Chapter 2

Basics of Affine Geometry

2.1 Affine Spaces

Suppose we have a particle moving in 3-space and that we want to describe the trajectory of this particle.

If one looks up a good textbook on dynamics, such as Greenwood [**?**], 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.

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 $\mathbf{ab}$ from a to b, the vector $\mathbf{ab} = (b_1 - a_1, b_2 - a_2, b_3 - a_3)$.

Note that

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

addition being understood as addition in $\mathbb{R}^3$.

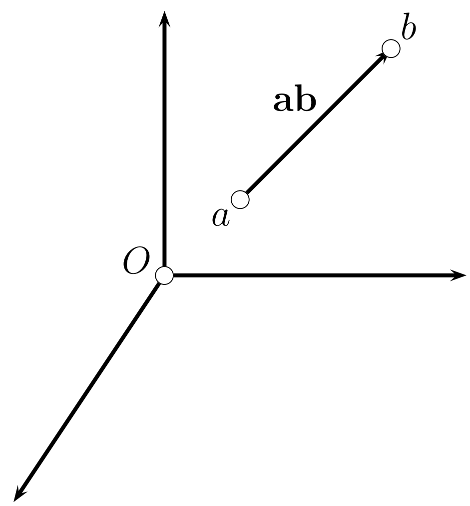

Figure 2.1: Points and free vectors

Then, in the standard frame, given a point $x = (x_1, x_2, x_3)$, the position of x is the vector $\mathbf{Ox} = (x_1, x_2, x_3)$, which coincides with the point itself.

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:

$\mathbf{Ox} = (x_1, x_2, x_3)$ in the frame $(O, (e_1, e_2, e_3))$, and

$\mathbf{\Omega x} = (x_1 - \omega_1, x_2 - \omega_2, x_3 - \omega_3)$ in the frame $(\Omega, (e_1, e_2, e_3))$.

This is because

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

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**. Inspired by physics, it is 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.

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

$$\mathbf{O\Omega} = (\omega_1, \omega_2, \omega_3),$$

given two points a and 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) of 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

$$\begin{aligned}(\lambda a_1 + \mu b_1 - (\lambda + \mu)\omega_1,\\ \lambda a_2 + \mu b_2 - (\lambda + \mu)\omega_2,\\ \lambda a_3 + \mu b_3 - (\lambda + \mu)\omega_3)\end{aligned}$$

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 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, $\mathbb{R}^3$, corresponding to points*, where we forget the vector space structure, and *the second copy, $\mathbb{R}^3$, 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.

This action $+\colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ satisfies some crucial properties. For example,

$$\begin{aligned} a + 0 &= a, \\ (a + u) + v &= a + (u + v), \end{aligned}$$

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

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

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).

Definition 2.1.1 An *affine space* is either 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 as $\mathbf{ab}$, or sometimes as $b - a$. Thus, we also write

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

(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

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

for all $a \in E$ and all $v \in \overrightarrow{E}$, since $\mathbf{a}(\mathbf{a} + \mathbf{v})$ is the unique vector such that $a + v = a + \mathbf{a}(\mathbf{a} + \mathbf{v})$.

Thus, $b = a + v$ is equivalent to $\mathbf{ab} = v$.

It is natural to think of all vectors as having the same origin, the null vector.

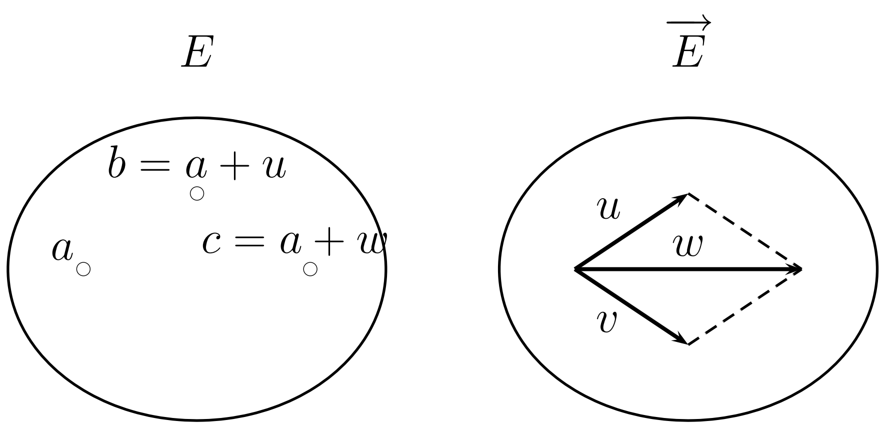

Figure 2.2: Intuitive picture of an affine space

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

$$u \mapsto a + u,$$

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

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

where $b \in E$.

