ALGÈBRE GÉOMÉTRIQUE

$\sim$

Université P. & M. Curie

.

Master M1 2010-2011

$\sim$

D. Bertrand

Plan

I .- Géométrie affine

1. Action d'un groupe
2. Espaces affines
3. Le groupe affine

II.- Géométrie projective

1. Espaces projectifs
2. Le groupe projectif

III.- Géométrie quadratique

1. Formes bilinéaires
2. Quadriques
3. Groupes d'isométries

IV.- Géométrie multilinéaire

1. Produits tensoriels
2. Algèbre tensorielle (et groupes algébriques)
3. Applications

Références

M. Audin : *Géométrie*; Edp Sciences, 2005 [trad. anglaise : *Geometry*, Springer, 2002]

M. Berger : *Géométrie*; Nathan, 1990

Y. Ladegaillerie : *Géométrie affine, projective, euclidienne et anallagmatique*; Ellipses, 2003.

J. Nekovar : *Algèbre géométrique*[1] ; Cours Univ. Paris 6, 2008.

---

[1] Du même auteur et dans la même collection, on consultera avec profit le *Manuel d'Anglais mathématique à usage des étudiants du Master*. On signale dans les pages qui suivent l'existence d'autres langues susceptibles d'exprimer les raisonnements mathématiques.

Table des matières

Chapitre I

Géométrie affine

Université de Paris VI
D. Bertrand

Master M1 2010-2011
Algèbre géométrique

CHAPITRE I

GÉOMÉTRIE AFFINE

I.1 Action d'un groupe.

I.1.1 Groupes agissant sur un ensemble

Soit G un groupe, dont on note $(g_1, g_2) \to g_1 g_2$ la loi de composition, e l'élément neutre, et $g \mapsto g^{-1}$ la loi d'inversion. Soit par ailleurs X un ensemble. Une action (à gauche) de G sur X est la donnée d'une application $G \times X \to X : (g, x) \mapsto g.x$ telle que :

$$\forall g_1, g_2 \in G, x \in X, (g_1 g_2).x = g_1.(g_2.x) \text{ et } e.x = x.$$

Pour tout $g \in G$, l'application $\tau(g) : x \mapsto g.x$ est alors une bijection de X sur X (d'inverse $x \mapsto g^{-1}.x$), et $\tau(g_1 g_2) = \tau(g_1) \circ \tau(g_2)$. Si $Perm(X) = \mathfrak{S}_X$ désigne le groupe des bijections de X sur X, il revient ainsi au même de se donner une action de G sur X, ou un morphisme de groupes τ de G dans $Perm(X)$.

Soit $x \in X$. L'orbite de x sous l'action de G est le sous-ensemble $G.x = \{g.x, g \in G\}$ de X. Le stabilisateur de x est le sous-groupe $G_x = \{g \in G, g.x = x\}$ de G ; pour $y = \gamma.x \in G.x$, les stabilisateurs G_x et $G_y = \gamma G_x \gamma^{-1}$ sont conjugués. L'application $g \mapsto g.x$ établit une bijection de l'ensemble G/G_x des classes à gauche modulo G_x vers l'orbite $G.x$ de x. En particulier, si G est un groupe d'ordre $|G|$ fini, on a $card(G.x) = \frac{|G|}{|G_x|}$.

Deux orbites sont distinctes ou confondues : $G.x = G.y \Leftrightarrow y \in G.x$, et on dit alors que x (ou y) est un représentant de l'orbite $G.x$. Ainsi, X est la reunion disjointe de ses différentes orbites. On note $G \backslash X$ l'ensemble des orbites (ou X/G quand il s'agit d'une action à droite). Si X et G sont finis et si $h = card(G \setminus X)$, on en déduit, en désignant par $x_1, ..., x_h$ un système de représentants des orbites, la "formule des classes" :

$$\frac{card(X)}{|G|} = \sum_{i=1}^{h} \frac{1}{|G_{x_i}|}.$$

On dit que G agit transitivement sur X s'il n'y a qu'une orbite ($h = 1$). On dit que G agit librement sur X si $\forall x \in X, G_x = \{e\}$; chaque orbite est alors en bijection avec G. Lorsque G agit librement et transitivement sur X, on dit que X est un espace homogène principal, ou encore, un torseur, sous G. En choisissant un élément x de X, on peut dans ce cas identifier $X = G.x$ et G, mais il n'y a en général aucun choix naturel pour x, et une telle identification doit être maniée avec précaution.

I.1.2 Espaces vectoriels

Dans tout ce cours, on désigne par K un corps, que (sauf mention du contraire) on suppose *commutatif*. On note $K^* = K - \{0\}$ le groupe multiplicatif de K. En général, K sera $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ ou un corps fini, par exemple $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, où p est un nombre premier.

Un espace vectoriel sur K est la donnée d'un groupe commutatif E (loi notée additivement, élément neutre $= \overrightarrow{0}$, éléments de E = vecteurs), muni d'une application de $K \times E$ dans E qui induit une action du groupe multiplicatif K^* :

$$K^* \times E \to E : (\lambda, \overrightarrow{u}) \mapsto \lambda \overrightarrow{u},$$

(donc telle que $\forall \lambda, \mu \in K^*, \overrightarrow{u} \in E : \lambda(\mu \overrightarrow{u}) = (\lambda\mu)\overrightarrow{u}$, et $1\overrightarrow{u} = \overrightarrow{u}$), et qui vérifie de plus, pour tout $\lambda, \mu \in K, \overrightarrow{u}, \overrightarrow{v} \in E$:

$$\lambda(\overrightarrow{u} + \overrightarrow{v}) = \lambda\overrightarrow{u} + \lambda\overrightarrow{v} \text{ et } (\lambda + \mu)\overrightarrow{u} = \lambda\overrightarrow{u} + \mu\overrightarrow{v}.$$

On en déduit (exercice) que $0\overrightarrow{u} = \overrightarrow{0}, (-1)\overrightarrow{u} = -\overrightarrow{u}$.

Si E, E' sont deux espaces vectoriels (sur K), on note $\mathcal{L}(E, E')$ le K-espace vectoriel formé par les applications K-linéaires de E dans E'. Par exemple, le dual $\mathcal{L}(E, K) := E^*$ de E est l'ensemble des formes K-linéaires sur E. Pour $E' = E$, $End(E) := \mathcal{L}(E) := \mathcal{L}(E, E)$ est un anneau. L'ensemble $GL(E) \subset \mathcal{L}(E)$ des automorphismes de E est le groupe des unités de l'anneau $End(E)$. Pour $E = K^n, E' = K^{n'}$, $\mathcal{L}(E, E')$ s'identifie à l'espace $Mat_{n',n}(K)$ des matrices $A = (a_{i,j}; 1 \leq i \leq n', 1 \leq j \leq n)$ à n' lignes et n colonnes, à coefficients dans K. Le groupe

$$GL(K^n) := GL_n(K) = \{g \in Mat_{n,n}(K), \det(g) \neq 0\}$$

s'appelle le groupe linéaire d'ordre n.

Si K' est une extension de K (typiquement, $K = \mathbb{R}$ et $K' = \mathbb{C}$), on désignera par $E_{K'}$ l'espace vectoriel sur K' déduit de E par extension des scalaires de K à K'. Pour une définition intrinsèque, voir le chapitre 4.

I.1.3 Espaces affines

Soient K un corps commutatif, et E un espace vectoriel sur K. On appelle espace affine sous E la donnée d'un ensemble non vide $\mathcal{E}$, muni d'une application $E \times \mathcal{E} \to \mathcal{E} : (\overrightarrow{u}, P) \mapsto\ P + \overrightarrow{u}$, vérifiant les 3 conditions suivantes :

$$\forall \overrightarrow{u}, \overrightarrow{v} \in E, \forall P \in \mathcal{E}, P + (\overrightarrow{u} + \overrightarrow{v}) = (P + \overrightarrow{u}) + \overrightarrow{v};$$

$$\forall P \in \mathcal{E}, P + \overrightarrow{0} = P;$$

$$\forall P, Q \in \mathcal{E}, \exists \text{ un unique } \overrightarrow{u} \in E \text{ tel que } Q = P + \overrightarrow{u}.$$

C'est donc la donnée d'une action sur $\mathcal{E}$ du groupe additif $G = E$ sous-jacent à l'espace vectoriel E, la 3e condition signifiant que cette action est transitive et libre. Comme E est commutatif, il n'y a ici pas besoin de préciser que l'action est à gauche (et cela justifie la notation $P + \overrightarrow{u}$ au lieu de $\overrightarrow{u} + P$.)

Soit $\mathcal{E}$ un espace affine sous l'espace vectoriel E. Pour tout couple ordonné (P, Q) de points de $\mathcal{E}$, on note $\overrightarrow{u} = \overrightarrow{PQ}$ l'unique vecteur de E associé à (P, Q) par la 3e condition : $Q = P + \overrightarrow{PQ}$. Alors (exercice), on a pour tous $P, Q, R, P', Q' \in \mathcal{E}$, les identités suivantes entre vecteurs de E :

$$\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR} \ \text{ (relation de Chasles)}$$

$$\overrightarrow{PP} = \overrightarrow{0},\ \overrightarrow{QP} = -\overrightarrow{PQ},$$

ainsi que la loi du parallèlogramme :

$$\overrightarrow{PQ} = \overrightarrow{P'Q'} \Leftrightarrow \overrightarrow{PP'} = \overrightarrow{QQ'}.$$

I.1.4 Espaces projectifs

Soit E un espace vectoriel sur K. Le groupe multiplicatif K^* agit librement sur l'ensemble $X = E - \{\overrightarrow{0}\}$. On note

$$\mathbb{P}(E) = (E - \{\overrightarrow{0}\})/K^*.$$

l'ensemble des orbites de X sous cette action de K^*, et on l'appelle l'espace projectif associé à E. L'orbite $K^*.\overrightarrow{x}$ d'un point $\overrightarrow{x}$ de X est l'ensemble des vecteurs non nuls de E colinéaires à $\overrightarrow{x}$. Par conséquent, $\mathbb{P}(E)$ s'identifie à l'ensemble des droites de E.

Soit n un entier ≥ 0. Lorsque $E = K^{n+1}$, on pose

$$\mathbb{P}(K^{n+1}) = \mathbb{P}_n(K).$$

Si K est un corps fini, de cardinal q, et si $E = K^{n+1}$, on a $card(X) = q^{n+1} - 1$. Comme chaque orbite a $q - 1$ éléments, on en déduit :

$$card(\mathbb{P}_n(K)) = \frac{q^{n+1} - 1}{q - 1} = q^n + ... + q + 1.$$

Par exemple, $\mathbb{P}_2(\mathbb{F}_2)$ a 7 éléments.

I.2 Espaces affines

I.2.1 Sous-espaces affines

Soient $\mathcal{E}$ un espace affine, d'espace vectoriel E. On dit qu'une partie $\mathcal{F}$ de $\mathcal{E}$ est un *sous-espace affine* de $\mathcal{E}$ s'il existe un point P de $\mathcal{F}$ et un sous-espace vectoriel F de E tel que

$$\mathcal{F} = \{P + \overrightarrow{u}, \overrightarrow{u} \in F\} := P + F.$$

Pour tout $Q \in \mathcal{F}$, on a alors $\mathcal{F} = Q + F$, et F est l'ensemble des vecteurs $\overrightarrow{QR}, Q, R \in \mathcal{F}$ de E. En particulier, F, qu'on appelle la *direction* de $\mathcal{F}$, est entièment déterminé par $\mathcal{F}$. L'application $F \times \mathcal{F} \to \mathcal{F} : (P, \overrightarrow{u}) \mapsto P + \overrightarrow{u}$ fait alors de $\mathcal{F}$ un espace affine sous l'espace vectoriel F. On appelle *(co-) dimension* de $\mathcal{F}$ la (co-) dimension de sa direction F. Un point, droite, plan, resp. hyperplan de $\mathcal{E}$ est un sous-espace affine de dimension $0, 1, 2$, resp. codimension 1.

Exemple : soit E un espace vectoriel. On peut le voir comme un espace affine sur lui-même, en définissant, pour tout élément $\overrightarrow{p}$ de $\mathcal{E} = E$ et tout vecteur $\overrightarrow{u}$ de E, l'élément $\overrightarrow{p} + \overrightarrow{u}$ de $\mathcal{E} = E$ comme cette somme, calculée dans E. Soit alors $f \in \mathcal{L}(E, E')$ une application linéaire de E vers un espace vectoriel E', et soit $F = \{\overrightarrow{u} \in E, f(\overrightarrow{u}) = \overrightarrow{0}\}$ le noyau de f. Pour tout $\overrightarrow{b} \in E'$, l'ensemble $\mathcal{F}_{\overrightarrow{b}}$ des solutions $\overrightarrow{v}$ de l'équation $f(\overrightarrow{v}) = \overrightarrow{b}$ est soit vide, soit un sous-espace affine de $\mathcal{E}$, de direction F.

Si $E = K^n$, tout sous-espace affine $\mathcal{F}$ de K^n, de dimension m et de direction F, est ainsi défini par un *système d'équations affines* de la forme

$A\overrightarrow{x} = A\overrightarrow{p}$, où $A \in Mat_{n-m,n}(K) \simeq \mathcal{L}(K^n, K^{n-m})$ est une matrice de rang $n-m$ à coefficients dans K, de noyau F, et $\overrightarrow{p}$ est un point quelconque de $\mathcal{F}$. En particulier, $\mathcal{F}$ est l'intersection de $n-m$ hyperplans affines de K^n.

Une façon différente de décrire $\mathcal{F}$ consiste à en donner un *paramétrage* : on choisit une base $\{\overrightarrow{a_1}, ..., \overrightarrow{a_m}\}$ de $F \subset K^n$, et un point $\overrightarrow{p}$ de $\mathcal{F} \subset K^n$. Alors, $\mathcal{F} = \{\overrightarrow{p} + t_1\overrightarrow{a_1} + ...t_m\overrightarrow{a_m}, (t_1, ..., t_m) \in K^m\}$.

Cette dernière méthode s'étend en un sens à tout espace affine $\mathcal{E}$. Choisissons-en, de façon non canonique, un point O, qu'on appellera origine de $\mathcal{E}$. Alors, l'application $\theta_O : \mathcal{E} \to E : P \mapsto \overrightarrow{OP}$ est une bijection, d'inverse $\overrightarrow{u} \mapsto O + \overrightarrow{u}$, et permet de munir $\mathcal{E}$ d'une structure non canonique d'espace vectoriel $\mathcal{E}_O$, appelée vectorialisé de $\mathcal{E}$ en O, par les formules : $P+Q = R$, où $\overrightarrow{OR} = \overrightarrow{OP} + \overrightarrow{OQ}, \lambda.P = \lambda\overrightarrow{OP}$. Son élément neutre est $\overrightarrow{0_{\mathcal{E}_O}} = O$. Pour E de dimension n, choisisson une base $\{\overrightarrow{a_1}, ..., \overrightarrow{a_n}\}$ de E. L'application $(t_1, ..., t_n) \in K^n \mapsto P = O + t_1\overrightarrow{a_1} + ...t_n\overrightarrow{a_n} \in \mathcal{E}$ est alors un isomorphisme de K^n sur $\mathcal{E}_O$. On dit que $(t_1, ..., t_n)$ sont les *coordonnées cartésiennes* du point P de $\mathcal{E}$ relativement à l'origine O et à la base $\overrightarrow{a_i}$ choisis.

Remarque (*prolongement vectoriel canonique d'un espace affine*) : soit $\mathcal{E}$ un espace affine, de direction E. On peut attacher, de façon cette fois canonique, à $\mathcal{E}$ un espace vectoriel $\hat{\mathcal{E}}$ contenant E comme hyperplan vectoriel, et $\mathcal{E}$ comme hyperplan affine. Ensemblistement, $\hat{\mathcal{E}}$ est la réunion disjointe de E et de symboles $\lambda x, \lambda \in K^*, x \in \mathcal{E}$, avec $1x = x$. On parvient à munir $\hat{\mathcal{E}}$ d'une structure d'espace vectoriel, telle que l'application $h = \hat{\mathcal{E}} \to K : h(\overrightarrow{v}) = 0$ si $\overrightarrow{v} \in E; h(\overrightarrow{v}) = \lambda$ si $\overrightarrow{v} = \lambda x$, soit une forme linéaire sur $\hat{\mathcal{E}}$. En particulier, $Ker(h) = E$, et $\mathcal{E}$ est l'hyperplan affine $h^{-1}(1)$ de $\hat{\mathcal{E}}$. Géométriquement :

Deux élements x et y de $\mathcal{E} \subset \hat{\mathcal{E}}$ ont pour différence $y - x$ dans l'espace vectoriel $\hat{\mathcal{E}}$ l'unique vecteur $\overrightarrow{u} = \overrightarrow{xy}$ de E tel que $y = x + \overrightarrow{u}$ dans $\mathcal{E}$. Cela justifie la notation $\overrightarrow{PQ} = Q - P$, parfois utilisée en géométrie affine.

I.2.2 Coordonnées barycentriques

Sauf mention du contraire, on suppose désormais E de dimension n finie.

Proposition I.2.1. *Soient $\mathcal{F} = P + F, \mathcal{G} = Q + G$ deux sous-espaces affines de l'espace affine $\mathcal{E}$.*

i) L'intersection $\mathcal{F} \cap \mathcal{G}$ de $\mathcal{F}$ et $\mathcal{G}$ est ou bien vide, ou bien un sous-espace affine, de direction $F \cap G$. Dans ce deuxième cas, $dim(\mathcal{F} \cap \mathcal{G}) = dim\mathcal{F} + dim\mathcal{G} - dim(F + G)$.

ii) Elle est non vide si et seulement si $\overrightarrow{PQ} \in F + G$. En particulier, elle est non vide si $E = F + G$, et réduite à un point si $E = F \oplus G$.

Preuve : i) si R est un point de $\mathcal{F} \cap \mathcal{G}$, alors $\mathcal{F} = R + F, \mathcal{G} = R + G$, donc $\mathcal{F} \cap \mathcal{G} = R + (F \cap G)$. Par ailleurs, pour tous sous-espaces vectoriels F, G de E, on a $dim(F + G) + dim(F \cap G) = dimF + dimG$.

ii) Un tel point R existe $\Leftrightarrow \exists \overrightarrow{u} \in F, \overrightarrow{v} \in G, R = P + \overrightarrow{u} = Q + \overrightarrow{v} \Leftrightarrow \overrightarrow{PQ} = \overrightarrow{u} - \overrightarrow{v} \in F + G$. Si F et G sont en somme directe dans E, leur intersection est de dimension nulle, et $dim(\mathcal{F} \cap \mathcal{G}) = 0$.

On dit que $\mathcal{F}$ et $\mathcal{G}$ sont *parallèles* si leurs directions coïncident : $F = G$. On dit que $\mathcal{G}$ est faiblement parallèle à $\mathcal{F}$ si $G \subset F$. On déduit de la Proposition précédente que :

si $\mathcal{F}$ et $\mathcal{G}$ sont parallèles, alors $\mathcal{F}$ et $\mathcal{G}$ sont soit disjoints, soit confondus ;
si $\mathcal{G}$ est faiblement parallèle à $\mathcal{F}$, alors $\mathcal{G} \cap \mathcal{F} = \emptyset$, ou $\mathcal{G} \subset \mathcal{F}$.

L'intersection $\bigcap_{i \in I} \mathcal{F}_i$ d'un nombre quelconque de sous-espace affines $\mathcal{F}_i$, $i \in I$, de $\mathcal{E}$ est vide ou un sous-espace affine de $\mathcal{E}$, de direction $\bigcap_{i \in I} F_i$. Étant donnée un partie $S = \{P_j, j \in J = [0, .., k]\}$ de $\mathcal{E}$, on peut donc parler du plus petit sous-espace affine $< S >$ de $\mathcal{E}$ contenant S. On appelle $< S >$ *le sous-espace affine de $\mathcal{E}$ engendré par S*. Sa direction est le sous-espace vectoriel de E engendré par les vecteurs $\overrightarrow{P_j P_{j'}}, j, j' \in J$. Les éléments de $< S >$ s'appellent les *combinaisons affines* des points $P_j, j \in J$. On dit que les points de S sont *affinement indépendants* si pour un choix (ou de façon équivalente, pour tout choix) d'indice $j \in J$, les vecteurs $\{\overrightarrow{P_j P_{j'}}, j' \in J, j' \neq j\}$ sont linéairement indépendants. Dans ce cas, le sous-espace affine $\mathcal{F} =< S >$ est de dimension $k = card(J) - 1$, et tout point P de $< S >$ s'écrit de façon unique sous la forme $P = P_0 + \lambda_1 \overrightarrow{P_0 P_1} + ... + \lambda_k \overrightarrow{P_0 P_k}, \lambda_j \in K$, de sorte qu'en posant $\lambda_0 = 1 - (\lambda_1 + ... + \lambda_k)$, on a pour tout point O de $\mathcal{E}$:

$$\overrightarrow{OP} = \sum_{j=0,\dots,k} \lambda_j \overrightarrow{OP_j}, \ (\lambda_0 + \lambda_1 + ... + \lambda_k = 1).$$

Les λ_j sont indépendants du choix de O (pour tout $O' \in \mathcal{E}$, $\overrightarrow{O'P} = \overrightarrow{O'O} + \overrightarrow{OP} = (\Sigma_{j=0,...,k}\lambda_j)\overrightarrow{O'O} + \Sigma_{j=0,...,k}\lambda_j\overrightarrow{OP_j} = \Sigma_{j=0,...,k}\lambda_k\overrightarrow{O'P_j}$). On les appelle *coordonnées barycentriques* du point $P \in < S >$ relativement aux points affinement indépendants $P_0, ..., P_k$, et on s'autorise à écrire :

$$P = \lambda_0 P_0 + \lambda_1 P_1 + ... + \lambda_k P_k, \ (\lambda_0 + \lambda_1 + ... + \lambda_k = 1).$$

Inversement, soient $P_0, ..., P_k$ des points de $\mathcal{E}$, qu'on ne suppose plus nécessairement affinement indépendants, et soient $p_0, ..., p_k$ des éléments de K de somme $p = p_0 + ... + p_k$ *non nulle*. Il existe alors un unique point P de $\mathcal{E}$ tel que $\Sigma_{j=0,...,k}p_j\overrightarrow{PP_j} = \overrightarrow{0}$. On appelle P le *barycentre* des points pondérés $(P_j, p_j), j \in J$ (et on parle d'isobarycentre si tous les poids p_j sont égaux, ce qui impose $card(J)$ premier à la caractéristique de K). Pour tout $O \in \mathcal{E}$, P vérifie $p\overrightarrow{OP} = \Sigma_j p_j \overrightarrow{OP_j}$, de sorte que si $P_0, ..., P_k$ sont affinement indépendants, les coordonnées barycentriques de P relativement aux P_j sont les quotients $\lambda_j = p_j/p$.

Proposition I.2.2. *Soient $(P_j, p_j), j \in J$ un ensemble fini de poids pondérés, et $J' \subset J$ une partie de J telle que $\Sigma_{j \in J'} p_j = p'$ soit non nul. Le barycentre P des $(P_j, p_j), j \in J$, ne change pas quand on remplace le sous-ensemble $(P_j, p_j), j \in J'$, par son barycentre P', affecté du poids p'.*

Preuve : $p'\overrightarrow{PP'} + \Sigma_{j \notin J'} p_j \overrightarrow{PP_j} = \Sigma_{j \in J} p_j \overrightarrow{PP_j} = \overrightarrow{0}$.

L'isobarycentre de deux points P_0, P_1, s'appelle le milieu du segment $[P_0P_1]$. Un triangle de $\mathcal{E}$ est la donnée de 3 points P_0, P_1, P_2 de $\mathcal{E}$ affinement indépendants. Ses médianes sont les droites joignant un sommet P_j au milieu M_j du côté opposé. On déduit de la Proposition précédente que les médianes concourent en un point M, isobarycentre des 3 sommets, et que $\overrightarrow{P_jM} = \frac{2}{3}\overrightarrow{P_jM_j}$ pour tout j.

Une *base, ou repère, affine de* $\mathcal{E}$ est une famille $S = \{P_j, j \in J\}$ de points de $\mathcal{E}$ affinement indépendants, tels que $< S > = \mathcal{E}$. On a alors nécessairement $card(J) = n + 1$, où $n = dimE$. Soient $\{P_0, ..., P_n\}$ un repère affine de $\mathcal{E}$, et P un point de $\mathcal{E}$ de coordonnées barycentriques $(\lambda_0, ..., \lambda_n)$ relativement à ce repère : $P = \Sigma_{j=0,...,n}\lambda_j P_j, \Sigma_{j=0,...,n}\lambda_j = 1$. Pour tout j, on a $\overrightarrow{P_jP} = \Sigma_{j'=0,...,n, j' \neq j}\lambda_{j'}\overrightarrow{P_jP_{j'}}$, de sorte que les coordonnées cartésiennes de P relativement à l'origine P_j de $\mathcal{E}$ et à la base $\{\overrightarrow{P_jP_{j'}}, j' \neq j\}$ de E sont données par les $\lambda_{j'}, j' \neq j$.

Exemple : soient A, B deux points de $\mathcal{E}$ affinement indépendants ($\Leftrightarrow$ distincts), et (AB) la droite affine qu'ils engendrent. Pour tout point C de (AB), il existe un unique scalaire λ tel que $C = \lambda A + (1-\lambda)B$. Si $C \neq A$ (c.-à-d. si $\lambda \neq 1$), $\overrightarrow{AB} = \frac{1}{1-\lambda}\overrightarrow{AC}, \overrightarrow{CB} = \frac{\lambda}{\lambda-1}\overrightarrow{CA}, ...$, et on écrit : $\frac{\overrightarrow{AB}}{\overrightarrow{AC}} = \frac{1}{1-\lambda}$, etc ; cette notation n'a pas de sens si A, B, C ne sont pas alignés. (Plus généralement, $\frac{\overrightarrow{A'B'}}{\overrightarrow{AB}} = \mu$ signifie que $B \neq A$, et que $\overrightarrow{A'B'} = \mu\overrightarrow{AB}$.)

I.2.3 Algèbre et géométrie

Nous allons maintenant exprimer quelques propriétés géométriques simples en termes de coordonnées barycentriques. On fixe un repère affine $\{P_0, ..., P_n\}$ de l'espace affine $\mathcal{E}$, de dimension n, et on note $(\lambda_j(Q), j = 0, ..., n)$, les coordonnées barycentriques d'un point Q dans ce repère, avec $\Sigma_{j=0,...,n}\lambda_j(Q) = 1$.

