Chapter 12

Embedding an Affine Space in a Vector Space

12.1 Embedding an Affine Space as a Hyperplane in a Vector Space: the "Hat Construction"

Assume that we consider the real affine space E of dimension 3, and that we have some affine frame $(a_0, (\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_2}))$.

With respect to this affine frame, every point $x \in E$ is represented by its coordinates (x_1, x_2, x_3), where

$$a = a_0 + x_1 \overrightarrow{v_1} + x_2 \overrightarrow{v_2} + x_3 \overrightarrow{v_3}.$$

A vector $\overrightarrow{u} \in \overrightarrow{E}$ is also represented by its coordinates (u_1, u_2, u_3) over the basis $(\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_2})$.

One way to *distinguish between points and vectors* is to *add a fourth coordinate*, and to agree that points are represented by (row) vectors $(x_1, x_2, x_3, 1)$ whose fourth coordinate is 1, and that vectors are represented by (row) vectors $(v_1, v_2, v_3, 0)$ whose fourth coordinate is 0.

This “programming trick” works actually very well. Of course, we are opening the door for strange elements such as $(x_1, x_2, x_3, 5)$, where the fourth coordinate is neither 1 nor 0.

The question is, *can we make sense of such elements and of such a construction*?

The answer is “yes”. We will present a construction in which an affine space $(E, \overrightarrow{E})$ is embedded in a vector space $\widehat{E}$, in which $\overrightarrow{E}$ is embedded as a hyperplane passing through the origin, and E itself is embedded as an affine hyperplane, defined as $\omega^{-1}(1)$, for some linear form $\omega\colon \widehat{E} \to \mathbb{R}$.

The vector space $\widehat{E}$ has the *universal property* that for any vector space $\overrightarrow{F}$ and any *affine* map $f\colon E \to \overrightarrow{F}$, there is a *unique linear map* $\widehat{f}\colon \widehat{E} \to \overrightarrow{F}$ extending $f\colon E \to \overrightarrow{F}$.

Some Simple Geometric Transformations

Given an affine space $(E, \overrightarrow{E})$, every $\overrightarrow{u} \in \overrightarrow{E}$ induces a mapping $t_u \colon E \to E$, called a *translation*, and defined such that $t_u(a) = a + \overrightarrow{u}$, for every $a \in E$.

Clearly, the set of translations is a vector space isomorphic to $\overrightarrow{E}$.

Given any point a and any scalar $\lambda \in \mathbb{R}$, we define the mapping $H_{a,\lambda} \colon E \to E$, called *dilatation (or central magnification, or homothety) of center a and ratio λ*, and defined such that

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

for every $x \in E$.

$H_{a,\lambda}(a) = a$, 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. The effect is a uniform dilatation (or contraction, if $\lambda < 1$).

When $\lambda = 0$, $H_{a,0}(x) = a$ for all $x \in E$, and $H_{a,0}$ is the constant affine map sending every point to a.

If we assume $\lambda \neq 1$, note that $H_{a,\lambda}$ is never the identity, and since a is a fixed-point, *$H_{a,\lambda}$ is never a translation*.

We now consider the set $\widehat{E}$ of geometric transformations from E to E, consisting of the union of the (disjoint) sets of translations and dilatations of ratio $\lambda \neq 1$.

We would like to give this set the structure of a vector space, *in such a way that both E and $\overrightarrow{E}$ can be naturally embedded into $\widehat{E}$*.

In fact, it will turn out that barycenters show up quite naturally too!

In order to "add" two dilatations H_{a_1,λ_1} and H_{a_2,λ_2}, it turns out that it is more convenient to consider dilatations of the form $H_{a,1-\lambda}$, where $\lambda \neq 0$.

To see this, let us see the effect of such a dilatation on a point $x \in E$: we have

$$H_{a,1-\lambda}(x) = a + (1 - \lambda)\overrightarrow{ax} = a + \overrightarrow{ax} - \lambda\overrightarrow{ax} = x + \lambda\overrightarrow{xa}.$$

For simplicity of notation, *let us denote $H_{a,1-\lambda}$ as $\langle a, \lambda\rangle$*. Then, we have

$$\langle a, \lambda\rangle(x) = x + \lambda\overrightarrow{xa}.$$

Lemma 12.1.1 *The set $\widehat{E}$ consisting of the disjoint union of the translations and the dilatations $H_{a,1-\lambda} = \langle a, \lambda \rangle$, $\lambda \in \mathbb{R}, \lambda \neq 0$, is a vector space under the following operations of addition and multiplication by a scalar:*

$$\langle a_1, \lambda_1 \rangle \,\widehat{+}\, \langle a_2, \lambda_2 \rangle = \lambda_1 \overrightarrow{a_2 a_1},$$

if $\lambda_1 + \lambda_2 = 0$*;*

$$\langle a_1, \lambda_1 \rangle \,\widehat{+}\, \langle a_2, \lambda_2 \rangle = \langle \frac{\lambda_1}{\lambda_1 + \lambda_2} a_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} a_2, \lambda_1 + \lambda_2 \rangle,$$

if $\lambda_1 + \lambda_2 \neq 0$*;*

$$\langle a, \lambda \rangle \,\widehat{+}\, \overrightarrow{u} = \langle a + \lambda^{-1} \overrightarrow{u}, \lambda \rangle;$$

$$\overrightarrow{u} \,\widehat{+}\, \overrightarrow{v} = \overrightarrow{u} + \overrightarrow{v};$$

$$\mu \cdot \langle a, \lambda \rangle = \langle a, \lambda\mu \rangle,$$

if $\mu \neq 0$, *and*

$$0 \cdot \langle a, \lambda \rangle = \overrightarrow{0},$$

$$\lambda \cdot \overrightarrow{u} = \lambda \overrightarrow{u}.$$

Furthermore, the map $\omega \colon \widehat{E} \to \mathbb{R}$ *defined such that*

$$\begin{aligned} \omega(\langle a, \lambda \rangle) &= \lambda, \\ \omega(\overrightarrow{u}) &= 0, \end{aligned}$$

is a linear form, $\omega^{-1}(0)$ *is a hyperplane isomorphic to* $\overrightarrow{E}$ *under the injective linear map* $i \colon \overrightarrow{E} \to \widehat{E}$ *such that* $i(\overrightarrow{u}) = t_u$ *(the translation associated with* $\overrightarrow{u}$*), and* $\omega^{-1}(1)$ *is an affine hyperplane isomorphic to* E *with direction* $i(\overrightarrow{E})$*, under the injective affine map* $j \colon E \to \widehat{E}$*, where* $j(a) = \langle a, 1 \rangle$*, for every* $a \in E$*.*