The composition of the first mapping with the second is

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

which, in view of (A3), yields u.

The composition of the second with the first mapping is

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

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 to $\overrightarrow{E}$ via the mapping $b \mapsto \mathbf{ab}$, we say that we consider E as the vector space obtained *by taking a as the origin in E*, and we denote it as 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.

For notational simplicity, we will often denote an affine space $\langle E, \overrightarrow{E}, +\rangle$ as $(E, \overrightarrow{E})$, or even as 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 substracted (but not added!).

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 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 as $\mathbb{A}^n$. In order to distinguish between the double role played by members of $\mathbb{R}^n$, points and vectors, we will denote points as row vectors, and vectors as 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

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

We will also use the convention that if $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, then the column vector associated with x is denoted as $\mathbf{x}$ (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 $n = 1, 2, 3$.

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 $\begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{R}^2$,

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

Affine spaces not already equipped with an obvious vector space structure arise in projective geometry. Indeed, the complement of a hyperplane in a projective space has an affine structure.

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

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

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

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

which is known as *Chasles' identity*.

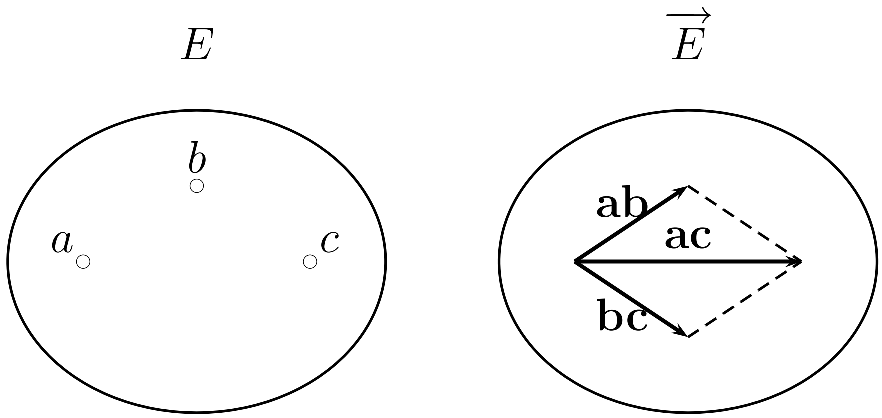

Figure 2.3: Points and corresponding vectors in affine geometry

2.2 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. The problem is that the sum $a + b$ may correspond to two different points depending on which coordinate system is used for its computation!

Thus, 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 2.2.1 *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

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

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

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

Thus, by lemma 2.2.1, 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

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

is independent of the choice of the origin $a \in E$.

The unique point x 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 also the unique point such that

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

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

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

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).

The figure below 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 $(c, 1)$.