Proposition I.2.3. *Avec ces notations, $n+1$ points $Q_0, Q_1, ..., Q_n$ de $\mathcal{E}$ sont contenus dans un hyperplan affine de $\mathcal{E}$ si et seulement si la matrice carrée $\big(\lambda_j(Q_i)\big)_{0\le i,j\le n}$ a un déterminant nul.*

Preuve : pour tout $i \neq 0$, les coordonnées de $\overrightarrow{Q_0Q_i} = \overrightarrow{P_0Q_i} - \overrightarrow{P_0Q_0}$ dans la base $\{\overrightarrow{P_0P_j}, j = 1, ..., n\}$ sont données par $\xi_{ij} = \lambda_j(Q_i) - \lambda_j(Q_0), j = 1, ..., n$. Or

$$det \begin{pmatrix} \lambda_0(Q_0) & \lambda_1(Q_0) & ... & \lambda_n(Q_0) \\ \lambda_0(Q_1) & \lambda_1(Q_1) & ... & \lambda_n(Q_1) \\ \vdots & . & . & \vdots \\ \lambda_0(Q_n) & \lambda_1(Q_n) & ... & \lambda_n(Q_n) \end{pmatrix} = det \begin{pmatrix} 1 & \lambda_1(Q_0) & ... & \lambda_n(Q_0) \\ 1 & \lambda_1(Q_1) & ... & \lambda_n(Q_1) \\ \vdots & . & . & \vdots \\ 1 & \lambda_1(Q_n) & ... & \lambda_n(Q_n) \end{pmatrix}$$

$$= det \begin{pmatrix} 1 & \lambda_1(Q_0) & ... & \lambda_n(Q_0) \\ 0 & \xi_{11} & ... & \xi_{1n} \\ \vdots & . & . & \vdots \\ 0 & \xi_{n1} & ... & \xi_{nn} \end{pmatrix} = det \begin{pmatrix} \xi_{11} & ... & \xi_{1n} \\ \vdots & . & \vdots \\ \xi_{n1} & ... & \xi_{nn} \end{pmatrix},$$

qui est donc le déterminant des vecteurs $\{\overrightarrow{Q_0Q_i}, i = 1, ..., n\}$ dans la base $\{\overrightarrow{P_0P_j}, j = 1, ..., n\}$ de E. Il est nul si et seulement si ces vecteurs sont linéairement dépendants, c'est-à-dire ssi les points $Q_0, Q_1, ..., Q_n$ sont affinement dépendants, ou de façon équivalente, ssi l'espace affine qu'ils engendrent est de dimension $< n$.

Supposons maintenant que $Q_1, ..., Q_n$ soient affinement indépendants, et soit $\mathcal{H}$ l'hyperplan affine de $\mathcal{E}$ qu'ils engendrent. Alors, un point Q de coordonnées barycentriques $(\lambda_j, j = 0, ..., n)$ comme supra appartient à $\mathcal{H}$ si et seulement si

$$\Sigma_{j=0,...,n} c_j \lambda_j = 0, \ \Sigma_{j=0,...,n} \lambda_j = 1, \ \exists j, j', c_j \neq c_{j'}, \qquad (*)$$

où les c_j sont les déterminants mineurs d'ordre n extraits de la matrice $\big(\lambda_j(Q_i), i = 1, ..., n; j = 0, ..., n\big) \in Mat_{n,n+1}(K)$; ces c_j ne sont pas tous égaux, sans quoi tous les points de $\mathcal{E}$ appartiendraient à $\mathcal{H}$. Inversement, tout $(n+1)$-uplet $(c_0, ..., c_n)$ d'éléments de K pas tous égaux définit via $(*)$ un hyperplan affine de $\mathcal{E}$, donné dans le vectorialisé $\mathcal{E}_{P_0}$ par l'équation affine $\Sigma_{j=1,...,n}(c_j - c_0)\lambda_j = -c_0$.

Proposition I.2.4. (avec $n = 2$) . *Soient D^0, D^1, D^2 trois droites affines dans le plan affine $\mathcal{E}$, d'équations (D^i) : $c_0^i\lambda_0 + c_1^i\lambda_1 + c_2^i\lambda_2 = 0$, où $\lambda_0 + \lambda_1 + \lambda_2 = 1$. Alors, $det\big(c_j^i\big)_{0\leq i,j\leq 2} = 0$ si et seulement si les trois droites sont concourantes ou parallèles.*

Preuve : si $D^0 \cap D^1 \cap D^2 \neq \emptyset$, la matrice (c_j^i) ne peut être inversible, sinon aucun point $(\lambda_0, \lambda_1, \lambda_2) \neq 0$ ne satisferait les trois équations homogènes ; si les trois droites sont parallèles, leurs équations dans le vectorialisé $\mathcal{E}_{P_0}$ montrent que (c_j^i) est de rang ≤ 2, donc de déterminant nul. Inversement, supposons que le noyau $N \subset K^3$ de la matrice (c_j^i) soit non trivial. Alors, ou bien N n'est pas contenu dans le plan W d'équation $\lambda_0 + \lambda_1 + \lambda_2 = 0$, auquel cas N rencontre le plan affine $\lambda_0 + \lambda_1 + \lambda_2 = 1$, et $D^0 \cap D^1 \cap D^2 \neq \emptyset$; ou bien $N \subset W$, auquel cas N est nécessairement de dimension 1 (sans quoi $c_0^i = c_1^i = c_2^i$, et ce même pour tout i), et alors les D^i admettent N comme direction, donc sont parallèles.

Remarque : fixons un repère affine $P_0, ..., P_n$ de $\mathcal{E}$. L'espace vectoriel $\hat{\mathcal{E}} = K^{n+1}$ formé par les $(n+1)$-uplets $(\lambda_0, ..., \lambda_n), \lambda_i \in K$, fournit alors une représentation concrète du prolongement vectoriel canonique de $\mathcal{E}$: sa structure vectorielle est celle (que l'on n'a pas) décrite au §2.1, Remarque, et la forme linéaire $h(\lambda_0, ..., \lambda_n) = \lambda_0 + ... + \lambda_n$, de noyau $Ker(h) := W \simeq E$, vérifie $h^{-1}(1) = \mathcal{E}$. Vus dans $\hat{\mathcal{E}}$, les $\{P_i, i = 0, .., n\}$ forment la base canonique de K^{n+1}, et les coordonnées barycentriques d'un point P de $\mathcal{E} \subset \hat{\mathcal{E}}$ sont ses coordonnées (au sens usuel) dans K^{n+1}. Un hyperplan $\hat{H}$ de $\hat{\mathcal{E}}$ rencontre $\mathcal{E}$ suivant un hyperplan affine $\mathcal{H}$ de $\mathcal{E}$ si et seulement si $\hat{H} \neq W$ est le noyau d'une forme linéaire $\Sigma_{j=0,\dots,n} c_j \lambda_j$, dont les coefficients c_j ne sont pas tous égaux. Ce point de vue fournit des preuves des énoncés précédents libres de tout choix d'une origine de $\mathcal{E}$.

Exercice : justifier, puis traduire, le texte suivant (cours de M. Candilera, Padoue)

Teorema di Menelao : *Se una retta taglia i tre lati del triangolo ABC nei punti L di BC, M di CA, ed N di AB, e se si scrive*

$$L = xB + x'C,\ M = yC + y'A,\ N = zA + z'B,$$

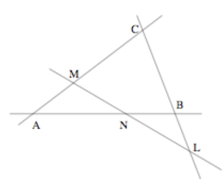
con $x + x' = y + y' = z + z' = 1$, allora $xyz = -x'y'z'$.

I coordinati baricentrici si rivelano uno strumento adattissimo per la dimostrazione del teorema "fratello gemello" del Teorema di Menelao, ossia il :

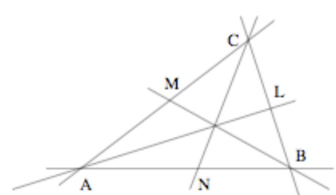
Teorema di Ceva : *Sia P un punto di un piano affine $\mathcal{E}$, e sia ABC un triangolo in $\mathcal{E}$. Siano L, M ed N i punti di intersezioni delle rette AP, BP e CP rispettivamente con i corrispondenti lati opposti del triangolo, di modo che*

$$L = xB + x'C,\ M = yC + y'A,\ N = zA + z'B,$$

con $x + x' = y + y' = z + z' = 1$. Allora $xyz = x'y'z'$.



Menelaüs



Ceva

I.3 Le groupe affine

I.3.1 Produits semi-directs

Soient A et H deux groupes. On dit qu'une action (à gauche) de H sur A est compatible avec la structure de groupe de A si l'homomorphisme correspondant $\tau : H \to Perm(A)$ est à valeurs dans le groupe $Aut(A)$ des automorphismes de groupe de A, autrement dit, en notant $(a,h) \mapsto {}^h a = \tau(h)(a)$ cette action, si l'on a : ${}^h(a_1 a_2) = {}^h a_1 \, {}^h a_2$ pour tous $h, a_1, a_2 \in H \times A \times A$. Dans ces conditions, on appelle *produit semi-direct* (externe) *de A par H relativement à $\tau : H \to Aut(A)$,* et on note $A \rtimes_\tau H$, l'ensemble $A \times H$, muni de la loi de composition interne

$$(a,h),(a',h') \to (a.{}^h a', h.h'),$$

qui en fait un groupe, d'élement neutre (e_A, e_H), et de loi d'inversion $(a,h)^{-1} = ({}^{h^{-1}}a^{-1}, h^{-1})$.

Lorsque l'action de H sur A est triviale (c'est-à-dire lorsque $\tau(h) = id_A$ pour tout $h \in H$), on retrouve le produit direct usuel, qu'on note simplement $A \times H$, de loi de groupe $(a,h).(a',h') = (aa', hh')$.

Revenons au cas général, et posons $G = A \rtimes_\tau H$. Les applications $i : A \to G : i(a) = (a, e_H)$ et $s : H \to G : s(h) = (e_A, h)$ sont des homomorphismes de groupes injectifs , qui permettent d'identifier A et H à des sous-groupes de G. Avec cette identification, A est distingué dans G, puisque pour tous (a,h,a') : $(a,h)a'(a,h)^{-1} = (a,h)(a',e_H)({}^{h^{-1}}a^{-1}, h^{-1}) = (a,h)(a' \, {}^{h^{-1}}a^{-1}, h^{-1}) = (a \, {}^h a' \, {}^h({}^{h^{-1}}a^{-1}), e_H) = (a \, {}^h a' a^{-1}, e_H) \in A$, et l'action de G sur A par conjugaison : $(g,a) \mapsto gag^{-1} := Int(g)(a)$ induit précisément sur $s(H) \subset G$ l'action donnée initialement de H sur A : $\forall (h,a) \in H \times A$, ${}^h a$ est égal à $s(h)as(h)^{-1}$, autrement dit : $\tau(h) = Int(s(h))$. En revanche, $s(H)$ n'est en général pas distingué dans G : selon un calcul similaire, il l'est si et seulement si l'action est triviale, c'est-à-dire si G est le produit direct. Enfin, l'application $\psi : A \times H \to G : (a,h) \mapsto ah$ (c-à-d. $i(a)s(h)$) est toujours une bijection ensembliste, mais n'est un isomorphisme de groupes (pour la structure de produit direct de $A \times H$) que si τ est triviale.

Inversement, partons d'un groupe G, et de deux sous-groupes A, H de G tels que l'application $\phi : A \times H \to G : (a,h) \mapsto ah$ soit une bijection ensembliste. Alors, $A \cap H = \{e_G\}$, et les sous-groupes A et H commutent[1] si et seulement s'ils sont tous deux distingués dans G ; dans ce cas, l'application ϕ est un isomorphisme de groupes (pour la structure de produit direct sur $A \times H$), et on dit que G est le produit direct (interne) de ses sous-groupes distingués A et H.

[1] On dit que deux parties U, V de G commutent si $uv = vu$ pour tout $(u,v) \in U \times V$.

Toujours sous l'hypothèse que ϕ est une bijection ensembliste, supposons seulement que A soit distingué dans G. Alors, H agit sur A par conjugaison : $\tau(h)(a) := hah^{-1}$, et la relation $aha'h' = a\,\tau(h)(a')\,hh'$ montre que ϕ^{-1} établit un isomorphisme de groupes de G vers le produit semi-direct $A \rtimes_\tau H$ de A par H pour cette action τ. On dit que G est le produit semi-direct (interne) de son sous-groupe distingué A par son sous-groupe H, et on note : $G = A \rtimes H$.

Soient enfin G un groupe, et A un sous-groupe distingué de G. L'ensemble G/A des classes à gauche modulo A est alors naturellement muni d'une structure de groupe. En général, il n'existe pas de sous-groupe de G dont les éléments forment un système de représentants de G/A, mais c'est le cas si (et seulement si) il existe un sous-groupe H de G tel que $G = A \rtimes H$.

Exemples : Considérons le sous-groupe $G = B_n(K)$ de $GL_n(K)$, formé par les matrices triangulaires supérieures inversibles d'ordre $n \geq 2$. Il admet comme sous-groupes le groupe abélien $H = D_n(K) \simeq (K^*)^n$ formé par les matrices diagonales inversibles, et le groupe $A = Unip_n(K)$ formé par les matrices triangulaires supérieures unipotentes $\begin{pmatrix} 1 & a_{1i} & a_{1n} \\ 0 & 1 & a_{jn} \\ 0 & 0 & 1 \end{pmatrix}$ (non abélien pour $n \geq 3$). $Unip_n(K)$ est distingué dans $B_n(K)$, et l'on a (exercice) : $B_n(K) = Unip_n(K) \rtimes D_n(K)$.

Pour $n = 2$, considérons le sous-groupe $GA_1(K)$ de $GL_2(K)$ formé des matrices de la forme $[v, u] := \begin{pmatrix} u & v \\ 0 & 1 \end{pmatrix}$, $u \in K^*, v \in K$. Son intersection avec $D_2(K)$ est un sous-groupe $H' = \{[0, u]\}$ isomorphe au groupe multiplicatif $(K^*, \times)$, tandis que son sous-groupe distingué $A' = Unip_2(K) = \{[v, 1]\}$ est isomorphe au groupe additif $(K, +)$. On a $GA_1(K) = A' \rtimes H'$. L'action de H' sur A' par conjugaison est donnée par $[0, u][v, 1][0, u^{-1}] = [uv, 1]$, de sorte qu'on peut aussi écrire : $GA_1(K) = K \rtimes_\tau K^*$, où l'action des éléments u de K^* sur les éléments v de K est donnée par : $\tau(u)(v) = uv$.

I.3.2 Applications affines

Soient $\mathcal{E}, \mathcal{F}$ deux espaces affines, de directions E, F. On dit que $f : \mathcal{E} \to \mathcal{F}$ est une *application affine* s'il existe une application linéaire $\overrightarrow{f} \in \mathcal{L}(E, F)$ et un point O de $\mathcal{E}$, tels que

$$\forall P \in \mathcal{E},\ \overrightarrow{f(O)f(P)} = \overrightarrow{f}(\overrightarrow{OP}),$$

ou de façon équivalente :

$$\forall \overrightarrow{u} \in E : f(O + \overrightarrow{u}) = f(O) + \overrightarrow{f}(\overrightarrow{u}).$$

Cette condition, qui exprime que f est une application linéaire du vectorialisé $\mathcal{E}_O$ vers le vectorialisé $\mathcal{F}_{f(O)}$, est indépendante du choix de l'origine O de $\mathcal{E}$: pour tout $O' \in \mathcal{E}$, elle entraîne en effet : $f(O' + \overrightarrow{u}) = f(O + \overrightarrow{OO'} + \overrightarrow{u}) = f(O) + \overrightarrow{f}(\overrightarrow{OO'} + \overrightarrow{u}) = f(O) + \overrightarrow{f}(\overrightarrow{OO'}) + \overrightarrow{f}(\overrightarrow{u}) = f(O') + \overrightarrow{f}(\overrightarrow{u})$. On voit ainsi de plus que l'application linéaire $\overrightarrow{f}$ attachée à f et à O est en fait indépendante du choix de O : on dit que $\overrightarrow{f}$ est la partie linéaire de f.

Proposition I.3.1. *Soit $f : \mathcal{E} \to \mathcal{F}$ une application affine. Alors :*

i) f préserve l'alignement et le parallèlisme : si A, B, C sont trois points de $\mathcal{E}$ alignés, $f(A), f(B), f(C)$ sont alignés ; si $\mathcal{W}_1$ et $\mathcal{W}_2$ sont deux sous-espaces affines de $\mathcal{E}$ parallèles, $f(\mathcal{W}_1)$ et $f(\mathcal{W}_2)$ sont des sous-espaces affines de $\mathcal{F}$ parallèles.

ii) f préserve les barycentres : si $P \in \mathcal{E}$ est le barycentre d'un système $\{(P_j, p_j)_{j\in J}\}$, le point $f(P) \in \mathcal{F}$ est le barycentre du système $\{(f(P_j), p_j)_{j\in J}\}$. En particulier, si $\mathcal{R} = \{P_j\}$ est un repère affine de $\mathcal{E}$, et si f est bijective, $f(\mathcal{R}) = \{f((P_j)\}$ est un repère affine de $\mathcal{F}$ et les coordonnées barycentriques de $f(P)$ relativement $f(\mathcal{R})$ sont égales à celle de P relativement à $\mathcal{R}$.

iii) f préserve les rapports : si $A, B \neq A, C \neq A \in \mathcal{E}$ sont alignés et si $f(A) \neq f(B)$, alors les quotients $\frac{\overrightarrow{f(A)f(B)}}{\overrightarrow{f(A)f(C)}}$ et $\frac{\overrightarrow{AB}}{\overrightarrow{AC}}$ ont un sens et sont égaux.

Preuve : i) L'image d'un sous-espace affine $\mathcal{W} = P + W$ de $\mathcal{E}$, de direction W, par l'application affine f est le sous-espace affine $f(\mathcal{W}) = f(P) + \overrightarrow{f}(W)$ de $\mathcal{F}$, de direction $\overrightarrow{f}(W)$. [NB : une réciproque, appelée "théorème fondamental de la géométrie affine", affirme que si $K = \mathbb{R}$ et $dim\ \mathcal{E} = dim\ \mathcal{F} \geq 2$, toute bijection de $\mathcal{E}$ sur $\mathcal{F}$ qui préserve l'alignement est une application affine.]

ii) De $\Sigma_{i=0,\ldots,k} p_i \overrightarrow{PP_i} = \overrightarrow{0}$, on tire que $\Sigma_{i=0,\ldots,k} p_i \overrightarrow{f(P)f(P_i)} = \Sigma_i p_i \overrightarrow{f}(\overrightarrow{PP_i}) = \overrightarrow{f}(\Sigma_{i=0,\ldots,k} p_i \overrightarrow{PP_i}) = \overrightarrow{0}$. [Exercice : réciproquement, toute application de $\mathcal{E}$ dans $\mathcal{F}$ qui préserve les barycentres est affine.]

iii) Prendre l'image par $\overrightarrow{f}$ de la relation $\overrightarrow{AB} = \alpha\overrightarrow{AC}, \alpha = \frac{\overline{AB}}{\overline{AC}} \in K^*$. Ou noter que f induit une bijection affine de la droite (AB) sur la droite $(f(A), f(B))$; les coordonnées barycentriques de $f(C)$ dans le repère $\{f(A), f(B)\}$ sont donc égales à celles de C dans le repère $\{A, B\}$. Mais d'après l'exemple concluant le §2.2, les quotients étudiés sont entièrement déterminés par ces coordonnées barycentriques.

Soient $f : \mathcal{E} \to \mathcal{F}$ et $g : \mathcal{F} \to \mathcal{G}$ deux applications affines. Alors, $g \circ f(P) = g(f(O)) + \overrightarrow{g}(\overrightarrow{f}(\overrightarrow{OP}))$ pour tout $P, O \in \mathcal{E}$, donc $g \circ f$ est une application affine de $\mathcal{E}$ dans $\mathcal{G}$ de partie linéaire $\overrightarrow{g \circ f} = \overrightarrow{g} \circ \overrightarrow{f}$. En particulier, *$f$ est bijective si et seulement si $\overrightarrow{f}$ l'est* : si f admet une application réciproque $g := f^{-1}$, celle-ci induit une application linéaire entre les vectorialisés $\mathcal{F}_{f(O)}$ et $\mathcal{E}_O$, donc est affine, et sa partie linéaire vérifie $\overrightarrow{g} \circ \overrightarrow{f} = id_E$, donc $\overrightarrow{f}$ est inversible. Si $\overrightarrow{f}$ admet une application réciproque $\overrightarrow{g} := \overrightarrow{f}^{-1}$, celle-ci est linéaire et l'application affine g, de partie linéaire $\overrightarrow{g}$, qui envoie $O' := f(O)$ sur O est réciproque de f. On dit alors que f est un isomorphisme affine. Quand $\mathcal{E} = \mathcal{F}$, on parle d'automorphisme (ou : tranformation) affine. Ainsi :

Definition I.3.2. *L'ensemble des automorphismes affine de de $\mathcal{E}$ forme un groupe, noté $GA(\mathcal{E})$ et appelé groupe affine de $\mathcal{E}$, et l'application canonique $\pi : f \mapsto \pi(f) := \overrightarrow{f}$ est un morphisme de groupes de $GA(\mathcal{E})$ sur $GL(E)$.*

Choisissons des repères affines $(P_0, .., P_n), (Q_0, ..., Q_m)$ de $\mathcal{E}, \mathcal{F}$, et identifions $\mathcal{E}$ à K^n, $\mathcal{F}$ à K^m par le choix des origines P_0, Q_0. Dans les coordonnées cartésiennes correspondantes, toute application affine f s'écrit sous la forme $\overrightarrow{x} \mapsto U\overrightarrow{x} + \overrightarrow{v}$, où $\overrightarrow{x} \in K^n, U \in Mat_{n,m}(K)$ représente $\overrightarrow{f}$, et $\overrightarrow{v} \in K^m$ représente $\overrightarrow{Q_0 f(P_0)}$. Il est commode d'écrire cette représentation de f sous la forme

$$\begin{pmatrix} \overrightarrow{x} \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} U & \overrightarrow{v} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \overrightarrow{x} \\ 1 \end{pmatrix} = \begin{pmatrix} U\overrightarrow{x} + \overrightarrow{v} \\ 1 \end{pmatrix}, \text{avec} \begin{pmatrix} U & \overrightarrow{v} \\ 0 & 1 \end{pmatrix} := [\overrightarrow{v}, U]$$

$\in Mat_{n+1,m+1}(K)$. Pour toute application affine $g = \mathcal{F} \to \mathcal{G}$, représentée par une matrice $[\overrightarrow{v'}, U'] \in Mat_{m+1,k+1}(K)$, où $dim\mathcal{G} = k$, la composée $g \circ f$ est représentée par la matrice

$$[\overrightarrow{v}', U'][\overrightarrow{v}, U] = [\overrightarrow{v}' + U'\overrightarrow{v}', U'U].$$

Le terme $U'U$ était prévu, et on voit que si $f(\Leftrightarrow U)$ est inversible, f^{-1} est représentée par $[\overrightarrow{v}, U]^{-1} = [-U^{-1}\overrightarrow{v}, U^{-1}]$.

Lorsque $\mathcal{E}$ est identifié à K^n comme supra, on pose $GA(K^n) = GA_n(K)$. Les formules précédentes, avec $m = n$, montrent que

$$GA_n(K) = K^n \rtimes_\tau GL_n(K),$$

pour l'action τ de $GL_n(K)$ sur K^n donnée par $\tau(U)(\overrightarrow{v}) = U\overrightarrow{v}$. Nous allons retrouver cette propriété de façon plus intrinsèque.

Soit $T(\mathcal{E}) = \{f \in GA(\mathcal{E}), \overrightarrow{f} = id_E\}$ le noyau du morphisme $\pi : GA(\mathcal{E}) \to GL(E)$. Un élément f de $T(\mathcal{E})$ vérifie $f(P) = P + \overrightarrow{v}$, où $\overrightarrow{v} = \overrightarrow{Of(O)}$ ne dépend que de f, et on dit que $f := t_{\overrightarrow{v}}$ est la *translation de vecteur* $\overrightarrow{v}$. L'application $\overrightarrow{v} \mapsto t_{\overrightarrow{v}}$ est un isomorphisme du groupe additif sous-jacent à l'espace vectoriel E vers $T(\mathcal{E})$. En tant que noyau, $A = T(\mathcal{E})$ est distingué dans $GA(\mathcal{E})$; plus précisément :

$$\forall \overrightarrow{v} \in E, f \in GA(\mathcal{E}),\ f t_{\overrightarrow{v}} f^{-1} = t_{\overrightarrow{f}(\overrightarrow{v})} \qquad (**)$$

car $f(f^{-1}(P) + \overrightarrow{v}) = f(f^{-1}(P)) + \overrightarrow{f}(\overrightarrow{v})$ pour tout P. Fixons alors un point O de $\mathcal{E}$, et considérons-en le stabilisateur $H := \big(GA(\mathcal{E})\big)_O = \{f \in GA(\mathcal{E}), f(O) = O\}$, qui s'identifie naturellement au groupe linéaire $GL(\mathcal{E}_O) \simeq GL(E)$ du vectorialisé de $\mathcal{E}$ en O, et ne rencontre A qu'en $id_{\mathcal{E}} = t_{\overrightarrow{0}}$. Mieux, la restriction à H du morphisme $\pi : G = GA(\mathcal{E}) \to GL(E) \simeq G/A$ est un isomorphisme de groupes. Par conséquent, tout élément de G s'écrit de façon unique sous la forme $ha', a' \in A, h \in H$, donc aussi sous la forme $ah, a = ha'h^{-1}$. Ainsi,

$$GA(\mathcal{E}) = T(\mathcal{E}) \rtimes \big(GA(\mathcal{E})\big)_O \simeq E \rtimes_\tau GL(E),$$

pour l'action τ de $GL(E)$ sur E donnée par $\tau(\overrightarrow{f})(\overrightarrow{v}) = \overrightarrow{f}(\overrightarrow{v})$.

Exemples

Outre les translations, caractérisées par la condition $\overrightarrow{f} = id_E$, plusieurs applications affines f de $\mathcal{E}$ dans lui-même jouent un rôle particulier. Notons tout d'abord que l'ensemble

$$Fix(f) = \{P \in \mathcal{E}, f(P) = P\}$$

des points fixes d'un tel endomorphisme f est soit vide, soit un sous-espace affine de $\mathcal{E}$ de direction $Ker(\overrightarrow{f} - id_E)$, et que $Fix(f)$ est réduit à un point si et seulement 1 n'est pas valeur propre de $\overrightarrow{f}$ (en revanche, le fait que 1 soit valeur propre de $\overrightarrow{f}$ n'entraîne a priori rien sur $Fix(f)$). Rappelons par ailleurs (cours de Licence) que si $\lambda_1, ..., \lambda_r$ sont des scalaires *distincts* tels que $\Pi_{i=1,..,r}(\overrightarrow{f} - \lambda_i id_E) = 0$, alors $Ker(\overrightarrow{f} - \lambda_i id_E) = Im\big(\Pi_{i' \neq i}(\overrightarrow{f} - \lambda_{i'} id_E)\big)$ pour tout i, et $E = \oplus_{i=1,...,r} Ker(\overrightarrow{f} - \lambda_i id_E)$. Enfin, on dit qu'une application ϕ d'un ensemble X dans lui-même est un involution si $\phi \circ \phi = id_X$; alors, ϕ est bijective, et $\phi^{-1} = \phi$.

(i) *Les homothéties* : ce sont les automorphismes affines f de $\mathcal{E}$ dont la partie linéaire $\overrightarrow{f} = \mu\, id_E$ est une homothétie vectorielle de rapport $\mu \in K^*, \mu \neq 1$. Un homothétie f admet donc un unique point fixe O_f, et s'écrit $f(P) = O_f + \mu\, \overrightarrow{O_f P}$

Soit f une homothétie. Pour tout sous-espace affine $\mathcal{W}$ de $\mathcal{E}$, $f(\mathcal{W})$ est parallèle à $\mathcal{W}$, car $W = \mu W = \overrightarrow{f}(W)$ est stable sous $\overrightarrow{f}$. Inversement, soit f un automorphisme affine tel que $f(\mathcal{H})$ soit parallèle à $\mathcal{H}$ pour tout hyperplan $\mathcal{H}$ de $\mathcal{E}$. Alors, les droites (intersections d'hyperplans) vérifient la même propriété et tous les vecteurs non nuls de E sont propres pour $\overrightarrow{f}$. On en déduit que $\overrightarrow{f}$ est une homothétie vectorielle $\mu\, id_E$, et f est une homothétie si $\mu \neq 1$, une translation sinon.

