CHAPTER 6

Affine morphisms

Contents

6.1 Affine morphisms

Definition 6.1. Let $\mathbf{X}$ and $\mathbf{Y}$ be two affine spaces with associated vector space $\mathbf{V}$ and $\mathbf{W}$ respectively. A map

$$\phi : \mathbf{X} \to \mathbf{Y}$$

is called *affine morphism* if there is a linear map

$$\mathrm{lin}(\phi) : \mathbf{V} \to \mathbf{W} \quad \text{such that} \quad \overrightarrow{\phi(A)\phi(B)} = \mathrm{lin}(\phi)(\overrightarrow{AB})$$

for any vector $\overrightarrow{AB} \in \mathbf{V}$. We call $\mathrm{lin}(\phi)$ the *linear map associated to the affine morphism* ϕ, or, shorter, *the linear part of* ϕ.

Proposition 6.2. Let $\mathbf{X}$ and $\mathbf{Y}$ be two affine space with associated vector space $\mathbf{V}$ and $\mathbf{W}$ respectively. Let $\mathcal{X}$ and $\mathcal{Y}$ be coordinate systems for $\mathbf{X}$ and $\mathbf{Y}$ respectively. Let $n = \dim(\mathbf{X})$ and $m = \dim(\mathbf{Y})$. A map $\phi : \mathbf{X} \to \mathbf{Y}$ is an affine morphism if there exists a matrix $\mathrm{M}_\phi \in \mathrm{Mat}_{m\times n}(\mathbf{K})$ and $\mathbf{b}_\phi \in \mathrm{Mat}_{m\times 1}(\mathbf{K})$ such that for any point $P \in \mathbf{X}$

$$[\phi(P)]_\mathcal{Y} = \mathrm{M}_\phi [P]_\mathcal{X} + \mathbf{b}_\phi.$$

If this is the case then $\mathrm{M}_\phi = [\mathrm{lin}(\phi)]_{\mathcal{Y},\mathcal{X}}$.

- In the context of vector spaces you discussed linear maps (which are also called *linear morphisms*). In the context of affine spaces the analog of linear maps are *affine morphisms*.
- If $\psi : \mathbf{V} \to \mathbf{W}$ is a linear map, then viewing $\mathbf{V}$ and $\mathbf{W}$ as affine spaces we have a map $\psi : \mathbf{V}_a \to \mathbf{W}_a$. It is the same map and it is an affine map. In other words, linear maps are affine maps.
- With the right interpretation, Proposition 6.2 is saying that an affine morphism is a map of the form
$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mapsto \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + b.$$

Definition 6.3. The set of all affine morphisms between the affine space $\mathbf{X}$ and the affine space $\mathbf{Y}$ is denoted by $\mathrm{Hom}_{\mathrm{aff}}(\mathbf{X},\mathbf{Y})$. The set of all $\phi \in \mathrm{Hom}_{\mathrm{aff}}(\mathbf{X},\mathbf{Y})$ which are bijective is denoted by $\mathrm{Iso}_{\mathrm{aff}}(\mathbf{X},\mathbf{Y})$. The elements of $\mathrm{Iso}_{\mathrm{aff}}(\mathbf{X})$ are called *affine isomorphisms*. The set of all affine morphisms from $\mathbf{X}$ to itself is denoted by $\mathrm{End}_{\mathrm{aff}}(\mathbf{X})$. The elements of $\mathrm{End}_{\mathrm{aff}}(\mathbf{X})$ are called *affine endomorphisms*. The set of all $\phi \in \mathrm{End}_{\mathrm{aff}}(\mathbf{X})$ which are bijective is denoted by $\mathrm{AGL}(\mathbf{X})$. The elements of $\mathrm{AGL}(\mathbf{X})$ are called *affine automorphisms*, or *affine transformations*.