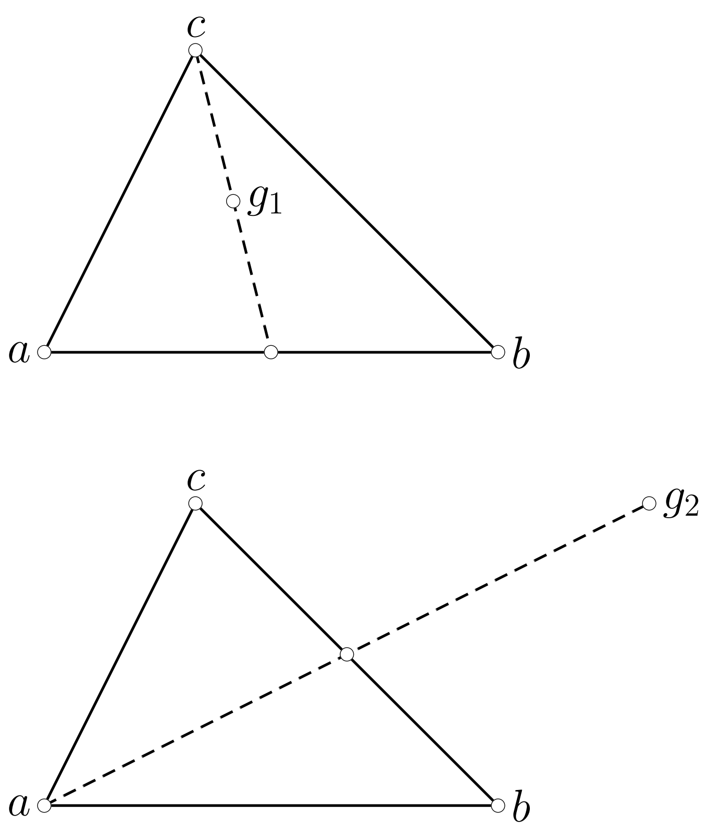

Figure 2.4: Barycenters, $g_1 = \frac{1}{4}a + \frac{1}{4}b + \frac{1}{2}c, \quad g_2 = -a + b + c.$

2.3 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 2.3.1 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 points $(a_i)_{i\in I}$ in V, for any family $(\lambda_i)_{i\in I}$ of scalars 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 2.3.1 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

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

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

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

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

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

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 $\mathbb{A}^2$, line which is parallel to the line U of equation $ax + by = c$.

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

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

where

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

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 $(x_0, y_0) \in U$,

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

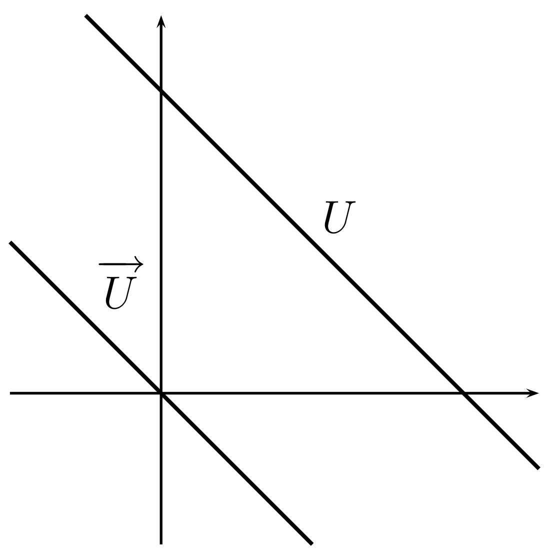

Figure 2.5: An affine line U and its direction

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 $A\mathbf{x} = b$.

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 also be characterized in terms of subspaces of $\overrightarrow{E}$.

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

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

Lemma 2.3.2 *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

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

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

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

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 $\overrightarrow{E}$, for any $a \in E$, the set $V = a + \overrightarrow{V}$ is an affine subspace.

The subspace $\overrightarrow{V}$ associated with an affine subspace V is called the *direction of V*.

It is clear that the map $+ \colon V \times \overrightarrow{V} \rightarrow V$ induced by $+ \colon E \times \overrightarrow{E} \rightarrow E$ confers to $\langle V, \overrightarrow{V}, + \rangle$ an affine structure.

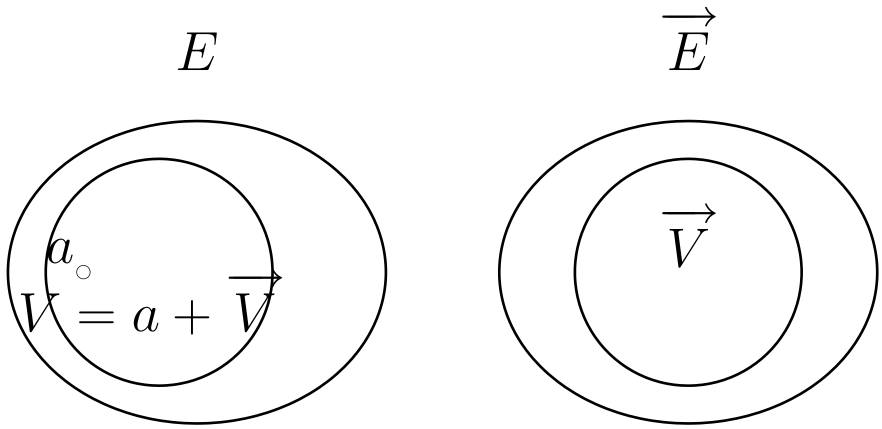

Figure 2.6: An affine subspace V and its direction $\overrightarrow{V}$

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 an *hyperplane*.

