2.6 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 as $\mathrm{GA}(E)$.

Recall that the group of bijective linear maps of the vector space $\overrightarrow{E}$ is denoted as $\mathrm{GL}(\overrightarrow{E})$. Then, the map $f \mapsto \overrightarrow{f}$ defines a group homomorphism $L\colon \mathrm{GA}(E) \to \mathrm{GL}(\overrightarrow{E})$. 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 $\mathrm{GL}(\overrightarrow{E})$, and is denoted as $\mathbb{R}^*\mathrm{id}_{\overrightarrow{E}}$.

The subgroup $\mathrm{DIL}(E) = L^{-1}(\mathbb{R}^*\mathrm{id}_{\overrightarrow{E}})$ of $\mathrm{GA}(E)$ 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 $\mathrm{DIL}(E)$ are called *affine dilatations (or dilations)*.

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

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

for every $x \in E$.

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" $\mathbf{ax}$ by λ. 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 2.6.1 *Given any affine space E, for any affine bijection $f \in GA(E)$, 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 $f = H_{c,\lambda}$.*

Clearly, if $\overrightarrow{f} = \mathrm{id}_{\overrightarrow{E}}$, the affine map f is a translation.

Thus, the group of affine dilatations $\mathrm{DIL}(E)$ 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.

2.7 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 Thalés.

Lemma 2.7.1 *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:*

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

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

The diagram below illustrates the theorem of Thalés.

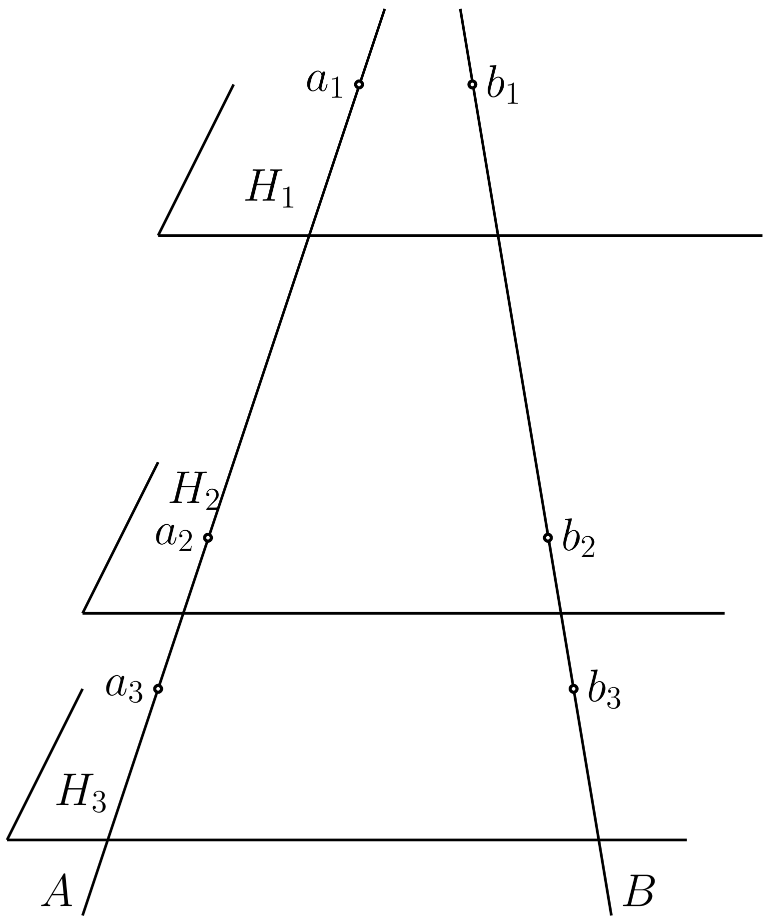

Figure 2.14: The theorem of Thalés

Lemma 2.7.2 *Given any affine space E, given any two distinct points $a, b \in E$, 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 $b' = D' \cap \langle c, b \rangle$.