- Notice that $\mathrm{AGL}(\mathbf{X}) = \mathrm{Iso}_{\mathrm{aff}}(\mathbf{X},\mathbf{X})$.
- Notice also that $\mathrm{AGL}(\mathbf{X})$ is a group with group law given by composition, $\phi \circ \psi$.
- Affine changes of coordinates are affine morphisms $\phi : \mathbf{X} \to \mathbf{X}$ since
$$[\phi(P)]_\mathcal{Y} = \mathrm{M}_{\mathcal{Y},\mathcal{X}} [P]_\mathcal{X} + [O]_\mathcal{Y}.$$
The linear part of an affine coordinate change is the base change of the associated vector spaces.
- One can show that any affine automorphisms is an affine change of coordinates, so one can view $\mathrm{AGL}(\mathbf{X})$ as the group of affine changes of coordinates.

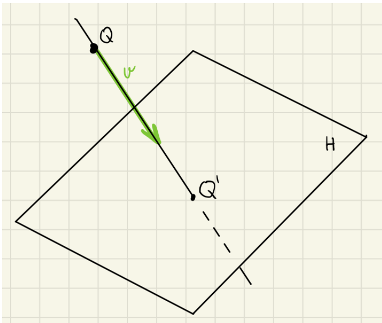
6.2 Parallel projections and reflections for hyperplanes

Let $\mathbf{A}$ be an affine space with associated vector space $\mathbf{V}$. All equations are with respect to a fixed coordinate system $\mathcal{K} = O\mathbf{e}_1 \ldots \mathbf{e}_n$ of $\mathbf{A}$.

6.2.1 Intersection of a line with a hyperplane

Consider a line $\ell \subseteq \mathbf{A}$ passing through a point $Q(q_1,\ldots,q_n)$ and having $\mathbf{v}(v_1,\ldots,v_n)$ as direction vector:

$$\ell = \{Q + t\mathbf{v} : t \in \mathbf{K}\}. \tag{6.1}$$



Consider a hyperplane $H \subseteq \mathbf{A}$ with associated vector subspace $\mathbf{W} \subseteq \mathbf{V}$ given by the Cartesian equation

$$H : \overbrace{\underbrace{a_1x_1 + \cdots + a_nx_n}_{(\mathrm{lin}\,\varphi)(x_1,\ldots,x_n)} + a_{n+1}}^{\varphi(x_1,\ldots,x_n)} = 0 \tag{6.2}$$

Notice that the hyperplane can be viewed as the zero-locus of an affine morphism $\varphi : \mathbf{A} \to \mathbf{K}_a$:

$$H = \varphi^{-1}(0) \quad \text{and} \quad \mathbf{W} = (\mathrm{lin}\,\varphi)^{-1}(0).$$

The intersection $\ell \cap H$ can be described as follows

$$\begin{aligned} Q + t\mathbf{v} \in \ell \cap H \quad &\Leftrightarrow \quad \varphi(Q + t\mathbf{v}) = 0 \\ &\Leftrightarrow \quad \varphi(Q) + \mathrm{lin}\,\varphi(t\mathbf{v}) = 0 \\ &\Leftrightarrow \quad \varphi(Q) + t\,\mathrm{lin}\,\varphi(\mathbf{v}) = 0. \end{aligned}$$

So, the intersection point (if it exists) is

$$Q' = Q - \frac{\varphi(Q)}{\mathrm{lin}\,\varphi(\mathbf{v})}\mathbf{v}. \tag{6.3}$$

6.2.2 Tensor products

Definition 6.4. Let $\mathbf{v}(v_1,\ldots,v_n)$ and $\mathbf{w}(w_1,\ldots,w_n)$ be two vectors. The *tensor product* $\mathbf{v} \otimes \mathbf{w}$ is the $n \times n$ matrix defined by $(\mathbf{v} \times \mathbf{w})_{ij} = v_iw_j$. In other words

$$\mathbf{v} \otimes \mathbf{w} = \mathbf{v} \cdot \mathbf{w}^t = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \cdot \begin{bmatrix} w_1, \ldots, w_n \end{bmatrix} = \begin{bmatrix} v_1w_1 & \ldots & v_1w_n \\ \vdots & & \vdots \\ v_nw_1 & \ldots & v_nw_n \end{bmatrix}.$$