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* $\mathbf{ab}$.

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).

We say that three points a, b, c are *collinear*, if the vectors $\mathbf{ab}$ and $\mathbf{ac}$ 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 $\mathbf{ac} = \lambda\mathbf{ab}$, and we define the ratio $\frac{\mathbf{ac}}{\mathbf{ab}} = \lambda$.

We say that four points a, b, c, d are *coplanar*, if the vectors $\mathbf{ab}, \mathbf{ac}$, and $\mathbf{ad}$, are linearly dependent.

Lemma 2.3.3 *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 $(a_i)_{i\in I}$.*

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

Remarks: 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, 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 $(1-\lambda)a + \lambda b$, $\lambda \in \mathbb{R}$.

Let us go back to the problem (see Chapter 1) of finding a plane that best fits a set of points, $\{p_1, \ldots, p_n\}$, in $\mathbb{A}^3$.

We explained that we can solve our problem in the least squares sense by minimizing

$$\sum_{i=1}^{n} (ax_i + by_i + cz_i + d)^2$$

But then, there is *something wrong with our formulation of the problem* because the least squares solution *is* $a = b = c = d = 0$!

We forgot to ensure that solutions of our problem must satisfy the condition

$$(a, b, c) \neq (0, 0, 0),$$

in order to have a well-defined plane.

The plane that we are looking for is not parallel to the x-axis ($a \neq 0$) or not parallel to the y-axis ($b \neq 0$) or not parallel to the z-axis ($c \neq 0$).

Say we have reasons to believe that this plane is not parallel to the z-axis. If we are wrong, in the least squares solution, one of the coefficients, a, b, will be very large.

If $c \neq 0$, then we may assume that our plane is given by an equation of the form

$$z = ax + by + d.$$

Then, our least squares problem is to minimize

$$\sum_{i=1}^{n} (ax_i + by_i + d - z_i)^2$$

This time, we get a solution $(a, b, 1, d)$, which is not trivial.

It turns out that solving in the least-squares sense may give too much weight to "outliers", that is, points clearly outside the best-fit plane. In this case, it is preferable to minimize (the ℓ_1-norm)

$$\sum_{i=1}^{n} |ax_i + by_i + d - z_i|.$$

This does not appear to be a linear problem but we can use a trick to convert this minimization problem into an LP!

Note that $|x| = \max\{x, -x\}$. So, our minimization problem is equivalent to the LP:

$$\begin{array}{rl} \text{minimize} & e_1 + \cdots + e_n \\ \text{subject to} & ax_i + by_i + d - z_i \leq e_i \\ & -(ax_i + by_i + d - z_i) \leq e_i \\ & 1 \leq i \leq n. \end{array}$$

Observe that the constraints are equivalent to

$$e_i \geq |ax_i + by_i + d - z_i|, \qquad 1 \leq i \leq n,$$

equality holding for an optimal solution.

2.4 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 $(\mathbf{a_i a_j})_{j \in (I - \{i\})}$ of vectors obtained by chosing any a_i as an origin.

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

Lemma 2.4.1 *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* $(\mathbf{a_i a_j})_{j \in (I - \{i\})}$ *is linearly independent for some* $i \in I$*, then* $(\mathbf{a_i a_j})_{j \in (I - \{i\})}$ *is linearly independent for every* $i \in I$*.*

Definition 2.4.2 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 $(\mathbf{a_i a_j})_{j\in(I-\{i\})}$ is linearly independent for some $i \in I$.