Finally, for every $a \in E$*, we have*

$$\widehat{E} = i(\overrightarrow{E}) \oplus \mathbb{R}j(a).$$

The following diagram illustrates the embedding of the affine space E into the vector space $\widehat{E}$, when E is an affine plane.

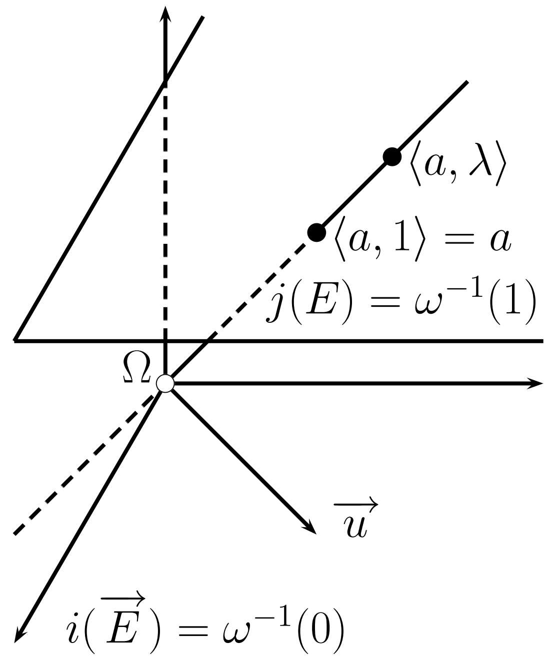

Figure 12.1: Embedding an affine space E into a vector space $\widehat{E}$

Note that $\widehat{E}$ is isomorphic to $\overrightarrow{E} \cup (E \times \mathbb{R}^*)$ (where $\mathbb{R}^* = \mathbb{R} - \{0\}$). Other authors, such as Ramshaw, use the notation E_* for $\widehat{E}$.

Ramshaw calls the linear form $\omega\colon \widehat{E} \to \mathbb{R}$ a *weight (or flavor)*, and he says that an element $z \in \widehat{E}$ such that $\omega(z) = \lambda$ is *λ-heavy (or has flavor λ)* ([**?**]).

The elements of $j(E)$ are 1-heavy and are called *points*, and the elements of $i(\overrightarrow{E})$ are 0-heavy and are called *vectors*.

In general, the λ-heavy elements all belong to the hyperplane $\omega^{-1}(\lambda)$ parallel to $i(\overrightarrow{E})$.

Thus, intuitively, we can thing of $\widehat{E}$ as a stack of parallel hyperplanes, one for each λ, a little bit like an infinite stack of very thin pancakes!

There are two privileged pancakes: one corresponding to E, for $\lambda = 1$, and one corresponding to $\overrightarrow{E}$, for $\lambda = 0$.

From now on, we will identify $j(E)$ and E, and $i(\overrightarrow{E})$ and $\overrightarrow{E}$.

We will also write λa instead of $\langle a, \lambda \rangle$, which we will call a *weighted point*, and write $1a$ just as a.

When we want to be more precise, we may also write $\langle a, 1 \rangle$ as $\overline{a}$ (as Ramshaw does).

In particular, when we consider the homogenized version $\widehat{\mathbb{A}}$ of the affine space $\mathbb{A}$ associated with the base field $\mathbb{R}$ considered as an affine space, we write $\overline{\lambda}$ for $\langle \lambda, 1 \rangle$, when viewing λ as a point in both $\mathbb{A}$ and $\widehat{\mathbb{A}}$, and simply λ, when viewing λ as a vector in $\mathbb{R}$ and in $\widehat{\mathbb{A}}$.

The elements of $\widehat{\mathbb{A}}$ are called *Bézier sites*, by Ramshaw.

Then, in view of the fact that

$$\langle a + \overrightarrow{u}, 1 \rangle = \langle a, 1 \rangle \mathbin{\widehat{+}} \overrightarrow{u},$$

and since we are identifying $a + \overrightarrow{u}$ with $\langle a + \overrightarrow{u}, 1 \rangle$ (under the injection j), in the simplified notation, the above reads as $a + \overrightarrow{u} = a \mathbin{\widehat{+}} \overrightarrow{u}$.

Thus, we go one step further, and denote $a \mathbin{\widehat{+}} \overrightarrow{u}$ as $a + \overrightarrow{u}$.

From lemma 12.1.1, for every $a \in E$, every element of $\widehat{E}$ can be written uniquely as $\overrightarrow{u} \mathbin{\widehat{+}} \lambda a$.

We also denote

$$\lambda a \mathbin{\widehat{+}} (-\mu) b$$

as

$$\lambda a \mathbin{\widehat{-}} \mu b.$$

Given any family $(a_i)_{i\in I}$ of points in E, and any family $(\lambda_i)_{i\in I}$ of scalars in $\mathbb{R}$, with finite support, it is easily shown by induction on the size of the support of $(\lambda_i)_{i\in I}$ that,

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

$$\sum_{i\in I}\langle a_i, \lambda_i\rangle = \sum_{i\in I}\lambda_i a_i,$$

where

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

for any $b \in E$, which, by lemma 2.2.1, is a vector independent of b, or

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

$$\sum_{i\in I}\langle a_i, \lambda_i\rangle = \langle\sum_{i\in I}\frac{\lambda_i}{\sum_{i\in I}\lambda_i}a_i, \sum_{i\in I}\lambda_i\rangle.$$

Thus, we see how *barycenters reenter the scene* quite naturally, and that in $\widehat{E}$, we can make sense of $\sum_{i\in I}\langle a_i, \lambda_i\rangle$, regardless of the value of $\sum_{i\in I}\lambda_i$.

