrightarrow{ab} \in \overrightarrow{E}$ for all $a \in M$ and all $b \in N$, and since $\overrightarrow{M} \cap \overrightarrow{N} = \{0\}$, the result follows from part (2) of the lemma.

We can now state the following lemma.

Lemma 2.16. *Given an affine space E and any two nonempty affine subspaces M and N, if S is the least affine subspace containing M and N, then the following properties hold:*

(1) If $M \cap N = \emptyset$, then

$$\dim(M) + \dim(N) < \dim(E) + \dim(\overrightarrow{M} + \overrightarrow{N})$$

and

$$\dim(S) = \dim(M) + \dim(N) + 1 - \dim(\overrightarrow{M} \cap \overrightarrow{N}).$$

(2) If $M \cap N \neq \emptyset$, then

$$\dim(S) = \dim(M) + \dim(N) - \dim(M \cap N).$$

Proof. The proof is not difficult, using Lemma 2.15 and Lemma 2.14, but we leave it as an exercise. □

2.12 Problems

2.1. Given a triangle (a,b,c), give a geometric construction of the barycenter of the weighted points $(a, \frac{1}{4})$, $(b, \frac{1}{4})$, and $(c, \frac{1}{2})$. Give a geometric construction of the barycenter of the weighted points $(a, \frac{3}{2})$, $(b, \frac{3}{2})$, and $(c, -2)$.

2.2. Given a tetrahedron (a,b,c,d) and any two distinct points $x, y \in \{a,b,c,d\}$, let let $m_{x,y}$ be the middle of the edge (x,y). Prove that the barycenter g of the weighted points $(a, \frac{1}{4})$, $(b, \frac{1}{4})$, $(c, \frac{1}{4})$, and $(d, \frac{1}{4})$ is the common intersection of the line segments $(m_{a,b}, m_{c,d})$, $(m_{a,c}, m_{b,d})$, and $(m_{a,d}, m_{b,c})$. Show that if g_d is the barycenter of the weighted points $(a, \frac{1}{3}), (b, \frac{1}{3}), (c, \frac{1}{3})$, then g is the barycenter of $(d, \frac{1}{4})$ and $(g_d, \frac{3}{4})$.

2.3. Let E be a nonempty set, and $\overrightarrow{E}$ a vector space and assume that there is a function $\Phi \colon E \times E \to \overrightarrow{E}$, such that if we denote $\Phi(a,b)$ by $\overrightarrow{ab}$, the following properties hold:

(1) $\overrightarrow{ab} + \overrightarrow{bc} = \overrightarrow{ac}$, for all $a, b, c \in E$;

(2) For every $a \in E$, the map $\Phi_a \colon E \to \overrightarrow{E}$ defined such that for every $b \in E$, $\Phi_a(b) = \overrightarrow{ab}$, is a bijection.

Let $\Psi_a \colon \overrightarrow{E} \to E$ be the inverse of $\Phi_a \colon E \to \overrightarrow{E}$.

Prove that the function $+ \colon E \times \overrightarrow{E} \to E$ defined such that

$$a + u = \Psi_a(u)$$

for all $a \in E$ and all $u \in \overrightarrow{E}$ makes $(E, \overrightarrow{E}, +)$ into an affine space.

Note. We showed in the text that an affine space $(E, \overrightarrow{E}, +)$ satisfies the properties stated above. Thus, we obtain an equivalent characterization of affine spaces.

2.4. Given any three points a, b, c in the affine plane $\mathbb{A}^2$, letting (a_1, a_2), (b_1, b_2), and (c_1, c_2) be the coordinates of a, b, c, with respect to the standard affine frame for $\mathbb{A}^2$, prove that a, b, c are collinear iff

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ 1 & 1 & 1 \end{vmatrix} = 0,$$

i.e., the determinant is null.

Letting (a_0, a_1, a_2), (b_0, b_1, b_2), and (c_0, c_1, c_2) be the barycentric coordinates of a, b, c with respect to the standard affine frame for $\mathbb{A}^2$, prove that a, b, c are collinear iff

$$\begin{vmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix} = 0.$$

Given any four points a, b, c, d in the affine space $\mathbb{A}^3$, letting (a_1, a_2, a_3), (b_1, b_2, b_3), (c_1, c_2, c_3), and (d_1, d_2, d_3) be the coordinates of a, b, c, d, with respect to the standard affine frame for $\mathbb{A}^3$, prove that a, b, c, d are coplanar iff

$$\begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ 1 & 1 & 1 & 1 \end{vmatrix} = 0,$$

i.e., the determinant is null.

Letting (a_0, a_1, a_2, a_3), (b_0, b_1, b_2, b_3), (c_0, c_1, c_2, c_3), and (d_0, d_1, d_2, d_3) be the barycentric coordinates of a, b, c, d, with respect to the standard affine frame for $\mathbb{A}^3$, prove that a, b, c, d are coplanar iff

$$\begin{vmatrix} a_0 & b_0 & c_0 & d_0 \\ a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \end{vmatrix} = 0.$$

2.5. The function $f\colon \mathbb{A} \to \mathbb{A}^3$ given by

$$t \mapsto (t, t^2, t^3)$$

defines what is called a *twisted cubic* curve. Given any four pairwise distinct values t_1, t_2, t_3, t_4, prove that the points $f(t_1), f(t_2), f(t_3)$, and $f(t_4)$ are not coplanar.
Hint. Have you heard of the Vandermonde determinant?

2.6. For any two distinct points $a, b \in \mathbb{A}^2$ of barycentric coordinates (a_0, a_1, a_2) and (b_0, b_1, b_2) with respect to any given affine frame (O, i, j), show that the equation of the line $\langle a, b\rangle$ determined by a and b is

$$\begin{vmatrix} a_0 & b_0 & x \\ a_1 & b_1 & y \\ a_2 & b_2 & z \end{vmatrix} = 0,$$

or, equivalently,

$$(a_1 b_2 - a_2 b_1)x + (a_2 b_0 - a_0 b_2)y + (a_0 b_1 - a_1 b_0)z = 0,$$

where (x, y, z) are the barycentric coordinates of the generic point on the line $\langle a, b\rangle$.