Soient f_1, f_2 sont deux homothéties, de points fixes O_1, O_2, de rapports μ_1, μ_2. Leur composée $f = f_2 \circ f_1$ est un translation si $\mu_1\mu_2 = 1$, tandis que si $\mu_1\mu_2 \neq 1$, c'est une homothétie de rapport $\mu_1\mu_2$. Dans ce cas, on a $O_f = O_1$ si $O_1 = O_2$ (et alors, f_1 et f_2 commutent) ; sinon, O_f se trouve sur la droite $\mathcal{D} = (O_1O_2)$, car $\mathcal{D}$ est stable sous f_1 et f_2, donc sous f, de sorte que f induit sur $\mathcal{D}$ une homothétie. Exercice : déterminer les coordonées barycentriques de O_f dans le repère affine $\{O_1, O_2\}$ de $\mathcal{D}$.

Une homothétie de rapport $-1 \neq 1$ s'appelle une symétrie centrale, et son point fixe le centre de la symétrie. Le composé de deux symétries centrales $f' \circ f$ est la translation de vecteur $2\overrightarrow{O_f O_{f'}}$.

(ii) *Les projections affines* : Soient $\mathcal{F}$ un sous-espace affine de $\mathcal{E}$, de direction F, et F' un supplémentaire de F dans $E = F \oplus F'$. D'après la Proposition 2.1.ii, pour tout point P de $\mathcal{E}$, les sous-espaces affines $\mathcal{F}$ et $P + F'$ se rencontrent en un unique point P'. L'application $p : P \mapsto P'$ s'appelle la projection sur $\mathcal{F}$ parallèlement à F'. C'est un endomorphisme affine de E, dont la partie linéaire est la projection vectorielle $\overrightarrow{p}$ de E sur F attachée à la décomposition $E = F \oplus F'$. On a $p \circ p = p$, et $Fix(p) = \mathcal{F}$. Inversement, tout endomorphisme affine p vérifiant $p \circ p = p$ est une projection affine, d'image le sous-espace affine $p(\mathcal{E}) := \mathcal{F}$ (de direction $\overrightarrow{p}(E) := F$), parallèlement à $Ker(\overrightarrow{p}) := F'$. En effet, on a $\overrightarrow{p}(\overrightarrow{p} - id_E) = 0$, donc $F = Im(\overrightarrow{p}) = Ker(\overrightarrow{p} - id_E)$ et $E = F \oplus F'$. Pour tout $P \in \mathcal{E}$, le point $P' = p(P)$ appartient à $\mathcal{F}$, et il reste à voir que $\overrightarrow{PP'} \in F'$. Mais $p(P') = p(p(P)) = p(P) = P'$, donc $\overrightarrow{p}(\overrightarrow{PP'}) = \overrightarrow{p(P)p(P')} = \overrightarrow{0}$, et $\overrightarrow{PP'} \in Ker\,\overrightarrow{p} = F'$.

(iii) *Les symétries affines* (on suppose ici que $car(K) \neq 2$) : reprenons les données $\mathcal{F} \subset \mathcal{E}, E = F \oplus F'$ et les notations P, P' de l'exemple (ii). L'application s qui associe à $P \in \mathcal{E}$ le point $s(P) = P + 2\overrightarrow{PP'}$ s'appelle la symétrie affine par rapport à $\mathcal{F}$ parallèlement à F'. C'est une involution de $\mathcal{E}$, qui vérifie $Fix(s) = \mathcal{F}$, et dont la partie linéaire est la symétrie vectorielle $\overrightarrow{s} : \overrightarrow{v} = \overrightarrow{u} + \overrightarrow{u'} \in F \oplus F' \mapsto \overrightarrow{s}(\overrightarrow{v}) = \overrightarrow{u} - \overrightarrow{u'}$. Inversement, toute involution affine s de $\mathcal{E}$ admet au moins un point fixe (considérer le milieu M d'un segment $[P, s(P)]$), et est une symétrie affine par rapport au sous-espace affine $\mathcal{F} := Fix(s)$, de direction $F := Ker(\overrightarrow{s} - id_E)$, parallèlement à $Ker(s + id_E) := F'$. En effet, on a $(\overrightarrow{s} + id_E)(\overrightarrow{s} - id_E) = 0$, donc $E = F \oplus F'$. Pour tout $P \in \mathcal{E}$, on a alors : $(\overrightarrow{s} + id_E)\overrightarrow{PM} = \overrightarrow{s(P)M} + \overrightarrow{PM} = \overrightarrow{0}$, donc $\overrightarrow{PM} \in F'$, M est le projeté $P' = p(P)$ de P, et $\overrightarrow{Ps(P)} = 2\overrightarrow{PP'}$.

De la relation de conjugaison $(**)$, on déduit qu'un automorphisme affine f commute avec une translation $t_{\overrightarrow{v}}$ si et seulement si $\overrightarrow{v} \in Ker(\overrightarrow{f} - id_E)$. En particulier, la symétrie affine s commute avec une translation $t_{\overrightarrow{u}}$ si et seulement si $\overrightarrow{u} \in F$. Si $\overrightarrow{u} \in F$ est non nul, la composée $S = t_{\overrightarrow{u}} \circ s = s \circ t_{\overrightarrow{u}}$ s'appelle une *symétrie glissée*. Bien que sa partie linéaire $\overrightarrow{S} = \overrightarrow{s}$ soit une symétrie vectorielle, S n'a aucun point fixe, et n'est donc pas une symétrie affine. En revanche, si $\overrightarrow{u}' \in F'$, $t_{\overrightarrow{u}'} \circ s$ est une involution donc une symétrie.

(iv) *Les affinités* : toujours avec les données $\mathcal{F}, E = F \oplus F'$ et la notation $P' = p(P)$, on appelle affinité de rapport $\mu \in K^*$, de base $\mathcal{F}$, parallèlement à F', l'automorphisme affine f_μ qui attache à $P \in \mathcal{E}$ le point $f(P) = P' + \mu\overrightarrow{P'P}$. Ainsi, f_{-1} est la symétrie s, et p correspond au cas limite $\mu = 0$. Si $\mu \neq 1$, on a toujours $Fix(f_\mu) = \mathcal{F}, F = Ker(\overrightarrow{f_\mu} - id_E), F' = Ker(\overrightarrow{f_\mu} - \mu\, id_E)$, de sorte que $\overrightarrow{f_\mu}$ est diagonalisable.

Quand $F = H$ est un hyperplan, et que $\mu \neq 1$, l'affinité f_μ s'appelle une *dilatation*. Alors, $Im(\overrightarrow{f_\mu} - id_E)$ est la droite $F' := D \not\subset H$.

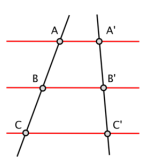
(v) *Les transvections affines* : une transvection vectorielle de E est un élément $\overrightarrow{f}$ de $GL(E)$ dont toutes les valeurs propres sont égales à 1, et tel que $Ker(\overrightarrow{f} - id_E)$ soit un hyperplan H de E. En particulier, $\overrightarrow{f}$ n'est pas diagonalisable, $(\overrightarrow{f} - id_E)^2 = 0$, et $Im(\overrightarrow{f} - id_E)$ est une droite $D \subset H$. Si $\overrightarrow{h} \in H$ désigne un vecteur non nul porté par D, il existe forme linéaire $\ell = \ell_{\overrightarrow{h}}$ sur E, de noyau H, telle que

$$\forall \overrightarrow{v} \in E, \ \overrightarrow{f}(\overrightarrow{v}) = \overrightarrow{v} + \ell(\overrightarrow{v})\overrightarrow{h}.$$

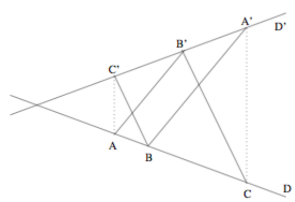
Si $\mathcal{H}$ est un hyperplan de $\mathcal{E}$ de direction H, on appelle transvection affine de base $\mathcal{H}$ tout automorphisme affine f tel que $Fix(f) = \mathcal{H}$ et dont la partie linéaire $\overrightarrow{f}$ soit une transvection vectorielle (pour laquelle on aura nécessairement $Ker(\overrightarrow{f} - id_E) = H$).

Tout automorphisme affine f de $\mathcal{E}$ tel que $Fix(f)$ soit un hyperplan est une dilatation ou une transvection affine.

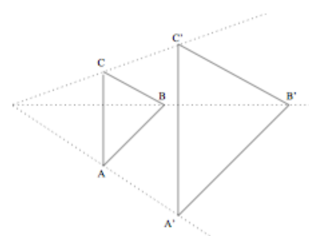
I.3.3 Algèbre et géométrie (suite)



Thalès



Pappus



Desargues

Théoreme I.3.3. (Thalès) *: soient $\mathcal{D}, \mathcal{D}'$ deux droites distinctes du plan affine $\mathcal{E}$, et A, B, C (resp. A', B', C') trois points distincts de $\mathcal{D}$ (resp. $\mathcal{D}'$), tels que $B \neq B', C \neq C'$.*

i) On suppose que les droites (BB'), (CC'), ainsi, si $A \neq A'$, que (AA'), sont parallèles. Alors, $\frac{\overrightarrow{AB}}{\overrightarrow{AC}} = \frac{\overrightarrow{A'B'}}{\overrightarrow{A'C'}}$, et ce rapport est égal à $\frac{\overrightarrow{BB'}}{\overrightarrow{CC'}}$ si $A = A'$.

ii) On suppose que $A = A'$, ou, si $A \neq A'$, que les droites (AA') et (BB') sont parallèles, et que $\frac{\overrightarrow{AB}}{\overrightarrow{AC}} = \frac{\overrightarrow{A'B'}}{\overrightarrow{A'C'}}$. Alors, (CC') est parallèle à (BB') .

Preuve i) Soient D' et Δ les droites vectorielles directrices, dans le plan vectoriel E, des droites affines $\mathcal{D}'$ et (BB'). Montrons que $D' \neq \Delta$: sinon, le point $B = B' + \overrightarrow{B'B}$ appartiendrait à $\mathcal{D}'$, le point $C = C' + \overrightarrow{C'C}$ également puisque $(CC')//(BB')$, et les droites $\mathcal{D}$ et $\mathcal{D}'$, admettant deux points B et C en commun, seraient confondues, ce qui contredit l'hypothèse. Le plan vectoriel E est donc la somme directe $D \oplus \Delta$, et on peut considérer la projection affine p de $\mathcal{E}$ sur $\mathcal{D}'$ parallèlement à Δ. L'hypothèse de (i) signifie que B', C', A' sont les images de B, C, A par p, et la Proposition 3.1.iii donne la 1ère égalité de la conclusion.

Dans le cas où $A = A'$, la 2ème égalité résulte alors de la relation de Chasles $\overrightarrow{B'C'} = \overrightarrow{B'A} + \overrightarrow{AC'}$, ou encore, de la considération de l'homothétie f_μ de centre $A = A'$, de rapport $\mu = \frac{\overrightarrow{AB}}{\overrightarrow{AC}} = \frac{\overrightarrow{A'B'}}{\overrightarrow{A'C'}}$: elle envoie C sur B, C' sur B', donc $\overrightarrow{BB'} = \overrightarrow{f_\mu}(\overrightarrow{CC'}) = \mu\overrightarrow{CC'}$.

ii) Reprenons les notations $\mathcal{D}', \Delta$. Le même raisonnement que supra, appliqué aux points A et B, montrer que $D' \neq \Delta$, et on peut de nouveau considérer la projection p, qui vérifie $p(A) = A', p(B) = B'$. En appliquant (i) à l'image $C'' = p(C)$ de C, on déduit de l'hypothèse de (ii) que $C' = C''$. Donc $\overrightarrow{CC'} = \overrightarrow{Cp(C)}$ appartient à Δ, et $(CC')//(BB')$.

Théoreme I.3.4. (Pappus) *: soient $\mathcal{D}, \mathcal{D}'$ deux droites distinctes du plan affine $\mathcal{E}$, et A, B, C (resp. A', B', C') trois points distincts de $\mathcal{D}$ (resp. $\mathcal{D}'$) - et tous distincts de $\mathcal{D} \cap \mathcal{D}'$ si ces droites sont concourantes -. On suppose que $(AB')//(BA')$ et $(BC')//(CB')$. Alors, $(CA')//(AC')$.*

Preuve : supposons d'abord que $\mathcal{D}$ et $\mathcal{D}'$ soient concourantes, et soit ϕ l'homothétie de centre $O = \mathcal{D} \cap \mathcal{D}'$ qui envoie A sur B. D'après Thalès (i), appliqué à $(O, A, B), (O, B', A')$, ϕ envoie B' sur A'. De même, l'homothétie ψ de centre O qui envoie B sur C envoie C' sur B'. Comme les homothéties ϕ et ψ ont même centre, elles commutent. Donc $\psi \circ \phi = \phi \circ \psi$ envoie A sur C, et C' sur A'. Comme $\phi \circ \psi$ est une homothétie, $\frac{\overrightarrow{OC'}}{\overrightarrow{OA'}} = \frac{\overrightarrow{OA}}{\overrightarrow{OC}}$, et par Thalès (ii), $(AC')//(CA')$.

Supposons maintenant que $\mathcal{D}$ et $\mathcal{D}'$ soient parallèles. En remplaçant ϕ (resp. ψ) par la translation de vecteur $\overrightarrow{AB}$ (resp. $\overrightarrow{BC}$), on trouve que les translations $t_{\overrightarrow{AB}}$ et $t_{\overrightarrow{C'A'}}$ coïncident. On conclut par la règle du parallèlogramme.

Théoreme I.3.5. (Desargues) *: soient (ABC) et $(A'B'C')$ deux triangles du plan affine $\mathcal{E}$. On suppose que $A \neq A', B \neq B', C \neq C'$, et que $(AB)//(A'B')$,*

$(BC)//(B'C')$ *et* $(CA)//(C'A')$*. Alors, les droites* $(AA'), (BB')$ *et* (CC') *sont parallèles ou concourantes.*

Preuve : Si les trois droites ne sont pas paralèles, deux au moins d'entre elles, soit (AA') et (BB') se coupent en un point O. D'après Thalès, l'homothétie ϕ de centre O qui envoie A sur A' envoie B sur B' ; la droite $\phi((AC))$ est la parallèle à (AC) passant par A', c-à- d. $(A'C')$. De même $\phi((BC)) = (B'C')$, et ϕ envoie le point $C = (AC) \cap (BC)$ sur $(A'C') \cap (B'C') = C'$. Par conséquent, les points O, C, C' sont alignés. et les droites $(AA'), (BB'), (CC')$ concourent en O.

Chapitre II

Géométrie projective

Université de Paris VI
D. Bertrand

Master M1 2010-2011
Algèbre géométrique

CHAPITRE II

GÉOMÉTRIE PROJECTIVE

II.1 Espaces projectifs.

II.1.1 Sous-espaces projectifs

Soient K un corps commutatif, et E un espace vectoriel de dimension finie sur K. Rappelons (chap. 1, §1.4) que l'espace projectif déduit de E est l'espace $\mathbb{P}(E) = (E - \{0\})/K^*$. Il est vide si $dim E = 0$, c-à-d. si $E = \{\overrightarrow{0}\}$. Sinon, ses éléments s'identifient aux droites (vectorielles) de E. Par définition, la *dimension de* $\mathbb{P}(E)$ est le nombre

$$dim\, \mathbb{P}(E) := dim E - 1.$$

Ainsi, pour $E = K^{n+1}$, $\mathbb{P}_n(K) := \mathbb{P}(K^{n+1})$ est de dimension n. On parle de point, droite, plan projectif si $dim E - 1 = 0, 1, 2$.

Un *sous-espace projectif de* $\mathbb{P}(E)$ est une partie de $\mathbb{P}(E)$ de la forme $\mathbb{P}(F)$, où F est un sous-espace vectoriel de E : c'est l'ensemble des droites vectorielles de E contenues dans F. Par définition, la codimension de $\mathbb{P}(F)$ dans $\mathbb{P}(E)$ est la codimension de F dans E, de sorte que

$$dim\, \mathbb{P}(F) + codim\, \mathbb{P}(F) = dim F - 1 + codim F = dim E - 1 = dim\, \mathbb{P}(E),$$

comme il se doit. Un hyperplan de $\mathbb{P}(E)$ est un sous-espace projectif de $\mathbb{P}(E)$ de codimension 1.

Proposition II.1.1. *Soient F, G deux sous-espaces vectoriels de E. Alors, $\mathbb{P}(F) \cap \mathbb{P}(G)$ est le sous-espace projectif $\mathbb{P}(F \cap G)$ de $\mathbb{P}(E)$, et l'on a : $dim\, (\mathbb{P}(F) \cap \mathbb{P}(G)) = dim\, \mathbb{P}(F) + dim\, \mathbb{P}(G) - dim\, \mathbb{P}(F + G)$. Ainsi, l'intersection de deux droites distinctes du plan projectif est toujours un point.*

Preuve : $\mathbb{P}(F) \cap \mathbb{P}(G)$ est l'ensemble des droites vectorielles de E contenues dans $F \cap G$, c-à-d. $\mathbb{P}(F \cap G)$. Sa dimension $dim(F \cap G) - 1$ vaut bien $dim F -$

$1 + dimG - 1 - (dim(F+G) - 1)$. Si $\Delta_1 = \mathbb{P}(\Pi_1)$ et $\Delta_2 = \mathbb{P}(\Pi_2)$ sont deux droites distinctes du plan projectif $\mathbb{P}(E)$, les plans vectoriels Π_1 et Π_2 de $E \simeq K^3$ sont distincts, donc $\Pi_1 + \Pi_2 = E$, et $\Delta_1 \cap \Delta_2$, de dimension 0, est un point. Plus généralement, dès que $\mathbb{P}(F)$ a pour dimension la codimension de $\mathbb{P}(G)$, leur intersection, de dimension ≥ 0, est non vide.

Outre cette propriété, qui montre que la notion de parallèlisme n'a pas de sens en géométrie projective, voici quelques applications de cette proposition.

• *Indépendance projective, repères* : l'intersection d'un nombre quelconque de sous-espaces projectifs de $\mathbb{P}(E)$ est encore un sous-espace projectif. Étant donné un ensemble $S = \{\mathbf{P}_j, j = 0,..,k\}$ de points $\mathbb{P}(E)$, on peut donc parler du plus petit sous-espace projectif $< S >$ de $\mathbb{P}(E)$ contenant S. On appelle $< S >$ *le sous-espace projectif de* $\mathbb{P}(E)$ *engendré par* S. Si $\mathbf{P}_j = \mathbb{P}(D_j)$, où les $D_j = K\overrightarrow{v_j}$ sont des droites vectorielles de E, on a $< S > = \mathbb{P}(F)$, où F désigne le sous-espace vectoriel de E engendré par les $\overrightarrow{v_j}$. On dit que les points $\mathbf{P}_j$ sont *projectivement indépendants* si les vecteurs $\overrightarrow{v_j}$ sont linéairement indépendants, c-à-d. si $dimF = k+1$. Par exemple, deux points distincts $\mathbf{P}_1 \neq \mathbf{P}_2$ sont projectivement indépendants : $K\overrightarrow{v_1} + K\overrightarrow{v_2}$ est alors un plan, et $(\mathbf{P}_1\mathbf{P}_2) := < \mathbf{P}_1, \mathbf{P}_2 >$ est une droite projective. Si $\mathbf{P}_3 \notin (\mathbf{P}_1\mathbf{P}_2)$, les points $\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3$ engendrent un plan, etc.

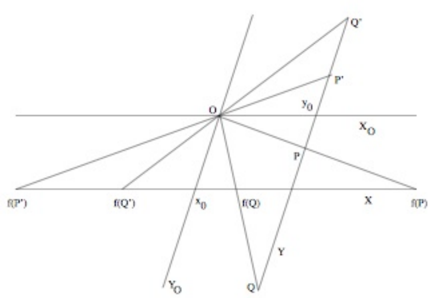
Si $dim\, \mathbb{P}(E) = n$, on dit qu'un $(n+2)$-uplet $\{\mathbf{P}_j, j = 0,..,n+1\}$ est un *repère projectif de* $\mathbb{P}(E)$ si $n+1$ quelconques de ces points sont projectivement indépendants, autrement dit si, pour tout j_0, l'ensemble $\{\mathbf{P}_j, j = 0,...,n+1, j \neq j_0\}$ engendre $\mathbb{P}(E)$. Comme on le montrera au §2.2, Prop. 2.2, le choix d'un repère projectif permet d'identifier $\mathbb{P}(E)$ et $\mathbb{P}_n(K)$, et d'utiliser des *coordonnées homogènes* pour repérer les points de $\mathbb{P}(E)$. Nous repoussons au §2.2 la traduction algébrique que ces coordonnées fournissent aux notions géométriques étudiées dans ce §1.

• *Projections centrales* : soient $\mathbf{O}$ un point de $\mathbb{P}(E)$, et $\mathbf{H}$ un hyperplan de $\mathbb{P}(E)$ ne passant pas par $\mathbf{O}$. Toute droite projective Δ de $\mathbb{P}(E)$ passant par $\mathbf{O}$ rencontre $\mathbf{H}$ en un unique point $\mathbf{p}_\Delta$. En effet, $\mathbf{O} = \mathbb{P}(D), \mathbf{H} = \mathbb{P}(H), \Delta = \mathbb{P}(\Pi)$, où D, H, Π, sont une droite, un hyperplan, un plan de l'espace vectoriel E. Les conditions $\mathbf{O} \notin \mathbf{H}, \mathbf{O} \in \Delta$ se transcrivent par $D \not\subset H, D \subset \Pi$, de sorte que $\Pi \cap H$ est une droite vectorielle de E, et $\mathbf{p}_\Delta := \Delta \cap \mathbf{H} = \mathbb{P}(\Pi \cap H)$ est un point projectif.

Pour tout point $\mathbf{P}$ de $\mathbb{P}(E)$ distinct de $\mathbf{O}$, la droite projective $\Delta_\mathbf{P} = (\mathbf{OP})$ est bien définie, et on peut considérer son intersection $\mathbf{p}_{\Delta_\mathbf{P}}$ avec $\mathbf{H}$.

L'application $\pi : \mathbb{P}(E) - \{\mathbf{O}\} \to \mathbf{H} : \mathbf{P} \mapsto \mathbf{p}_{\mathbf{\Delta_P}}$ s'appelle la *projection de centre* $\mathbf{O}$ *sur* $\mathbf{H}$. Elle est surjective. Si $\mathbf{H'}$ est un second hyperplan de de $\mathbb{P}(E)$ ne passant pas par $\mathbf{O}$, la restriction $\pi_{|\mathbf{H'}} : \mathbf{H'} \to \mathbf{H}$ de π à $\mathbf{H'}$ établit une bijection de $\mathbf{H'}$ sur $\mathbf{H}$, qu'on appelle *perspective* (projective) de centre $\mathbf{O}$ de $\mathbf{H'}$ sur $\mathbf{H}$.

Remarque : en géométrie affine, étant donnés un point O d'un espace affine $\mathcal{E}$, et un hyperplan $\mathcal{H}$ ne passant pas par O, soit $\mathcal{H}_O$ l'hyperplan parallèle à $\mathcal{H}$ passant par O. On appelle projection centrale de centre O sur $\mathcal{H}$ l'application $f : \mathcal{E} - \mathcal{H}_O \to \mathcal{H}$ qui, à un point P de $\mathcal{E}$ n'appartenant pas $\mathcal{H}_O$ attache le point d'intersection $p_{(OP)}$ de (OP) avec $\mathcal{H}$. La restriction $f_{|\mathcal{H}'}$ de f à un hyperplan $\mathcal{H}'$ ne passant pas par O s'appelle une perspective affine. Elle n'est pas définie aux points d'intersection de $\mathcal{H}'$ avec $\mathcal{H}_O$, et elle ne respecte pas le parallèlisme. Après avoir étudié la section suivante (mais sans en garder les notations), le lecteur pourra interpréter les hyperplans $\mathcal{H}, \mathcal{H}'$ comme des ouverts affines d'hyperplans projectifs $\mathbb{P}(H), \mathbb{P}(H')$ d'un espace projectif $\mathbb{P}(E)$, et la perspective projective $\pi_{|\mathbf{H'}}$ comme un prolongement "continu"[1], et partout défini, de la perspective affine $f_{|\mathcal{H}'}$.



Ici, $\mathcal{E}$ est un plan.
$X = \mathcal{H}$ // $X_O = \mathcal{H}_O$,
et $Y = \mathcal{H}'$ sont des droites.
$\mathcal{H}' \cap \mathcal{H}_O$ est un point y_0.
Où va $f(P)$ quand $P \to y_0$?

[1] Bien entendu, cette notion nécessite de munir $\mathbb{P}(E)$ d'une topologie. Pour E de dimension finie et $K = \mathbb{R}$ ou $\mathbb{C}$, un partie de $\mathbb{P}(E)$ est ouverte si la réunion des droites de E, privées de $\overrightarrow{0}$, qui la constituent est un ouvert de $E - \{\overrightarrow{0}\}$. On vérifie alors que $\mathbb{P}(E)$ est compact, que les ouverts affines sont effectivement des ouverts, et que les perspectives projectives sont des homéomorphismes. Topologiquement, $\mathbb{P}_1(\mathbb{R})$ est un cercle, $\mathbb{P}_1(\mathbb{C})$ est une sphère (bord d'une boule de $\mathbb{R}^3$), tandis que $\mathbb{P}_2(\mathbb{R})$ n'est pas simplement connexe. Pour visualiser ce dernier, voir http ://images.math.cnrs.fr/L-infini-est-une-droite-comme-les.html , par E. Brugallé et J. Marché.

II.1.2 Du projectif à l'affine, et inversement

Les éléments de la droite projective $\mathbb{P}_1(K)$ sont :

- les droites D de $E = K^2$ d'équation $y = \alpha x$, qu'on peut repérer par leur pente $\alpha \in K$, c-à-d. par l'ordonnée de leur point d'intersection P_D avec la droite affine $\mathcal{H}$ de K^2 d'équation $x = 1$.

- la droite H de K^2 d'équation $x = 0$ (l'axe des ordonnées, de pente infinie), qu'on peut repérer par le symbole ∞'.

Ainsi, $\mathbb{P}_1(K) \simeq K \cup \infty'$. Mais cette décomposition de la droite projective en réunion d'une droite affine et d'un point n'a rien de canonique : $\mathbb{P}_1(K)$ est aussi formé des droites de K^2 d'équation $x = \beta y$, qu'on peut repérer par le scalaire $\beta \in K$, et de la droite de K^2 d'équation $y = 0$, qu'on peut repérer par le symbole ∞, d'où $\mathbb{P}_1(K) = K \cup \infty$. En dehors des points (distincts) ∞ et ∞', le lien entre les droites affines K fournies par ces deux décompositions est donné par la relation $\beta = 1/\alpha$, qui n'est pas une transformation affine.