When $\sum_{i\in I}\lambda_i = 1$, the element $\sum_{i\in I}\langle a_i, \lambda_i\rangle$ belongs to the hyperplane $\omega^{-1}(1)$, and thus, it is a point.

When $\sum_{i\in I}\lambda_i = 0$, the linear combination of points $\sum_{i\in I}\lambda_i a_i$ is a vector, and when $I = \{1,\ldots,n\}$, we allow ourselves to write

$$\lambda_1 a_1 \mathbin{\widehat{+}} \cdots \mathbin{\widehat{+}} \lambda_n a_n,$$

where some of the occurrences of $\widehat{+}$ can be replaced by $\widehat{-}$, as

$$\lambda_1 a_1 + \cdots + \lambda_n a_n,$$

where the occurrences of $\widehat{-}$ (if any) are replaced by $-$.

In fact, we have the following slightly more general property, which is left as an exercise.

Lemma 12.1.2 *Given any affine space $(E, \overrightarrow{E})$, for any family $(a_i)_{i \in I}$ of points in E, for any family $(\lambda_i)_{i \in I}$ of scalars in $\mathbb{R}$, with finite support, and any family $(\overrightarrow{v_j})_{j \in J}$ of vectors in $\overrightarrow{E}$ also with finite support, and with $I \cap J = \emptyset$, the following properties hold:*

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

$$\sum_{i \in I} \langle a_i, \lambda_i \rangle \mathbin{\widehat{+}} \sum_{j \in J} \overrightarrow{v_j} = \sum_{i \in I} \lambda_i a_i + \sum_{j \in J} \overrightarrow{v_j},$$

where

$$\sum_{i \in I} \lambda_i a_i = \sum_{i \in I} \lambda_i \overrightarrow{b a_i}$$

for any $b \in E$, which, by lemma 2.2.1, is a vector independent of b, or

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

$$\sum_{i \in I} \langle a_i, \lambda_i \rangle \mathbin{\widehat{+}} \sum_{j \in J} \overrightarrow{v_j}$$

$$= \langle \sum_{i \in I} \frac{\lambda_i}{\sum_{i \in I} \lambda_i} a_i + \sum_{j \in J} \frac{\overrightarrow{v_j}}{\sum_{i \in I} \lambda_i}, \sum_{i \in I} \lambda_i \rangle.$$

The above formulae show that we have *some kind of extended barycentric calculus*.

Operations on weighted points and vectors were introduced by H. Grassmann, in his book published in 1844! This calculus is helpful in dealing with rational curves.

There is also a nice relationship between affine frames in $(E, \overrightarrow{E})$ and bases of $\widehat{E}$, stated in the following lemma.

Lemma 12.1.3 *Given any affine space $(E, \overrightarrow{E})$, for any affine frame $(a_0, (\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m}))$ for E, the family $(\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m}, a_0)$ is a basis for $\widehat{E}$, and for any affine frame $(a_0, \ldots, a_m)$ for E, the family $(a_0, \ldots, a_m)$ is a basis for $\widehat{E}$.*

Furthermore, given any element $\langle x, \lambda \rangle \in \widehat{E}$, if

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

over the affine frame $(a_0, (\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m}))$ in E, then the coordinates of $\langle x, \lambda \rangle$ over the basis

$$(\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m}, a_0)$$

in $\widehat{E}$, are

$$(\lambda x_1, \ldots, \lambda x_m, \lambda).$$

For any vector $\overrightarrow{v} \in \overrightarrow{E}$*, if*

$$\overrightarrow{v} = v_1\overrightarrow{a_0a_1} + \cdots + v_m\overrightarrow{a_0a_m}$$

over the basis

$$(\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m})$$

in $\overrightarrow{E}$*, then over the basis*

$$(\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m}, a_0)$$

in $\widehat{E}$*, the coordinates of* $\overrightarrow{v}$ *are*

$$(v_1, \ldots, v_m, 0).$$

For any element $\langle a, \lambda \rangle$, *where* $\lambda \neq 0$, *if the barycentric coordinates of* a *w.r.t. the affine basis* $(a_0, \ldots, a_m)$ *in* E *are* $(\lambda_0, \ldots, \lambda_m)$ *with* $\lambda_0 + \cdots + \lambda_m = 1$, *then the coordinates of* $\langle a, \lambda \rangle$ *w.r.t. the basis* $(a_0, \ldots, a_m)$ *in* $\widehat{E}$ *are*

$$(\lambda\lambda_0, \ldots, \lambda\lambda_m).$$

If a vector $\overrightarrow{v} \in \overrightarrow{E}$ *is expressed as*

$$\begin{aligned}\overrightarrow{v} = v_1\overrightarrow{a_0a_1} + \cdots + v_m\overrightarrow{a_0a_m} \\ = -(v_1 + \cdots + v_m)a_0 + v_1a_1 + \cdots + v_ma_m,\end{aligned}$$

with respect to the affine basis $(a_0, \ldots, a_m)$ *in* E, *then its coordinates w.r.t. the basis* $(a_0, \ldots, a_m)$ *in* $\widehat{E}$ *are*

$$(-(v_1 + \cdots + v_m), v_1, \ldots, v_m).$$

The following diagram shows the basis $(\overrightarrow{a_0a_1}, \overrightarrow{a_0a_2}, a_0)$ corresponding to the affine frame (a_0, a_1, a_2) in E.