Prove that the equation of a line in barycentric coordinates is of the form

$$ux + vy + wz = 0,$$

where $u \neq v$ or $v \neq w$ or $u \neq w$. Show that two equations

$$ux + vy + wz = 0 \quad \text{and} \quad u'x + v'y + w'z = 0$$

represent the same line in barycentric coordinates iff $(u', v', w') = \lambda(u, v, w)$ for some $\lambda \in \mathbb{R}$ (with $\lambda \neq 0$).

A triple (u, v, w) where $u \neq v$ or $v \neq w$ or $u \neq w$ is called a system of *tangential coordinates* of the line defined by the equation

$$ux + vy + wz = 0.$$

2.7. Given two lines D and D' in $\mathbb{A}^2$ defined by tangential coordinates (u, v, w) and (u', v', w') (as defined in Problem 2.6), let

$$d = \begin{vmatrix} u & v & w \\ u' & v' & w' \\ 1 & 1 & 1 \end{vmatrix} = vw' - wv' + wu' - uw' + uv' - vu'.$$

(a) Prove that D and D' have a unique intersection point iff $d \neq 0$, and that when it exists, the barycentric coordinates of this intersection point are

$$\frac{1}{d}(vw' - wv', wu' - uw', uv' - vu').$$

(b) Letting (O, i, j) be any affine frame for $\mathbb{A}^2$, recall that when $x + y + z = 0$, for any point a, the vector

$$x\overrightarrow{aO} + y\overrightarrow{ai} + z\overrightarrow{aj}$$

is independent of a and equal to

$$y\overrightarrow{Oi} + z\overrightarrow{Oj} = (y, z).$$

The triple (x, y, z) such that $x + y + z = 0$ is called the *barycentric coordinates* of the vector $y\overrightarrow{Oi} + z\overrightarrow{Oj}$ w.r.t. the affine frame (O, i, j).

Given any affine frame (O, i, j), prove that for $u \neq v$ or $v \neq w$ or $u \neq w$, the line of equation

$$ux + vy + wz = 0$$

in barycentric coordinates (x, y, z) (where $x + y + z = 1$) has for direction the set of vectors of barycentric coordinates (x, y, z) such that

$$ux + vy + wz = 0$$

(where $x + y + z = 0$).

Prove that D and D' are parallel iff $d = 0$. In this case, if $D \neq D'$, show that the common direction of D and D' is defined by the vector of barycentric coordinates

$$(vw' - wv', wu' - uw', uv' - vu').$$

(c) Given three lines D, D', and D'', at least two of which are distinct and defined by tangential coordinates (u, v, w), (u', v', w'), and (u'', v'', w''), prove that D, D', and D'' are parallel or have a unique intersection point iff

$$\begin{vmatrix} u & v & w \\ u' & v' & w' \\ u'' & v'' & w'' \end{vmatrix} = 0.$$

2.8. Let (A, B, C) be a triangle in $\mathbb{A}^2$. Let M, N, P be three points respectively on the lines BC, CA, and AB, of barycentric coordinates $(0, m', m'')$, $(n, 0, n'')$, and $(p, p', 0)$, w.r.t. the affine frame (A, B, C).

(a) Assuming that $M \neq C$, $N \neq A$, and $P \neq B$, i.e., $m'n''p \neq 0$, show that

$$\frac{\overrightarrow{MB}}{\overrightarrow{MC}}\frac{\overrightarrow{NC}}{\overrightarrow{NA}}\frac{\overrightarrow{PA}}{\overrightarrow{PB}} = -\frac{m''np'}{m'n''p}.$$

(b) Prove *Menelaus's theorem*: The points M, N, P are collinear iff

$$m''np' + m'n''p = 0.$$

When $M \neq C$, $N \neq A$, and $P \neq B$, this is equivalent to

$$\frac{\overrightarrow{MB}}{\overrightarrow{MC}}\frac{\overrightarrow{NC}}{\overrightarrow{NA}}\frac{\overrightarrow{PA}}{\overrightarrow{PB}} = 1.$$

(c) Prove *Ceva's theorem*: The lines AM, BN, CP have a unique intersection point or are parallel iff

$$m''np' - m'n''p = 0.$$

When $M \neq C$, $N \neq A$, and $P \neq B$, this is equivalent to

$$\frac{\overrightarrow{MB}}{\overrightarrow{MC}}\frac{\overrightarrow{NC}}{\overrightarrow{NA}}\frac{\overrightarrow{PA}}{\overrightarrow{PB}} = -1.$$

2.9. This problem uses notions and results from Problems 2.6 and 2.7. In view of (a) and (b) of Problem 2.7, it is natural to extend the notion of barycentric coordinates of a point in $\mathbb{A}^2$ as follows. Given any affine frame (a, b, c) in $\mathbb{A}^2$, we will say that the barycentric coordinates (x, y, z) of a point M, where $x + y + z = 1$, are the *normalized barycentric coordinates* of M. Then, any triple (x, y, z) such that $x + y + z \neq 0$ is also called a system of barycentric coordinates for the point of normalized barycentric coordinates

$$\frac{1}{x+y+z}(x, y, z).$$

With this convention, the intersection of the two lines D and D' is either a point or a vector, in both cases of barycentric coordinates

$$(vw' - wv', wu' - uw', uv' - vu').$$

When the above is a vector, we can think of it as a point at infinity (in the direction of the line defined by that vector).