Proposition 6.5. The map $\mathbf{K}^n \times \mathbf{K}^n \to \mathrm{Mat}_{n\times n}(\mathbf{K})$ given by $(\mathbf{v},\mathbf{w}) \mapsto \mathbf{v}\otimes\mathbf{w}$ has the following properties:

1. It is linear in both arguments,
2. $(\mathbf{v}\otimes\mathbf{w})^t = \mathbf{w}\otimes\mathbf{v}$.

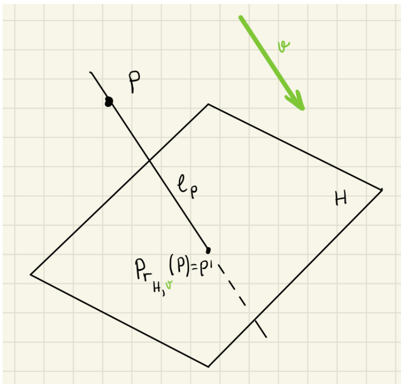
- The tensor product of two vectors is also called outer product - to be compared with the inner product which is

$$\langle \mathbf{v},\mathbf{w}\rangle = \mathbf{v}^t \cdot \mathbf{w} = \begin{bmatrix} v_1,\dots,v_n \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = v_1 w_n + \dots + v_n w_n.$$

- It is easy to show that for three vectors $\mathbf{u},\mathbf{v},\mathbf{w} \in \mathbf{V}$ we have

$$(\mathbf{u}\otimes\mathbf{v})\mathbf{w} = (\mathbf{v}\cdot\mathbf{w})\mathbf{u}.$$

6.2.3 Parallel projection on a hyperplane



Definition 6.6. Let H be a hyperplane and let $\mathbf{v}$ be a vector in $\mathbf{V}$ which is not parallel to H. For any point $P \in \mathbf{A}$ there is a unique line ℓ_P passing through P and having $\mathbf{v}$ as direction vector. The line ℓ_P is not parallel to H, hence, it intersects H in a unique point P'. We denote P' by $\mathrm{Pr}_{H,\mathbf{v}}(P)$ and call it the *projection of the point P on the hyperplane H parallel to* $\mathbf{v}$. This gives a map

$$\mathrm{Pr}_{H,\mathbf{v}} : \mathbf{A} \to \mathbf{A}$$

called, the *projection on the hyperplane H parallel to* $\mathbf{v}$.

- With respect to the reference frame $\mathcal{K}$, the hyperplane H has an equation as in (6.2).
- By (6.3), $\mathrm{Pr}_{H,\mathbf{v}}(P) = P - \frac{\varphi(P)}{\mathrm{lin}\,\varphi(\mathbf{v})}\mathbf{v}$.

- Hence, if we denote by $p_1',\dots,p_n'$ the coordinates of the projected point $\mathrm{Pr}_{H,\mathbf{v}}(P)$ then

$$\begin{cases} p_1' = p_1 + \mu v_1 \\ \quad\vdots \\ p_n' = p_n + \mu v_n \end{cases} \quad \text{where} \quad \mu = -\frac{a_1p_1 + \cdots + a_np_n + a_{n+1}}{a_1v_1 + \cdots + a_nv_n}.$$

- In matrix form, we can rearrange this as follows

$$\begin{bmatrix} p_1' \\ p_2' \\ \vdots \\ p_n' \end{bmatrix} = \underbrace{\frac{1}{\operatorname{lin}\varphi(\mathbf{v})}\begin{bmatrix} \sum_{i=1}^{n,i\neq 1} a_iv_i & -a_2v_1 & -a_3v_1 & \dots & -a_nv_1 \\ -a_1v_2 & \sum_{i=1}^{n,i\neq 2} a_iv_i & -a_3v_2 & \dots & -a_nv_2 \\ \vdots & & & & \vdots \\ -a_1v_n & -a_2v_n & \dots & -a_{n-1}v_n & \sum_{i=1}^{n,i\neq n} a_iv_i \end{bmatrix}}_{\operatorname{lin}\mathrm{Pr}_{H,\mathbf{v}}}\begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix} - \frac{a_{n+1}}{\operatorname{lin}\varphi(\mathbf{v})}\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