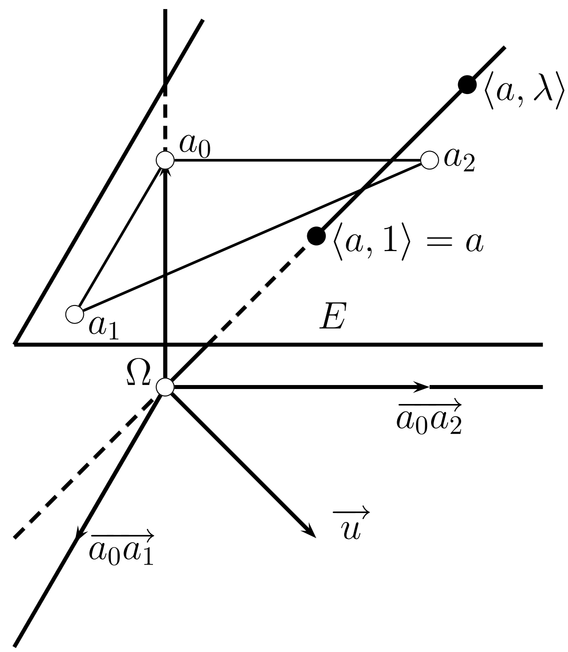

Figure 12.2: The basis $(\overrightarrow{a_0a_1}, \overrightarrow{a_0a_2}, a_0)$ in $\widehat{E}$

If $(x_1, \ldots, x_m)$ are the coordinates of x w.r.t. to the affine frame $(a_0, (\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m}))$ in E, then, $(x_1, \ldots, x_m, 1)$ are the coordinates of x in $\widehat{E}$, i.e., the last coordinate is 1, and if $\overrightarrow{u}$ has coordinates $(u_1, \ldots, u_m)$ with respect to the basis $(\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_m})$ in $\overrightarrow{E}$, then $\overrightarrow{u}$ has coordinates $(u_1, \ldots, u_m, 0)$ in $\widehat{E}$, i.e., the last coordinate is 0.

The following diagram shows the affine frame (a_0, a_1, a_2) in E viewed as a basis in $\widehat{E}$.

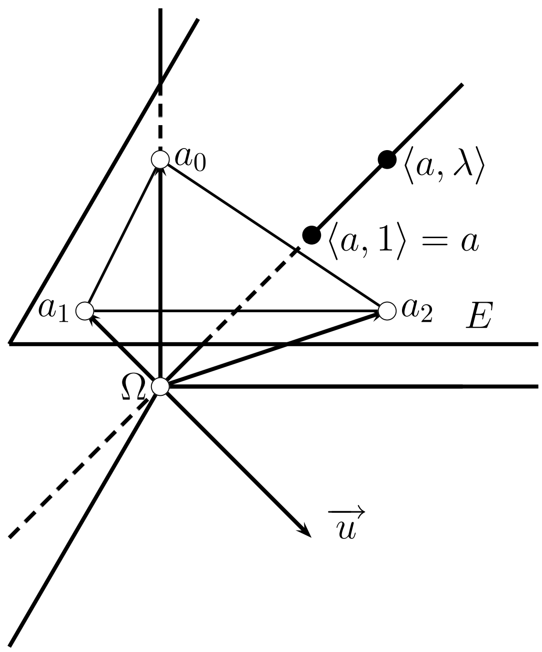

Figure 12.3: The basis (a_0, a_1, a_2) in $\widehat{E}$

Now that we have defined $\widehat{E}$ and investigated the relationship between affine frames in E and bases in $\widehat{E}$, we can give one more construction of a vector space $\mathcal{F}$ from E and $\overrightarrow{E}$, that will allow us to "visualize" in a much more intuitive fashion the structure of $\widehat{E}$ and of its operations $\widehat{+}$ and $\cdot$.

Definition 12.1.4 Given any affine space $(E, \overrightarrow{E})$, we define the vector space $\mathcal{F}$ as the direct sum $\overrightarrow{E} \oplus \mathbb{R}$, where $\mathbb{R}$ denotes the field $\mathbb{R}$ considered as a vector space (over itself). Denoting the unit vector in $\mathbb{R}$ as $\overrightarrow{1}$, since $\mathcal{F} = \overrightarrow{E} \oplus \mathbb{R}$, every vector $\overrightarrow{v} \in \mathcal{F}$ can be written as $\overrightarrow{v} = \overrightarrow{u} + \lambda\overrightarrow{1}$, for some unique $\overrightarrow{u} \in \overrightarrow{E}$, and some unique $\lambda \in \mathbb{R}$. Then, for any choice of an origin Ω_1 in E, we define the map $\widehat{\Omega}: \widehat{E} \to \mathcal{F}$, as follows:

$$\widehat{\Omega}(\theta) = \begin{cases} \lambda(\overrightarrow{1} + \overrightarrow{\Omega_1 a}) & \text{if } \theta = \langle a, \lambda \rangle,\ a \in E,\ \lambda \neq 0; \\ \overrightarrow{u} & \text{if } \theta = \overrightarrow{u},\ \overrightarrow{u} \in \overrightarrow{E}. \end{cases}$$

The idea is that, once again, viewing $\mathcal{F}$ as an affine space under its canonical structure, *E is embedded in $\mathcal{F}$ as the hyperplane $H = \overrightarrow{1} + \overrightarrow{E}$*, with direction $\overrightarrow{E}$, the hyperplane $\overrightarrow{E}$ in $\mathcal{F}$.

Then, every point $a \in E$ is in bijection with the point $A = \overrightarrow{1} + \overrightarrow{\Omega_1 a}$, in the hyperplane H.

Denoting the origin $\overrightarrow{0}$ of the canonical affine space $\mathcal{F}$ as Ω, the map $\widehat{\Omega}$ maps a point $\langle a, \lambda \rangle \in E$ to a point in $\mathcal{F}$, as follows: $\widehat{\Omega}(\langle a, \lambda \rangle)$ is the point on the line passing through both the origin Ω of $\mathcal{F}$ and the point $A = \overrightarrow{1} + \overrightarrow{\Omega_1 a}$ in the hyperplane $H = \overrightarrow{1} + \overrightarrow{E}$, such that

$$\widehat{\Omega}(\langle a, \lambda \rangle) = \lambda \overrightarrow{\Omega A} = \lambda(\overrightarrow{1} + \overrightarrow{\Omega_1 a}).$$

The following lemma shows that $\widehat{\Omega}$ is an isomorphism of vector spaces.