Let (D_0, D_0'), (D_1, D_1'), and (D_2, D_2') be three pairs of six distinct lines, such that the four lines belonging to any union of two of the above pairs are neither parallel nor concurrent (have a common intersection point). If D_0 and D_0' have a unique intersection point, let M be this point, and if D_0 and D_0' are parallel, let M denote a nonnull vector defining the common direction of D_0 and D_0'. In either case, let (m, m', m'') be the barycentric coordinates of M, as explained at the beginning of the problem. We call M the *intersection* of D_0 and D_0'. Similarly, define $N = (n, n', n'')$ as the intersection of D_1 and D_1', and $P = (p, p', p'')$ as the intersection of D_2 and D_2'.

Prove that

$$\begin{vmatrix} m & n & p \\ m' & n' & p' \\ m'' & n'' & p'' \end{vmatrix} = 0$$

iff either

(i) (D_0, D_0'), (D_1, D_1'), and (D_2, D_2') are pairs of parallel lines; or
(ii) the lines of some pair (D_i, D_i') are parallel, each pair (D_j, D_j') (with $j \neq i$) has a unique intersection point, and these two intersection points are distinct and determine a line parallel to the lines of the pair (D_i, D_i'); or
(iii) each pair (D_i, D_i') $(i = 0, 1, 2)$ has a unique intersection point, and these points M, N, P are distinct and collinear.

2.10. Prove the following version of *Desargues's theorem*. Let A, B, C, A', B', C' be six distinct points of $\mathbb{A}^2$. If no three of these points are collinear, then the lines AA', BB', and CC' are parallel or collinear iff the intersection points M, N, P (in the sense of Problem 2.7) of the pairs of lines $(BC, B'C')$, $(CA, C'A')$, and $(AB, A'B')$ are collinear in the sense of Problem 2.9.

2.11. Prove the following version of *Pappus's theorem*. Let D and D' be distinct lines, and let A, B, C and A', B', C' be distinct points respectively on D and D'. If these points are all distinct from the intersection of D and D' (if it exists), then the intersection points (in the sense of Problem 2.7) of the pairs of lines (BC', CB'), (CA', AC'), and (AB', BA') are collinear in the sense of Problem 2.9.

2.12. The purpose of this problem is to prove *Pascal's theorem* for the nondegenerate conics. In the affine plane $\mathbb{A}^2$, a *conic* is the set of points of coordinates (x, y) such that

$$\alpha x^2 + \beta y^2 + 2\gamma xy + 2\delta x + 2\lambda y + \mu = 0,$$

where $\alpha \neq 0$ or $\beta \neq 0$ or $\gamma \neq 0$. We can write the equation of the conic as

$$(x, y, 1) \begin{pmatrix} \alpha & \gamma & \delta \\ \gamma & \beta & \lambda \\ \delta & \lambda & \mu \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0.$$

If we now use barycentric coordinates (x, y, z) (where $x + y + z = 1$), we can write

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Let

$$B = \begin{pmatrix} \alpha & \gamma & \delta \\ \gamma & \beta & \lambda \\ \delta & \lambda & \mu \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

(a) Letting $A = C^\top BC$, prove that the equation of the conic becomes

$$X^\top AX = 0.$$

Prove that A is symmetric, that $\det(A) = \det(B)$, and that $X^\top AX$ is homogeneous of degree 2. The equation $X^\top AX = 0$ is called the *homogeneous equation* of the conic.

We say that a conic of homogeneous equation $X^\top AX = 0$ is *nondegenerate* if $\det(A) \neq 0$, and *degenerate* if $\det(A) = 0$. Show that this condition does not depend on the choice of the affine frame.

(b) Given an affine frame (A,B,C), prove that any conic passing through A,B,C has an equation of the form

$$ayz + bxz + cxy = 0.$$

Prove that a conic containing more than one point is degenerate iff it contains three distinct collinear points. In this case, the conic is the union of two lines.

(c) Prove *Pascal's theorem*. Given any six distinct points A,B,C, A',B', C', if no three of the above points are collinear, then a nondegenerate conic passes through these six points iff the intersection points M,N,P (in the sense of Problem 2.7) of the pairs of lines (BC',CB'), (CA',AC') and (AB',BA') are collinear in the sense of Problem 2.9.

Hint. Use the affine frame (A,B,C), and let (a,a',a''), (b,b',b''), and (c,c',c'') be the barycentric coordinates of A',B',C' respectively, and show that M,N,P have barycentric coordinates

$$(bc, cb', c''b), \quad (c'a, c'a', c''a'), \quad (ab'', a''b', a''b'').$$

2.13. The *centroid* of a triangle (a,b,c) is the barycenter of $(a, \frac{1}{3})$, $(b, \frac{1}{3})$, $(c, \frac{1}{3})$. If an affine map takes the vertices of triangle $\Delta_1 = \{(0,0),(6,0),(0,9)\}$ to the vertices of triangle $\Delta_2 = \{(1,1),(5,4),(3,1)\}$, does it also take the centroid of Δ_1 to the centroid of Δ_2? Justify your answer.

2.14. Let E be an affine space over $\mathbb{R}$, and let $(a_1,\ldots,a_n)$ be any $n \geq 3$ points in E. Let $(\lambda_1,\ldots,\lambda_n)$ be any n scalars in $\mathbb{R}$, with $\lambda_1 + \cdots + \lambda_n = 1$. Show that there must be some i, $1 \leq i \leq n$, such that $\lambda_i \neq 1$. To simplify the notation, assume that $\lambda_1 \neq 1$. Show that the barycenter $\lambda_1 a_1 + \cdots + \lambda_n a_n$ can be obtained by first determining the barycenter b of the $n-1$ points $a_2,\ldots,a_n$ assigned some appropriate weights, and then the barycenter of a_1 and b assigned the weights λ_1 and $\lambda_2 + \cdots + \lambda_n$. From this, show that the barycenter of any $n \geq 3$ points can be determined by repeated computations of barycenters of two points. Deduce from the above that a nonempty subset V of E is an affine subspace iff whenever V contains any two points $x,y \in V$, then V contains the entire line $(1-\lambda)x + \lambda y$, $\lambda \in \mathbb{R}$.