Plus généralement, soient E un espace vectoriel, H un hyperplan vectoriel de E, et $\mathcal{H}$ un hyperplan affine de l'espace affine attaché à E (au sens du I.2.1, exemple), de direction H et distinct de H. D'après I, Prop. 2.1.ii, une droite vectorielle D de E non contenue dans H rencontre $\mathcal{H}$ en un unique point P_D ; inversément, par tout point P de $\mathcal{H}$ passe une unique droite vectorielle D_P de E, et cette droite n'est pas contenue dans H. Par conséquent, l'application

$$P \mapsto D_P : \mathcal{H} \simeq \mathbb{P}(E) - \mathbb{P}(H)$$

établit une bijection de l'espace affine $\mathcal{H}$ avec le complémentaire dans $\mathbb{P}(E)$ du sous-espace projectif $\mathbb{P}(H)$ formé par les droites vectorielles de E contenues dans H, et permet de munir ce complémentaire d'une structure d'espace affine sous H, qu'on identifie à $\mathcal{H}$. On a donc une décomposition

$$\mathbb{P}(E) = \mathcal{H} \cup \mathbb{P}(H),$$

dans laquelle l'espace affine $\mathcal{H}$ est de même dimension $dim\mathcal{H} = dimH = dimE - 1 = dim\,\mathbb{P}(E)$ que $\mathbb{P}(E)$. On dit que $\mathbb{P}(E) - \mathbb{P}(H)$ est *l'ouvert affine* complémentaire de $\mathbb{P}(H)$ dans $\mathbb{P}(E)$.

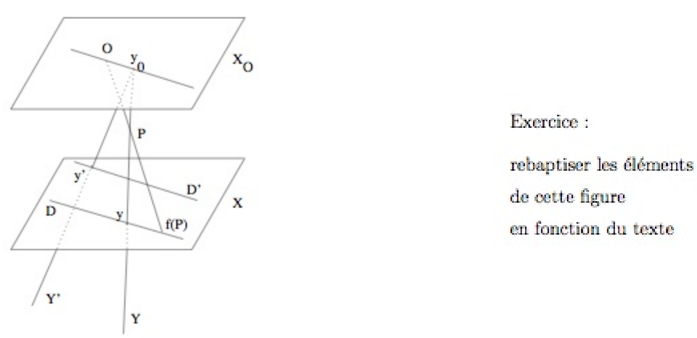
Pour $\mathbb{P}(H)$ fixé, on peut remplacer $\mathcal{H}$ par un autre hyperplan affine $\tilde{\mathcal{H}} \neq H$ de E de direction H. Cela modifie la structure d'espace affine sous H de l'ouvert affine $\mathbb{P}(E) - \mathbb{P}(H) \simeq \tilde{\mathcal{H}}$, mais de façon négligeable. En effet, l'application $f : P_D \rightsquigarrow \tilde{P}_D$ de $\mathcal{H}$ dans $\tilde{\mathcal{H}}$ à laquelle conduit la construction précédente est la restriction à $\mathcal{H}$ d'une homothétie vectorielle f_μ de E, de sorte que pour tout $\overrightarrow{v} \in H$, $f(P_D +_{\mathcal{H}} \overrightarrow{v}) = \tilde{P}_D +_{\mathcal{H}'} \mu\overrightarrow{v}$. Les notions de sous-espaces affines, parallèlisme, etc, dont $\mathcal{H}$ et $\tilde{\mathcal{H}}$ munissent l'ouvert affine $\mathbb{P}(E) - \mathbb{P}(H)$ sont donc les mêmes, et nous ferons l'abus d'omettre le choix de $\mathcal{H}$ dans la définition de sa structure d'espace affine. En fait, $\mathbb{P}(E) - \mathbb{P}(H)$ est de façon canonique un espace affine sous l'espace vectoriel $Hom(E/H, H)$, et notre abus consiste simplement à identifier la droite vectorielle E/H à K.

Remarque : en démarrant avec l'espace vectoriel $E = K^{n+1}$, et en itérant le procédé, on obtient des décompositions de $\mathbb{P}(E)$ en réunions disjointes

$$\mathbb{P}_n(K) = K^n \cup K^{n-1} \cup ... \cup K \cup \{\text{un point}\},$$

d'espaces affines de dimensions $n = dim\, \mathbb{P}_n(K), n-1, ..., 0$. Si K est un corps de cardinal q, on retrouve ainsi la formule (dite d'intégration motivique) $card\ (\mathbb{P}_n(K)) = q^n + q^{n-1} + ... + q + 1$ du chapitre I, §1.4.

Revenons à la décomposition $\mathbb{P}(E) = \mathcal{H} \cup \mathbb{P}(H)$ attachée à un hyperplan $\mathbb{P}(H)$ (et à un choix de $\mathcal{H}$), et considérons un sous-espace projectif $\mathbf{W} = \mathbb{P}(W)$, de dimension w, de $\mathbb{P}(E)$ non contenu dans $\mathbb{P}(H)$, c-à-d. tel que le sous-espace vectoriel W, de dimension $w+1$, de E ne soit pas contenu dans H. Alors $\mathcal{W} = W \cap \mathcal{H}$ est un sous-espace affine de $\mathcal{H}$ de même dimension w que $\mathbf{W}$. Inversement, pour tout sous-espace affine $\mathcal{W}$, de dimension w, de $\mathcal{H}$, l'espace vectoriel W engendré par $\mathcal{W}$ dans l'espace vectoriel E est de dimension $w+1$, et définit un sous-espace projectif $\mathbf{W} := \mathbb{P}(W)$ de dimension w de $\mathbb{P}(E)$ non contenu dans $\mathbb{P}(H)$. Ainsi, l'application $\mathbf{W} \rightsquigarrow \mathcal{W}$ établit une bijection entre l'ensemble des sous-espaces projectifs de $\mathbb{P}(E)$ non contenus dans $\mathbb{P}(H)$ et l'ensemble des sous-espaces affines de $\mathcal{H}$.



Cette bijection respecte les dimensions et l'alignement des points, mais pas les intersections : deux droites $\Delta_1 = \mathbb{P}(\Pi_1)$ et $\Delta_2 = \mathbb{P}(\Pi_2)$ de $\mathbb{P}(E)$ non contenues dans $\mathbb{P}(H)$ peuvent se rencontrer en un point ϖ de $\mathbb{P}(H)$; dans ce cas, $\Pi_1 \cap \Pi_2$ est une droite vectorielle D contenue dans H, avec $\varpi = \mathbb{P}(D)$, et les droites affines $\mathcal{D}_i = \Pi_i \cap \mathcal{H}, i = 1, 2$, de $\mathcal{H}$, toutes deux de direction D, sont parallèles.

Chacun des points $\varpi = \mathbb{P}(D)$ de $\mathbb{P}(H)$ peut ainsi être vu comme le point de rencontre "à l'infini" de toutes les droites (parallèles) de $\mathcal{H}$ de direction D. De ce fait, on dit que la décompositions $\mathbb{P}(E) = \mathcal{H} \cup \mathbb{P}(H)$ correspond au *choix de* $\mathbb{P}(H)$ *comme hyperplan à l'infini de* $\mathbb{P}(E)$. *Tout hyperplan* de $\mathbb{P}(E)$ peut servir d'hyperplan à l'infini, il n'y a pas de choix canonique. Et pour des choix d'hyperplans à l'infini $\mathbb{P}(H) \neq \mathbb{P}(H')$ distincts, les ouverts affines $\mathcal{H}, \mathcal{H}'$ qu'ils définissent ont des structures affines totalement différentes. Par exemple, les droites $\mathcal{D}_1$ et $\mathcal{D}_2$ de $\mathcal{H}$ étudiées ci-dessus sont parallèles (au sens de la structure affine de $\mathcal{H}$), mais si le point ϖ de $\mathbb{P}(H)$ n'appartient pas à $\mathbb{P}(H')$ (c-à-d. si $\varpi \in \mathcal{H}'$), elles définissent dans l'espace affine $\mathcal{H}'$ des droites $\mathcal{D}'_1, \mathcal{D}'_2$ qui se rencontrent au point ϖ de $\mathcal{H}'$. Un énoncé sur les droites parallèles se transcrit alors sans plus de justification en un nouvel énoncé, concernant des droites concourantes. Voir le §1.3 pour quelques applications de ce principe. Auparavant, nous allons, en sens inverse, associer à tout espace affine $\mathcal{E}$ un espace projectif $\hat{\mathbb{P}}(\mathcal{E})$ dont $\mathcal{E}$ est un ouvert affine ; la construction sera cette fois canonique.

Projectivisé d'un espace affine

Soit E un espace vectoriel (qui va jouer le rôle précédemment dévolu à H), de dimension n, et soit $\mathcal{E}$ un espace affine sous l'espace vectoriel E. Son prolongement vectoriel canonique $\hat{\mathcal{E}}$ (voir I, §2.1, Remarque) est un espace vectoriel de dimension $n + 1$, qui contient E comme sous-espace vectoriel, et $\mathcal{E} \neq E$ comme sous-espace affine de direction E. On dispose donc de la décomposition

$$\mathbb{P}(\hat{\mathcal{E}}) = \mathcal{E} \cup \mathbb{P}(E),$$

qui permet de voir naturellement $\mathcal{E}$ comme l'ouvert affine complémentaire de l'hyperplan $\mathbb{P}(E)$ de l'espace projectif $\mathbb{P}(\hat{\mathcal{E}}) := \hat{\mathbb{P}}(\mathcal{E})$. On dit que $\hat{\mathbb{P}}(\mathcal{E})$ est le projectivisé de l'espace affine $\mathcal{E}$, et que $\mathbb{P}(E)$, qu'on note $\mathcal{E}_\infty$, est *l'hyperplan*[2]

[2] Noter l'article défini. L'absence de cet article dans certaines langues les rend d'un usage délicat en mathématiques. Une comparaison des productions mathématiques grecques et romaines suffira pour s'en convaincre.

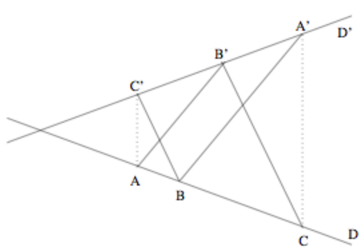
à l'infini de $\mathcal{E}$. Les dimensions $dim\ \mathcal{E} = dimE = dim\hat{\mathcal{E}} - 1 = dim\ \hat{\mathbb{P}}(\mathcal{E})$ de $\mathcal{E}$ et de son projectivisé sont égales. Si $\mathcal{E} = \mathcal{D}$ est une droite, $\mathcal{D}_\infty$ est un point, qu'on note plutôt $\infty_\mathcal{D}$, et qu'on appelle le point à l'infini de la droite affine $\mathcal{D}$.

Si $\mathcal{F}$ est un sous-espace affine de $\mathcal{E}$, son projectivisé $\mathbb{P}(\hat{\mathcal{F}})$ s'identifie naturellement à un sous-espace projectif de $\mathbb{P}(\hat{\mathcal{E}})$, avec $\mathcal{F}_\infty \subset \mathcal{E}_\infty$. Ainsi, pour toute droite $\mathcal{D}$ de $\mathcal{E}$, le point à l'infini $\infty_\mathcal{D}$ de $\mathcal{D}$ est un point de $\mathcal{E}_\infty$, et deux droites $\mathcal{D}_1, \mathcal{D}_2$ sont parallèles si et slt si $\infty_{\mathcal{D}_1} = \infty_{\mathcal{D}_2}$ dans $\mathcal{E}_\infty$.

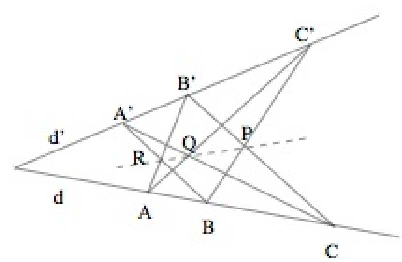
N-B. : si K est infini, toute configuration finie de points d'un espace projectif est contenue dans un ouvert affine convenable. C'est là qu'on les dessine, et nous abandonnerons souvent les notations en lettres grasses utilisées pour désigner les éléments d'un espace projectif.

II.1.3 Retour à la géométrie affine

Le principe de changement d'hyperplan à l'infini évoqué plus haut fournit des théorèmes de Pappus et de Desargues (I. §3.3) les versions équivalentes suivantes.



Pappus affine

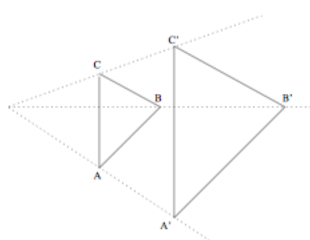


Pappus "projectif"

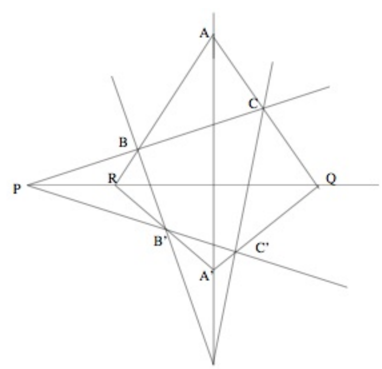
Théoreme II.1.2. (Pappus) *Soient d, d' deux droites distinctes d'un plan projectif, et A, B, C (resp. A', B', C') trois points distincts situés sur d (resp. d'), mais distincts de $d \cap d'$. On suppose que les droites (BC') et $(B'C)$ se coupent en un point P, (CA') et $(C'A)$ en Q, et (AB') et $(A'B)$ en R. Alors, les points P, Q, R sont alignés.*

Preuve : la figure de droite est dessinée dans un plan affine, que nous commençons par projectiviser en un plan projectif $\mathbb{P}$. Choisissons comme nouvelle droite à l'infini de $\mathbb{P}$ la droite projective $\Delta = (PQ)$ (ou, comme on dit plus brièvement : envoyons les points P et Q à l'infini). Les droites (AC') et $(A'C)$ se rencontrent en le point Q de cette droite à l'infini. Vues dans le plan affine $\mathcal{H}$ complementaire de Δ dans $\mathbb{P}$, elles sont donc parallèles. De même, (BC') et (CB'), qui se coupent en $P \in \Delta$, sont parallèles dans $\mathcal{H}$. D'après la version affine du théorème de Pappus (I. Théorème 3.3), les droites affines (AB') et $(A'B)$ de $\mathcal{H}$ sont parallèles. Par conséquent, le point d'intersection R des droites projectives (AB') et $(A'B)$ de $\mathbb{P}$ est situé sur $\Delta = (PQ)$, et les points P, Q, R sont bien alignés.

Inversement, le Théorème 3.3 du chapitre I est conséquence de cet énoncé : soit $\mathbb{P}$ le projectivisé du plan affine $\mathcal{E}$ de la figure de gauche et $\mathcal{E}_\infty$ sa droite à l'infini. Les droites (AB') et (BA') de $\mathcal{E}$ sont parallèles, donc se rencontrent dans $\mathbb{P}$ en un point R de $\mathcal{E}_\infty$. De même, (BC') et (CB') se rencontrent dans $\mathbb{P}$ en un point P de $\mathcal{H}_\infty$, distinct de R puisque $C \neq A$. D'après l'énoncé projectif, le point $Q = (AC') \cap (A'C)$ est situé sur la droite (PR), qui est la droite à l'infini de $\mathcal{E}$, donc $(AC')//(A'C)$ dans $\mathcal{E}$.



Desargues affine



Desargues "projectif"

Théoreme II.1.3. (Desargues) *Soient A, B, C (resp. A', B', C') trois points projectivement indépendants d'un plan projectif, avec $A \neq A', B \neq B', C \neq C'$. Alors les points $P = (BC) \cap (B'C'), Q = (CA) \cap (C'A'), R = (AB) \cap (A'B')$ sont alignés si et seulement si les droites $(AA'), (BB'), (CC')$ sont concourantes.*

La preuve de l'implication $\Rightarrow$ est laissée au plaisir du lecteur. Pour $\Leftarrow$, on commencera par établir une réciproque élémentaire du Théorème 3.5 du chapitre I. Voir aussi l'exercice du §2.3 du présent chapitre.

II.2 Le groupe projectif

II.2.1 Applications projectives

Soient E, E' deux espaces vectoriels, $\mathbb{P}(E), \mathbb{P}(E')$ les espaces projectifs associés, et $f \in \mathcal{L}(E, E')$ une application linéaire non identiquement nulle de E dans E'. Toute droite vectorielle D de E admet pour image sous f :

- ou bien le point $\overrightarrow{0}_{E'}$ si $D \subset Ker(f)$;

- ou bien une droite $f(D)$ de E',

de sorte que f induit une application, dite projective :

$$\mathbb{P}(f) : \mathbb{P}(E) - \mathbb{P}(Ker(f)) \to \mathbb{P}(E') : \mathbb{P}(D) \mapsto \mathbb{P}(f)(D) := \mathbb{P}(f(D))$$

définie sur le complémentaire dans $\mathbb{P}(E)$ du sous-espace projectif $\mathbb{P}(Kerf)$, à valeurs dans $\mathbb{P}(E')$. En particulier, $\mathbb{P}(f)$ est définie sur tout $\mathbb{P}(E)$ si et seulement si f est injective, et (exercice) cela équivaut à demander que $\mathbb{P}(f)$ soit injective sur son domaine de définition. Si f est un isomorphisme, $\mathbb{P}(f)$ est une bijection de $\mathbb{P}(E)$ sur $\mathbb{P}(E')$, d'application réciproque $\mathbb{P}(f^{-1})$; on dit alors que c'est une *transformation projective*, ou encore une *homographie.* La loi $\mathbb{P}(f) \circ \mathbb{P}(g) = \mathbb{P}(f \circ g)$ munit ainsi l'ensemble des transformations projectives de $\mathbb{P}(E)$ dans lui-même d'une structure de groupe, noté $G\mathbb{P}(E)$, et appelé le *groupe projectif* de $\mathbb{P}(E)$.

Si $\mathbb{P}(W)$ est un sous-espace projectif de $\mathbb{P}(E)$ ne rencontrant pas $\mathbb{P}(Kerf)$, c-à-d. si $W \cap Kerf = \{\overrightarrow{0}_E\}$, $\mathbb{P}(f)$ est définie sur tout $\mathbb{P}(W)$, et l'envoie injectivement sur le sous-espace projectif $\mathbb{P}(f)(\mathbb{P}(W)) = \mathbb{P}(f(W))$ de $\mathbb{P}(E')$, de même dimension que $\mathbb{P}(W)$. Autrement dit, $\mathbb{P}(f)$ induit une homographie de $\mathbb{P}(W)$ sur son image. Exemple : les perspectives projectives introduites au §1.1 (et généralisées dans l'exemple ci-dessous).

Les applications projectives $\mathbb{P}(f)$ respectent l'alignement : si $P_i = \mathbb{P}(D_i)$, $i = 1, 2, 3$, sont trois points de $\mathbb{P}(E) - \mathbb{P}(Ker(f))$ alignés et distincts, les droites D_1, D_2, D_3 engendrent dans E un plan $W \not\subset Ker(f)$. Si $W \cap Ker(f) = \{\overrightarrow{0}_E\}$, $f(W)$ est un plan, et les points $\mathbb{P}(f)(P_i)$ sont situés sur la droite

$\mathbb{P}(f(W))$; sinon, $f(W)$ est une droite, et les trois points $\mathbb{P}(f)(P_i)$ coïncident avec le point $\mathbb{P}(f(W))$.

Exemple (projections et perspectives généralisées) : Soient F, W deux sous-espaces vectoriels supplémentaires de dimensions > 0 de E. La projection vectorielle $f : E \to W$ attachée à la décomposition $E = F \oplus W$ a pour noyau F, et définit une application projective surjective : $\mathbb{P}(f) : \mathbb{P}(E) - \mathbb{P}(F) \to \mathbb{P}(W)$, appelée projection de centre $\mathbb{P}(F)$ sur $\mathbb{P}(W)$. Pour tout point $P = \mathbb{P}(D)$ de $\mathbb{P}(E)$ hors de $\mathbb{P}(F)$, le sous-espace projectif $\Phi_P := \mathbb{P}(D + F)$ a pour dimension $dimF = dim\mathbb{P}(D) + dim\mathbb{P}(F)$; d'après la Proposition 1.1, Φ_P rencontre donc $\mathbb{P}(W)$ en un unique point $p_{\Phi_P} = \mathbb{P}((D+F)\cap W)$. Mais la projection vectorielle f vérifie : $f(D) = f((D+F)\cap W) = (D+F)\cap W$, donc $p_{\Phi_P} = \mathbb{P}(f)(P)$. Si F est une droite, $\mathbb{P}(F) = \mathbf{O}$ et $\mathbb{P}(W) = \mathbf{H}$ sont un point et un hyperplan : on retrouve la projection centrale du §1.1, qui est donc une application projective. En particulier, la perspective $\pi_{|\mathbf{H}'} : \mathbf{H}' \to \mathbf{H}$ qu'elle induit sur tout hyperplan $\mathbf{H}'$ ne passant pas par $\mathbf{O}$ est une homographie. Idem, en dimension supérieure, pour la restriction de $\mathbb{P}(f)$ à tout $\mathbb{P}(W')$ ne rencontrant pas $\mathbb{P}(F)$.

Soit E un espace vectoriel de dimension $n + 1$. L'ensemble $\{\lambda \mathbf{I}_{n+1}, \lambda \in K^*\} \subset GL_{n+1}(K)$ des matrices carrées scalaires inversibles forme un sous-groupe de $GL_{n+1}(K)$ isomorphe à K^*, et dont les éléments commutent avec tous les éléments de $GL_{n+1}(K)$ (en fait, c'en est tout le centre). Le quotient

$$PGL_{n+1}(K) = GL_{n+1}(K)/K^*$$

est donc naturellement muni d'une structure de groupe. De façon intrinsèque, les homothéties $\lambda.id_E, \lambda \in K^*$, de E forment le centre de $GL(E)$, et on pose $PGL(E) = GL(E)/K^*.id_E$. La partie (i) de la proposition suivante justifie que ces quotients s'appellent aussi des groupes projectifs. La partie (ii) montre que les groupes affines s'identifient aux sous-groupes de groupes projectifs laissant stables un hyperplan projectif.

Proposition II.2.1. *soit E un espace vectoriel de dimension > 0.*

i) L'application $\mathbb{P} : GL(E) \to G\mathbb{P}(E) : f \mapsto \mathbb{P}(f)$ induit un isomorphisme de groupes

$$PGL(E) \simeq G\mathbb{P}(E).$$

ii) Soient $\mathcal{E}$ un espace affine sous E, et $\mathcal{E}_\infty = \mathbb{P}(E)$ l'hyperplan à l'infini de son projectivisé $\mathbb{P}(\hat{\mathcal{E}})$. Toute transformation affine f de $\mathcal{E}$ se prolonge de *façon unique en une transformation projective $\mathbb{P}(\hat{f})$ de $\mathbb{P}(\hat{\mathcal{E}})$, et l'application $\hat{\mathbb{P}} : f \mapsto \mathbb{P}(\hat{f})$ établit un isomorphisme de groupe*

$$GA(\mathcal{E}) \simeq \{\varphi \in G\mathbb{P}(\hat{\mathcal{E}}), \varphi(\mathcal{E}_\infty) = \mathcal{E}_\infty.\}$$

Preuve : i) puisque $\mathbb{P}(f \circ g) = \mathbb{P}(f) \circ \mathbb{P}(g)$, $\mathbb{P}$ est un morphisme de groupe, qui est surjectif par définition. Son noyau est constitué des automorphismes de E admettant tous les vecteurs de E comme vecteurs propres, et est donc réduit aux homothéties.

Pour $E = K^{n+1}$, prendre garde aux indices : c'est $PGL_{n+1}(K)$ qui représente le groupe projectif de $\mathbb{P}_n(K)$.

ii) Identifions $\hat{\mathcal{E}}$ à $E \oplus K = E \times K$ en fixant une origine P_0 de $\mathcal{E}$, de sorte que $\mathcal{E} = E \times \{1\}, P_0 = (\overrightarrow{0}_E, 1)$. Les éléments f de $GA(\mathcal{E})$ se mettent alors sous la forme $[\overrightarrow{v}, U] = \begin{pmatrix} U & \overrightarrow{v} \\ 0 & 1 \end{pmatrix}$, où $\overrightarrow{v} \in E, U = \overrightarrow{f} \in GL(E)$ (notations suivant la Definition 3.2 du chapitre I). Notons $\hat{f}$ l'élément de $GL(\hat{\mathcal{E}})$ défini par cette matrice. Pour tout $P = P_0 + \overrightarrow{P_0P} \in \mathcal{E}$, c-à-d. pour $P = (\overrightarrow{P_0P}, 1) \in \hat{\mathcal{E}}$, le point $f(P) = f(P_0) + \overrightarrow{f}(\overrightarrow{P_0P}) = P_0 + \overrightarrow{v} + U.\overrightarrow{P_0P}$ coïncide avec le point $\hat{f}(P) = (U.\overrightarrow{P_0P} + \overrightarrow{v}, 1)$ de $\hat{\mathcal{E}}$. Donc $\hat{f}$ prolonge bien f à $\hat{\mathcal{E}}$, et puisque $\mathcal{E}$ engendre l'espace vectoriel $\hat{\mathcal{E}}$, c'est le seul prolongement linéaire possible. L'application $f \mapsto \hat{f}$ est un morphisme de groupes.

Dans ces conditions, $\mathbb{P}(\hat{f}) \in G\mathbb{P}(\hat{\mathcal{E}})$ induit l'application f sur l'ouvert affine $\mathcal{E}$ de $\mathbb{P}(\hat{\mathcal{E}})$ formé par les droites vectorielles $(0_{\hat{E}}P), P \in \mathcal{E}$, et pour toute transformation projective $\mathbb{P}(g)$ de $\mathbb{P}(\hat{\mathcal{E}})$ prolongeant f, l'application linéaire $g^{-1} \circ \hat{f}$ est une homothétie. Ainsi, f admet un unique prolongement projectif $\mathbb{P}(\hat{f}) := \hat{\mathbb{P}}(f)$ à $\mathbb{P}(\hat{\mathcal{E}})$, et l'application $\hat{\mathbb{P}}$ ainsi définie est un morphisme de groupes.

Enfin, un élément $\varphi = \mathbb{P}(g)$ de $G\mathbb{P}(\hat{\mathcal{E}})$ laisse stable l'hyperplan $\mathcal{E}_\infty = \mathbb{P}(E)$ si et seulement si $g \in GL(\hat{\mathcal{E}})$ laisse stable $E \times 0$, c'-à- d. si et seulement s'il est de la forme $g = \begin{pmatrix} U' & \overrightarrow{v}' \\ 0 & \lambda \end{pmatrix}$, où $\lambda \in K^*$; quitte à remplacer g par $\lambda^{-1}g$, ce qui ne modifie pas $\mathbb{P}(g)$, on peut en plus supposer que $\lambda = 1$. On en déduit que $\hat{\mathbb{P}}$ admet pour image le stabilisateur décrit dans l'énoncé. Comme $\hat{\mathbb{P}}$ est par définition injective, c'est un isomorphisme.

II.2.2 Coordonnés homogènes

Soit E un espace vectoriel de dimension $n+1$ sur K. Le choix d'une base $\overrightarrow{e_0}, ..., \overrightarrow{e_n}$ de E identifie E à K^{n+1}, et les vecteurs de $E - \{\overrightarrow{0}\}$ a des $(n+1)$-uplets de scalaires *non tous nuls*. Deux vecteurs $\overrightarrow{v} = (X_0, ..., X_n)$, $\overrightarrow{v}' = (X'_0, ..., X'_n)$ non nuls engendrent la même droite D de E, définissant donc le même point $P = \mathbb{P}(D)$ de $\mathbb{P}(E) \simeq \mathbb{P}_n(K)$, si et seulement si

$$\exists \lambda \in K^*,\ X'_0 = \lambda X_0,\ X'_1 = \lambda X_1, ...,\ X'_n = \lambda X_n.$$

On dit que chacun de ces $(n+1)$-uplets est un *système de coordonnées homogènes* du point P, et on écrit :

$$P = (X_0 : X_1 : .... : X_n) = (X'_0 : X'_1 : ... : X'_n).$$

Par exemple, un point P de $\mathbb{P}_1(K) = \mathbb{P}(K^2)$ est représenté par $(X_0 : X_1)$, avec $(X_0, X_1) \neq (0, 0)$, donc par $(1 : \alpha)$, avec $\alpha = X_1/X_0$ si $X_0 \neq 0$, mais aussi par $(\beta : 1)$, avec $\beta = X_0/X_1$ si $X_1 \neq 0$. Le point $(0 : X_1) = (0 : 1)$ est le point noté ∞' au §1.2, tandis que $(X_0 : 0) = (1 : 0) = \infty$. Pour $P \notin \{\infty, \infty'\}$, on a $\alpha\beta = 1$.