Lemma 12.1.5 *Given any affine space $(E, \overrightarrow{E})$, for any choice Ω_1 of an origin in E, the map $\widehat{\Omega}\colon \widehat{E} \rightarrow \mathcal{F}$ is a linear isomorphism between $\widehat{E}$ and the vector space $\mathcal{F}$ of definition 12.1.4. The inverse of $\widehat{\Omega}$ is given by*

$$\widehat{\Omega}^{-1}(\overrightarrow{u} + \lambda\overrightarrow{1}) = \begin{cases} \langle \Omega_1 + \lambda^{-1}\overrightarrow{u}, \lambda\rangle) & \text{if } \lambda \neq 0; \\ \overrightarrow{u} & \text{if } \lambda = 0. \end{cases}$$

The following diagram illustrates the embedding of the affine space E into the vector space $\mathcal{F}$, when E is an affine plane.

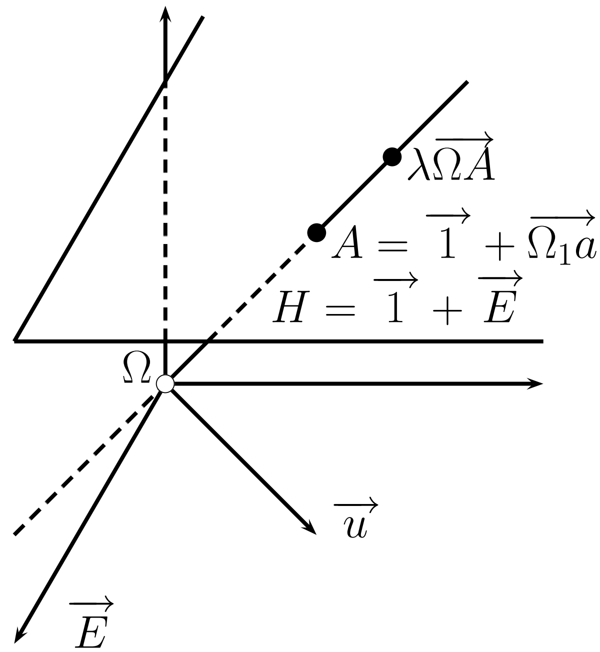

Figure 12.4: Embedding an affine space E into a vector space $\mathcal{F}$

We now consider the universal property of $\widehat{E}$. Other authors, such as Ramshaw, use the notation f_* for $\widehat{f}$.

First, we define rigorously the notion of homogenization of an affine space.

Definition 12.1.6 Given any affine space $(E, \overrightarrow{E})$, an *homogenization (or linearization) of* $(E, \overrightarrow{E})$, is a triple $\langle \mathcal{E}, j, \omega \rangle$, where $\mathcal{E}$ is a vector space, $j\colon E \to \mathcal{E}$ is an injective affine map with associated injective linear map $i\colon \overrightarrow{E} \to \mathcal{E}$, $\omega\colon \mathcal{E} \to \mathbb{R}$ is a linear form, such that $\omega^{-1}(0) = i(\overrightarrow{E})$, $\omega^{-1}(1) = j(E)$, and for every vector space $\overrightarrow{F}$ and every affine map $f\colon E \to \overrightarrow{F}$, there is a *unique* linear map $\widehat{f}\colon \mathcal{E} \to \overrightarrow{F}$ extending f, i.e. $f = \widehat{f} \circ j$, as in the following diagram:

$$\begin{tikzcd} E \arrow[r, "j"] \arrow[dr, "f"'] & \mathcal{E} \arrow[d, "\widehat{f}"] \\ & \overrightarrow{F} \end{tikzcd}$$

Thus, $j(E) = \omega^{-1}(1)$ is an affine hyperplane with direction $i(\overrightarrow{E}) = \omega^{-1}(0)$.

Lemma 12.1.7 *Given any affine space* $(E, \overrightarrow{E})$ *and any vector space* $\overrightarrow{F}$, *for any affine map* $f\colon E \to \overrightarrow{F}$, *there is a unique linear map* $\widehat{f}\colon \widehat{E} \to \overrightarrow{F}$ *extending* f, *such that*

$$\widehat{f}(\overrightarrow{u} \mathbin{\widehat{+}} \lambda a) = \lambda f(a) + \overrightarrow{f}(\overrightarrow{u}),$$

for all $a \in E$, *all* $\overrightarrow{u} \in \overrightarrow{E}$, *and all* $\lambda \in \mathbb{R}$, *where* $\overrightarrow{f}$ *is the linear map associated with* f. *In particular, when* $\lambda \neq 0$, *we have*

$$\widehat{f}(\overrightarrow{u} \mathbin{\widehat{+}} \lambda a) = \lambda f(a + \lambda^{-1}\overrightarrow{u}).$$

Lemma 12.1.7 shows that $\langle \widehat{E}, j, \omega \rangle$, is an homogenization of $(E, \overrightarrow{E})$. As a corollary, we obtain the following lemma.