- It is possible to give a more compact description of the above matrix form if we use tensor products:

$$[\mathrm{Pr}_{H,\mathbf{v}}(P)]_{\mathcal{K}} = \frac{1}{\operatorname{lin}\varphi(\mathbf{v})}\left((\mathbf{v}^t\cdot\mathbf{a})\,\mathrm{Id}_n - \underbrace{\mathbf{v}\cdot\mathbf{a}^t}_{\mathbf{v}\otimes\mathbf{a}}\right)\cdot[P]_{\mathcal{K}} - \frac{a_{n+1}}{\operatorname{lin}\varphi(\mathbf{v})}[\mathbf{v}]_{\mathcal{K}}$$

  where $\mathbf{a} = (a_1,\dots,a_n)^t$.

- If we further notice that $\operatorname{lin}\varphi(\mathbf{v}) = \mathbf{v}^t\cdot\mathbf{a}$ then

$$[\mathrm{Pr}_{H,\mathbf{v}}(P)]_{\mathcal{K}} = \left(\mathrm{Id}_n - \frac{\mathbf{v}\cdot\mathbf{a}^t}{\mathbf{v}^t\cdot\mathbf{a}}\right)\cdot[P]_{\mathcal{K}} - \frac{a_{n+1}}{\mathbf{v}^t\cdot\mathbf{a}}[\mathbf{v}]_{\mathcal{K}}$$

- Parallel projections on hyperplanes are affine morphisms. Obviously, they are not bijective, so

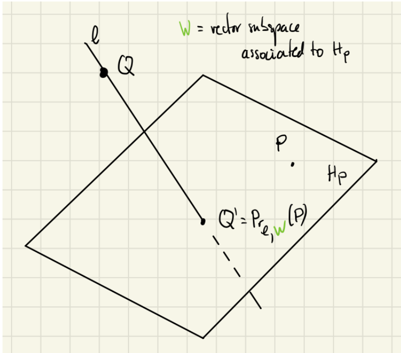
$$\mathrm{Pr}_{H,\mathbf{v}} \in \mathrm{End}_{\mathrm{aff}}(\mathbf{A}) \quad \text{but} \quad \mathrm{Pr}_{H,\mathbf{v}} \notin \mathrm{AGL}(\mathbf{A}).$$

6.2.4 Parallel projection on a line

Definition 6.7. Let ℓ be a line and let $\mathbf{W}$ be an $(n-1)$-dimensional vector subspace in $\mathbf{V}$ which is not parallel to ℓ. For any point $P\in\mathbf{A}$ there is a unique hyperplane H_P passing through P and having $\mathbf{W}$ as associated vector subspace. The hyperplane H_P is not parallel to ℓ, hence, it intersects ℓ in a unique point P'. We denote P' by $\mathrm{Pr}_{\ell,\mathbf{W}}(P)$ and call it the *projection of the point P on the line ℓ parallel to* $\mathbf{W}$. This gives a map

$$\mathrm{Pr}_{\ell,\mathbf{W}} : \mathbf{A} \to \mathbf{A}$$

called, the *projection on the line ℓ parallel to* $\mathbf{W}$.