La propriété pour un point P de $\mathbb{P}_n(K)$ que X_0 soit nul est indépendante du choix des coordonnées homogènes : on aura alors $X'_0 = \lambda X_0 = 0$, et réciproquement, car $\lambda \neq 0$. Elle entraîne que l'une au moins des autres coordonnées $X_1, ..., X_n$ soit non nulle. En fait, l'ensemble

$$\{P \in \mathbb{P}_n(K), X_0 = 0\} \simeq \{(X_1 : ... : X_n), (X_1, ..., X_n) \in K^n - \{\overrightarrow{0}\}\} = \mathbb{P}_{n-1}(K)$$

est l'hyperplan projectif $\mathbb{P}(H_0)$ de $\mathbb{P}_n(K)$ défini par l'hyperplan vectoriel H_0 de $E = K^{n+1}$ d'équation $X_0 = 0$. Comme $X_0 \neq 0$ sur l'ouvert affine $\mathcal{H}_0$ complémentaire de $\mathbb{P}(H_0)$ dans $\mathbb{P}_n(K)$, on peut, en multipliant par $\lambda = 1/X_0$, représenter ses points par $(1 : x_1 = \frac{X_1}{X_0} : ... : x_n = \frac{X_n}{X_0})$, et identifier $\mathcal{H}_0$ au sous-espace affine $\overrightarrow{e_0} + H_0$ de E, muni des coordonnées cartésiennes $(x_1, ..., x_n)$ attachées à l'origine $\overrightarrow{e_0}$ de $\mathcal{H}_0$ et à la base $\overrightarrow{e_1}, ..., \overrightarrow{e_n}$ de H_0. Ce passage des coordonnées homogènes aux coordonnéees cartésiennes traduit donc le passage du projectif à l'affine décrit au §1.2. Remarquons qu'en effectuant cette construction pour $i = 0, ..., n$, on obtient $n+1$ ouverts affines $\mathcal{H}_i$, dont la réunion (non disjointe) *recouvre* l'espace projectif $\mathbb{P}_n(K)$ tout entier. Voici par exemple les formules de passage de l'ouvert affine $\mathcal{H}_0$ à l'ouvert affine $\mathcal{H}_1$ de $\mathbb{P}_2(K)$: l'isomorphisme $\mathcal{H}_0 \simeq K^2$ est donné par

$(X_0 \neq 0 : X_1 : X_2) \mapsto (x = \frac{X_1}{X_0}, y = \frac{X_2}{X_0})$; l'isomorphisme $\mathcal{H}_1 \simeq K^2$ est donné par $(X_0 : X_1 \neq 0 : X_2) \mapsto (u = \frac{X_0}{X_1}, v = \frac{X_2}{X_1})$. Sur $\mathcal{H}_0 \cap \mathcal{H}_1$, on a donc $(u,v) = (\frac{1}{x}, \frac{y}{x}), (x,y) = (\frac{1}{u}, \frac{v}{u})$.

Soit maintenant $\mathcal{E}$ un espace affine de dimension n, muni d'une origine et d'une base, qui l'identifient à K^n. Ses points sont repérés par des coordonnées cartésiennes $(x_1, ..., x_n)$. Écrivons son prolongement vectoriel $\hat{\mathcal{E}}$ sous la forme $K \oplus K^n$, repéré par les coordonnées cartésiennes $(x_0, x_1, ..., x_n)$. Le plongement de $\mathcal{E}$ dans son projectivisé $\mathbb{P}(\hat{\mathcal{E}})$ est alors donné par $(x_1, ..., x_n) \mapsto (1 : x_1 : ... : x_n)$, égal à $(\lambda : \lambda x_1 : ... : \lambda x_n)$ pour tout $\lambda \neq 0$, et $\mathcal{E}_\infty = \mathbb{P}(\{x_0 = 0\})$.

Exemple : dans le plan affine $\mathcal{E} = xOy$, la projection centrale de centre $O = (0,0)$ sur la droite $\mathcal{H} : y = 1$, s'écrit : $(x,y) \mapsto (\frac{x}{y}, 1)$ et induit sur la droite $\mathcal{H}' : x = 1$, la perspective affine $(1,y) \mapsto (\frac{1}{y}, 1)$. Dans son projectivisé $\mathbb{P}(\hat{\mathcal{E}})$, repéré par les coordonnées homogènes $(X_0 : X_1 : X_2)$, avec $x = \frac{X_1}{X_0}, y = \frac{X_2}{X_0}$ sur $\mathcal{E}$, la projection centrale de centre $O = (1 : 0 : 0)$ sur la droite H : $X_0 - X_2 = 0$, s'écrit : $(X_0 : X_1 : X_2) \mapsto (X_2 : X_1 : X_2)$, et induit sur la droite $H' : X_0 - X_1 = 0$, la perspective projective $(X_0 : X_0 : X_2) \mapsto (X_2 : X_0 : X_2)$.

Équations des sous-espaces

Soient $\mathbb{P}(H)$ un hyperplan de $\mathbb{P}_n(K)$, et $L(X_0, ..., X_n) = a_0 X_0 + ... + a_n X_n$ une forme linéaire sur K^{n+1} telle que $Ker(L) = H$. Quel que soit le choix de coordonnées projectives $(c_0 : ... : c_n)$ d'un point P de $\mathbb{P}_n(K)$, la relation $P \in \mathbb{P}(H)$ équivaut à la condition $L(c_0, ..., c_n) = 0$. On écrit cette condition sous la forme $L(P) = 0$ (tout en gardant à l'esprit que L n'est pas une fonction de P), et on dit que L est une équation de $\mathbb{P}(H)$. Les autres équations de $\mathbb{P}(H)$ sont de la forme $\lambda . L$, où $\lambda \in K^*$. Si $E^* = \mathcal{L}(E, K)$ désigne l'espace dual de E, ces équations définisssent donc un unique point de $(E^* - \{0\})/K^*$, c'est-à-dire un point de l'espace projectif $\mathbb{P}(E^*)$. Nous revenons plus bas sur la "dualité" qui en résulte entre l'ensemble des hyperplans de $\mathbb{P}(E)$ et l'ensemble des points de $\mathbb{P}(E^*)$. Notons pour l'instant que si $\mathbb{P}(H) \neq \mathbb{P}(H_0)$, c'est-à-dire si $a_1, ..., a_n$ ne sont pas tous nuls, $\mathbb{P}(H)$ induit sur l'ouvert affine $\mathcal{H}_0 : X_0 \neq 0$ un hyperplan affine, donné dans les coordonnées cartésiennes $(x_1, ..., x_n)$ de $\mathcal{H}_0$ par l'équation affine $\ell(x_1, ... x_n) := a_1 x_1 + ... + a_n x_n + a_0 = 0$. On dit que que la forme affine $\ell(\frac{X_1}{X_0}, ..., \frac{X_n}{X_0}) := \frac{1}{X_0} L(X_0, ..., X_n)$ est la *déshomogénéisée par rapport à* X_0 de la forme linéaire L.

Inversement, un hyperplan $\mathcal{H}$, de direction H, de l'espace affine $\mathcal{E} = K^n$ admet une équation affine de la forme $\ell(x_1, ..., x_n) := a_1 x_1 + ... + a_n x_n +$

$a_0 = 0$, dont la partie linéaire, non identiquement nulle, est une équation de H. Son projectivisé $\mathbb{P}(\hat{\mathcal{H}})$ est défini dans $\mathbb{P}(\hat{\mathcal{E}})$ par l'équation homogène $L(X_0, ..., X_n) := a_0X_0 + a_1X_1 + ... + a_nX_n = 0$. On dit que $L(X_0, ..., X_n) = X_0\ell(\frac{X_1}{X_0}, ..., \frac{X_n}{X_0})$ est l'*homogénéisée* de la forme affine ℓ.

Ces constructions s'étendent sans difficulté aux systèmes d'équations linéaires homogènes $L_1 = .. = L_k = 0$ définissant un sous-espace projectif $\mathbb{P}(W)$ de $\mathbb{P}_n(K)$: le sous-espace affine $\mathcal{W} = \mathbb{P}(W) \cap \mathcal{H}_0$ est défini dans les coordonnées cartésiennes de $\mathcal{H}_0$ par les équations déshomogénéisées relativement à X_0 : $\ell_1 = ... = \ell_k = 0$. On a $codim\mathbb{P}(W) = k$ si et seulement si les formes $L_1, ..., L_k$ sont linéairement indépendantes dans E^* ; $\mathbb{P}(W) \not\subset \mathbb{P}(H_0)$ (c-à-d. $\mathcal{W} \neq \emptyset$) si et seulement si X_0 n'appartient pas au sous-espace vectoriel de E^* engendré par les L_i ; et $codim\mathcal{W} = k$ si et seulement si $L_1, ..., L_k$ et X_0 sont linéairement indépendantes, ce qui équivaut à dire que les parties linéaires $L_i(0, X_1, ..., X_n)$ des formes affines $\ell_i, i = 1, ...k$, sont linéairement indépendantes. Ces parties linéaires, qui sont des éléments du dual H_0^* de H_0, forment, un système d'équations homogènes de $\mathbb{P}(W) \cap \mathbb{P}(H_0)$; celui-ci n'est le sous-espace à l'infini $\mathcal{W}_\infty$ de $\mathcal{W}$ que quand $\mathcal{W} \neq \emptyset$.

Paramétrage des sous-espaces.

Comme en vectoriel et en affine, un autre façon de représenter les sous-espaces projectifs de $\mathbb{P}_n(K)$ consiste à les paramétrer. Soient $\mathbb{P}(W)$ un sous-espace projectif de dimension $m \leq n$, et $P_0, ..., P_m$ une famille de $m + 1$ points de $\mathbb{P}(W)$ projectivement indépendants. Alors, $\mathbb{P}(W) =< P_0, ..., P_m >$. Pour tout $i = 0, ..., m$, soit $\overrightarrow{v}_i = {}^t(c_{0i}, ..., c_{ni})$ un vecteur de K^{n+1} tel que $P_i = \mathbb{P}(K\overrightarrow{v_i})$, c-à -d. tel que $(c_{0i} : ... : c_{ni})$ est un système de coordonnées projectives du point P_i. Par définition de l'indépendance projective, les $\overrightarrow{v_i}$ sont linéairement indépendants. Tout point P de $\mathbb{P}(W)$ s'écrit alors $\mathbb{P}(K\overrightarrow{v})$, où $\overrightarrow{v} = \lambda_0\overrightarrow{v_0} + ... + \lambda_m\overrightarrow{v_m}, (\lambda_0, ..., \lambda_m) \in K^{m+1}$, est unique à homothétie près, d'où une bijection

$$\psi : \mathbb{P}_m(K) \simeq \mathbb{P}(W) : (\lambda_0 : ... : \lambda_m) \mapsto \mathbb{P}(\Sigma_{i=0,...,m}\lambda_i\overrightarrow{v_i}).$$

De plus, ψ est une homographie. En effet, on a $\psi = \mathbb{P}(f)$, où f désigne l'application linéaire de K^{m+1} dans $W \subset K^{n+1}$ de matrice représentative $\begin{pmatrix} c_{00} & . & c_{0m} \\ \vdots & . & \vdots \\ c_{n0} & . & c_{nm} \end{pmatrix}$. La relation $rang \begin{pmatrix} c_{00} & . & c_{0m} & X_0 \\ \vdots & . & \vdots & . \\ c_{n0} & . & c_{nm} & X_n \end{pmatrix} = m + 1 \Leftrightarrow P = (X_0 : ... : X_n) \in \mathbb{P}(W)$ permet alors de retrouver un système d'équations homogènes définissant $\mathbb{P}(W)$, en considérant les mineurs d'ordre maximal $(= m+2)$ de cette deuxième matrice.

On prendra garde que la notation $\psi(\lambda_0 : ... : \lambda_m)$" = "$\lambda_0 P_0 + ... + \lambda_m P_m$ est abusive : ψ dépend non seulement des points P_i, mais aussi des systèmes de coordonnées projectives *individuels* qu'on a choisis pour les représenter. Pour la même raison, la donnée de $n+1$ points projectivement indépendants d'un espace projectif $\mathbb{P}(E)$ de dimension n ne suffit pas à l'identifier à $\mathbb{P}_n(K)$. Le $(n+2)$-ième point introduit dans la définition d'un repère projectif au §1.1 répond précisément à ce problème :

Proposition II.2.2. *Soient $\mathbb{P}(E)$ un espace projectif de dimension $n \geq 1$, et $(P_0, ..., P_{n+1})$ un repère projectif de $\mathbb{P}(E)$. Il existe une unique homographie $\phi : \mathbb{P}(E) \simeq \mathbb{P}_n(K)$ telle que*

$$\forall i = 0, ..., n,\ \phi(P_i) = (0 : ... : 1_{(i)} : ... : 0), \text{ et } \phi(P_{n+1}) = (1 : .... : 1).$$

Preuve Pour tout $i = 0, ..., n$, soit $\overrightarrow{v}_i$ un générateur de la droite de E associée à P_i. Il existe un (unique) isomorphisme $f' : E \simeq K^{n+1}$ envoyant $(\overrightarrow{v_0}, ..., \overrightarrow{v_n})$ sur la base canonique $(\overrightarrow{e_0}, ..., \overrightarrow{e_n})$ de K^{n+1}. Comme $(P_0, ..., P_{n+1})$ est un repère projectif, aucune des coordonnées c_i de $f'(\overrightarrow{v_{n+1}}) = \Sigma_{i=0,...,n} c_i \overrightarrow{e_i}$ n'est nulle. Alors, $\phi = \mathbb{P}(f)$, où $f = diag(c_0^{-1}, ..., c_n^{-1}) \circ f'$, répond à la question. Si $\mathbb{P}(g)$ est une autres solution, l'automorphisme $g \circ f^{-1}$ de K^{n+1} admet pour vecteurs propres $\overrightarrow{e_0}, ..., \overrightarrow{e_n}$ (sa matrice représentative est donc diagonale) et $\overrightarrow{e_0} + ... + \overrightarrow{e_n}$. Elle est donc scalaire, et $\mathbb{P}(g) = \mathbb{P}(f)$.

Le birapport

Nous identifions ici $\mathbb{P}_1(K)$ à $K \cup \infty := \hat{K}$, en décrétant que $\infty = \{X_1 = 0\}$ et que les points de K sont repérés par X_0/X_1. Alors $\{\infty = (1 : 0)\ ,\ 0 = (0 : 1)\ , 1 = (1, 1)\}$ est un repère projectif de $\mathbb{P}_1(K)$. De façon générale, trois points distincts d'une droite projective en forment un repère projectif. Pour $n = 1$, la proposition précédente valide donc l'importante :

Definition II.2.3. *: soit Δ une droite projective, et A, B, C trois points distincts de Δ, de sorte qu'il existe une unique homographie $\phi : \Delta \to \hat{K}$ telle que $\phi(A) = \infty, \phi(B) = 0, \phi(C) = 1$. Pour tout point D de Δ, on appelle birapport des points ordonnés A, B, C, D l'élément*

$$[A, B, C, D] := \phi(D) \in K \cup \infty.$$

On dit que (A, B, C, D) forme une division harmonique *si* $[A, B, C, D] = -1$.

Calculons ce birapport quand $\Delta = \mathcal{D} \cup \infty_{\mathcal{D}}$ est la projectivisée d'une droite affine $\mathcal{D}$, identifiée à K par le choix d'une coordonnée cartésienne z, et contenant les points A, B, C, repérés par leurs affixes $z_A, z_B, z_C \in K$. L'homographie $\phi : \mathbb{P}(\Delta) \simeq \mathbb{P}_1(K) \to \hat{K} \simeq \mathbb{P}_1(K)$ s'identifie alors à un élément $\mathbb{P}(g)$ de $PGL_2(K)$, représenté par une matrice $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, avec $ad - bc \neq 0$, et $\phi(X_0 : X_1) = (aX_0 + bX_1 : cX_0 + dX_1)$. En particulier, un point $z \sim (z, 1)$ de K s'envoie sur le point $\phi(z) = (az + b : cz + d) \sim \frac{az+b}{cz+d}$ de $\hat{K}$, et $\phi(\infty_{\mathcal{D}}) = \frac{a}{c} \in \hat{K}$. Dans ces notations $\frac{u}{v}$, u et v ne sont jamais tous deux nuls, et on convient de $\frac{u}{0} = \infty$. Ainsi, $\phi(\infty_{\mathcal{D}}) = \infty$ si et seulement si $c = 0$. Les conditions $\phi(A) = \infty, \phi(B) = 0, \phi(C) = 1$ entraînent alors que $\phi(z) = \frac{z_C - z_A}{z_C - z_B} \times \frac{z - z_B}{z - z_A} := \frac{z_C - z_A}{z_C - z_B} : \frac{z - z_A}{z - z_B}$, d'où

$$\phi(z_D) = [A, B, C, D] = \frac{z_C - z_A}{z_C - z_B} : \frac{z_D - z_A}{z_D - z_B} = \frac{\overrightarrow{CA}}{\overrightarrow{CB}} : \frac{\overrightarrow{DA}}{\overrightarrow{DB}}$$

si $D \neq \infty_{\mathcal{D}}$, et $\phi(\infty_{\mathcal{D}}) = \frac{z_C - z_A}{z_C - z_B} = \frac{\overrightarrow{CA}}{\overrightarrow{CB}}$.

En particulier (pour $car K \neq 2$), $A, B, C, \infty_{\mathcal{D}}$ sont en division harmonique si et seulement si C est le milieu du segment $[AB]$ de la droite affine $\mathcal{D}$.

Proposition II.2.4. *: soit $\theta : \Delta \to \Delta'$ une bijection entre deux droites projectives. Alors, θ est une homographie si et seulement si elle préserve les birapports.*

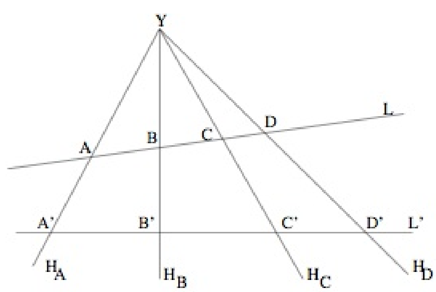
Preuve . $\Rightarrow$: si θ est une homographie, et si les points distincts A, B, C, et un point D de Δ ont pour image par θ les points, encore distincts, A', B', C' et le point D' de Δ', il existe une unique homographie $\phi : \Delta \to \hat{K}$ (resp. $\psi = \Delta' \to \hat{K}$ envoyant le repère projectif (A, B, C) (resp. (A', B', C')) sur $(\infty, 0, 1)$. L'homographie $\psi \circ \phi^{-1}$ de $\hat{K}$ fixant $(\infty, 0, 1)$, c'est l'identité. Donc $\psi(D') = \phi(D)$, et $[A', B', C', D'] = \psi(D') = \phi(D) = [A, B, C, D]$.

$\Leftarrow$: identifiant Δ et Δ' à $\mathbb{P}_1(K)$ au moyen d'homographies arbitraires, on se ramène à étudier une bijection θ de $\mathbb{P}_1(K)$ sur lui-même. D'après la proposition 2.1, il existe une homographie ϕ de $\mathbb{P}_1(K)$ envoyant $0, 1, \infty$ sur leurs images par la bijection θ. L'hypothèse et le sens $\Rightarrow$ montrent que la bijection $\theta' = \phi^{-1} \circ \theta$, qui fixe $\infty, 0, 1$, préserve le birapport. Pour tout $z \neq 0, 1, \infty$, on a donc $\theta'(z) = [\infty, 0, 1, \theta'(z)] = [\theta'(\infty), \theta'(0), \theta'(1), \theta'(z)] = [\infty, 0, 1, z] = z$, donc θ' est l'identité, et $\theta = \phi$ est une homographie.

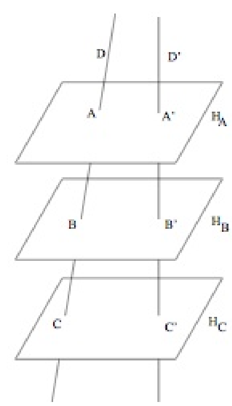
On déduit de cet énoncé les identités $[C, D, A, B] = [A, B, C, D] = [B, A, C, D]^{-1} = [A, B, D, C]^{-1} = 1 - [A, C, B, D]$. Ainsi, les 24 permutations de $\mathfrak{S}_4$ conduisent à (au plus) 6 valeurs de birapports, et on peut, dans la définition d'une division harmonique, intervertir A et B, ainsi que C et D (et bien sûr, $\{A, B\}$ et $\{C, D\}$).

II.2.3 Retour à la géométrie

Soient $\mathbf{H}_1$ et $\mathbf{H}_2$ deux hyperplans distincts d'un espace projectif $\mathbb{P}(E)$, d'équa-tions respectives $L_1 = 0, L_2 = 0$ (L_1 et L_2 sont donc des éléments linéairement indépendants du dual E^* de E). On appelle *faisceau d'hyperplans* de $\mathbb{P}(E)$ engendré par $\mathbf{H}_1$ et $\mathbf{H}_2$ l'ensemble $\mathfrak{F}$ des hyperplans $\mathbf{H}$ de $\mathbb{P}(E)$ d'équations $L = 0$, où $L = a_1L_1 + a_2L_2, (a_1, a_2) \in K^2$, parcourt le plan de E^* engendré par L_1 et L_2. Dans ces conditions, $\mathbf{H}_1 \cap \mathbf{H}_2$ est un sous-espace projectif $\mathbf{W}$ de codimension 2 dans $\mathbb{P}(E)$ défini par le système $L_1 = L_2 = 0$, et un hyperplan $\mathbf{H}$ appartient au faisceau $\mathfrak{F}$ si et seulement s'il contient $\mathbf{W}$, qu'on appelle le centre du faisceau. Par exemple, dans le plan projectif, un faisceau de droites est constitué par l'ensemble des droites projectives passant par un point donné.



Thalès projectif



Thalès affine ($\frac{\overrightarrow{CA}}{\overrightarrow{CB}} = \frac{\overrightarrow{C'A'}}{\overrightarrow{C'B'}}$)

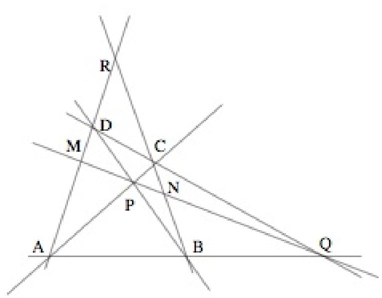
Théoreme II.2.5. (Thalès projectif) : *soient* $\mathbf{H}_i, i = 1, ..., 4$, *quatre hyperplans d'un faisceau d'hyperplans* $\mathfrak{F}$ *de* $\mathbb{P}(E)$, *et* Δ, Δ' *deux droites de* $\mathbb{P}(E)$. *On suppose que* Δ *(resp.* Δ'*) rencontre les hyperplans en quatre points distincts* A_i *(resp.* A'_i*),* $i = 1, .., 4$. *Alors,* $[A_1, A_2, A_3, A_4] = [A'_1, A'_2, A'_3, A'_4]$.

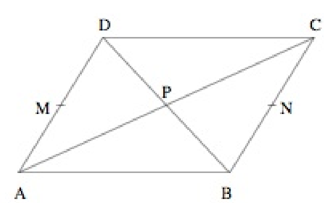
Preuve : soient $\mathbf{W} = \mathbb{P}(W)$ le centre du faisceau $\mathfrak{F}$, $\Delta = \mathbb{P}(\Pi), \Delta' = \mathbb{P}(\Pi')$. Comme $W \oplus \Pi' = E = W \oplus \Pi$, on peut considérer la projection centrale généralisée de centre $\mathbf{W}$ sur Δ', et sa restriction à Δ. Cette perspective généralisée est une homographie θ de Δ sur Δ', qui envoie les A_i sur les A'_i. On conclut par la proposition 2.4. Ainsi, le birapport des points d'intersection est indépendant de la sécante Δ. On l'appelle le *birapport des 4 hyperplans du faisceau*, et on le note $[\mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3, \mathbf{H}_4]$.

Envoyons à l'infini l'un des hyperplans. Dans l'ouvert affine complémentaire, les hyperplans restant sont parallèles, leurs intersections avec les droites Δ, Δ' sont des triplets de points, les 4es points devenant leurs points à l'infini. La relation $[A, B, C, \infty_{\mathcal{D}}] = \frac{\overrightarrow{CA}}{\overrightarrow{CB}}$ entraîne alors la généralisation suivante du théorème de Thalès : dans les notations de la figure de droite, $\frac{\overrightarrow{CA}}{\overrightarrow{CB}} = \frac{\overrightarrow{C'A'}}{\overrightarrow{C'B'}}$, même si Δ et Δ' ne sont pas coplanaires.

Dans le même esprit, soit $ABCD$ un quadrilatère propre quelconque d'un plan affine (c-à-d. : soit (A, B, C, D) un repère projectif d'un plan projectif). Ses 4 côtés et ses 2 diagonales définissent 6 droites, d'où, (au moins en projectif) 3 nouveaux points d'intersection : P (pour les diagonales), et Q, R (pour les couples de côté opposés). Soient M, N les points d'intersection de (PQ) avec $(AD), (BC)$. Alors,

$$[M, N, P, Q] = -1 \qquad \text{car} \qquad \frac{\overrightarrow{PM}}{\overrightarrow{PN}} = -1$$





Théoreme II.2.6. (du quadrilatère) *: M, N, P, Q sont en division harmonique.*

Preuve : envoyons la droite (QR) à l'infini. Le quadrilatère devient un parallélogramme, donc P est le milieu du segment $[MN]$, tandis que $Q = \infty_{(MN)}$. Ainsi $[M, N, P, \infty_{(MN)}] = -1$. Mais le birapport ne dépend que des points sur la droite, pas du choix d'un point à l'infini, donc $[M, N, P, Q] = -1$.

Remarque : considérons le faisceau $\mathfrak{F}$ engendré par les droites (RA) et (RB), et soit Q un point du plan projectif hors de ces droites. Il existe une unique droite Δ_Q de $\mathfrak{F}$ telle que $((RA), (RB), \Delta_Q, (RQ)) = -1$. On dit que Δ_Q est la *polaire* de Q par rapport aux droites (RA) et (RB). La figure de gauche en fournit une construction : tracer deux sécantes depuis Q, donnant les points $B, A; C, D$, et poser $P = (AC) \cap (BD)$. Alors, $\Delta_Q = (RP)$.