Lemma 12.1.8 *Given two affine spaces E and F and an affine map $f\colon E \to F$, there is a unique linear map $\widehat{f}\colon \widehat{E} \to \widehat{F}$ extending f, as in the diagram below,*

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[d, "j"'] & F \arrow[d, "j"] \\ \widehat{E} \arrow[r, "\widehat{f}"] & \widehat{F} \end{tikzcd}$$

such that

$$\widehat{f}(\overrightarrow{u} \mathbin{\widehat{+}} \lambda a) = \overrightarrow{f}(\overrightarrow{u}) \mathbin{\widehat{+}} \lambda f(a),$$

for all $a \in E$, all $\overrightarrow{u} \in \overrightarrow{E}$, and all $\lambda \in \mathbb{R}$, where $\overrightarrow{f}$ is the linear map associated with f. In particular, when $\lambda \neq 0$, we have

$$\widehat{f}(\overrightarrow{u} \mathbin{\widehat{+}} \lambda a) = \lambda f(a + \lambda^{-1}\overrightarrow{u}).$$

From a practical point of view, lemma 12.1.8 shows us how to homogenize an affine map to turn it into a linear map between the two homogenized spaces.

Assume that E and F are of finite dimension, and that $(a_0, (\overrightarrow{u_1}, \ldots, \overrightarrow{u_n}))$ is an affine basis of E, with origin a_0, and $(b_0, (\overrightarrow{v_1}, \ldots, \overrightarrow{v_m}))$ is an affine basis of F, with origin b_0.

Then, with respect to the two bases $(\overrightarrow{u_1}, \ldots, \overrightarrow{u_n}, a_0)$ in $\widehat{E}$ and $(\overrightarrow{v_1}, \ldots, \overrightarrow{v_m}, b_0)$ in $\widehat{F}$, a linear map $h\colon \widehat{E} \to \widehat{F}$ is given by an $(m+1) \times (n+1)$ matrice A.

If this linear map h is equal to the homogenized version $\widehat{f}$ of an affine map f, since

$$\widehat{f}(\overrightarrow{u} \mathbin{\widehat{+}} \lambda a) = \overrightarrow{f}(\overrightarrow{u}) \mathbin{\widehat{+}} \lambda f(a),$$

since over the basis $(\overrightarrow{u_1}, \ldots, \overrightarrow{u_n}, a_0)$ in $\widehat{E}$, points are represented by vectors whose last coordinate is 1, and vectors are represented by vectors whose last coordinate is 0, the last row of the matrix $A = M(\widehat{f})$ with respect to the given bases is

$$(0, 0, \ldots, 0, 1),$$

with m occurrences of 0, the last column contains the coordinates

$$(\mu_1, \ldots, \mu_m, 1)$$

of $f(a_0)$ with respect to the basis $(\overrightarrow{v_1}, \ldots, \overrightarrow{v_m}, b_0)$,
the submatrix of A obtained by deleting the last row and the last column is the matrix of the linear map $\overrightarrow{f}$ with respect to the bases $(\overrightarrow{u_1}, \ldots, \overrightarrow{u_n})$ and $(\overrightarrow{v_1}, \ldots, \overrightarrow{v_m})$, and since

$$f(a_0 + \overrightarrow{u}) = \widehat{f}(\overrightarrow{u} \mathbin{\widehat{+}} a_0),$$

given any $x \in E$ and $y \in F$, with coordinates
$(x_1, \ldots, x_n, 1)$ and $(y_1, \ldots, y_m, 1)$,
for $X = (x_1, \ldots, x_n, 1)^\top$ and $Y = (y_1, \ldots, y_m, 1)^\top$, we have

$$y = f(x) \quad \text{iff} \quad Y = AX.$$

For example, consider the following affine map $f\colon \mathbb{A}^2 \to \mathbb{A}^2$ defined as follows:

$$
\begin{aligned}
y_1 &= ax_1 + bx_2 + \mu_1, \\
y_2 &= cx_1 + dx_2 + \mu_2.
\end{aligned}
$$

The matrix of $\widehat{f}$ is

$$
\begin{pmatrix}
a & b & \mu_1 \\
c & d & \mu_2 \\
0 & 0 & 1
\end{pmatrix}
$$

and we have

$$
\begin{pmatrix} y_1 \\ y_2 \\ 1 \end{pmatrix} = \begin{pmatrix}
a & b & \mu_1 \\
c & d & \mu_2 \\
0 & 0 & 1
\end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix}
$$

In $\widehat{E}$, we have

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} a & b & \mu_1 \\ c & d & \mu_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

which means that the homogeneous map $\widehat{f}$ is is obtained from f by “adding the variable of homogeneity x_3”:

$$\begin{aligned} y_1 &= ax_1 + bx_2 + \mu_1 x_3, \\ y_2 &= cx_1 + dx_2 + \mu_2 x_3, \\ y_3 &= x_3. \end{aligned}$$

We now show how to homogenize multiaffine maps.

Lemma 12.1.9 *Given any affine space E and any vector space $\overrightarrow{F}$, for any m-affine map $f\colon E^m \to \overrightarrow{F}$, there is a unique m-linear map $\widehat{f}\colon (\widehat{E})^m \to \overrightarrow{F}$ extending f, such that, if*

$$f(a_1 + \overrightarrow{v_1}, \ldots, a_m + \overrightarrow{v_m}) = f(a_1, \ldots, a_m) + \sum_{\substack{S \subseteq \{1,\ldots,m\},\, k = \mathrm{card}(S) \\ S = \{i_1,\ldots,i_k\},\, k \geq 1}} f_S(\overrightarrow{v_{i_1}}, \ldots, \overrightarrow{v_{i_k}}),$$

for all $a_1 \ldots, a_m \in E$, and all $\overrightarrow{v_1}, \ldots, \overrightarrow{v_m} \in \overrightarrow{E}$, where the f_S are uniquely determined multilinear maps (by lemma **??***), then*

$$\widehat{f}(\overrightarrow{v_1} \,\widehat{+}\, \lambda_1 a_1, \ldots, \overrightarrow{v_m} \,\widehat{+}\, \lambda_m a_m) = \lambda_1 \cdots \lambda_m f(a_1, \ldots, a_m) + \sum_{\substack{S \subseteq \{1,\ldots,m\},\, k = \mathrm{card}(S) \\ S = \{i_1,\ldots,i_k\},\, k \geq 1}} \Big(\prod_{\substack{j \in \{1,\ldots,m\} \\ j \notin S}} \lambda_j \Big) f_S(\overrightarrow{v_{i_1}}, \ldots, \overrightarrow{v_{i_k}}),$$

for all $a_1 \ldots, a_m \in E$, all $\overrightarrow{v_1}, \ldots, \overrightarrow{v_m} \in \overrightarrow{E}$, and all $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$. Furthermore, for $\lambda_i \neq 0$, $1 \leq i \leq m$, we have

$$\widehat{f}(\overrightarrow{v_1} \,\widehat{+}\, \lambda_1 a_1, \ldots, \overrightarrow{v_m} \,\widehat{+}\, \lambda_m a_m) = \lambda_1 \cdots \lambda_m f(a_1 + \lambda_1^{-1} \overrightarrow{v_1}, \ldots, a_m + \lambda_m^{-1} \overrightarrow{v_m}).$$