2.15. Assume that K is a field such that $2 = 1+1 \neq 0$, and let E be an affine space over K. In the case where $\lambda_1 + \cdots + \lambda_n = 1$ and $\lambda_i = 1$, for $1 \leq i \leq n$ and $n \geq 3$, show that the barycenter $a_1 + a_2 + \cdots + a_n$ can still be computed by repeated computations of barycenters of two points.

Finally, assume that the field K contains at least three elements (thus, there is some $\mu \in K$ such that $\mu \neq 0$ and $\mu \neq 1$, but $2 = 1 + 1 = 0$ is possible). Prove that the barycenter of any $n \geq 3$ points can be determined by repeated computations of barycenters of two points. Prove that a nonempty subset V of E is an affine subspace iff whenever V contains any two points $x, y \in V$, then V contains the entire line $(1-\lambda)x + \lambda y$, $\lambda \in K$.
Hint. When $2 = 0$, $\lambda_1 + \cdots + \lambda_n = 1$ and $\lambda_i = 1$, for $1 \leq i \leq n$, show that n must be odd, and that the problem reduces to computing the barycenter of three points in two steps involving two barycenters. Since there is some $\mu \in K$ such that $\mu \neq 0$ and $\mu \neq 1$, note that μ^{-1} and $(1-\mu)^{-1}$ both exist, and use the fact that

$$\frac{-\mu}{1-\mu} + \frac{1}{1-\mu} = 1.$$

2.16. (i) Let (a, b, c) be three points in $\mathbb{A}^2$, and assume that (a, b, c) are not collinear. For any point $x \in \mathbb{A}^2$, if $x = \lambda_0 a + \lambda_1 b + \lambda_2 c$, where $(\lambda_0, \lambda_1, \lambda_2)$ are the barycentric coordinates of x with respect to (a, b, c), show that

$$\lambda_0 = \frac{\det(\overrightarrow{xb}, \overrightarrow{bc})}{\det(\overrightarrow{ab}, \overrightarrow{ac})}, \qquad \lambda_1 = \frac{\det(\overrightarrow{ax}, \overrightarrow{ac})}{\det(\overrightarrow{ab}, \overrightarrow{ac})}, \qquad \lambda_2 = \frac{\det(\overrightarrow{ab}, \overrightarrow{ax})}{\det(\overrightarrow{ab}, \overrightarrow{ac})}.$$

Conclude that $\lambda_0, \lambda_1, \lambda_2$ are certain signed ratios of the areas of the triangles (a, b, c), (x, a, b), (x, a, c), and (x, b, c).

(ii) Let (a, b, c) be three points in $\mathbb{A}^3$, and assume that (a, b, c) are not collinear. For any point x in the plane determined by (a, b, c), if $x = \lambda_0 a + \lambda_1 b + \lambda_2 c$, where $(\lambda_0, \lambda_1, \lambda_2)$ are the barycentric coordinates of x with respect to (a, b, c), show that

$$\lambda_0 = \frac{\overrightarrow{xb} \times \overrightarrow{bc}}{\overrightarrow{ab} \times \overrightarrow{ac}}, \qquad \lambda_1 = \frac{\overrightarrow{ax} \times \overrightarrow{ac}}{\overrightarrow{ab} \times \overrightarrow{ac}}, \qquad \lambda_2 = \frac{\overrightarrow{ab} \times \overrightarrow{ax}}{\overrightarrow{ab} \times \overrightarrow{ac}}.$$

Given any point O not in the plane of the triangle (a, b, c), prove that

$$\lambda_1 = \frac{\det(\overrightarrow{Oa}, \overrightarrow{Ox}, \overrightarrow{Oc})}{\det(\overrightarrow{Oa}, \overrightarrow{Ob}, \overrightarrow{Oc})}, \quad \lambda_2 = \frac{\det(\overrightarrow{Oa}, \overrightarrow{Ob}, \overrightarrow{Ox})}{\det(\overrightarrow{Oa}, \overrightarrow{Ob}, \overrightarrow{Oc})},$$

and

$$\lambda_0 = \frac{\det(\overrightarrow{Ox}, \overrightarrow{Ob}, \overrightarrow{Oc})}{\det(\overrightarrow{Oa}, \overrightarrow{Ob}, \overrightarrow{Oc})}.$$

(iii) Let (a, b, c, d) be four points in $\mathbb{A}^3$, and assume that (a, b, c, d) are not coplanar. For any point $x \in \mathbb{A}^3$, if $x = \lambda_0 a + \lambda_1 b + \lambda_2 c + \lambda_3 d$, where $(\lambda_0, \lambda_1, \lambda_2, \lambda_3)$ are the barycentric coordinates of x with respect to (a, b, c, d), show that

$$\lambda_1 = \frac{\det(\overrightarrow{ax}, \overrightarrow{ac}, \overrightarrow{ad})}{\det(\overrightarrow{ab}, \overrightarrow{ac}, \overrightarrow{ad})}, \ \lambda_2 = \frac{\det(\overrightarrow{ab}, \overrightarrow{ax}, \overrightarrow{ad})}{\det(\overrightarrow{ab}, \overrightarrow{ac}, \overrightarrow{ad})}, \ \lambda_3 = \frac{\det(\overrightarrow{ab}, \overrightarrow{ac}, \overrightarrow{ax})}{\det(\overrightarrow{ab}, \overrightarrow{ac}, \overrightarrow{ad})},$$

and

$$\lambda_0 = \frac{\det(\overrightarrow{xb}, \overrightarrow{bc}, \overrightarrow{bd})}{\det(\overrightarrow{ab}, \overrightarrow{ac}, \overrightarrow{ad})}.$$

Conclude that $\lambda_0, \lambda_1, \lambda_2, \lambda_3$ are certain signed ratios of the volumes of the five tetrahedra (a,b,c,d), (x,a,b,c), (x,a,b,d), (x,a,c,d), and (x,b,c,d).