La dualité projective

Soient E un espace vectoriel de dimension ≥ 2, et $E^* = \mathcal{L}(E, K)$ son dual. Comme on l'a vu au §2.2, Équations, la loi qui, à un hyperplan $\mathbf{H} = \mathbb{P}(H)$ de l'espace projectif $\mathbb{P}(E)$, attache l'ensemble $\mathbf{h} = \{\lambda.L, \lambda \in K^*\}$ des formes linéaire non nulles qui s'annulent sur $\mathbb{P}(H)$, s'interprète comme une bijection

$$\delta_1 : Hyp(E) \to \mathbb{P}(E^*) : \mathbf{H} \leftrightarrow \mathbf{h}$$

de l'ensemble $Hyp(E)$ des hyperplans de $\mathbb{P}(E)$ (ou de E, cela revient au même) vers l'ensemble des points de $\mathbb{P}(E^*)$. Dans cette transformation, les hyperplans du faisceau $\mathfrak{F}$ engendré par deux hyperplans $\mathbf{H}_1 \neq \mathbf{H}_2$, d'équations $L_1 = 0, L_2 = 0$, admettent pour équations les combinaisons linéaires de L_1 et de L_2, et correspondent donc aux points de la droite $\mathfrak{f}$ engendrée par $\mathbf{h}_1$ et $\mathbf{h}_2$ dans $\mathbb{P}(E^*)$. Autrement dit, si $\mathbf{W} = \mathbf{H}_1 \cap \mathbf{H}_2$ désigne le centre du faisceau $\mathfrak{F}$, δ_1 induit une bijection

$$\{\mathbf{H} \in Hyp(\mathbb{P}(E)), \mathbf{H} \supset \mathbf{H}_1 \cap \mathbf{H}_2\} \leftrightarrow \{\mathbf{h} \in \mathbb{P}(E^*), \mathbf{h} \in (\mathbf{h}_1\mathbf{h}_2))\} \qquad (*)$$

entre l'ensemble des hyperplans passsant par $\mathbf{H}_1 \cap \mathbf{H}_2$ et la droite $\mathfrak{f} =< \mathbf{h}_1, \mathbf{h}_2 >$ de $\mathbb{P}(E^*)$. Par exemple, quand $\mathbb{P}(E)$ est un plan, la bijection δ_1 fait passer des droites de $\mathbb{P}(E)$ aux point de $\mathbb{P}(E^*)$, et une famille de droites concourantes devient une famille de points alignés. On exprime $(*)$ en disant que δ_1 *préserve les relations d'incidence.*

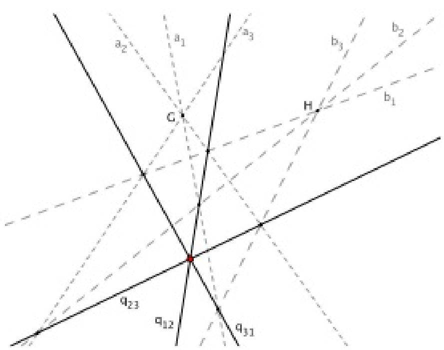
Plus généralement, soit $m \in [1, n]$ un entier. La loi qui, à un sous-espace projectif $\mathbf{W} = \mathbb{P}(W)$ de $\mathbb{P}(E)$ de codimension m, attache l'ensemble $\mathbf{w} = \mathbb{P}(\{L \in E^*, W \subset KerL\})$ des classes dans $\mathbb{P}(E^*)$ de formes linéaire non nulles s'annulant sur $\mathbb{P}(H)$, s'interprète comme une bijection

$$\delta_m : H^m(\mathbb{P}(E)) \to H_{m-1}(\mathbb{P}(E^*)) : \mathbf{W} \leftrightarrow \mathbf{w}$$

de l'ensemble $H^m(\mathbb{P}(E)$ des sous-espaces projectifs de codimension m de $\mathbb{P}(E)$ vers l'ensemble $H_{m-1}(\mathbb{P}(E^*))$ des sous-espaces projectifs de dimension $m-1$ de $\mathbb{P}(E^*)$. La collection δ des applications $\delta_m, m = 1, ..., n$, s'appelle la *dualité projective*. Elle renverse les inclusions ($\mathbf{W} \supset \mathbf{W'} \Rightarrow \mathbf{w} \subset \mathbf{w'}$), et préserve donc les relations d'incidence. Par exemple, quand $\mathbb{P}(E)$ est un plan, δ_2 envoie des points alignés de $\mathbb{P}(E)$ sur des droites concourantes de $\mathbb{P}(E^*)$. On voit d'ailleurs, en identifiant E à son bidual E^{**}, que $\delta_{m,E\to E^*}$ admet $\delta_{n-m+1,E^*\to E^{**}}$ pour application réciproque.

Puisque la dualité projective préserve les relations d'incidence, tout théorème Th dont les hypothèses et les conclusions sont des relations d'incidence dans $\mathbb{P}(E)$ fournira par dualité un théorème d'incidence Th^* dans $\mathbb{P}(E^*)$, et donc dans $\mathbb{P}(E)$ puisque E^* et E sont (non canoniquement) isomorphes. On n'a pas toujours de la chance (voir (i) ci-dessous), mais l'énoncé auquel on aboutit ainsi est en général nouveau !

Exercice : justifier, puis traduire, le texte suivant (cours de U. Stuhler, Göttingen).



Der

Brianchonsche

Satz

i) Der Satz von Desargues ist selbstdual.

ii) Die Dualisierung des Satzes von Pappos führt uns in $\mathbb{P}_2(K)$ auf den

Satz von Brianchon : *Für jede Auswahl von paarweise verschiedenen Geraden* $a_1, a_2, a_3, b_1, b_2, b_3$ *gilt : verlaufen* a_1, a_2, a_3 *durch einen gemeinsamen Punkt* G *und* b_1, b_2, b_3 *durch einen gemeinsamen Punkt* H, *wobei* $a_i \neq (GH) \neq b_i$ $(i = 1, 2, 3)$, *so verlaufen die Geraden* $q_{12} = <(a_1 \cap b_2), (b_1 \cap a_2)>$ *(d.h. die projektive Gerade durch die Punkte* $a_1 \cap b_2$ *und* $b_1 \cap a_2$*),* $q_{23} = <(a_2 \cap b_3), (b_2 \cap a_3)>$, $q_{31} = <(a_3 \cap b_1), (b_3 \cap a_1)>$ *durch einen gemeinsamen Punkt.*

La notion de tangence

Les procédés de (dés)homogénéisation vus au §2.2 s'étendent aux "variétés algébriques" de degré quelconque. Ainsi, un polynôme homogène $F(X_0, ..., X_n)$ de degré $d \geq 1$ vérifie l'identité $F(\lambda X_0, ..., \lambda X_n) \equiv \lambda^d F(X_0, ..., X_n)$, de sorte que l'hypersurface projective

$$Z(F) := \{P = (X_0 : ... : X_n) \in \mathbb{P}_n, F(X_0, ..., X_n) = 0\}$$

formée par les zéros[3] de F dans $\mathbb{P}_n$ est bien définie (comme dans le cas linéaire, on ne peut pas parler de la valeur de F en P, mais l'expression $F(P) = 0$ a un sens). Supposons que $F \neq a_0 X_0^d$. Alors, $Z(F)$ induit sur l'ouvert affine $\mathcal{H}_0 \simeq K^n$ une hypersurface affine $\mathcal{Z}(f)$ d'équation $f(x_1, ..., x_n) = 0$, où f est le polynôme (en général non homogène) non constant de degré total $\delta \leq d$ en n variables défini par $f(\frac{X_1}{X_0}, ..., \frac{X_n}{X_0}) := \frac{1}{X_0^d} F(X_0, ..., X_n)$. Le complémentaire de $\mathcal{Z}(f)$ dans $Z(F)$ est formé par les points d'intersection de $Z(F)$ avec l'hyperplan à l'infini $\mathbb{P}(H_0) \simeq \mathbb{P}_{n-1}(K)$: ce sont les zéros du polynôme en n variables $\Phi(X_1, ..., X_n) := F(0, X_1, ..., X_n)$, qui est non nul (et homogène de degré d) si $Z(F) \not\supset \mathbb{P}(H_0)$, c-à-d. si X_0 ne divise pas F dans l'anneau $K[X_0, ..., X_n]$.

Inversement, soit $f(x_1, ..., x_n)$ un polynôme de degré total $\delta \geq 1$ en n variables. Son lieu des zéros $\mathcal{Z}(f)$ dans l'espace affine $\mathcal{E} = K^n$ est une hypersurface affine. On appelle projectivisée de $\mathcal{Z}(f)$ l'hypersurface $\hat{\mathbb{P}}(\mathcal{Z}(f)) := Z(F)$ de $\mathbb{P}(\hat{\mathcal{E}}) = \mathbb{P}_n(K)$ définie par le polynôme homogène $F(X_0, ..., X_n) = X_0^\delta f(\frac{X_1}{X_0}), .., \frac{X_n}{X_0})$, de degré δ en $n+1$ variables (et jamais divisible par X_0). On obtient $Z(F)$ en adjoignant à $\mathcal{Z}(f)$ ses "points à l'infini", définis comme supra par l'annulation du polynôme Φ, soit : $Z(F) = \mathcal{Z}(f) \cup Z(\Phi)$.

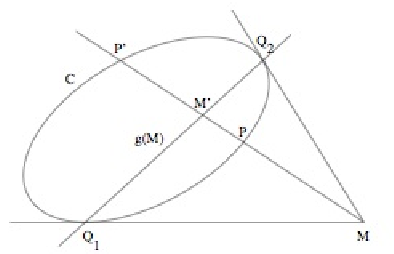
Les hypersurfaces de degré 2 s'appellent des quadriques (des coniques si $n = 2$). Elles font l'objet du chapitre III, §2. Pour $K = \mathbb{R}$ et $n = 2$, les polynômes homogènes F de degré 2 de $\mathbb{R}[X_0, X_1, X_2]$ sont

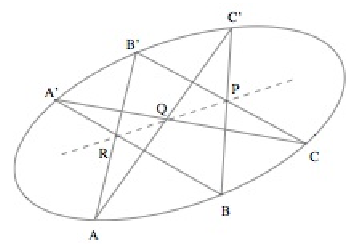
[3] Cette expression cache une subtilité importante. Nous considérons les zéros de F (c-à-d. les points d'annulation de F) non seulement dans $\mathbb{P}_n(K)$, mais aussi dans $\mathbb{P}_n(K')$ pour tout corps commutatif K' contenant K, et nous tenons compte de leurs multiplicités. Ainsi pour $K = \mathbb{R}$, les polynômes $F_1(X_0, X_1, X_2) = X_0^2 + X_1^2 + X_2^2$ et $F_2(X_0, X_1, X_2) = X_0^2 + X_1^2 + 2X_2^2$ ne définissent pas les mêmes hypersurfaces : bien que dans $\mathbb{P}_2(\mathbb{R})$, $Z(F_1) = Z(F_2)$ soient tous deux l'ensemble vide, $Z(F_1) \neq Z(F_2)$ dans $\mathbb{P}_2(\mathbb{C})$. De même, $F_1(X_0, X_1 X_2) = X_1$ et $F_2(X_0, X_1 X_2) = X_1^2$, bien qu'ayant le même lieu de zéros dans $\mathbb{P}_2(K)$ (l'hyperplan H_1), définissent des hypersurfaces $Z(F_1), Z(F_2)$ distinctes : les zéros de F_2 sont doubles, et on dit que $Z(F_2)$ est l'hyperplan H_1, compté deux fois. On peut aussi les distinguer en considérant les zéros de F_1 et F_2 à coordonnées dans la K-algèbre $\tilde{K} = K[t]/(t^2)$.

- ou bien irréductibles dans l'anneau $\mathbb{C}[X_0, X_1, X_2]$;

- ou bien produit de deux formes linéaires (éventuellement complexes).

Dans ce deuxième cas, $Z(F)$ est la réunion de deux droites (éventuellement confondues, ou complexes), et on dit que la conique $Z(F)$ est dégénérée. On verra au chapitre III les propriétés suivantes :





$[P, P', M, M'] = -1$

P, Q, R sont alignés (Théorème de Pascal)

Quand $Z(F) = \Delta \cup \Delta'$ est une conique dégénérée, la figure de droite redonne le théorème de Pappus. Nous donnons maintenant une définition formelle des "tangentes" qui interviennent dans la figure de gauche.

Soit $\mathcal{Z} = \mathcal{Z}(f)$ une hypersurface affine de l'espace affine K^n. Le théorème des fonctions implicites conduit à dire qu'un point $P = (\gamma_1, ..., \gamma_n)$ de $\mathcal{Z}(f)$ est *lisse* si l'une au moins des dérivées partielles de f en P est non nulle : $\exists i \in [1, ..., n], f'_{x_i}(P) := \frac{\partial f}{\partial x_i}(\gamma_1, ..., \gamma_n) \neq 0$. L'équation

$$\Sigma_{i=1,\ldots,n} f'_{x_i}(P)(x_i - \gamma_i) = 0$$

définit alors un hyperplan affine $T_P(\mathcal{Z})$, qu'on appelle *l'hyperplan tangent à $\mathcal{Z}$ en P.*

Soit maintenant $Z := Z(F) \subset \mathbb{P}_n(K)$ une hypersurface projective de degré d. Pour tout $j = 0, ..., n$, son intersection avec l'ouvert affine $\mathcal{H}_j : X_j \neq 0$, si elle est non vide, est l'hypersurface affine $\mathcal{Z}(f^{[j]})$, où $f^{[j]}(\frac{X_0}{X_j}, ..., \frac{X_n}{X_j}) := \frac{1}{X_j^d} F(X_0, ..., X_n)$ est le polynôme en n variables déshomogénéisé de F par rapport à X_j. Les dérivées partielles $F'_{X_i} = \frac{\partial F}{\partial X_i}, i = 0, ..., n$, de F, sont des polynômes homogènes de degré $d - 1$, et on vérifie, en considérant chacun des monômes de F, que pour tout $i \neq j$,

$$f^{[j]\prime}_{x_i}(\frac{X_0}{X_j}, ..., \frac{X_n}{X_j}) = \frac{1}{X_j^{d-1}} F'_{X_i}(X_0, ..., X_n).$$

En prenant par exemple $j = 0$, et en posant $f = f^{[0]}, \mathcal{Z} = \mathcal{Z}(f^{[0]})$, on a :

Proposition II.2.7. *Soient $Z = Z(F)$ une hypersurface de $\mathbb{P}_n(K)$, et $P = (c_0 : ... : c_n)$ un point de l'hypersurface affine $\mathcal{Z} = Z \cap \mathcal{H}_0$. Alors,*

i) P est lisse sur $\mathcal{Z}$ si et seulement si $F'_{X_i}(P) \neq 0$ pour un au moins des indices $i = 0, ..., n$;

ii) dans ce cas, le projectivisé $\hat{\mathbb{P}}(T_P(\mathcal{Z}))$ de l'hyperplan tangent à $\mathcal{Z}$ en P admet pour équation homogène :

$$\sum_{i=0,...,n} F'_{X_i}(c_0, ..., c_n)\ X_i = 0$$

Preuve : soit d le degré du polynôme homogène F. Pour donner sens à cet énoncé, on doit d'abord noter que $Z \cap \mathcal{H}_0$ étant par hypothèse non vide, c'est bien une hypersurface, donnée par $\mathcal{Z} = \mathcal{Z}(f^{[0]})$, puis, pour (i), que comme les dérivées partielles F'_{X_i} sont des polynômes homogènes, les expressions $F'_{X_i}(P) \neq 0$ sont licites, et enfin, pour (ii), que comme elles sont toutes de degré $d-1$, le $(n+1)$-uplet $(F'_{X_i}(c_0, ..., c_n) = \frac{1}{\lambda^{d-1}} F'_{X_i}(\lambda c_0, ..., \lambda c_n), i = 0, ..., n)$ définit un hyperplan indépendant du choix des coordonnées homogènes de P. Rappelons par ailleurs la *relation d'Euler* : pour tout polynôme homogène F de degré d,

$$\sum_{i=0,...,n} F'_{X_i}(X_0, ..., X_n).X_i = d\ F(X_0, ..., X_n),$$

qui montre déjà que cet hyperplan passe bien par le point $P \in Z(F)$.

i) Puisque $P \in \mathcal{H}_0$ on peut le repérer par les coordonnées $(1 : \gamma_1, ..., \gamma_n)$, où $\gamma_i = \frac{c_i}{c_0}$. Alors, P est lisse sur $\mathcal{Z}$ si et slt s'il existe $i \in [1, ..., n]$ tel que $0 \neq f'_{x_i}(\gamma_1, ..., \gamma_n)$. Mais cette expression vaut $\frac{1}{c_0^{d-1}} F'_{X_i}(c_0, ..., c_n)$. Il reste à noter que si $F'_{X_i}(P) = 0$ pour tout $i = 1, .., n$, la relation d'Euler et l'hypothèse $c_0 \neq 0$ entraîne qu'on a aussi $F'_{X_0}(P) = 0$.

ii) L'équation du projectivisé du plan tangent à $\mathcal{Z}$ en P, d'équation affine $\Sigma_{i=1,...,n} f'_{x_i}(\gamma_1, ..., \gamma_n)(x_i - \gamma_i) = 0$, en est l'homogénéisée. Après multiplication par c_0^{d-1}, elle s'écrit : $\Sigma_{i=1,...,n} F'_{X_i}(c_0, ..., c_n)(X_i - \frac{c_i}{c_0} X_0) = 0$. D'après Euler, $\Sigma_{i=1,...,n} F'_{X_i}(c_0, ..., c_n) c_i = -F'_{X_0}(c_0, ..., n) c_0$, d'où l'équation recherchée.

Soit enfin $P = (c_0 : ... : c_n)$ un point quelconque de $Z(F)$. Il appartient à un au moins des ouverts affines $\mathcal{H}_j, j = 0, ..., n$, puisque ceux-ci recouvrent $\mathbb{P}_n(K)$. Si j_1 et j_2 sont deux tels indices, c-à-d. si P appartient à $\mathcal{Z}(f^{[j_1]})$ et à $\mathcal{Z}(f^{[j_2]})$, la proposition précédente, montre que :

i) P est lisse sur $\mathcal{Z}(f^{[j_1]})$ si et seulement s'il est lisse sur $\mathcal{Z}(f^{[j_2]})$: cela équivaut à la condition

$$\exists i \in [0, ..., n], F'_{X_i}(P) \neq 0;$$

ii) alors, les hyperplans tangents à $\mathcal{Z}(f^{[j_1]})$ et à $\mathcal{Z}(f^{[j_2]})$ en P sont les intersections avec $\mathcal{H}_{j_1}$ et $\mathcal{H}_{j_2}$ du même hyperplan de $\mathbb{P}_n(K)$

$$T_P(Z) : \sum_{i=0,...,n} F'_{X_i}(c_0, ..., c_n)\ X_i = 0.$$

Dans ces conditions, on dit que P est *lisse* sur l'hypersurface projective $Z = Z(F)$ s'il vérifie la condition (i), et on appelle *hyperplan tangent à Z en P* l'hyperplan projectif $T_P(Z)$ défini par (ii).

Exemple : soit $F = a_0 X_0^2 + a_1 X_1^2 + a_2 X_2^2 \in \mathbb{R}[X_0, X_1, X_2]$. Si $Z(F)$ est une conique non dégénérée ($\Leftrightarrow a_0 a_1 a_2 \neq 0$), tous les points P de $Z(F)$ sont lisses, et pour tout $M \notin Z(F)$, le faisceau de droites de centre M contient exactement deux tangentes à $Z(F)$ (éventuellement complexes).

Chapitre III

Géométrie quadratique

Université de Paris VI
D. Bertrand

Master M1 2010-2011
Algèbre géométrique

CHAPITRE III

GÉOMÉTRIE QUADRATIQUE

Dans tout ce chapitre, K est un corps commutatif de caractéristique $\neq 2$, et les K-espaces vectoriels sont de dimension finie. On suppose parfois K muni d'une involution $\sigma \neq id_K$. Alors, $K = K_0 \oplus K_0\omega$, où $\sigma(\omega) = -\omega$ et $K_0 = \{\lambda \in K, \sigma(\lambda) = \lambda\}$. Exemples : $K = \mathbb{C}, \sigma =$ la conjugaison complexe, $K_0 = \mathbb{R}$; ou $K = \mathbb{F}_5[T]/(T^2-2) := \mathbb{F}_{25}, \sigma(\lambda) = \lambda^5, \omega =$ classe de T, $K_0 = \mathbb{F}_5$.

III.1 Formes bilinéaires.

III.1.1 Vocabulaire

Soient E, E' deux espaces vectoriels sur K. Une application $b : E \times E \to E'$ est dite bilinéaire (resp. sesquilinéaire) si pour tout $y \in E$, l'application $x \mapsto b(x,y)$ est linéaire et pour tout $x \in E$, l'application $y \mapsto b(x,y)$ est linéaire (resp. σ-linéaire : $\forall \lambda \in K, b(x, \lambda y) = \sigma(\lambda) b(x,y)$). Lorsque $E' = K$, on parle de forme bi- resp. sesqui-linéaire.

Soient b une forme sesquilinéaire, et $\mathcal{B} = \{e_1, ..., e_n\}, n = dimE$, une base de E, dans laquelle x, y sont représentés par des vecteurs colonnes $X = {}^t(x_1, ..., x_n), Y$. Alors, $b(x,y) = {}^tXB\sigma(Y) \in K$, où $B = (b_{i,j} = b(e_i, e_j), 1 \leq i,j \leq n)$ est la matrice représentative de b dans la base $\mathcal{B}$. Si $\mathcal{B}' = (e_1, ..., e_n).P, P \in GL_n(K)$, désigne une autre base, b est représentée dans $\mathcal{B}'$ par la matrice $B' = {}^tPB\sigma(P)$. Idem (sans les σ) pour une forme bilinéaire : dans ce cas, on a $det(B') = det(B)(det(P))^2$, et la classe de $det(B)$ dans $K/(K^*)^2$ ne dépend pas de la base choisie ; on l'appelle le *discriminant* $discr(b)$ de la forme bilinéaire b.

Une forme bilinéaire b fournit deux applications linéaires naturelles de E vers son dual E^* :

$$\phi_b : E \to E^* : x \mapsto \{\phi_b(x) : y \mapsto b(x,y)\},$$

et $\phi_b^t : x \mapsto \{\phi_b^t(x) : y \mapsto b(y,x)\}$, qui est la transposée de ϕ_b. Elles sont données, relativement à la base $\mathcal{B}$ et à sa base duale $\mathcal{B}^* = (e_1^*, ..., e_n^*)$, par les matrices ${}^tB$ et B. Ces applications (ou matrices) ont le même rang, qu'on appelle *le rang de* b. On dit que *la forme* b *est non dégénérée si* $rang(b) = n$, c-à-d. si $det(B) \neq 0$, c-à-d. si ϕ_b (ou, de façon équivalente, ϕ_b^t) est un isomorphisme[1]. Les noyaux de ϕ_b et de ϕ_b^t ont même dimension, égale à $n - rang(b)$. Sous les hypothèses qu'on va maintenant faire sur b, ils coïncident, et s'appellent alors *le noyau de* b.

On suppose désormais que l'une des conditions suivantes est satisfaite :

- b est *bilinéaire symétrique*[2] : $\forall x, y \in E, b(x,y) = b(y,x)$, c-à-d. : $\phi_b = \phi_b^t$;
- b est *bilinéaire antisymétrique* : $\forall x, y, b(x,y) = -b(y,x)$, c-à-d. : $\phi_b = -\phi_b^t$;
- b est *hermitienne* : b sesquilin. et $\forall x, y, b(x,y) = \sigma(b(y,x))$: $\phi_b = \sigma \circ \phi_b^t$.

ce qui, en termes matriciels, se traduit par ${}^tB = B$ (matrice symétrique), ${}^tB = -B$ (matrice antisymétrique) ou ${}^tB = \sigma(B)$ (matrice hermitienne).

Dans ces conditions, on dit que deux parties X, Y de E sont*orthogonales* relativement à b, notation : $X \perp Y$, si $\forall x \in X, y \in Y, b(x,y) = 0$. Alors $Y \perp X$. L'orthogonal $E^\perp$ de E est le noyau $Ker(\phi_b) = Ker(\phi_b^t)$ de b. On a donc $dimE^\perp + rang(b) = dimE$, et b est non dégénérée si et slt si $E^\perp = \{0\}$. Pour toute partie S de E, l'ensemble $S^\perp = \{y \in E, \forall x \in S, x \perp y\}$ est un sous-espace vectoriel de E, qui vérifie $S \subset (S^\perp)^\perp$. Une somme directe directe $E = E_1 \oplus ... \oplus E_k$ est dite orthogonale si $E_i \perp E_j$ pour tout $i \neq j$.

Si b est une forme bilinéaire symétrique ou hermitenne, on appelle *forme quadratique* ou quadratique hermitienne associée à b l'application $q = q_b : E \to K : x \mapsto q(x) := b(x,x)$. Les formules $b(x,y) = \frac{1}{4}(q(x+y) - q(x-y))$ (cas bilinéaire symétrique) et $b(x,y) = \frac{1}{4}(q(x+y) - q(x-y)) + \frac{1}{4\omega}(q(\omega x + y) - q(\omega x - y))$ (cas hermitien) permettent de retrouver b à partir de q. On dit que b est la forme polaire de q, et on appelle rang de q celui de b. Dans le cas bilinéaire symétrique, la formule $q(x+h) = q(x) + 2b(x,h) + q(h)$ permet d'interpréter $2b(x,.) = 2\phi_b(x)$ comme la differentielle de q en x.

On dit qu'une forme bilinéaire b est alternée si $\forall x \in E, b(x,x) = 0$. Comme $carK \neq 2$, cela équivaut à dire que b est antisymétrique (tandis que

[1] On pourra donner des énoncé similaires pour b sesquilinéaire, en introduisant le K-espace vectoriel $E^{*,\sigma}$ des formes σ-linéaires sur E. Par ailleurs, tout élément ϕ de $\mathcal{L}(E, E^*)$ fournit deux formes bilinéaires $b_\phi(x,y) = \phi(x)(y), b_{\phi^t}(x,y) = \phi(y)(x)$, et on peut ainsi retrouver b à partir de ϕ_b ou de ϕ_b^t. Quelle est la version sesquilinéaire ?

[2] En anglais : *symmetric*. Mais un seul m en français.

pour $car K = 2$, b est antisymétrique si et slt si elle est symétrique).

Un vecteur non nul $x \in E$ est dit *isotrope* si $q(x) = 0$. La droite qu'il porte est alors formée de vecteurs isotropes. L'ensemble $C(q) = \{x \in E, q(x) = 0\}$ s'appelle le *cône isotrope* de q. Un sous-espace vectoriel F de E est dit *totalement isotrope* si $\forall x, y \in F, b(x, y) = 0$, c-à-d. si $F \perp F$, ou encore $F \subset F^{\perp}$. Par exemple, $E^{\perp}$ (ou, plus généralement, dans le cas symétrique ou hermitien, tout sous-espace vectoriel F de E contenu dans $C(q)$) est totalement isotrope. *L'indice* $\nu(b)$ de b (ou : $\nu(q)$ de q dans le cas symétrique ou hermitien) est la dimension maximale des sous-espaces totalement isotropes de E.

Proposition III.1.1. *: supposons que b soit* non-dégénérée. *Pour tous sous-espaces vectoriels F, G de E, on a :*

i) $dimF + dimF^{\perp} = dimE$; en particulier, $\nu(b) \leq \frac{1}{2} dimE$.

ii) $(F^{\perp})^{\perp} = F$

iii) $(F + G)^{\perp} = F^{\perp} \cap G^{\perp}$ et $(F \cap G)^{\perp} = F^{\perp} + G^{\perp}$;

iv) si $F \cap F^{\perp} = \{0\}$, alors $E = F \perp F^{\perp}$.