- With respect to the reference frame $\mathcal{K}$, the vector subspace $\mathbf{W}$ is given by a homogeneous equation

$$\overbrace{a_1x_1 + a_2x_2 + \cdots + a_nx_n}^{\psi(x_1,\ldots,x_n)} = 0$$

- For a given point $P \in \mathbf{A}$, the equation of H_P is

$$H_P : a_1x_1 + a_2x_2 + \cdots + a_nx_n - \psi(P) = 0$$

- Hence, if we denote by $p'_1,\ldots,p'_n$ the coordinates of the projected point $\mathrm{Pr}_{\ell,\mathbf{W}}(P)$ then

$$\begin{cases} p'_1 = q_1 + v_1\mu \\ \quad\vdots \\ p'_n = q_n + v_n\mu \end{cases} \quad \text{where} \quad \mu = -\frac{\psi(Q) - \psi(P)}{\psi(\mathbf{v})}$$

  where $Q(q_1,\ldots,q_n)$ is a point on ℓ and $\mathbf{v}(v_1,\ldots,v_n)$ is a direction vector for ℓ (as in (6.1)).

- In matrix form we can rearrange this as follows

$$[\mathrm{Pr}_{\ell,\mathbf{W}}(P)]_{\mathcal{K}} = \underbrace{\frac{1}{\psi(\mathbf{v})}\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}\begin{bmatrix} a_1 & \ldots & a_n \end{bmatrix}}_{\mathrm{lin}\,\mathrm{Pr}_{\ell,\mathbf{W}} = \frac{1}{\psi(\mathbf{v})}\mathbf{v}\otimes\mathbf{a}}\begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix} + \left(\begin{bmatrix} q_1 \\ \vdots \\ q_n \end{bmatrix} - \frac{\psi(Q)}{\psi(\mathbf{v})}\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}\right) = \frac{\mathbf{v}\cdot\mathbf{a}^t}{\mathbf{v}^t\cdot\mathbf{a}}[P]_{\mathcal{K}} + \left([Q]_{\mathcal{K}} - \frac{\mathbf{v}\cdot[Q]_{\mathcal{K}}}{\mathbf{v}^t\cdot\mathbf{a}}\mathbf{v}\right).$$

  where $\mathbf{a} = (a_1,\ldots,a_n)^t$.

- Parallel projections on lines are affine morphisms. Obviously, they are not bijective, so

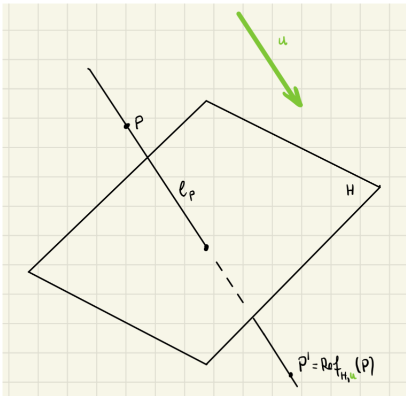
$$\mathrm{Pr}_{\ell,\mathbf{W}} \in \mathrm{End}_{\mathrm{aff}}(\mathbf{A}) \quad \text{but} \quad \mathrm{Pr}_{\ell,\mathbf{W}} \notin \mathrm{AGL}(\mathbf{A}).$$

6.2.5 Parallel reflection in a hyperplane

Definition 6.8. Let H be a hyperplane and let $\mathbf{v}$ be a vector in $\mathbf{V}$ which is not parallel to H. For any point $P \in \mathbf{A}$ there is a unique point P' such that $\mathrm{Pr}_{H,\mathbf{v}}(P)$ is the midpoint of the segment $[PP']$. We denote P' by $\mathrm{Ref}_{H,\mathbf{v}}(P)$ and call it the *reflection of the point P in the hyperplane H parallel to* $\mathbf{v}$. This gives a map

$$\mathrm{Ref}_{H,\mathbf{v}} : \mathbf{A} \to \mathbf{A}$$

called, the *reflection in the hyperplane H parallel to* $\mathbf{v}$.



- With respect to the reference frame $\mathcal{K}$, the hyperplane H has an equation as in (6.2).