(iv) Let $(a_0, \ldots, a_m)$ be $m+1$ points in $\mathbb{A}^m$, and assume that they are affinely independent. For any point $x \in \mathbb{A}^m$, if $x = \lambda_0 a_0 + \cdots + \lambda_m a_m$, where $(\lambda_0, \ldots, \lambda_m)$ are the barycentric coordinates of x with respect to $(a_0, \ldots, a_m)$, show that

$$\lambda_i = \frac{\det(\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_{i-1}}, \overrightarrow{a_0x}, \overrightarrow{a_0a_{i+1}}, \ldots, \overrightarrow{a_0a_m})}{\det(\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_{i-1}}, \overrightarrow{a_0a_i}, \overrightarrow{a_0a_{i+1}}, \ldots, \overrightarrow{a_0a_m})}$$

for every i, $1 \leq i \leq m$, and

$$\lambda_0 = \frac{\det(\overrightarrow{xa_1}, \overrightarrow{a_1a_2}, \ldots, \overrightarrow{a_1a_m})}{\det(\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_i}, \ldots, \overrightarrow{a_0a_m})}.$$

Conclude that λ_i is the signed ratio of the volumes of the simplexes $(a_0, \ldots, x, \ldots a_m)$ and $(a_0, \ldots, a_i, \ldots a_m)$, where $0 \leq i \leq m$.

2.17. With respect to the standard affine frame for the plane $\mathbb{A}^2$, consider the three geometric transformations f_1, f_2, f_3 defined by

$$\begin{aligned}
x' &= -\frac{1}{4}x - \frac{\sqrt{3}}{4}y + \frac{3}{4}, \quad y' = \frac{\sqrt{3}}{4}x - \frac{1}{4}y + \frac{\sqrt{3}}{4}, \\
x' &= -\frac{1}{4}x + \frac{\sqrt{3}}{4}y - \frac{3}{4}, \quad y' = -\frac{\sqrt{3}}{4}x - \frac{1}{4}y + \frac{\sqrt{3}}{4}, \\
x' &= \frac{1}{2}x, \quad y' = \frac{1}{2}y + \frac{\sqrt{3}}{2}.
\end{aligned}$$

(a) Prove that these maps are affine. Can you describe geometrically what their action is (rotation, translation, scaling)?

(b) Given any polygonal line L, define the following sequence of polygonal lines:

$$\begin{aligned}
S_0 &= L, \\
S_{n+1} &= f_1(S_n) \cup f_2(S_n) \cup f_3(S_n).
\end{aligned}$$

Construct S_1 starting from the line segment $L = ((-1,0),(1,0))$.

Can you figure out what S_n looks like in general? (You may want to write a computer program.) Do you think that S_n has a limit?

2.18. In the plane $\mathbb{A}^2$, with respect to the standard affine frame, a point of coordinates (x,y) can be represented as the complex number $z = x + \mathrm{i}y$. Consider the set of geometric transformations of the form

$$z \mapsto az + b,$$

where a, b are complex numbers such that $a \neq 0$.

(a) Prove that these maps are affine. Describe what these maps do geometrically.

(b) Prove that the above set of maps is a group under composition.

(c) Consider the set of geometric transformations of the form

$$z \mapsto az + b \quad \text{or} \quad z \mapsto a\bar{z} + b,$$

where a, b are complex numbers such that $a \neq 0$, and where $\bar{z} = x - \mathrm{i}y$ if $z = x + \mathrm{i}y$. Describe what these maps do geometrically. Prove that these maps are affine and that this set of maps is a group under composition.

2.19. Given a group G, a subgroup H of G is called a *normal subgroup* of G iff $xHx^{-1} = H$ for all $x \in G$ (where $xHx^{-1} = \{xhx^{-1} \mid h \in H\}$).

(i) Given any two subgroups H and K of a group G, let

$$HK = \{hk \mid h \in H, \, k \in K\}.$$

Prove that every $x \in HK$ can be written in a unique way as $x = hk$ for $h \in H$ and $k \in K$ iff $H \cap K = \{1\}$, where 1 is the identity element of G.

(ii) If H and K are subgroups of G, and H is a normal subgroup of G, prove that HK is a subgroup of G. Furthermore, if $G = HK$ and $H \cap K = \{1\}$, prove that G is isomorphic to $H \times K$ under the multiplication operation

$$(h_1, k_1) \cdot (h_2, k_2) = (h_1 k_1 h_2 k_1^{-1}, \, k_1 k_2).$$

When $G = HK$, where H, K are subgroups of G, H is a normal subgroup of G, and $H \cap K = \{1\}$, we say that G is the *semidirect product of H and K*.

(iii) Let $(E, \overrightarrow{E})$ be an affine space. Recall that the *affine group of E*, denoted by $\mathbf{GA}(E)$, is the set of affine bijections of E, and that the *linear group of $\overrightarrow{E}$*, denoted by $\mathbf{GL}(\overrightarrow{E})$, is the group of bijective linear maps of $\overrightarrow{E}$. The map $f \mapsto \overrightarrow{f}$ defines a group homomorphism $L\colon \mathbf{GA}(E) \to \mathbf{GL}(\overrightarrow{E})$, and the kernel of this map is the set of translations on E, denoted as $T(E)$. Prove that $T(E)$ is a normal subgroup of $\mathbf{GA}(E)$.

(iv) For any $a \in E$, let

$$\mathbf{GA}_a(E) = \{f \in \mathbf{GA}(E) \mid f(a) = a\},$$

the set of affine bijections leaving a fixed. Prove that that $\mathbf{GA}_a(E)$ is a subgroup of $\mathbf{GA}(E)$, and that $\mathbf{GA}_a(E)$ is isomorphic to $\mathbf{GL}(\overrightarrow{E})$. Prove that $\mathbf{GA}(E)$ is isomorphic to the direct product of $T(E)$ and $\mathbf{GA}_a(E)$.