Definition 2.4.2 is reasonable, since by Lemma 2.4.1, the independence of the family $(\mathbf{a_i a_j})_{j\in(I-\{i\})}$ 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 2.4.3 *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 $(\mathbf{a_0a_1}, \ldots, \mathbf{a_0a_m})$ is linearly independent.*

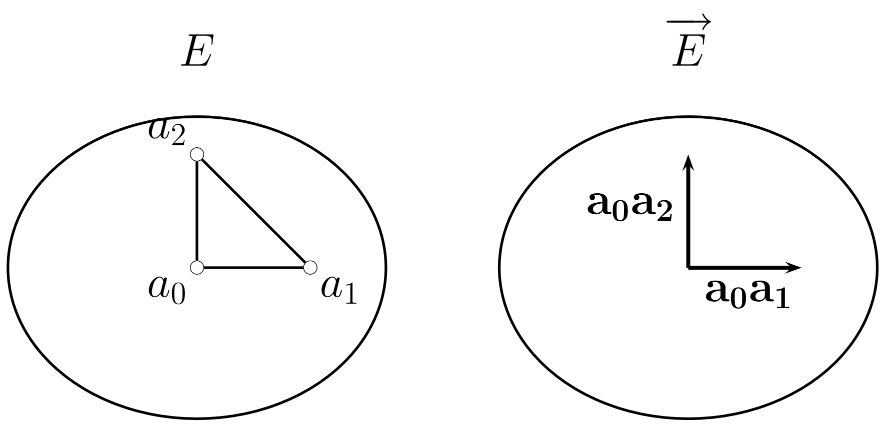

Figure 2.7: Affine independence and linear independence

Lemma 2.4.3 suggests the notion of affine frame.

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 $(\mathbf{a_0a_1}, \ldots, \mathbf{a_0a_m})$ 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 $\overrightarrow{E}$.

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

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

The scalars $(x_1, \ldots, x_m)$ are *coordinates* with respect to $(a_0, (\mathbf{a_0a_1}, \ldots, \mathbf{a_0a_m}))$. Since

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

$x \in E$ 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 $1 \leq i \leq m$.

The scalars $(\lambda_0, \ldots, \lambda_m)$ are also certain kinds of *coordinates* with respect to $(a_0, \ldots, a_m)$.

Definition 2.4.4 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 $(\mathbf{a_0a_1}, \ldots, \mathbf{a_0a_m})$ is a basis of $\overrightarrow{E}$. The pair $(a_0, (\mathbf{a_0a_1}, \ldots, \mathbf{a_0a_m}))$ is also called an *affine frame with origin a_0*.

Then, every $x \in E$ can be expressed as

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

for a unique family $(x_1, \ldots, x_m)$ of scalars, called the *coordinates of x w.r.t. the affine frame $(a_0, (\mathbf{a_0a_1}, \ldots, \mathbf{a_0a_m}))$*.

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 $1 \leq i \leq m$.

An affine frame is called an *affine basis* by some authors. The figure below shows affine frames and their convex hulls for $|I| = 0, 1, 2, 3$.

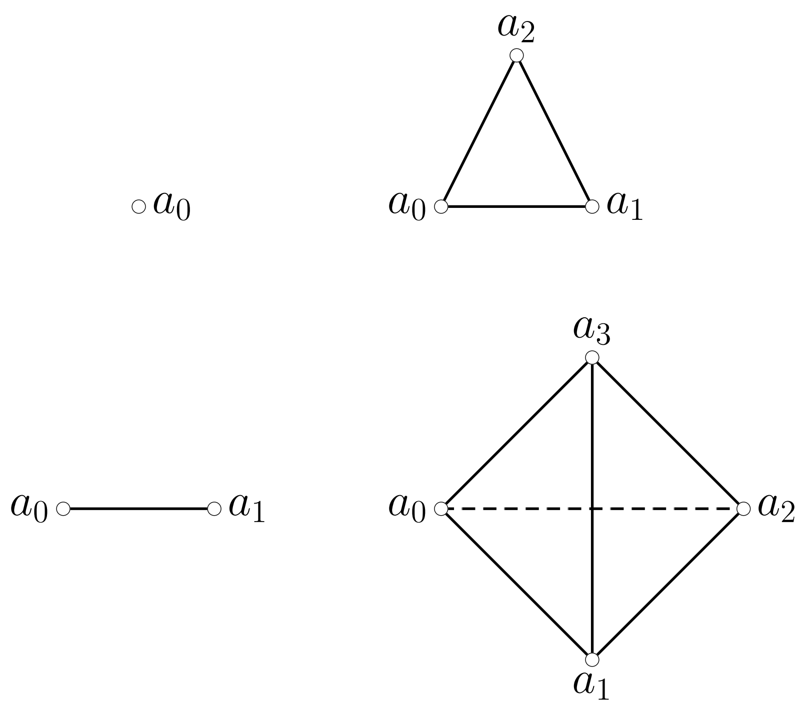

Figure 2.8: Examples of affine frames and their convex hulls.