12.2 Differentiating Affine Polynomial Functions Using Their Homogenized Polar Forms, Osculating Flats

Let $\delta = \overrightarrow{1}$, the unit (vector) in $\mathbb{R}$. When dealing with derivatives, it is also more convenient to denote the vector $\overrightarrow{ab}$ as $b - a$.

For any $\overline{a} \in \mathbb{A}$, the derivative $\mathrm{D}F(\overline{a})$ is the limit,

$$\lim_{t\to 0,\, t\neq 0} \frac{F(\overline{a} + t\delta) - F(\overline{a})}{t},$$

if it exists.

However, since $\widehat{F}$ agrees with F on $\mathbb{A}$, we have

$$F(\overline{a} + t\delta) - F(\overline{a}) = \widehat{F}(\overline{a} + t\delta) - \widehat{F}(\overline{a}),$$

and thus, we need to see what is the limit of

$$\frac{\widehat{F}(\overline{a} + t\delta) - \widehat{F}(\overline{a})}{t},$$

when $t \to 0$, $t \neq 0$, with $t \in \mathbb{R}$.

Recall that since $F\colon \mathbb{A} \to \mathcal{E}$, where $\mathcal{E}$ is an affine space, the derivative $\mathrm{D}F(\overline{a})$ of F at $\overline{a}$ is a *vector* in $\overrightarrow{\mathcal{E}}$, and not a point in $\mathcal{E}$.

However, the structure of $\widehat{\mathcal{E}}$ takes care of this, since $\widehat{F}(\overline{a}+t\delta) - \widehat{F}(\overline{a})$ is indeed a vector (remember our convention that $-$ is an abbreviation for $\widehat{-}$).

Since

$$\widehat{F}(\overline{a}+t\delta) = \widehat{f}(\underbrace{\overline{a}+t\delta,\ldots,\overline{a}+t\delta}_{m}),$$

where $\widehat{f}$ is the homogenized version of the polar form f of F, and $\widehat{F}$ is the homogenized version of F, since

$$\widehat{F}(\overline{a}+t\delta) - \widehat{F}(\overline{a}) = \widehat{f}(\underbrace{\overline{a}+t\delta,\ldots,\overline{a}+t\delta}_{m}) - \widehat{f}(\underbrace{\overline{a},\ldots,\overline{a}}_{m}),$$

by multilinearity and symmetry, we have

$$\widehat{F}(\overline{a}+t\delta) - \widehat{F}(\overline{a}) =$$

$$m\,t\,\widehat{f}(\underbrace{\overline{a},\ldots,\overline{a}}_{m-1},\delta) + \sum_{k=2}^{k=m} \binom{m}{k}\, t^k\, \widehat{f}(\underbrace{\overline{a},\ldots,\overline{a}}_{m-k},\underbrace{\delta,\ldots,\delta}_{k}),$$

and thus,

$$\lim_{t\to 0,\, t\neq 0} \frac{\widehat{F}(\overline{a}+t\delta) - \widehat{F}(\overline{a})}{t} = m\widehat{f}(\underbrace{\overline{a},\ldots,\overline{a}}_{m-1},\delta).$$

However, since $\widehat{F}$ extends F on $\mathbb{A}$, we have $\mathrm{D}F(\overline{a}) = \mathrm{D}\widehat{F}(\overline{a})$, and thus, we showed that

$$\mathrm{D}F(\overline{a}) = m\widehat{f}(\underbrace{\overline{a}, \ldots, \overline{a}}_{m-1}, \delta).$$

This shows that the derivative of F at $\overline{a} \in \mathbb{A}$ can be computed by evaluating the homogenized version $\widehat{f}$ of the polar form f of F, by replacing just one occurrence of $\overline{a}$ in $\widehat{f}(\overline{a}, \ldots, \overline{a})$ by δ.

More generally, we have the following useful lemma.

Lemma 12.2.1 *Given an affine polynomial function $F\colon \mathbb{A} \to \mathcal{E}$ of polar degree m, where $\mathcal{E}$ is a normed affine space, the k-th derivative $\mathrm{D}^k F(\overline{a})$ can be computed from the homogenized polar form $\widehat{f}$ of F as follows, where $1 \leq k \leq m$:*

$$\mathrm{D}^k F(\overline{a}) = m(m-1)\cdots(m-k+1)\,\widehat{f}(\underbrace{\overline{a},\ldots,\overline{a}}_{m-k},\underbrace{\delta,\ldots,\delta}_{k}).$$

Since coefficients of the form $m(m-1)\cdots(m-k+1)$ occur a lot when taking derivatives, following Knuth, it is useful to introduce the *falling power* notation. We define the *falling power* $m^{\underline{k}}$, as

$$m^{\underline{k}} = m(m-1)\cdots(m-k+1),$$

for $0 \leq k \leq m$, with $m^{\underline{0}} = 1$, and with the convention that $m^{\underline{k}} = 0$ when $k > m$.

Using the falling power notation, the previous lemma reads as

$$\mathrm{D}^k F(\overline{a}) = m^{\underline{k}}\,\widehat{f}(\underbrace{\overline{a},\ldots,\overline{a}}_{m-k},\underbrace{\delta,\ldots,\delta}_{k}).$$

We also get the following explicit formula in terms of control points.