- Since $\mathrm{Pr}_{H,\mathbf{v}}(P)$ is the midpoint of the segment $[PP']$, with respect to $\mathcal{K}$ we have

$$[P]_{\mathcal{K}} - \frac{\varphi(P)}{\mathrm{lin}\,\varphi(\mathbf{v})}[\mathbf{v}]_{\mathcal{K}} = \frac{[P]_{\mathcal{K}} + [P']_{\mathcal{K}}}{2} \quad \Leftrightarrow \quad 2[P]_{\mathcal{K}} - 2\frac{\varphi(P)}{\mathrm{lin}\,\varphi(\mathbf{v})}[\mathbf{v}]_{\mathcal{K}} = [P]_{\mathcal{K}} + [P']_{\mathcal{K}}.$$

  Therefore

$$[\mathrm{Ref}_{H,\mathbf{v}}(P)]_{\mathcal{K}} = [P]_{\mathcal{K}} - 2\frac{\varphi(P)}{\mathrm{lin}\,\varphi(\mathbf{v})}[\mathbf{v}]_{\mathcal{K}}.$$

- As in the case of $\mathrm{Pr}_{H,\mathbf{v}}$, it is possible to give a compact description of the matrix form if we use tensor products:

$$[\mathrm{Pr}_{H,\mathbf{v}}(P)]_{\mathcal{K}} = \left(\mathrm{Id}_n - 2\frac{\mathbf{v} \otimes \mathbf{a}}{\mathbf{v}^t \cdot \mathbf{a}}\right) \cdot [P]_{\mathcal{K}} - 2\frac{a_{n+1}}{\mathbf{v}^t \cdot \mathbf{a}}[\mathbf{v}]_{\mathcal{K}}$$

  where $\mathbf{a} = (a_1, \ldots, a_n)^t$.

- Parallel reflections in hyperplanes are affine morphisms. Obviously, they are bijective, so

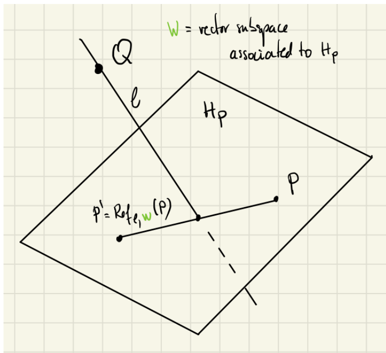
$$\mathrm{Ref}_{H,\mathbf{v}} \in \mathrm{AGL}(\mathbf{A}) \subseteq \mathrm{End}_{\mathrm{aff}}(\mathbf{A}).$$

6.2.6 Parallel reflection in a line

Definition 6.9. Let ℓ be a line and let $\mathbf{W}$ be an $(n-1)$-dimensional vector subspace in $\mathbf{V}$ which is not parallel to ℓ. For any point $P \in \mathbf{A}$ there is a unique point P' such that $\mathrm{Pr}_{\ell,\mathbf{W}}(P)$ is the midpoint of the segment $[PP']$. We denote P' by $\mathrm{Ref}_{\ell,\mathbf{W}}(P)$ and call it the *reflection of the point P in the line ℓ parallel to* $\mathbf{W}$. This gives a map

$$\mathrm{Ref}_{\ell,\mathbf{W}} : \mathbf{A} \to \mathbf{A}$$

called, the *reflection in the line ℓ parallel to* $\mathbf{W}$.



- With respect to the reference frame $\mathcal{K}$, the vector subspace $\mathbf{W}$ is given by a homogeneous equation
$$\overbrace{a_1x_1 + a_2x_2 + \cdots + a_nx_n}^{\psi(x_1,\ldots,x_n)} = 0$$

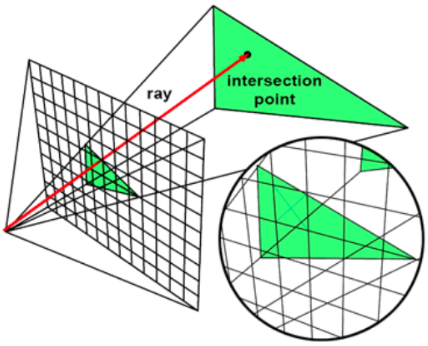
- Since $\mathrm{Pr}_{H,\mathbf{v}}(P)$ is the midpoint of the segment $[PP']$, with respect to $\mathcal{K}$ we have
$$[\mathrm{Ref}_{\ell,\mathbf{W}}(P)]_{\mathcal{K}} = 2[\mathrm{Pr}_{\ell,\mathbf{W}}(P)]_{\mathcal{K}} - [P]_{\mathcal{K}}$$
as in the case of reflections in hyperplanes.