A family of two points (a, b) in E is affinely independent iff $\mathbf{ab} \neq 0$, 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 **ab** and **ac** are linearly independent, which means that a, b, and c are not on a 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 **ab**, **ac**, and **ad** are linearly independent, which means that a, b, c, and d are not in a same plane (they are not coplanar). In this case, a, b, c, and d, are the vertices of a tetrahedron.

Given $n + 1$ affinely independent points $(a_0, \ldots, a_n)$ in E, we can consider the set of points $\lambda_0 a_0 + \cdots + \lambda_n a_n$, 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).

Barycentric coordinates are very convenient for finding the equation of a line passing through two given points or for determining the intersection of two lines.

For example, given 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, the equation of the line $\langle a, b \rangle$ determined by a and b is

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

or equivalently

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

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

The above can be generalized to find the equation of the plane determined by three affinely independent points in $\mathbb{A}^3$ and more generally, of the hyperplane determined by n affinely independent points in $\mathbb{A}^n$.

If a line, D, is given by the equation

$$ax + by + c = 0$$

with $a \neq 0$ or $b \neq 0$, with respect to (nonbarycentric) coordinates (x, y), then in barycentric coordinates, (z, x, y), as $x + y + z = 1$, the line D is given by

$$(a + c)x + (b + c)y + cz = 0.$$

Observe that $a + c = b + c = c$ iff $a = b = 0$, which is impossible, by hypothesis.

Thus, in barycentric coordinates, a line is given by an equation

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

where $u \neq w$ or $v \neq w$ or $u \neq v$.

The triple (u, v, w) is called a set of *tangential coordinates* for the line D.

Then, given two lines D and D' in $\mathbb{A}^2$ defined by tangential coordinates (u, v, w) and (u', v', w') let

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

It can be shown 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

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

It can also be shown that D and D' are parallel iff $d = 0$.

More is true. 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''), it can be shown that D, D', and D'' are parallel or have a unique intersection point iff

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

The set

$$\{a_0 + \lambda_1 \mathbf{a_0 a_1} + \cdots + \lambda_n \mathbf{a_0 a_n} \mid \text{where } 0 \leq \lambda_i \leq 1 \ (\lambda_i \in \mathbb{R})\},$$

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*.

A parallelotope is shown in figure 2.9: it consists of the points inside of the parallelogram (a_0, a_1, a_2, d), including its boundary.

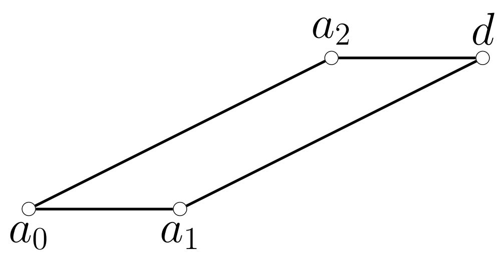

Figure 2.9: A parallelotope

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 $0 \leq \lambda \leq 1$ ($\lambda \in \mathbb{R}$).

2.5 Affine Maps

Corresponding to *linear maps*, we have the notion of an affine map.

Definition 2.5.1 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 $(a_i)_{i\in I}$ of points in E, for every family $(\lambda_i)_{i\in I}$ of scalars such that $\sum_{i\in I} \lambda_i = 1$, we have

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

In other words, f preserves affine combinations (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 $h\colon \overrightarrow{E} \to \overrightarrow{E'}$, the map $f\colon E \to E'$ defined such that

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

is an affine map.

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').

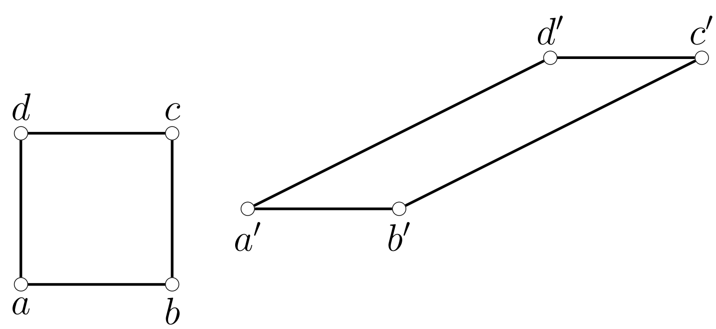

Figure 2.10: The effect of a shear

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

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

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').