Lemma 12.2.2 *Given an affine polynomial function $F\colon \mathbb{A} \to \mathcal{E}$ of polar degree m, where $\mathcal{E}$ is a normed affine space, for any $\overline{r}, \overline{s} \in \mathbb{A}$, with $r \neq s$, the k-th derivative $\mathrm{D}^k F(\overline{r})$ can be computed from the polar form f of F as follows, where $1 \leq k \leq m$:*

$$\mathrm{D}^k F(\overline{r}) = \frac{m^{\underline{k}}}{(s-r)^k} \sum_{i=0}^{i=k} \binom{k}{i} (-1)^{k-i} f(\underbrace{\overline{r}, \ldots, \overline{r}}_{m-i}, \underbrace{\overline{s}, \ldots, \overline{s}}_{i}).$$

If F is specified by the sequence of $m+1$ control points $b_i = f(\overline{r}^{\,m-i}\, \overline{s}^{\,i})$, $0 \leq i \leq m$, the above lemma shows that the k-th derivative $\mathrm{D}^k F(\overline{r})$ of F at $\overline{r}$, depends only on the $k+1$ control points $b_0, \ldots, b_k$

In terms of the control points $b_0, \ldots, b_k$, the formula of lemma **??** reads as follows:

$$\mathrm{D}^k F(\overline{r}) = \frac{m^{\underline{k}}}{(s-r)^k} \sum_{i=0}^{i=k} \binom{k}{i} (-1)^{k-i}\, b_i.$$

In particular, if $b_0 \neq b_1$, then $\mathrm{D}F(\overline{r})$ is the velocity vector of F at b_0, and it is given by

$$\mathrm{D}F(\overline{r}) = \frac{m}{s-r}\,\overrightarrow{b_0b_1} = \frac{m}{s-r}\,(b_1 - b_0),$$

the last expression making sense in $\widehat{\mathcal{E}}$.

In terms of the de Casteljau diagram

$$\mathrm{D}F(\bar{t}) = \frac{m}{s-r}(b_{1,m-1} - b_{0,m-1}).$$

Similarly, the acceleration vector $\mathrm{D}^2F(\overline{r})$ is given by

$$\mathrm{D}^2F(\overline{r}) = \frac{m(m-1)}{(s-r)^2}\,(\overrightarrow{b_0b_2} - 2\overrightarrow{b_0b_1}) =$$
$$\frac{m(m-1)}{(s-r)^2}\,(b_2 - 2b_1 + b_0),$$

the last expression making sense in $\widehat{\mathcal{E}}$.

Later on when we deal with surfaces, it will be necessary to generalize the above results to directional derivatives. However, we have basically done all the work already.

Let us assume that E and $\mathcal{E}$ are normed affine spaces, and consider a map $F\colon E \to \mathcal{E}$.

Recall from definition **??**, that if A is any open subset of E, for any $a \in A$, for any $\overrightarrow{u} \neq \overrightarrow{0}$ in $\overrightarrow{E}$, the *directional derivative of F at a w.r.t. the vector $\overrightarrow{u}$*, denoted as $\mathrm{D}_u F(a)$, is the limit, if it exists,

$$\lim_{t \to 0, t \in U, t \neq 0} \frac{F(a + t\overrightarrow{u}) - F(a)}{t},$$

where $U = \{t \in \mathbb{R} \mid a + t\overrightarrow{u} \in A\}$.

If $F\colon E \to \mathcal{E}$ is a polynomial function of degree m, with polar form the symmetric multiaffine map $f\colon E^m \to \mathcal{E}$, then

$$F(a + t\overrightarrow{u}) - F(a) = \widehat{F}(a + t\overrightarrow{u}) - \widehat{F}(a),$$

where $\widehat{F}$ is the homogenized version of F, that is, the polynomial map $\widehat{F}\colon \widehat{E} \to \widehat{\mathcal{E}}$ associated with the homogenized version $f\colon (\widehat{E})^m \to \widehat{\mathcal{E}}$ of the polar form $f\colon E^m \to \mathcal{E}$ of $F\colon E \to \mathcal{E}$.

Thus, $\mathrm{D}_u F(a)$ exists iff the limit

$$\lim_{t\to 0,\, t\neq 0} \frac{\widehat{F}(a + t\overrightarrow{u}) - \widehat{F}(a)}{t}$$

exists, and in this case, this limit is $\mathrm{D}_u F(a) = \mathrm{D}_u \widehat{F}(a)$. We get

$$\mathrm{D}_u F(a) = m\widehat{f}(\underbrace{a, \ldots, a}_{m-1}, \overrightarrow{u}).$$

By a simple, induction, we can prove the following lemma.

Lemma 12.2.3 *Given an affine polynomial function $F\colon E \to \mathcal{E}$ of polar degree m, where E and $\mathcal{E}$ are normed affine spaces, for any k nonzero vectors $\overrightarrow{u_1}$, ..., $\overrightarrow{u_k} \in \overrightarrow{E}$, where $1 \leq k \leq m$, the k-th directional derivative $\mathrm{D}_{u_1} \ldots \mathrm{D}_{u_k} F(a)$ can be computed from the homogenized polar form $\widehat{f}$ of F as follows:*

$$\mathrm{D}_{u_1} \ldots \mathrm{D}_{u_k} F(a) = m^{\underline{k}}\, \widehat{f}(\underbrace{a, \ldots, a}_{m-k}, \overrightarrow{u_1}, \ldots, \overrightarrow{u_k}).$$

If E has finite dimension,

$$\mathrm{D}^k F(a)(\overrightarrow{u_1}, \ldots, \overrightarrow{u_k}) = m^{\underline{k}}\, \widehat{f}(\underbrace{a, \ldots, a}_{m-k}, \overrightarrow{u_1}, \ldots, \overrightarrow{u_k}).$$