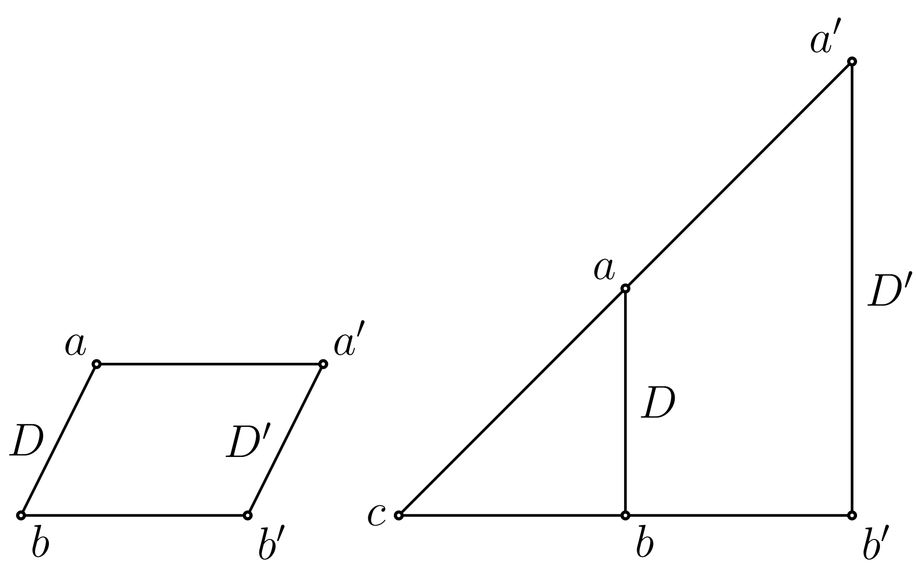

Figure 2.15: Affine Dilatations

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

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

There is a converse to Pappus' theorem, which yields a fancier version of Pappus' theorem, but it is easier to prove it using projective geometry.

Lemma 2.7.3 *Given any affine plane E, given any two distinct lines D and D', 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.*

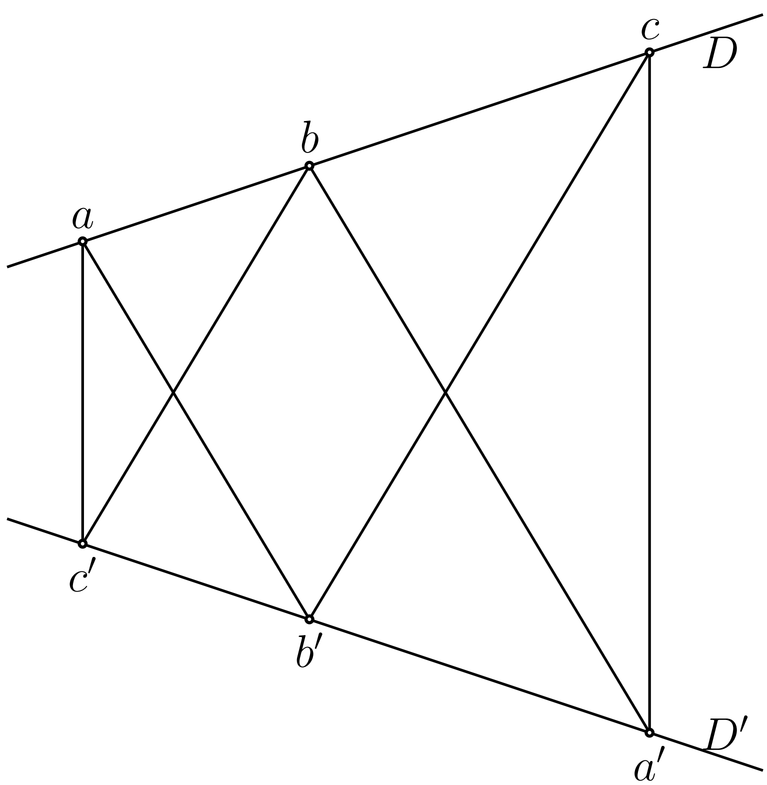

Figure 2.16: Pappus' theorem (affine version)

We now prove an affine version of Desargues' theorem.

Lemma 2.7.4 *Given any affine space E, 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).*

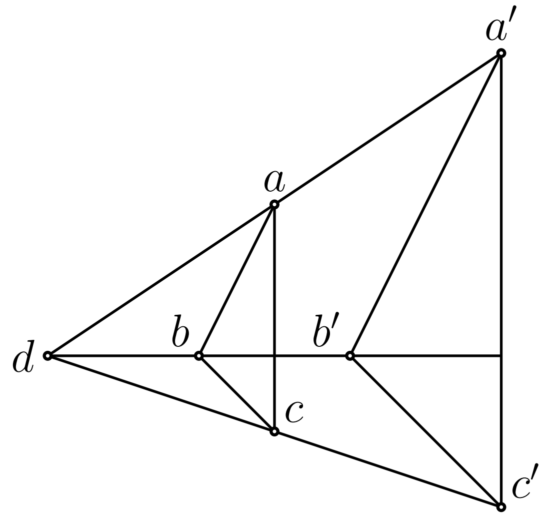

Figure 2.17: Desargues' theorem (affine version)

There is a fancier version of Desargues' theorem, but it is easier to prove it using projective geometry.

Desargues' 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.

2.8 Affine Hyperplanes

In section 2.3, 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.

Recall the characterization of hyperplanes in terms of linear forms.

Given a vector space E over a field K, a linear map $f\colon E \to K$ is called a *linear form*. The set of all linear forms $f\colon E \to K$ is a vector space called the *dual space of* E, and denoted as E^*.