Hint. Note that if $u = \overrightarrow{f(a)a}$ and t_u is the translation associated with the vector u, then $t_u \circ f \in \mathbf{GA}_a(E)$ (where the translation t_u is defined such that $t_u(a) = a + u$ for every $a \in E$).

(v) Given a group G, let $\mathbf{Aut}(G)$ denote the set of homomorphisms $f\colon G \to G$. Prove that the set $\mathbf{Aut}(G)$ is a group under composition (called the *group of* *automorphisms of* G). Given any two groups H and K and a homomorphism $\theta\colon K \to \mathbf{Aut}(H)$, we define $H \times_\theta K$ as the set $H \times K$ under the multiplication operation

$$(h_1,k_1)\cdot(h_2,k_2) = (h_1\theta(k_1)(h_2),\, k_1k_2).$$

Prove that $H \times_\theta K$ is a group.
Hint. The inverse of (h,k) is $(\theta(k^{-1})(h^{-1}),\, k^{-1})$.

Prove that the group $H \times_\theta K$ is the semidirect product of the subgroups $\{(h,1) \mid h \in H\}$ and $\{(1,k) \mid k \in K\}$. The group $H \times_\theta K$ is also called the *semidirect product of* H *and* K *relative to* θ.
Note. It is natural to identify $\{(h,1) \mid h \in H\}$ with H and $\{(1,k) \mid k \in K\}$ with K.

If G is the semidirect product of two subgroups H and K as defined in (ii), prove that the map $\gamma\colon K \to \mathbf{Aut}(H)$ defined by conjugation such that

$$\gamma(k)(h) = khk^{-1}$$

is a homomorphism, and that G is isomorphic to $H \times_\gamma K$.

(vi) Define the map $\theta\colon \mathbf{GL}(\overrightarrow{E}) \to \mathbf{Aut}(\overrightarrow{E})$ as follows: $\theta(f) = f$, where $f \in \mathbf{GL}(\overrightarrow{E})$ (note that θ can be viewed as an inclusion map). Prove that $\mathbf{GA}(E)$ is isomorphic to the semidirect product $\overrightarrow{E} \times_\theta \mathbf{GL}(\overrightarrow{E})$.

(vii) Let $\mathbf{SL}(\overrightarrow{E})$ be the subgroup of $\mathbf{GL}(\overrightarrow{E})$ consisting of the linear maps such that $\det(f) = 1$ (the *special linear group of* $\overrightarrow{E}$), and let $\mathbf{SA}(E)$ be the subgroup of $\mathbf{GA}(E)$ (the *special affine group of* E) consisting of the affine maps f such that $\overrightarrow{f} \in \mathbf{SL}(\overrightarrow{E})$. Prove that $\mathbf{SA}(E)$ is isomorphic to the semidirect product $\overrightarrow{E} \times_\theta \mathbf{SL}(\overrightarrow{E})$, where $\theta\colon \mathbf{SL}(\overrightarrow{E}) \to \mathbf{Aut}(\overrightarrow{E})$ is defined as in (vi).

(viii) Assume that $(E, \overrightarrow{E})$ is a Euclidean affine space, as defined in Chapter 6. Let $\mathbf{SO}(\overrightarrow{E})$ be the *special orthogonal group of* $\overrightarrow{E}$, as defined in Definition 6.6 (the isometries with determinant $+1$), and let $\mathbf{SE}(E)$ be the subgroup of $\mathbf{SA}(E)$ (the *special Euclidean group of* E) consisting of the affine isometries f such that $\overrightarrow{f} \in \mathbf{SO}(\overrightarrow{E})$. Prove that $\mathbf{SE}(E)$ is isomorphic to the semidirect product $\overrightarrow{E} \times_\theta \mathbf{SO}(\overrightarrow{E})$, where $\theta\colon \mathbf{SO}(\overrightarrow{E}) \to \mathbf{Aut}(\overrightarrow{E})$ is defined as in (vi).

2.20. The purpose of this problem is to study certain affine maps of $\mathbb{A}^2$.

(1) Consider affine maps of the form

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Prove that such maps have a unique fixed point c if $\theta \neq 2k\pi$, for all integers k. Show that these are rotations of center c, which means that with respect to a frame with origin c (the unique fixed point), these affine maps are represented by rotation matrices.

(2) Consider affine maps of the form

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} \lambda \cos\theta & -\lambda \sin\theta \\ \mu \sin\theta & \mu \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Prove that such maps have a unique fixed point iff $(\lambda+\mu)\cos\theta \neq 1+\lambda\mu$. Prove that if $\lambda\mu = 1$ and $\lambda > 0$, there is some angle θ for which either there is no fixed point, or there are infinitely many fixed points.

(3) Prove that the affine map

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} 8/5 & -6/5 \\ 3/10 & 2/5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

has no fixed point.

(4) Prove that an arbitrary affine map

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

has a unique fixed point iff the matrix

$$\begin{pmatrix} a_1 - 1 & a_2 \\ a_3 & a_4 - 1 \end{pmatrix}$$

is invertible.

2.21. Let $(E, \overrightarrow{E})$ be any affine space of finite dimension. For every affine map $f \colon E \to E$, let $\mathrm{Fix}(f) = \{a \in E \mid f(a) = a\}$ be the set of fixed points of f.

(i) Prove that if $\mathrm{Fix}(f) \neq \emptyset$, then $\mathrm{Fix}(f)$ is an affine subspace of E such that for every $b \in \mathrm{Fix}(f)$,

$$\mathrm{Fix}(f) = b + \mathrm{Ker}\,(\overrightarrow{f} - \mathrm{id}).$$

(ii) Prove that $\mathrm{Fix}(f)$ contains a unique fixed point iff
$\mathrm{Ker}\,(\overrightarrow{f} - \mathrm{id}) = \{0\}$, i.e., $\overrightarrow{f}(u) = u$ iff $u = 0$.
Hint. Show that

$$\overrightarrow{\Omega f(a)} - \overrightarrow{\Omega a} = \overrightarrow{\Omega f(\Omega)} + \overrightarrow{f}(\overrightarrow{\Omega a}) - \overrightarrow{\Omega a},$$

for any two points $\Omega, a \in E$.