Preuve : Pour i), soient B une matrice représentative de b et $\{f_1, ..., f_k\}$ des vecteurs de K^n représentant une base de F. Alors, $F^{\perp}$ s'identifie à l'ensemble de solutions $y \in K^n$ du système $\{\sigma({}^t f_i.B).y = 0, i = 1, ..., k\}$. Comme B est inversible, ce système est de rang $k = dimF$. Par conséquent, $codimF^{\perp} = k$. Les autres énoncés sont des exercices.

Corollary III.1.2. *(On ne suppose plus b non dégénérée.) Soit F un sous-espace vectoriel supplémentaire de $E^{\perp}$ dans E. Alors, $E = F \perp E^{\perp}$, et la restriction $b_{|F}$ de b à F est non-dégénérée..*

Preuve : la somme directe $E = F \oplus E^{\perp}$ est orthogonale puisque $E^{\perp} \perp E$. Dans une base de E respectant cette décomposition, la matrice représentative de b est donnée par $B = \begin{pmatrix} B_1 & 0 \\ 0 & 0 \end{pmatrix}$, où B_1 représente $b_{|F}$. Alors, $rg(b_{|F}) = rg(B_1) = rg(B) = dimE - dimE^{\perp} = dimF$, donc $b_{|F}$ est non-dégénérée.

III.1.2 "Diagonalisation"

Soient q, q' deux formes quadratiques (resp. quadratiques hermitiennes) sur E, de formes polaires b, b'. On dit que q et q' sont *équivalentes* s'il existe un automorphisme $u \in GL(E)$ tel que $q' = q \circ u$, ou de façon équivalente : $\forall x, y, b'(x,y) = b(u(x), u(y))$. En termes des matrices représentatives B, B', U de b, b', u dans une base donnée de E, cela revient à demander que $B' = {}^tUBU$ (resp. $B' = {}^tUB\sigma(U)$). On a une notion similaire d'équivalence pour les couples b, b' de formes alternées.

Proposition III.1.3. *i) Soit q une forme quadratique (resp. quadratique hermitienne) sur E, de rang r. Il existe des éléments $d_1, ..., d_r$ de K^* (resp. de K_0^*) tels que q est équivalente à la forme $q'(x) = d_1x_1^2 + ... + d_rx_r^2$ (resp. $d_1x_1\sigma(x_1)+...+d_rx_r\sigma(x_r)$). Autrement dit, E admet une* base orthogonale *(c-à-d. formée de vecteurs deux à deux orthogonaux), et dont les $n-r$ derniers forment une base de $E^\perp$. Ou encore, si $B = {}^tB$ (resp. $\sigma({}^tB)$) désigne une matrice représentative de q, il existe $U \in GL_n(K)$ telle que ${}^tUBU$ (resp. ${}^tUB\sigma(U)$) $= \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_r & 0 \\ 0 & 0 & 0 \end{pmatrix}$.*

ii) Soit b une forme bilinéaire alternée sur E, de matrice représentative $B = -{}^tB$. Alors, b est equivalente à $b'(x,y) = \Sigma_{i=1,..,m} x_iy_{m+i} - y_ix_{m+i}$. Autrement dit, il existe $U \in GL_n(K)$ telle que ${}^tUBU = \begin{pmatrix} 0 & \mathbf{I}_m & 0 \\ -\mathbf{I}_m & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. En particulier, le rang $r = 2m$ de b est pair.

Preuve i) Une preuve algorithmique a été vue en L3. En voici une autre. Le corollaire précédent permet de se ramener au cas où b est non dégénérée. Pour $e_1 \in E$ non isotrope, $E_1 = \{e_1\}^\perp$ est un hyperplan ne contenant pas e_1, donc $E = K.e_1 \perp E_1$, et la restriction de b à E_1 est non dégénérée. On itère.

ii) On se ramène de nouveau au cas non-dégénéré. Il existe alors deux vecteurs e_1, ϵ_1 tels que $b(e_1, \epsilon_1) = 1$. Ils sont forcément linéairement indépendants, et $b(\epsilon_1, e_1) = -1$, de sorte que la restriction de b au plan P_1 qu'ils engendrent, de matrice représentative $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, est non-dégénérée. On en déduit que $P_1 \cap P_1^\perp = \{0\}$, donc $E = P_1 \perp P_1^\perp$, et que la restriction de b à $E_1 = P_1^\perp$ est non-dégénérée. On peut donc itérer.

Remarques : i) Bien que le résulat de (i) soient souvent décrit comme une "diagonalisation" de la forme q, *il ne fournit pas une diagonalisation de la matrice* B, ce qui reviendrait à trouver une matrice $U \in GL_n(K)$ telle que $U^{-1}BU$ soit diagonale ; les scalaires d_i de cet énoncé ne sont d'ailleurs pas uniques, et n'ont donc pas de rapport avec les valeurs propres de B. Bref, on prendra garde à ne pas confondre (i) avec le "théorème spectral", en vertu duquel les matrices symétriques réelles (resp. hermitiennes complexes) sont effectivement diagonalisables - et qui plus est, par une matrice $U = {}^tU^{-1}$ orthogonale (resp. $U = \sigma({}^tU^{-1})$ unitaire).

ii) Un *espace symplectique* est un couple (E, b) où b est une forme alternée non dégénérée. Tout plan symplectique est équivalent à l'espace vectoriel K^2, muni de la forme alternée $b(x, y) = x_1y_2 - x_2y_1 = det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}$, et le resultat de (ii) exprime que E s'écrit comme une somme directe orthogonale de $E^\perp$ et de plans symplectiques. ll entraîne par ailleurs que tout espace symplectique (E, b) est de dimension $n = 2m$ paire, où $m = \nu(b)$, et admet une décomposition en somme directe (non orthogonale) de deux sous-espaces $W_1 =< e_1, ..., e_m >$, $W_2 =< \epsilon_1, ..., \epsilon_m >$ totalement isotropes de dimension $n/2$. De tels sous-espaces s'appellent des lagrangiens de (E, b), et les bases correspondantes $\{e_1, \epsilon_1, ..., e_m, \epsilon_m\}$ des bases symplectiques.

iii) Ne pas confondre (ii) avec les énoncés suivants. Un *espace hyperbolique* est un couple (E, b), où E est de dimension $n = 2m$ paire, et b est une forme bilinéaire symétrique non dégénéree d'indice $\nu(b) = m$. Tout plan hyperbolique (E, b) est équivalent à l'e-v K^2, muni de la forme bilinéaire de matrice représentative $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ (ou de façon équivalente $B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$). Tout espace hyperbolique est somme directe orthogonale de plans hyperboliques, et admet donc une décomposition en somme directe (non orthogonale) de deux sous-espaces W_1, W_2 totalement isotropes de dimension $n/2$. De tels sous-espaces s'appellent aussi des lagrangiens de (E, b).

III.2 Quadriques

Soient $\mathbb{P}(E)$ un espace projectif sur K de dimension n, et $q \neq 0$ une forme quadratique sur E, de forme polaire b. Dans une base de E, q s'exprime comme un polynôme homogène $Q(X_0, ..., X_n) = {}^tXBX$ de degré 2 en $n+1$ variables, et l'ensemble $Z(Q) := \Gamma$ des zéros de Q (au sens de la Note [(3)] du chapitre 2, §2.3) est bien défini. On l'appelle la quadrique définie par q. Ainsi, les points de Γ à coordonnées dans K correspondent exactement aux droites vectorielles du cône isotrope $C(q)$ de q.

Les multiples non nuls $\lambda q, \lambda \in K^*$, de q définissent la même quadrique, de sorte que la donnée de Γ équivaut à celle d'un point de l'espace projectif $\mathbb{P}(\mathcal{Q}(E))$, où $\mathcal{Q}(E)$ désigne le K-espace vectoriel, de dimension $\frac{(n+1)(n+2)}{2}$, formé par les formes quadratiques sur E (on retrouvera cet espace vectoriel sous le nom de code $Sym^2(E^*)$ au chapitre 4). Un changement de base fournit une nouvelle expression $Q'(X_0, ..., X_n)$ de q, où Q et Q' sont des formes quadratiques sur K^{n+1} équivalentes, et on dit que les quadriques correspondantes de $\mathbb{P}_n(K)$ sont équivalentes. Une classe d'équivalence de quadriques de $\mathbb{P}_n(K)$ correspond donc à une "classe d'équivalence projective" de matrices symétriques B , où $B \sim B'$ signifie : $\exists \lambda \in K^*, U \in GL_{n+1}(K), B' = \lambda {}^tUBU$.

Hyperplan tangent et tangentes

Soit $P \in \mathbb{P}(E)$ un point de la quadrique Γ définie par q. En coordonnées, si $P = (\gamma_0 : ... : \gamma_n)$ est représenté par un vecteur $\gamma = (\gamma_0, ..., \gamma_n)$ de E, $Q(\gamma_0, ..., \gamma_n) = 0$. Pour tout $i = 0, ..., n$, les dérivées partielles de Q s'écrivent : $Q'_{X_i}(\gamma_0, ..., \gamma_n) = 2b(\gamma, e_i)$, où $e_i = (0, .., 1_{(i)}, ...0)$. Donc P est lisse sur Γ (au sens du chapitre 2, proposition 2.7) si et seulement si ses représentants γ n'appartiennent pas au noyau de b, et pour un tel point P, l'hyperplan tangent à Γ en P admet pour équation

$$T_P(\Gamma) : b(\gamma, X) = 0.$$

Ainsi, $\mathbb{P}(E^{\perp_{(b)}}) \subset \Gamma$ est l'ensemble des points non lisses (aussi appelés : singuliers) de Γ. On dit que la quadrique Γ est *non-singulière* (ou : *lisse*) si la forme bilinéaire b est non-dégénérée ; tous ses points sont alors lisses.

Soient P, P' deux points distincts de $\mathbb{P}(E)$, représentés par des vecteurs γ, γ', et $\Delta = \mathbb{P}(K\gamma \oplus K\gamma') =< P, P' >$ la droite projective qu'ils engendrent. En dehors de P', les points d'intersections de Γ et de Δ correspondent aux solutions en t de l'équation

$$Q(\gamma + t\gamma') := t^2 Q(\gamma') + 2tb(\gamma, \gamma') + Q(\gamma) = 0.$$

Supposons que P soit un point de Γ, et qu'il soit lisse. Alors, $b(\gamma, \gamma') = 0$ si et slt si la droite Δ est contenue dans l'hyperplan tangent $T_P(\Gamma)$; on dit alors que Δ *est tangente à* Γ *en* P. Dans ce cas, ou bien Δ est toute entière contenue dans Γ ($\Leftrightarrow P' \in \Gamma$), ou bien $\Delta \cap \Gamma$ est réduit à P, compté deux fois (car $t = 0$ est alors une racine double de l'équation). Ainsi, si Γ est une quadrique non singulière, une droite Δ de $\mathbb{P}(E)$ non contenue dans Γ rencontre Γ :

- soit en deux points distincts, et elle n'y est alors pas tangente à Γ ;

- soit en un seul point, et elle y est alors tangente à Γ ;

- soit en aucun point ; ce dernier cas ne peut d'ailleurs pas se produire si $K = \mathbb{C}$ (plus généralement si K est algébriquement clos), l'équation précédente ayant forcément une solution si $P' \notin \Gamma$.

Si maintenant P est un point singulier de Γ, alors ou bien Δ est toute entière contenue dans Γ, ou bien elle ne rencontre (deux fois) Γ qu'au point P. En particulier, toute quadrique singulière non réduite à un point contient au moins une droite. Mais la réciproque est fausse : pour $n \geq 3$, toutes les quadriques de $\mathbb{P}_n(\mathbb{C})$ contiennent des droites.

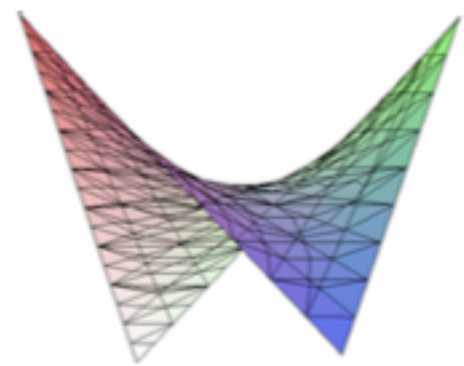
Classification

Décrivons tout d'abord les quadriques de $\mathbb{P}_1(K)$: une forme quadratique $q \neq 0$ en deux variables est projectivement équivalente à $X_0^2 + aX_1^2$, où $a \in K$ représente le discriminant $discr(q) \in K/(K^*)^2$ de q. Si les éléments de K ne sont pas tous des carrés, il y a donc trois classes d'équivalence de quadriques ponctuelles : un point double ($r = 1$), deux points distincts de $\mathbb{P}_1(K)$, ou un couple de points "conjugués" de $K(\sqrt{a})$.

Pour $K = \mathbb{C}$ ou $\mathbb{R}$, la proposition 1.3 fournit la classification suivante des quadriques projectives de $\mathbb{P}_n(K), n \geq 2$:

i) si $K = \mathbb{C}$ (ou plus généralement, si K est algébriquement clos) : la forme quadratique q est (projectivement ou pas) équivalente à $X_0^2 + ... + X_{r-1}^2$, où r est le rang de q. Les classes d'équivalence de quadriques sont donc en bijection avec les entiers $r \in [1, n+1]$. Comme $X_i^2 + X_j^2 = (X_i + \sqrt{-1}X_j)(X_i - \sqrt{-1}X_j)$, toute quadrique lisse de $\mathbb{P}_3(K)$ admet aussi pour équation $X_0X_1 - X_2X_3 = 0$, et contient donc les familles de droites $\Delta_\lambda : X_0 - \lambda X_2 = \lambda X_1 - X_3 = 0; \Delta_\mu : X_0 - \mu X_4 = \mu X_1 - X_2 = 0$. Ces droites correspondent à des couples de lagrangiens de la forme quadratique q.

ii) si $K = \mathbb{R}$: la forme quadratique q est équivalente à $\Sigma_{i=0,..,s-1}X_i^2 - \Sigma_{i=s,...,s+t-1}X_i^2$, où $s + t = r$ est le rang de q. Le couple (s, t) ne dépend que de la classe d'équivalence de q (Sylvester), et s'appelle la *signature* de q. Les classes d'équivalence de quadriques sont donc en bijection avec les couples non ordonnés $\{s, t\}$. L'indice $\nu(q)$ d'une forme quadratique non dégénérée q est égal à $min(s, t)$. Ainsi, la condition $\{s, t\} \neq \{0, n+1\}$ suffit à assurer que les quadriques associées ont des points réels. La quadrique de $\mathbb{P}_3(\mathbb{R})$ de signature $(2, 2)$, qui admet pour équation $X_0X_1 - X_2X_3 = 0$, contient comme supra deux familles de droites (réelles) ; on l'appelle paraboloïde hyperbolique.



Signalons que si K est un corps fini, on démontre qu'une forme quadratique q de rang r est équivalente à $X_0^2+...+X_{r-2}^2+aX_{r-1}^2$ où $a \in K^*$, et que toute quadrique de $\mathbb{P}_n(K), n \geq 2$, admet des points à coordonnées dans K.

Les quadriques de $\mathbb{P}_2(K)$ s'appellent des *coniques*. Pour $K = \mathbb{C}$, il y a trois types de coniques : une droite double ($r = 1, X_0^2 = 0$), la réunion de deux droites ($r = 2, X_0X_1 = 0$), une conique non singulière ($r = 3$). Pour $K = \mathbb{R}$, le cas $r = 2$ donne naissance à deux types : la réunion de deux droites réelles ($X_0X_1 = 0$), ou de deux droites complexes conjuguées ($X_0^2+X_1^2 = 0$) ; le cas $r = 3$ à deux types, suivant que la conique admet des points réels ($X_0^2 + X_1^2 - X_2^2 = 0$) ou pas ($X_0^2 + X_1^2 + X_2^2 = 0$).

Remarque : on dit parfois que la quadrique Γ est impropre si elle contient un hyperplan $\mathbf{H} = \mathbb{P}(KerL)$, où $0 \neq L \in E^*$. Cela équivaut à dire que Q est divisible par L dans $K[X_0, ...X_n]$, soit $Q = LL'$, et Γ est alors la réunion de deux hyperplans $\mathbf{H}, \mathbf{H}'$ distincts, ou l'hyperplan $\mathbf{H}$ compté deux fois. Une telle quadrique n'est pas lisse (tous les points de $\mathbf{H} \cap \mathbf{H}'$ sont singuliers), mais pas inversement : par exemple, la quadrique de $\mathbb{P}_3(K)$ d'équation $X_0^2+X_1^2-X_2^2 = 0$ n'est pas lisse, mais ne contient pas de plan. En revanche, comme on vient de le voir, une conique Γ de $\mathbb{P}_2(\mathbb{C})$ est propre (au sens : Γ n'est pas une droite double ou la reunion de deux droites) si et slt si elle est lisse (au sens : la forme quadratique q est non dégénérée).

Retour à la géométrie

Soit Γ une conique propre de $\mathbb{P}_2(K)$ admettant un point M défini sur K, et soit $\mathcal{F}_M$ le faisceau de droites de centre M. Rappelons que $\mathcal{F}_M$ s'identifie par dualité à une droite $M^* \simeq \mathbb{P}_1(K)$ de l'espace projectif dual. D'après le §2.1, l'application de $\mathcal{F}_M$ dans Γ qui, à une droite Δ de $\mathcal{F}_M$ associe le point $P = \Gamma \cap \Delta$ (avec $P = M$ si D est la tangente à Γ en P) établit une bijection ϕ_M de M^* sur l'ensemble des points de Γ définis sur K. En d'autres termes, ϕ_M fournit un *paramétrage* de Γ, dont la preuve qui suit donne une expression explicite dans un choix convenable de coordonnées.

En convenant de remplacer la droite (MP_i) par la tangente $T_{P_i}(\Gamma)$ si $M = P_i$, on a :

Théoreme III.2.1. ("Chasles-Steiner") *soient* Γ *une conique propre du plan projectif* $\mathbb{P}_2(K)$*, et* $P_1, ..., P_4$ *quatre points distincts de* Γ*. Pour tout point* M *de* Γ*, le birapport des quatre droites* MP_1, MP_2, MP_3, MP_4 *ne dépend pas de* M*. On l'appelle birapport* $[P_1, P_2, P_3, P_4]_\Gamma$ *des quatre points sur la conique.*

Preuve : soit $M' \neq M$ un autre point de Γ, et $\phi_{M'}$ le paramétrage de Γ correspondant. Nous allons montrer que $\phi_{M'}^{-1} \circ \phi_M : M^* \to M'^*$ est une homographie. Par ailleurs (exercice), le birapport de 4 droites d'un faisceau est égal à celui de leurs images par la dualité δ_1. Comme les homographies préservent le birapport, le théorème sera établi.

Choisissons comme repère projectif du plan les points $M = (0 : 1 : 0)$, $M' = (1 : 0 : 0)$, un troisième point $N' = (1 : 1 : 1)$ de Γ, et le point d'intersection $N = (0 : 0 : 1)$ des tangentes à Γ en M et M', et déterminons l'équation $a_0X_0^2 + a_1X_1^2 + a_2X_2^2 + 2b_0X_1X_2 + 2b_1X_0X_2 + 2b_2X_0X_1 = 0$ de Γ dans ce repère. La tangente $(MN) = \{X_0 = 0\}$ à Γ en M a pour équation : $a_1X_1 + b_2X_0 + b_0X_2 = 0$, donc $a_1 = b_0 = 0$. En considérant la tangente $(M'N) = \{X_1 = 0\}$, on voit de même que $a_0 = b_1 = 0$. De $N' \in \Gamma$, on déduit alors que l'équation de Γ est $X_0X_1 - X_2^2 = 0$.

Dans ce repère, une droite Δ du faisceau $\mathcal{F}_M$ a pour équation $\lambda X_0 + \mu X_2 = 0$, correspondant au point $(\lambda : \mu)$ de $\mathbb{P}_1(K) \simeq M^*$, et rencontre Γ aux points $M = (0 : 1 : 0)$ et $P = (\mu^2 : \lambda^2 : -\lambda\mu)$. Ainsi,

$$\phi_M\big((\lambda : \mu)\big) = (\mu^2 : \lambda^2 : -\lambda\mu),$$

ou en coordonnées affines, avec $t = -\frac{\mu}{\lambda}$: $\phi_M(t) = (t^2 : 1 : t)$. On trouve de même, en repérant les droites de $\mathcal{F}_{M'}$ par les équations $X_1 - t'X_2 = 0$, que $\phi_{M'}(t') = (1 : t'^2 : t')$, d'où $\phi_{M'}^{-1} \circ \phi_M(t) = 1/t$. Ainsi, $\phi_{M'}^{-1} \circ \phi_M$ est bien une homographie.

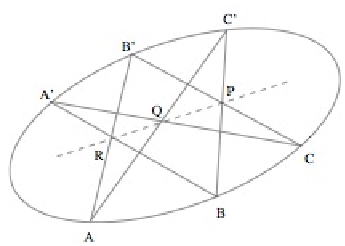
Nous sommes maintenant en mesure d'établir l'un des énoncés du chapitre II, p. 47 :

Exercice : Justifier, puis traduire, le texte suivant (cours de H. Umemura, Nagoya) :

Pascal の定理． 射影平面 $P_2(R)$ 内の非特異円錐曲線 Γ 上の相異なる 6 点 A, B, C, A', B', C' を考える．3 点 P, Q, R を次のように定める:

$$P = (BC') \cap (B'C), \quad Q = (CA') \cap (C'A), \quad R = (AB') \cap (A'B).$$

このとき、3 点 P, Q, R は同一直線上にある．



Indication : supposer par l'absurde que le point $P' = (RQ) \cap (B'C)$ soit différent de P. Posant $x = (AB') \cap (A'C), y = (AC') \cap (B'C)$, vérifier que

$$[A, R, x, B'] = [A'A, A'B, A'C, A'B'] = [C'A, C'B, C'C, C'B'] = [y, P, C, B'].$$

Puis considérer la perspective de centre Q de $(B'A)$ sur $(B'C)$, qui envoie A sur y, R sur P', x sur C et B' sur B'. En déduire que

$$[A, R, x, B'] = [y, P', C, B'],$$

et conclure.

Polarité.

Dans ce paragraphe, on fixe une forme quadratique non dégénérée q sur l'espace vectoriel E, de forme polaire b. Soient Γ la quadrique lisse correspondante, $\phi_b : E \to E^*$ l'isomorphisme de E vers son dual associé à b, et $\phi := \mathbb{P}(\phi_b) : \mathbb{P}(E) \to \mathbb{P}(E^*)$ la transformation projective que définit ϕ_b. Pour tout point M de $\mathbb{P}(E)$, le point $\phi(M)$ de $\mathbb{P}(E^*)$ fournit par dualité un hyperplan de E, qu'on appelle l'hyperplan *polaire* de M par rapport à Γ, et qu'on note $M^{\perp_\Gamma}$, ou plus simplement $M^\perp$. Si $M = \mathbb{P}(K.\gamma)$, où $\gamma \in E$, on a

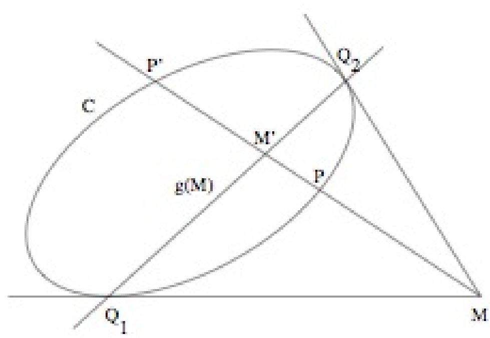
$$M^\perp = \{P' = \mathbb{P}(K.\gamma') \in \mathbb{P}(E), b(\gamma, \gamma') = 0.\}$$

Autrement dit, $M^\perp = \mathbb{P}((K.\gamma)^{\perp_b})$. De même, pour tout hyperplan $\mathbf{H} = \mathbb{P}(H)$ de $\mathbb{P}(E)$ (c'est-à -dire pour tout point de $\mathbb{P}(E^*)$), l'orthogonal $H^\perp$ de H dans E relativement à b est une droite de E, qui définit un point $\mathbf{H}^\perp = \mathbb{P}(H^\perp)$ de $\mathbb{P}(E)$, qu'on appelle le *pôle* de H par rapport à Γ. On a $(M^\perp)^\perp = M, (\mathbf{H}^\perp)^\perp = \mathbf{H}$.

Si $M \in \Gamma$, son hyperplan polaire, d'équation $b(\gamma, X) = 0$, est l'hyperplan tangent à Γ en M (et le pôle de $T_M(\Gamma)$ est le point M). Par conséquent, l'ensemble des hyperplans de $\mathbb{P}(E)$ tangents à Γ est l'image Γ^* de Γ dans $\mathbb{P}(E^*)$ par la transformation projective $\phi = \mathbb{P}(\phi_b)$. C'est donc une quadrique Γ^*, appelée *quadrique duale* de Γ, et l'on a $(\Gamma^*)^* = \Gamma$.

Proposition III.2.2. *i) Si dim$(E) = 2$, et si $\Gamma = \{P, P'\}$, deux points $M \notin \Gamma$ et M' de la droite $\mathbb{P}(E)$ sont pôles l'un de l'autre par rapport à Γ si et seulement si $[P, P', M, M'] = -1$.*

ii) Si dim$(E) = 3$, la droite polaire $M^\perp = g(M)$ d'un point $M \notin \Gamma$ du plan $\mathbb{P}(E)$ par rapport à la conique Γ est donnée par la construction de la figure ci-dessous, et l'on a : $[P, P', M, M'] = -1$.



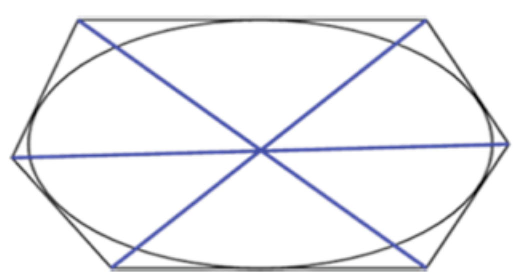
Preuve : i) Par hypothèse, Γ contient deux points, donc q est équivalente à X_0X_1 et on peut supposer que $P = (1 : 0) = \infty, P' = (0 : 1) = 0$ sur $\mathbb{P}_1(K) \simeq K \cup \infty$. Alors, les points $M = (x_0 : x_1) = (x : 1)$ et $M' = (y_0 :$

$y_1) = (y:1)$ sont en polarité si et seulement si $x_0y_1+x_1y_0 = 0$, soit $x+y = 0$, autrement dit ssi P' est le milieu du segment $[MM']$ de la droite affine K, soit $[M, M', P', P] = -1 = [M, M', P, P']$.

ii) Vérifions tout d'abord que le faisceau $\mathcal{F}_M$ des droites passant par M possède bien deux tangentes à Γ (tout au moins si K est algébriquement clos) . Ce faisceau donne par dualité une droite Δ^* de $\mathbb{P}(E^*)$, et les tangentes recherchées sont par définition les points d'intersection de la droite Δ^* avec la conique duale Γ^*. Comme $(\Gamma^*)^* = \Gamma$, Δ^* est tangente à Γ^* si et slt si M appartient à Γ, de sorte que sous l'hypothèse $M \notin \Gamma$ de la proposition, on peut tirer de M deux tangentes distinctes[3] $(MQ_1), (MQ_2)$ à Γ.

Comme dans la preuve du théorème 2.1, choisissons alors comme repère projectif du plan $Q_1 = (0:1:0), Q_2 = (1:0:0), M = (0:0:1)$, et $P = (1:1:1)$. L'équation de Γ devient $X_0X_1 - X_2^2 = 0$, et la droite $(MP) : \{X_0 - X_1 = 0\}$ coupe $(Q_1Q_2) : \{X_2 = 0\}$ en $M' = (1:1:0)$ et recoupe Γ en $P' = (1:1:-1)$. Ainsi, (Q_1Q_2) admet bien $(0:0:1) = M$ pour pôle, donc est la polaire de M, et $[P, P', M, M'] = [1, -1, 0, \infty] = -1$.