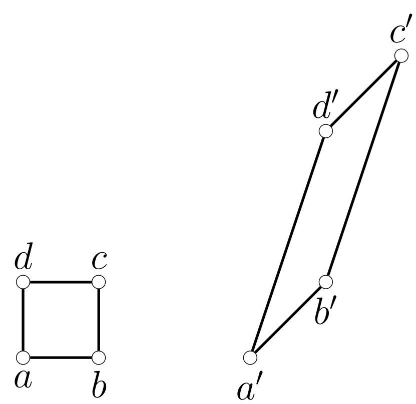

Figure 2.11: The effect of an affine map

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.

Lemma 2.5.2 *Given an affine map* $f: E \to E'$, *there is a unique linear map* $\overrightarrow{f}: \overrightarrow{E} \to \overrightarrow{E'}$, *such that*

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

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

The unique linear map $\overrightarrow{f}: \overrightarrow{E} \to \overrightarrow{E'}$ given by lemma 2.5.2 is 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

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

for all $a, x \in E$.

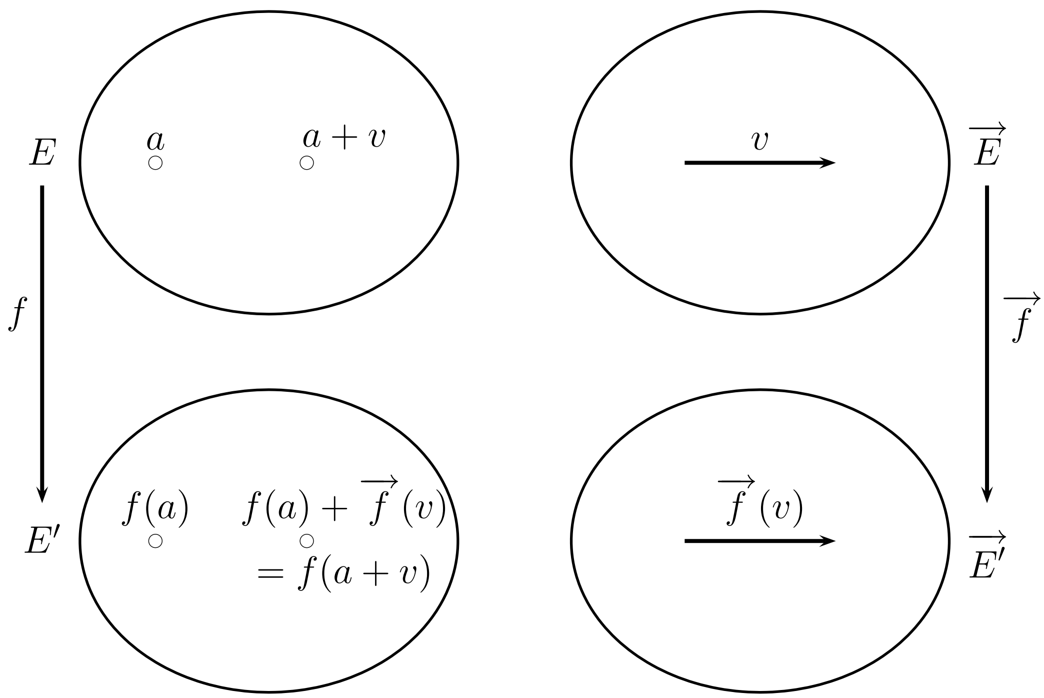

Figure 2.12: An affine map f and its associated linear map $\overrightarrow{f}$

Lemma 2.5.2 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 $\overrightarrow{f}\colon \overrightarrow{E} \to \overrightarrow{E'}$, such that

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

for all $v \in \overrightarrow{E}$ (just let $b = f(a)$, for any $a \in E$).

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, the image of a plane is either a point, a line, or a plane.

Affine maps for which $\overrightarrow{f}$ is the identity map are called *translations*. Indeed, if $\overrightarrow{f} = \mathrm{id}$, it is easy to show that for any two points $a, x \in E$,

$$f(x) = x + \mathbf{af}(\mathbf{a}).$$

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

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

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 $\overrightarrow{f} : \overrightarrow{E} \to \overrightarrow{E'}$ is injective, and that $f\colon E \to E'$ is surjective iff $\overrightarrow{f} : \overrightarrow{E} \to \overrightarrow{E'}$ is surjective.