2.22. Given two affine spaces $(E, \overrightarrow{E})$ and $(F, \overrightarrow{F})$, let $\mathscr{A}(E,F)$ be the set of all affine maps $f \colon E \to F$.

(i) Prove that the set $\mathscr{A}(E, \overrightarrow{F})$ (viewing $\overrightarrow{F}$ as an affine space) is a vector space under the operations $f+g$ and λf defined such that

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

for all $a \in E$.

(ii) Define an action

$$+\colon \mathscr{A}(E,F)\times\mathscr{A}(E,\overrightarrow{F})\to\mathscr{A}(E,F)$$

of $\mathscr{A}(E,\overrightarrow{F})$ on $\mathscr{A}(E,F)$ as follows: For every $a\in E$, every $f\in\mathscr{A}(E,F)$, and every $h\in\mathscr{A}(E,\overrightarrow{F})$,

$$(f+h)(a)=f(a)+h(a).$$

Prove that $(\mathscr{A}(E,F),\mathscr{A}(E,\overrightarrow{F}),+)$ is an affine space.
Hint. Show that for any two affine maps $f,g\in\mathscr{A}(E,F)$, the map $\overrightarrow{fg}$ defined such that

$$\overrightarrow{fg}(a)=\overrightarrow{f(a)g(a)}$$

(for every $a\in E$) is affine, and thus $\overrightarrow{fg}\in\mathscr{A}(E,\overrightarrow{F})$. Furthermore, $\overrightarrow{fg}$ is the unique map in $\mathscr{A}(E,\overrightarrow{F})$ such that

$$f+\overrightarrow{fg}=g.$$

(iii) If $\overrightarrow{E}$ has dimension m and $\overrightarrow{F}$ has dimension n, prove that $\mathscr{A}(E,\overrightarrow{F})$ has dimension $n+mn=n(m+1)$.

2.23. Let $(c_1,\ldots,c_n)$ be $n\geq 3$ points in $\mathbb{A}^m$ (where $m\geq 2$). Investigate whether there is a closed polygon with n vertices $(a_1,\ldots,a_n)$ such that c_i is the middle of the edge (a_i,a_{i+1}) for every i with $1\leq i\leq n-1$, and c_n is the middle of the edge (a_n,a_0).
Hint. The parity (odd or even) of n plays an important role. When n is odd, there is a unique solution, and when n is even, there are no solutions or infinitely many solutions. Clarify under which conditions there are infinitely many solutions.

2.24. Given an affine space E of dimension n and an affine frame $(a_0,\ldots,a_n)$ for E, let $f\colon E\to E$ and $g\colon E\to E$ be two affine maps represented by the two $(n+1)\times(n+1)$ matrices

$$\begin{pmatrix}A & b\\ 0 & 1\end{pmatrix}\quad\text{and}\quad\begin{pmatrix}B & c\\ 0 & 1\end{pmatrix}$$

w.r.t. the frame $(a_0,\ldots,a_n)$. We also say that f and g are represented by (A,b) and (B,c).

(1) Prove that the composition $f\circ g$ is represented by the matrix

$$\begin{pmatrix}AB & Ac+b\\ 0 & 1\end{pmatrix}.$$

We also say that $f\circ g$ is represented by $(A,b)(B,c)=(AB,Ac+b)$.

(2) Prove that f is invertible iff A is invertible and that the matrix representing f^{-1} is

$$\begin{pmatrix}A^{-1} & -A^{-1}b\\ 0 & 1\end{pmatrix}.$$

We also say that f^{-1} is represented by $(A,b)^{-1} = (A^{-1}, -A^{-1}b)$. Prove that if A is an orthogonal matrix, the matrix associated with f^{-1} is

$$\begin{pmatrix} A^\top & -A^\top b \\ 0 & 1 \end{pmatrix}.$$

Furthermore, denoting the columns of A by $A_1, \ldots, A_n$, prove that the vector $A^\top b$ is the column vector of components

$$(A_1 \cdot b, \ldots, A_n \cdot b)$$

(where $\cdot$ denotes the standard inner product of vectors).

(3) Given two affine frames $(a_0, \ldots, a_n)$ and $(a'_0, \ldots, a'_n)$ for E, any affine map $f \colon E \to E$ has a matrix representation (A,b) w.r.t. $(a_0, \ldots, a_n)$ and $(a'_0, \ldots, a'_n)$ defined such that $b = \overrightarrow{a'_0 f(a_0)}$ is expressed over the basis $(\overrightarrow{a'_0a'_1}, \ldots, \overrightarrow{a'_0a'_n})$, and a_{ij} is the ith coefficient of $f(\overrightarrow{a_0a_j})$ over the basis $(\overrightarrow{a'_0a'_1}, \ldots, \overrightarrow{a'_0a'_n})$. Given any three frames $(a_0, \ldots, a_n)$, $(a'_0, \ldots, a'_n)$, and $(a''_0, \ldots, a''_n)$, for any two affine maps $f \colon E \to E$ and $g \colon E \to E$, if f has the matrix representation (A,b) w.r.t. $(a_0, \ldots, a_n)$ and $(a'_0, \ldots, a'_n)$ and g has the matrix representation (B,c) w.r.t. $(a'_0, \ldots, a'_n)$ and $(a''_0, \ldots, a''_n)$, prove that $g \circ f$ has the matrix representation $(B,c)(A,b)$ w.r.t. $(a_0, \ldots, a_n)$ and $(a''_0, \ldots, a''_n)$.