- One deduces the matrix form as in the previous cases
$$\mathrm{Ref}_{\ell,\mathbf{W}}(p) = \Big(\frac{2}{\psi(\mathbf{v})}\mathbf{v} \otimes \mathbf{a} - \mathrm{Id}_n\Big)\begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix} + 2\Big(\begin{bmatrix} q_1 \\ \vdots \\ q_n \end{bmatrix} - \frac{\psi(Q)}{\psi(\mathbf{v})}\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}\Big).$$
where $\mathbf{a} = (a_1, \ldots, a_n)^t$ and where $Q(q_1, \ldots, q_n)$ is a point on ℓ and $\mathbf{v}(v_1, \ldots, v_n)$ is a direction vector for ℓ (as in (6.1)).

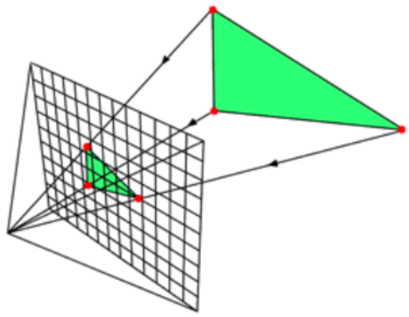
- Parallel reflections in lines are affine morphisms. Obviously, they are bijective, so
$$\mathrm{Ref}_{\ell,\mathbf{W}} \in \mathrm{AGL}(\mathbf{A}) \subseteq \mathrm{End}_{\mathrm{aff}}(\mathbf{A}).$$

6.3 Connections to reality

There are two main types of algorithms which project a 3D scene in computer graphics: ray-tracing algorithms and rasterization algorithms. Ray-tracing algorithms intersect rays with planes determined by the triangles in the scene while rasterization algorithms try to project the triangles on the display screen.



(a) Ray-tracing.



(b) Rasterization.

It is clearly much more efficient to construct a projection map like $\mathrm{Pr}_{H,\ell}$ and project all the triangles, however, this only works for parallel projections. In order to simulate perspective, rasterization algorithms use a projective transformation before using a projection like the one described in the previous paragraphs. Ray-tracing algorithms are conceptually much simpler but they are much more resource intensive.

In cryptography, two classical substitution ciphers are: the Hill cipher and the affine cipher. In both cases you view your code as points in some affine space $\mathbf{X} = \mathbf{A}^n(\mathbb{F}_p)$. In order to encrypt the message you apply an affine transformation $\phi \in \mathrm{AGL}(\mathbf{X})$ and in order to decrypt the message you apply ϕ^{-1}. So, in such a situation, we do arithmetics in some finite field $\mathbf{K} = \mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$ (or, more general, in a finite ring). Letters are mapped to natural numbers

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Number	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

and we calculate modulo p. For example, consider the message

plaintext	A	F	F	I	N	E	C	I	P	H	E	R
x	0	5	5	8	13	4	2	8	15	7	4	17

In order to encrypt x we can use an invertible affine map

$$\varphi : \mathbf{A}^4(\mathbb{F}_{29}) \to \mathbf{A}^4(\mathbb{F}_{29}) \quad \varphi(\begin{bmatrix} x \\ y \\ z \\ v \end{bmatrix}) = A \cdot \begin{bmatrix} x \\ y \\ z \\ v \end{bmatrix} + b$$

three times. Such ciphers are insecure nowadays.