An affine map $f\colon E \to E'$ is constant iff $\overrightarrow{f} : \overrightarrow{E} \to \overrightarrow{E'}$ 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, for any other affine space F, 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 $0 \leq i \leq m$.

The following diagram illustrates the above result when $m = 2$.

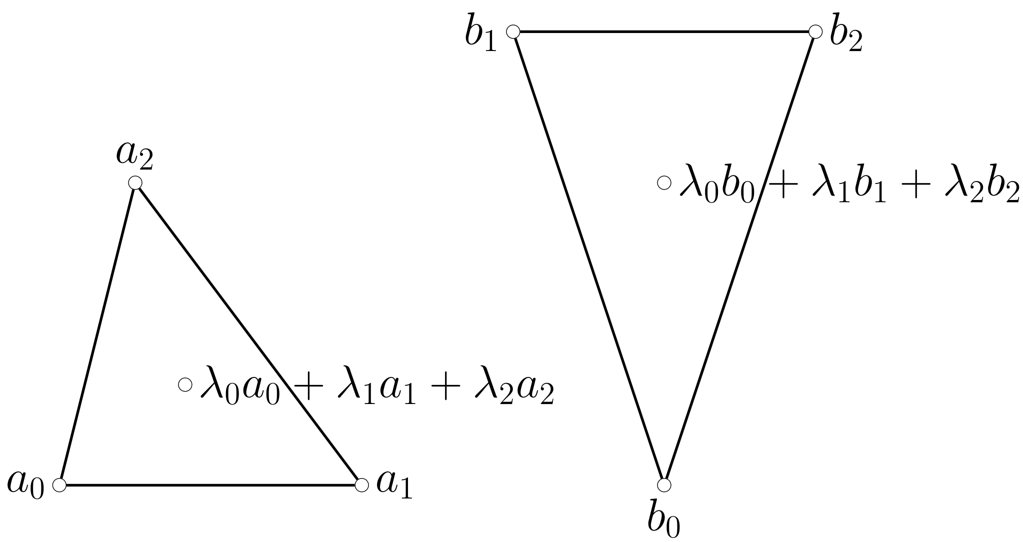

Figure 2.13: An affine map mapping a_0, a_1, a_2 to b_0, b_1, b_2.

Using affine frames, affine maps can be represented in terms of matrices.

We explain how an affine map $f: E \to E$ is represented with respect to a frame $(a_0, \ldots, a_n)$ in E.

Since

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

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

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

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

$$\begin{aligned} x &= x_1 \mathbf{a_0 a_1} + \cdots + x_n \mathbf{a_0 a_n}, \\ \mathbf{a_0 f(a_0)} &= b_1 \mathbf{a_0 a_1} + \cdots + b_n \mathbf{a_0 a_n}, \\ \mathbf{a_0 f(a_0 + x)} &= y_1 \mathbf{a_0 a_1} + \cdots + y_n \mathbf{a_0 a_n}, \end{aligned}$$

if $A = (a_{i\,j})$ is the $n \times n$-matrix of the linear map $\overrightarrow{f}$ over the basis $(\mathbf{a_0 a_1}, \ldots, \mathbf{a_0 a_n})$, letting x, y, and b denote the column vectors of components $(x_1, \ldots, x_n)$, $(y_1, \ldots, y_n)$, and $(b_1, \ldots, b_n)$,

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

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

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

can also be written as

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

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 $1/2$ 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

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

has no fixed point, even though

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

and the second matrix is a rotation of angle θ such that $\cos\theta = \frac{4}{5}$ and $\sin\theta = \frac{3}{5}$.

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 [**?**] or Samuel [**?**]).

The above theorem is often referred to (pompously!) as the *fundamental theorem of affine geometry*

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

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

provided that $\beta \neq 1$, i.e. that $b \neq c$. 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 $-\mathrm{ratio}(a, b, c) = \frac{\mathbf{ba}}{\mathbf{bc}}$. Since affine maps preserves barycenters, it is clear that affine maps preserve the ratio of three points.