(4) Given two affine frames $(a_0, \ldots, a_n)$ and $(a'_0, \ldots, a'_n)$ for E, there is a unique affine map $h \colon E \to E$ such that $h(a_i) = a'_i$ for $i = 0, \ldots, n$, and we let (P, ω) be its associated matrix representation with respect to the frame $(a_0, \ldots, a_n)$. Note that $\omega = \overrightarrow{a_0a'_0}$, and that p_{ij} is the ith coefficient of $\overrightarrow{a'_0a'_j}$ over the basis $(\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_n})$. Observe that (P, ω) is also the matrix representation of id_E w.r.t. the frames $(a'_0, \ldots, a'_n)$ and $(a_0, \ldots, a_n)$, **in that order**. For any affine map $f \colon E \to E$, if f has the matrix representation (A,b) over the frame $(a_0, \ldots, a_n)$ and the matrix representation (A',b') over the frame $(a'_0, \ldots, a'_n)$, prove that

$$(A',b') = (P,\omega)^{-1}(A,b)(P,\omega).$$

Given any two affine maps $f \colon E \to E$ and $g \colon E \to E$, where f is invertible, for any affine frame $(a_0, \ldots, a_n)$ for E, if $(a'_0, \ldots, a'_n)$ is the affine frame image of $(a_0, \ldots, a_n)$ under f (i.e., $f(a_i) = a'_i$ for $i = 0, \ldots, n$), letting (A,b) be the matrix representation of f w.r.t. the frame $(a_0, \ldots, a_n)$ and (B,c) be the matrix representation of g w.r.t. the frame $(a'_0, \ldots, a'_n)$ (**not** the frame $(a_0, \ldots, a_n)$), prove that $g \circ f$ is represented by the matrix $(A,b)(B,c)$ w.r.t. the frame $(a_0, \ldots, a_n)$.

Remark: Note that this is the **opposite** of what happens if f and g are both represented by matrices w.r.t. the "fixed" frame $(a_0, \ldots, a_n)$, where $g \circ f$ is represented by the matrix $(B,c)(A,b)$. The frame $(a'_0, \ldots, a'_n)$ can be viewed as a "moving" frame. The above has applications in robotics, for example to rotation matrices expressed in terms of Euler angles, or "roll, pitch, and yaw."

2.25. (a) Let E be a vector space, and let U and V be two subspaces of E such that they form a direct sum $E = U \oplus V$. Recall that this means that every vector $x \in E$ can be written as $x = u + v$, for some unique $u \in U$ and some unique $v \in V$. Define the function $p_U : E \to U$ (resp. $p_V : E \to V$) so that $p_U(x) = u$ (resp. $p_V(x) = v$), where $x = u + v$, as explained above. Check that that p_U and p_V are linear.

(b) Now assume that E is an affine space (nontrivial), and let U and V be affine subspaces such that $\overrightarrow{E} = \overrightarrow{U} \oplus \overrightarrow{V}$. Pick any $\Omega \in V$, and define $q_U : E \to \overrightarrow{U}$ (resp. $q_V : E \to \overrightarrow{V}$, with $\Omega \in U$) so that

$$q_U(a) = p_{\overrightarrow{U}}(\overrightarrow{\Omega a}) \quad \text{(resp.} \quad q_V(a) = p_{\overrightarrow{V}}(\overrightarrow{\Omega a})), \quad \text{for every } a \in E.$$

Prove that q_U does not depend on the choice of $\Omega \in V$ (resp. q_V does not depend on the choice of $\Omega \in U$). Define the map $p_U : E \to U$ (resp. $p_V : E \to V$) so that

$$p_U(a) = a - q_V(a) \quad \text{(resp.} \quad p_V(a) = a - q_U(a)), \quad \text{for every } a \in E.$$

Prove that p_U (resp. p_V) is affine.

The map p_U (resp. p_V) is called the *projection onto U parallel to V (resp. projection onto V parallel to U)*.

(c) Let $(a_0, \ldots, a_n)$ be $n+1$ affinely independent points in $\mathbb{A}^n$ and let $\Delta(a_0, \ldots, a_n)$ denote the convex hull of $(a_0, \ldots, a_n)$ (an n-simplex). Prove that if $f : \mathbb{A}^n \to \mathbb{A}^n$ is an affine map sending $\Delta(a_0, \ldots, a_n)$ inside itself, i.e.,

$$f(\Delta(a_0, \ldots, a_n)) \subseteq \Delta(a_0, \ldots, a_n),$$

then f has some fixed point $b \in \Delta(a_0, \ldots, a_n)$, i.e., $f(b) = b$.
Hint: Proceed by induction on n. First, treat the case $n = 1$. The affine map is determined by $f(a_0)$ and $f(a_1)$, which are affine combinations of a_0 and a_1. There is an explicit formula for some fixed point of f. For the induction step, compose f with some suitable projections.

References

1. Michael Artin. *Algebra*. Prentice-Hall, first edition, 1991.
2. Marcel Berger. *Géométrie 1*. Nathan, 1990. English edition: Geometry 1, Universitext, Springer-Verlag.
3. Marcel Berger. *Géométrie 2*. Nathan, 1990. English edition: Geometry 2, Universitext, Springer-Verlag.
4. H.S.M. Coxeter. *Introduction to Geometry*. Wiley, second edition, 1989.
5. Jean H. Gallier. *Curves and Surfaces in Geometric Modeling: Theory and Algorithms*. Morgan Kaufmann, first edition, 1999.
6. Donald T. Greenwood. *Principles of Dynamics*. Prentice-Hall, second edition, 1988.
7. D. Hilbert and S. Cohn-Vossen. *Geometry and the Imagination*. Chelsea Publishing Co., 1952.
8. Serge Lang. *Algebra*. Addison-Wesley, third edition, 1993.
9. Dan Pedoe. *Geometry, A Comprehensive Course*. Dover, first edition, 1988.