Hyperplanes are precisely the Kernels of nonnull linear forms.

Lemma 2.8.1 *Let E be a vector space. The following properties hold:*

(a) Given any nonnull linear form $f \in E^*$, its kernel $H = \operatorname{Ker} f$ is a hyperplane.

(b) For any hyperplane H in E, there is a (nonnull) linear form $f \in E^*$ such that $H = \operatorname{Ker} f$.

(c) Given any hyperplane H in E and any (nonnull) linear form $f \in E^*$ such that $H = \operatorname{Ker} f$, for every linear form $g \in E^*$, $H = \operatorname{Ker} g$ iff $g = \lambda f$ for some $\lambda \neq 0$ in K.

Going back to an affine space E, given an affine map $f \colon E \to \mathbb{R}$, we also denote $f^{-1}(0)$ as $\operatorname{Ker} f$, and we call it the *kernel* of f. Recall that an (affine) hyperplane is an affine subspace of codimension 1.

Affine hyperplanes are precisely the Kernels of nonconstant affine forms.

Lemma 2.8.2 *Let E be an affine space. The following properties hold:*

(a) Given any nonconstant affine form $f: E \to \mathbb{R}$, its kernel $H = \mathrm{Ker}\, f$ is a hyperplane.

(b) For any hyperplane H in E, there is a nonconstant affine form $f: E \to \mathbb{R}$ such that $H = \mathrm{Ker}\, f$. For any other affine form $g: E \to \mathbb{R}$ 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 $f: E \to \mathbb{R}$ 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 $a \in E$, for some $\lambda \in \mathbb{R}$.

2.9 Intersection of Affine Spaces

In this section, we take a closer look at the intersection of affine subspaces.

First, we need a result of linear algebra.

Lemma 2.9.1 *Given a vector space E and any two subspaces M and N, we have the Grassmann relation:*

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

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

Lemma 2.9.2 *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 $\mathbf{ab} \in \overrightarrow{M} + \overrightarrow{N}$ for some $a \in M$ and some $b \in N$.

(2) $M \cap N$ consists of a single point iff $\mathbf{ab} \in \overrightarrow{M} + \overrightarrow{N}$ 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 $\overrightarrow{S} = \overrightarrow{M} + \overrightarrow{N} + K\mathbf{ab}$ (the vector space $\overrightarrow{E}$ is defined over the field K).

Remarks: (1) The proof of Lemma 2.9.2 shows that if $M \cap N \neq \emptyset$ then $\mathbf{ab} \in \overrightarrow{M} + \overrightarrow{N}$ for all $a \in M$ and all $b \in N$.

(2) Lemma 2.9.2 (2) 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.

Lemma 2.9.3 *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

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

Chapter 3

Properties of Convex Sets: A Glimpse

3.1 Convex Sets

Convex sets play a very important role in geometry. In this chapter, we state some of the "classics" of convex affine geometry: Carathéodory's theorem, Radon's theorem, and Helly's theorem.

These theorems share the property that they are easy to state, but they are deep, and their proof, although rather short, requires a lot of creativity.

Given an affine space E, recall 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 $0 \leq \lambda \leq 1$ ($\lambda \in \mathbb{R}$).

The notation $[a, b]$ is often used to denote the line segment between a and b, that is,

$$[a, b] = \{c \in E \mid c = (1 - \lambda)a + \lambda b,\ 0 \leq \lambda \leq 1\},$$

and thus, a set V is convex if $[a, b] \subseteq V$ for any two points $a, b \in V$ ($a = b$ is allowed).

The empty set is trivially convex, every one-point set $\{a\}$ is convex, and the entire affine space E is of course convex.

It is obvious that the intersection of any family (finite or infinite) of convex sets is convex.

Then, given any (nonempty) subset S of E, there is a smallest convex set containing S denoted as $\mathcal{C}(S)$ (or $\mathrm{conv}(S)$) and called the *convex hull of* S (namely, the intersection of all convex sets containing S).

Lemma 3.1.1 *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 convex combinations $\sum_{i\in I} \lambda_i a_i$ (where $\sum_{i\in I} \lambda_i = 1$ and $\lambda_i \geq 0$) is the convex hull of $(a_i)_{i\in I}$.*

In view of lemma 3.1.1, it is obvious that any affine subspace of E is convex.

Convex sets also arise in terms of hyperplanes. Given a hyperplane H, if $f\colon E \to \mathbb{R}$ is any nonconstant affine form defining H (i.e., $H = \mathrm{Ker}\, f$), we can define the two subsets

$$
\begin{aligned}
H_+(f) &= \{a \in E \mid f(a) \geq 0\},\\
H_-(f) &= \{a \in E \mid f(a) \leq 0\},
\end{aligned}
$$

called *(closed) half spaces associated with f*.

Observe that if $\lambda > 0$, then $H_+(\lambda f) = H_+(f)$, but if $\lambda < 0$, then $H_+(\lambda f) = H_-(f)$, and similarly for $H_-(\lambda f)$.

However, the set $\{H_+(f), H_-(f)\}$ only depends on the hyperplane H, and the choice of a specific f defining H amounts to the choice of one of the two half-spaces.

For this reason, we will also say that $H_+(f)$ and $H_-(f)$ are the (closed) half spaces associated with H.

Clearly,

$$H_+(f) \cup H_-(f) = E \quad \text{and} \quad H_+(f) \cap H_-(f) = H.$$

It is immediately verified that $H_+(f)$ and $H_-(f)$ are convex.

Bounded convex sets arising as the intersection of a finite family of half-spaces associated with hyperplanes play a major role in convex geometry and topology (they are called *convex polytopes*).

It is natural to wonder whether lemma 3.1.1 can be sharpened in two directions:

(1) is it possible have a fixed bound on the number of points involved in the convex combinations?

(2) Is it necessary to consider convex combinations of all points, or is it possible to only consider a subset with special properties?

The answer is yes in both cases. In case 1, assuming that the affine space E has dimension m, *Carathéodory's theorem* asserts that it is enough to consider convex combinations of $m+1$ points.

In case 2, the theorem of Krein and Milman asserts that a convex set which is also compact is the convex hull of its extremal points (see Berger [**?**] or Lang [**?**]).

First, we will prove Carathéodory's theorem. The following technical (and dull!) lemma plays a crucial role in the proof.

Lemma 3.1.2 *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* $\sum_{i\in I} \lambda_i \mathbf{xa_i} = 0$ *for every* $x \in E$.

Theorem 3.1.3 *Given any affine space E of dimension m, for any (nonempty) family $S = (a_i)_{i \in L}$ in E, the convex hull $\mathcal{C}(S)$ of S is equal to the set of convex combinations of families of $m + 1$ points of S.*

Proof. By lemma 3.1.1,

$$\mathcal{C}(S) = \{\sum_{i \in I} \lambda_i a_i \mid a_i \in S,\ \sum_{i \in I} \lambda_i = 1,\ \lambda_i \geq 0,\ I \subseteq L,\ I \text{ finite}\}.$$

We would like to prove that

$$\mathcal{C}(S) = \{\sum_{i \in I} \lambda_i a_i \mid a_i \in S,\ \sum_{i \in I} \lambda_i = 1,\ \lambda_i \geq 0,\ I \subseteq L,\ |I| = m + 1\}.$$

We proceed by contradiction. If the theorem is false, there is some point $b \in \mathcal{C}(S)$ such that b can be expressed as a convex combination $b = \sum_{i \in I} \lambda_i a_i$, where $I \subseteq L$ is a finite set of cardinality $|I| = q$ with $q \geq m + 2$, and b cannot be expressed as any convex combination
$b = \sum_{j \in J} \mu_j a_j$ of strictly less than q points in S
(with $|J| < q$).

We shall prove that b can be written as a convex combination of $q-1$ of the a_i. Since E has dimension m and $q \geq m+2$, the points $a_1, \ldots, a_q$ must be affinely dependent, and we use lemma 3.1.2. $\square$

If S is a finite (of infinite) set of points in the affine plane $\mathbb{A}^2$, theorem 3.1.3 confirms our intuition that $\mathcal{C}(S)$ is the union of triangles (including interior points) whose vertices belong to S.

Similarly, the convex hull of a set S of points in $\mathbb{A}^3$ is the union of tetrahedra (including interior points) whose vertices belong to S.

We get the feeling that triangulations play a crucial role, which is of course true!

We conlude this short section by stating two other classics of convex geometry. We begin with *Radon's theorem*.

Theorem 3.1.4 *Given any affine space E of dimension m, for every subset X of E, if X has at least $m+2$ points, then there is a partition of X into two nonempty disjoint subsets X_1 and X_2 such that the convex hulls of X_1 and X_2 have a nonempty intersection.*

Finally, we state a version of *Helly's theorem*.

Theorem 3.1.5 *Given any affine space E of dimension m, for every family $\{K_1, \ldots, K_n\}$ of n convex subsets of E, if $n \geq m+2$ and the intersection $\bigcap_{i \in I} K_i$ of any $m+1$ of the K_i is nonempty (where $I \subseteq \{1, \ldots, n\}$, $|I| = m+1$), then $\bigcap_{i=1}^{n} K_i$ is nonempty.*

An amusing corollary of Helly's theorem is the following result. Consider $n \geq 4$ parallel line segments in the affine plane $\mathbb{A}^2$. If every three of these line segments meet a line, then all of these line segments meet a common line.