Comme la dualité, la polarité préserve les relations d'incidence. En voici une application classique : six points A, B, C, A', B', C' d'une conique lisse Γ admettent pour polaires par rapport à Γ les tangentes a, b, c, a', b', c' à Γ en ces points. Le pôle de $(A'B')$ est le point $a' \cap b'$, le pôle de (BC) est le point $b \cap c$, la diagonale joignant ces points est la polaire du point $(A'B') \cap (BC)$. D'après Pascal, les trois points de ce type sont alignés. Par conséquent, leurs polaires, c-à-d. les trois diagonales de ce type, sont concourantes (théorème de Brianchon, dual du théorème de Pascal).



[3] Si $M = (0:0:1)$ et si $\{X_0 = 0\}$ n'est pas tangente à Γ, leurs équations sont données par $X_1 = tX_0$, où t est tel que le discriminant de la forme quadratique $q_t(X_0, X_2) := Q(X_0, tX_0, X_2)$ s'annule.

Quadriques affines

Soit $\Gamma = Z(Q)$ une quadrique de $\mathbb{P}_n(K)$ ne contenant pas l'hyperplan $\mathbf{H}_0$ d'équation $X_0 = 0$, c-à-d. telle que X_0 ne divise pas Q dans l'anneau $K[X_0, ..., X_n]$. D'après le chapitre II, p. 46, Γ induit sur l'ouvert affine $\mathcal{H}_0 \simeq K^n$ complémentaire de $\mathbf{H}_0$ une hypersurface $\mathcal{Z}(f)$ d'équation $f(x_1, ..., x_n) = 0$, où $f(\frac{X_1}{X_0}, ..., \frac{X_n}{X_0}) = \frac{1}{X_0^2} Q(X_0, ..., X_n)$ est un polynôme de degré total δ égal à 2 . Plus précisément, la forme quadratique en n variables $\Phi(X_1, ...X_n) = F(0, X_1, ..., X_n)$ coïncide avec la somme des monômes de degré 2 de f, et on dit que $\mathcal{Z}(f)$ est une quadrique de l'espace affine K^n, de forme initiale Φ. On a :

$$f(x_1, ..., x_n) = \Phi(x_1, ..., x_n) + L(x_1, ..., x_n) + c,$$

où L est une forme linéaire en n variables et $c \in K$.

En effectuant des changements de coordonnées affines, ce qui ne change pas le rang ρ de Φ, on peut mettre l'équation de toute quadrique affine $\mathcal{Z}(f)$ de $\mathbb{C}^n$, sous l'une des formes suivantes :

$$\Sigma_{i=1,...,\rho}\ x_i^2 = 0 \quad (1 \le \rho \le n)\,;$$
$$\Sigma_{i=1,...,\rho}\ x_i^2 - 1 = 0 \quad (1 \le \rho \le n)\,;$$
$$\Sigma_{i=1,...,\rho}\ x_i^2 + x_{\rho+1} = 0 \quad (1 \le \rho \le n-1).$$

Pour $n = 2$ et $K = \mathbb{R}$, on peut mettre l'équation de toute conique affine réelle $\mathcal{C}$ sous l'une des formes : $x_1^2 = 0; x_1^2 - x_2^2 = 0, x_1^2 + x_2^2 = 0, x_1^2 - 1 = 0, x_1^2 + 1 = 0; x_1^2 + x_2^2 + 1 = 0$, et les classiques $x_1^2 + x_2^2 - 1 = 0$ (ellipse), $x_1^2 - x_2^2 + 1 = 0$ (hyperbole), et $x_1^2 + x_2 = 0$ (parabole).

Inversement, toute quadrique affine $\mathcal{C} = \mathcal{Z}(f)$ donne lieu à une quadrique projective $\Gamma = Z(F)$ par le procédé d'homogénisation décrit au chapitre II. Dans le cas $n = 2$, les paraboles $\mathcal{C}$ sont caractérisées par le fait que la droite à l'infini $\Delta = \{X_0 = 0\}$ est tangente à Γ. Lorsque l'espace affine $\mathbb{R}^2$ est muni de sa métrique usuelle, les cercles $\mathcal{C}$ sont caractérisés par le fait que l'intersection de Γ avec la droite à l'infini est formée des points complexes $P_+ = (0 : i : 1), P_- = (0 : i : 1)$, qu'on appelle les *points cycliques* à l'infini. En appliquant la Proposition 2.2.ii aux points de la droite à l'infini $M = (0 : x_1 : x_2), P = P_+, P' = P_-$ et $M' = M^{\perp_\Gamma} \cap \Delta = (0 : y_1 : y_2)$, on voit que $[P_+, P_-, M, M'] = -1$ si et seulement si $x_1 y_1 + x_2 y_2 = 0$. Ainsi, deux droites D, D' du plan euclidien sont orthogonales si et seulement si leurs points à l'infini $\infty_D, \infty_{D'}$ forment avec les points cycliques P_+, P_- une division harmonique sur la droite à l'infini Δ :

$$D \perp D' \quad \Leftrightarrow \quad [P_+, P_-, \infty_D, \infty_{D'}] = -1.$$

III.3 Groupes d'isométries

III.3.1 Les "groupes classiques"

Soit E un espace vectoriel de dimension n muni d'une forme b vérifiant l'une des trois hypothèses du §1.1, et non dégénérée. On dit qu'une application de E dans E est une *isométrie* si

$$\forall x, y \in E \ , \ b(u(x), u(y)) = b(x, y).$$

Dans les cas symétrique ou hermitien, cela revient à demander que $\forall x \in E, q(u(x)) = q(x)$, où q désigne la forme quadratique ou quadratique hermitienne associée à b.

Proposition III.3.1. *Soit u une isométrie de (E, b). Alors, u est un automorphisme K-linéaire de E.*

Preuve : dans le cas symétrique ou hermitien, E admet une base $\{e_1, ..., e_n\}$ orthogonale (cf. Prop. 1.3.i), avec $q(e_i) := d_i \neq 0$ pour tout i puisque b est non dégénérée, et tout $x \in E$ s'écrit $\Sigma_{i=1,...,n} x_i e_i$, où $d_i x_i = b(x, e_i)$. Comme u est une isométrie, $\{u(e_1), ..., u(e_n)\}$ forme également une système orthogonal, avec $q(u(e_i)) = q(e_i) \neq 0$, donc (exercice) est une base de E, et $u(x) = \Sigma_{i=1,...,n} d_i^{-1} b(u(x), u(e_i)) u(e_i) = \Sigma_i d_i^{-1} b(x, e_i) u(e_i) = \Sigma_i x_i u(e_i)$, ce qui montre que u est bien linéaire. Comme elle applique une base sur une base, c'est de plus un automorphisme.

Dans le cas antisymétrique, choisissons la base symplectique $e_1, \epsilon_1, ..., e_m, \epsilon_m$ donnée par la proposition 1.3.ii. Son image par u vérifie les mêmes relations d'orthogonalité. On en déduit que c'est aussi une base de E, et pour tout $x \in E$, on a : $u(x) = \Sigma_{i=1,...,m} b(u(x), u(\epsilon_i)) u(e_i) - b(u(x), u(e_i)) u(\epsilon_i)$. Le même raisonnement montre que u est linéaire, et que c'est un automorphisme.

Les isométries forment donc un sous-groupe de $GL(E)$, qu'on note $O(q)$ dans le cas bilinéaire symétrique (*groupe orthogonal*), $U(q)$ dans le cas hermitien (*groupe unitaire*), $Sp(b)$ dans le cas bilinéaire antisymétrique (*groupe symplectique*), ou simplement $O(E)$ lorsque b est donnée. Pour tout $u \in$

$\mathcal{L}(E,E)$, de transposée ${}^tu \in \mathcal{L}(E^*,E^*)$ (remplacer E^* par $E^{*\sigma}$ dans le cas hermitien), on note $u^* = \phi_b^{-1} \circ {}^tu \circ \phi_b$ *l'adjoint* de u, c-à-d. l'unique endomorphisme de E tel que $b(u^*x,y) = b(x,uy)$ pour tout $x,y \in E$. Alors, u est une isométrie si et slt si $u^*u = id_E$. Lorsque b est le produit scalaire usuel, de matrice représentative $\mathbf{I}_n$ sur K^n, ou la forme symplectique usuelle sur K^{2m}, de matrice $\mathbf{J}_{2m} = \begin{pmatrix} 0 & \mathbf{I}_m \\ -\mathbf{I}_m & 0 \end{pmatrix}$, on désigne ces groupes par

$$O_n(K) = \{P \in GL_n(K), {}^tPP = \mathbf{I}_n\}, U_n(K) = \{P \in GL_n(K), \sigma({}^tP)P = \mathbf{I}_n\},$$

$Sp_{2m}(K) = \{P \in GL_{2m}(K), {}^tP\mathbf{J}_{2m}P = \mathbf{J}_{2m}\}$. Les matrices de $O(n) := O_n(\mathbb{R})$ sont dites orthogonales, celles de $U(n) := U_n(\mathbb{C})$ unitaires. Lorsque $E = \mathbb{R}^n$, et b est de signature $(s, t = n-s)$, on pose $O(E) = O(s,t)$. Par exemple, $O(3,1)$ est le groupe de Lorentz des physiciens.

Soit $SL(E) = \{u \in GL(E), det(u) = 1\}$ (noté $SL_n(K)$ quand $E = K^n$) le groupe spécial linéaire. Son intersection avec chacun de groupes supra en fournit des sous-groupes normaux, notés $SO(q)$ (d'indice 2 dans $O(q)$, car $(det(u))^2 = 1$ pour u orthogonale), $SU(q)$, etc. Le groupe $Sp(b)$ est déjà contenu dans $SL(E)$ (ce n'est pas immédiat, mais pour $m = 1$, on vérifie facilement que $Sp_2(K) = SL_2(K)$.)

Remarque 1 (*algèbres de Lie*) : lorsque $K = \mathbb{R}$ ou $\mathbb{C}$, chacun de ces groupes G est naturellement muni de la topologie induite par celle de l'espace vectoriel $Mat_{n,n}(K)$, et on peut parler de la composante connexe G^+ de $\mathbf{I}_n$ dans G. On verra plus loin que $O(n)^+ = SO(n)$. En revanche, $U(n)$ est connexe. On peut également considérer l'application exponentielle, définie par la série normalement convergente

$$exp : Mat_{n,n}(K) \to GL_n(K) : X \mapsto exp(X) = \Sigma_{n\geq 0}\frac{1}{n!}X^n$$

et associer à G son *algèbre de Lie*

$$Lie(G) := \{X \in Mat_{n,n}(K), \forall t \in \mathbb{R}, exp(tX) \in G\},$$

qui est un $\mathbb{R}$-sous-espace vectoriel de $Mat_{n,n}(K)$, muni de la loi de composition interne $(X,Y) \mapsto [X,Y] := XY - YX \in Lie(G)$. Par exemple,

$Lie(SL_n(K)) = \{X \in Mat_{n,n}(K), tr(X) = 0\}$ (matrices de trace nulle) ;
$Lie(O(n)) = \{X \in Mat_{n,n}(\mathbb{R}), {}^tX + X = 0\}$ (X est antisymétrique) ;

$Lie(U(n)) = \{X \in Mat_{n,n}(\mathbb{C}),\ {}^tX + \overline{X} = 0\}$ (X est “antihermitienne”) ; Puisque exp est continue, $Lie(G^+) = Lie(G)$. La dimension du $\mathbb{R}$-espace vectoriel $Lie(G)$ s'appelle la *dimension* du groupe G.

Remarque 2 (*quaternions*) : ils fournissent de nouvelles familles de groupes, et des isomorphismes entre certains des groupes précédents, en basse dimension. Le corps (non commutatif) des quaternions de Hamilton est $\mathbb{H} = \mathbb{R} \oplus \mathbb{R}i \oplus \mathbb{R}j \oplus \mathbb{R}k$, muni de la loi de multiplication définie par $\mathbb{R}$-bilinéarité à partir des relations $i^2 = j^2 = k^2 = -1, ij = -ji = k, jk = -kj = i, ki = -ik = j$ (on vérifie qu'elle est bien associative, et que tout élément non nul de $\mathbb{H}$ est inversible). Soit σ l'anti-involution (c-à-d. $\sigma(h_1h_2) = \sigma(h_2)\sigma(h_1)$) de $\mathbb{H}$ définie par $\sigma(x+yi+zj+tk) = x-yi-zj-tk$, de sorte que $q(h) := \sigma(h).h := ||h||^2$ fournit une structure euclidienne à $\mathbb{H} \simeq \mathbb{R}^4$, et soit $U := U_1(\mathbb{H})$ le sous-groupe du groupe multiplicatif $\mathbb{H}^*$ formé par les quaternions de norme 1. On peut voir $\mathbb{H} = \mathbb{C} \oplus \mathbb{C}j \simeq \mathbb{C}^2$ comme un $\mathbb{C}$-espace vectoriel de dimension 2, que $||.||$ munit d'une structure hermitienne. Alors, l'application $\rho : U \to GL_2(\mathbb{C})$ qui attache à $u \in U$ l'automorphisme $\mathbb{C}$-linéaire de $\mathbb{H}$: $h \to \rho(u)(h) := hu^{-1}$ est un homomorphisme de groupes injectif. Comme $||hu|| = ||h||.||u|| = ||h||$ pour tout $h \in \mathbb{H}$, son image est contenue dans le groupe unitaire $U(2)$. En fait, pour $u^{-1} = \alpha + \beta j \in \mathbb{C} \oplus \mathbb{C}j$ de norme $\alpha\overline{\alpha} + \beta\overline{\beta} = 1$, la matrice représentative de $\rho(u)$ dans la base $\{1, j\}$ de $\mathbb{H}$ est donnée par $\begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix} = \begin{pmatrix} \overline{\alpha} & \overline{\beta} \\ -\beta & \alpha \end{pmatrix}^{-1}$, de sorte que ρ établit un isomorphisme de U sur $U(2)$.

L'espace $\mathbb{R}i \oplus \mathbb{R}j \oplus \mathbb{R}k$ des quaternions purs, muni de $||.||$, s'identifie à l'espace euclidien usuel $\mathbb{R}^3$. Pour tout $u \in U$, l'application $Int(u) : h \mapsto uhu^{-1}$ de $\mathbb{H}$ induit une isométrie de $\mathbb{R}^3$, d'où un homomorphisme de groupe $\pi : U \to SO(3)$, appliquant u sur $\pi(u) = (Int(u))|_{\mathbb{R}^3}$, de noyau $U \cap$ centre de $\mathbb{H}^* = \{\pm 1\}$. La théorie des algèbres de Lie permet de montrer que π est surjective, de sorte que $SO(3) \simeq SU(2)/\{\pm \mathbf{I}_2\}$.

Dans le même ordre d'idée, considérons $\mathbb{H}^n$ comme un espace vectoriel à droite sur $\mathbb{H}$, et soit $GL_n(\mathbb{H})$ le groupe des automorphismes $\mathbb{H}$-linéaires de $\mathbb{H}^n$. Avec l'identification $\mathbb{H}^n = \mathbb{C}^n + j\mathbb{C}^n$, on a $GL_n(\mathbb{H}) = \{P \in GL_{2n}(\mathbb{C}), P\mathbf{J}_{2n} = \mathbf{J}_{2n}\overline{P}\}$. Son sous-groupe $U_n(\mathbb{H}) = U(2n) \cap Sp_{2n}(\mathbb{C})$ coïncide avec à $U \simeq SU(2)$ pour $n = 1$, mais c'est en général un nouveau groupe.

III.3.2 Géométrie euclidienne

Nous nous restreignons désormais au cas euclidien : $K = \mathbb{R}$, et b est un produit scalaire sur $E \simeq \mathbb{R}^n$, c'est-à-dire une forme bilinéaire symétrique définie positive (= de signature $(n, 0)$). Alors, E admet des bases $\{e_1, ..., e_n\}$ orthonormées (= orthogonales et pour tout i, $q(e_i) = 1$), et on peut sans restreindre la généralité supposer que $q(x) = x_1^2 + ... + x_n^2 := ||x||^2$, ce qui identifie $O(E)$ à $O(n)$. Les éléments de $SO(n) = \{u \in O(n), det(u) = 1\}$ $= O(n)^+$ s'appellent les *rotations* de E. Le complémentaire de $SO(n)$ dans $O(n)$ est noté $O(n)^-$. Pour n impair, il contient $-\mathbf{I}_n$, de sorte que $O(2m+1) \simeq SO(2m + 1) \times \{\pm 1\}$; pour n pair, il contient des éléments d'ordre 2, mais aucun n'est central, donc on a seulement : $O(2m) \simeq SO(2m) \rtimes \{\pm 1\}$.

Dans le cas $n = 2$, le groupe $SO(2) = \{R_\theta = \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix}, \theta \in \mathbb{R}/2\pi\mathbb{Z}\}$ est commutatif (et isomorphe à $U(1) = \{z \in \mathbb{C}^*, |z| = 1\}$). Tout élément de $O(2)^-$ est une symétrie orthogonale S_D par rapport à une droite vectorielle D de $\mathbb{R}^2$, et $O(2) = SO(2) \rtimes_\tau \{\pm 1\}$, où -1, représentable par n'importe quelle symétrie $S_D = S_D^{-1}$, agit sur $SO(2)$ par $\tau(-1)(R_\theta) = R_{-\theta}$ (c-à-d. $S_D R_\theta S_D^{-1} = R_{-\theta}$). La composée $S_{D'} \circ S_D$ de deux symétries dont les droites forment un angle $(D, D') \equiv \theta \bmod (\pi\mathbb{Z})$ est la rotation $R_{2\theta}$.

Soit W un sous-espace vectoriel de E, d'orthogonal $W^\perp$. La relation $E = W \oplus W^\perp$ permet de définir la *projection orthogonale* π_W sur W (c-à-d. parallèlement à $W^\perp$) et de même, la *symétrie orthogonale* σ_W par rapport à W ; si W est un hyperplan H, cette symétrie s'appelle la *réflexion* τ_H par rapport à H. Toute symétrie orthogonale σ_W est une isométrie, et vérifie $\sigma_W^2 = id_E$. Inversement :

Lemme III.3.2. *i) Pour toute isométrie u d'un espace vectoriel euclidien E, on a : $E = Ker(u - id_E) \perp Im(u - id_E)$*

ii) Pour tout automorphisme u d'ordre 2 d'un espace vectoriel E , on a $E = Ker(u - id_E) \oplus Im(u - id_E)$, et $Im(u - id_E) = Ker(u + id_E)$.

iii) Tout élément u d'ordre 2 de $O(n)$ est une symétrie orthogonale.

Preuve : i) D'une part, $Ker(u - id_E)$ et $Im(u - id_E)$ sont de dimensions complémentaires. D'autre part, pour tout $(x \in Ker(u - id_E), z \in E)$: $b(x, u(z) - z) = b(u^{-1}x, z) - b(x, z) = b(x, z) - b(x, z) = 0$, d'où $Ker(u - id_E) \perp Im(u - id_E)$.

ii) Montrer, plus généralement, que si $P, Q \in K[T]$ sont des polynômes premiers entre eux tels que $P(u)Q(u) = 0$, on a $E = W \oplus W'$, où $W =$

$Ker(P(u)) = Im(Q(u)), W' = Ker(Q(u)) = Im(P(u))$. On pourra partir d'une relation de Bézout : il existe $A, B \in K[T]$ tels que $AP + BQ = 1$.

iii) Posant $W = Ker(u - id_E)$, on voit que $u = \begin{pmatrix} id_W & 0 \\ 0 & -id_{W^\perp} \end{pmatrix} = \sigma_W$.

Proposition III.3.3. *Soit u une isométrie sur $E \simeq \mathbb{R}^n$.*

i) Il existe des entiers $a, b, c \geq 0$, avec $a + b + 2c = n$, et une base orthonormée de E, dans laquelle u est représentée par une matrice du type $\begin{pmatrix} \mathbf{I}_a & 0 & 0 & 0 \\ 0 & -\mathbf{I}_b & 0 & 0 \\ 0 & 0 & R_{\theta_1} & 0 \\ 0 & 0 & 0 & R_{\theta_c} \end{pmatrix}$*, où $a + b + 2c = n$, et pour tout i, $\theta_i \notin \pi\mathbb{Z}$.*

ii) Il existe un entier $p \leq n$ et des reflexions $\tau_{H_1}, ..., \tau_{H_p}$, telles que $u = \tau_{H_p} \circ ... \circ \tau_{H_1}$. En particulier :

iii) tout élément de $O(2)^-$ est une reflexion ; les points fixes d'un élément $u \neq id_E$ de $SO(3)$ forment une droite, qu'on appelle l'axe de la rotation u ; $SO(n)$ est connexe par arcs, donc connexe.

La partie (i) signifie que pour tout $U \in O(n)$, il existe $P \in O(n)$ telle que $P^{-1}UP$ soit une matrice du type indiqué. Ou encore : il existe une décomposition orthogonale $E = V^+ \perp V^- \perp \Pi_1 \perp ... \perp \Pi_c$, en sous espaces stables sous u, tels que $u|V^+ = id, u|V^- = -id$, les Π_i sont des plans , et u y induit des rotations d'angles $\theta_i \not\equiv 0, \pi \mod 2\pi$.

Preuve i) Par récurrence sur n. Le cas $n = 1$, resp. 2, est trivial, resp. classique. Pour $n \geq 3$, soient $\lambda_1, ..., \lambda_m, m \leq n$ les valeurs propres de u dans $\mathbb{C}$, c'est-à-dire les racines de $det(u - \lambda id_E) = 0$. Si l'une, soit λ_1, est réelle, il lui correspond un vecteur propre $v_1 \in E$, qu'on peut supposer de norme 1 et qui vérifie $q(v_1) = q(u(v_1)) = \lambda^2 q(v_1)$, donc $\lambda = \pm 1$. Comme les isométries préservent la relation d'orthogonalité, l'hyperplan $E_1 := (\mathbb{R}v_1)^\perp$ est alors stable sous u, et on peut appliquer l'hypothèse de récurrence à $u_1 = u_{|E_1}$.

Si en revanche aucune valeur propre n'est réelle, deux au moins, soit λ_1 et $\lambda_2 = \overline{\lambda_1}$, sont complexes conjuguées. Considérons alors le complexifié $E_\mathbb{C}$ de E, sur lequel b s'étend en un produit scalaire hermitien $b_\mathbb{C}$, et notons $u_\mathbb{C}$ l'automorphisme $\mathbb{C}$-linéaire que u induit sur $E_\mathbb{C}$. Il est unitaire, et admet λ_1, λ_2 parmi ses valeurs propres. Donc $|\lambda_1| = |\lambda_2| = 1$, et on peut écrire $\lambda_1 = e^{i\theta_1}, \lambda_2 = e^{-i\theta_1}$, où $\theta_1 \not\equiv 0, \pi$. Les vecteurs propres $v_1, v_2 \in E_\mathbb{C}$ correspondant sont linéairement indépendants sur $\mathbb{C}$ (ils sont même orthogonaux pour $b_\mathbb{C}$, puisque $\lambda_1^{-1} b_\mathbb{C}(v_1, v_2) = b_\mathbb{C}(u_\mathbb{C}^{-1}(v_1), v_2) = b_\mathbb{C}(v_1, u_\mathbb{C}(v_2)) = \overline{\lambda_2} b_\mathbb{C}(v_1, v_2) =$

$\lambda_1 b_{\mathbb{C}}(v_1, v_2)$, tandis que $\lambda_1^{-1} \neq \lambda_1$). De plus, $E_{\mathbb{C}}$ admet une involution σ semilinéaire ($\forall (\lambda, v) \in \mathbb{C} \times E_{\mathbb{C}}, \sigma(\lambda v) = \overline{\lambda} v$), qui vérifie $Ker(\sigma - id_{E_{\mathbb{C}}}) = E$ et qui commute à $u_{\mathbb{C}}$. On peut donc choisir $v_2 = \sigma(v_1)$ comme vecteur propre pour λ_2. Dans ces conditions, le plan $\Pi_{\mathbb{C}}$ engendré par v_1, v_2 dans $E_{\mathbb{C}}$ contient les vecteurs $w_1 = \frac{1}{2}(v_1 + \sigma(v_1)), w_2 = \frac{1}{2i}(v_1 - \sigma(v_1))$, qui appartiennent à E. Ainsi, $\Pi_{\mathbb{C}} \cap E$ est un plan réel Π_1 de E stable sous u. Les valeurs propres de l'isométrie $u^1 := u_{|\Pi_1}$ sont $e^{i\theta_1}, e^{-i\theta_1}$, donc u^1 est la rotation d'angle θ_1 (qui est représentée par R_{θ_1} dans toute base orthonormée de Π_1), et le sous-espace $E_1 = \Pi_1^{\perp}$ de E, de dimension $n-2$, est stable sous u. On peut donc appliquer l'hypothèse de récurrence à $u_1 = u_{|E_1}$.

ii) C'est un corollaire de (i) : chaque R_{θ_i} peut s'écrire comme composée de deux reflexions $\tau_{D_i^+} \circ \tau_{D_i^-}$ du plan Π_i, qui sont induites par les reflexions $\tau_{H_i^{\pm}}$ de E d'hyperplans $H_i^{\pm} = D_i^{\pm} \oplus_{j \neq i} \Pi_j \oplus V^+ \oplus V^-$. Toute droite D de V^- fournit de même une reflexion τ_H relative à l'hyperplan $H = D^{\perp}$ de E. En choisissant b droites de V' orthogonales, on obtient b hyperplans $H_1, ..., H_b$ tels que $u = \tau_{H_1} \circ ... \circ \tau_{H_b} \circ \tau_{H_1^+} \circ \tau_{H_1^-} ... \circ \tau_{H_c^+} \circ \tau_{H_c^-}$, d'où les $p = b + 2c \leq n$ réflexions recherchées.

iii) Les décompositions ci-dessus ne sont pas uniques, mais la parité de b et de p ne dépend que de u, puisque $det(u) = (-1)^b = (-1)^p$. Pour $n = 2$ et $det(u) = -1$, cela impose $p = 1$, et u est une réflexion (en fait, nous avons utilisé cette propriété de $O(2)^-$ au démarrage de la récurrence de (i)). Pour $n = 3$ et $det(u) = 1$, cela impose $p = 2$, $u = \tau_{H_2} \circ \tau_{H_1}$: si $H_1 = H_2$, $u = id_E$ (et $a = 3$) ; sinon, $a = 1$, donc $Ker(u - id_E)$ est une droite Δ, qui contient, donc coïncide avec la droite $H_1 \cap H_2$, et u est la rotation d'axe Δ, d'angle deux fois l'angle de (H_1, H_2). Enfin, tout élément u de $SO(n)$ admet un $b = 2b'$ pair, pour lequel $-\mathbf{I}_b$ s'écrit comme la matrice à b' blocs diagonaux $\begin{pmatrix} R_\pi & 0 \\ 0 & R_\pi \end{pmatrix}$. Après conjugaison par une matrice de changement de base P, l'application $t \in [0, 1] \mapsto (R_{t\pi}, ..., R_{t\pi}, R_{t\theta_1}, ..., R_{t\theta_c})$ fournit alors un chemin continu reliant $\mathbf{I}_n$ à u dans $SO(n)$. Donc $SO(n)$ est contenu dans la composante connexe $O(n)^+$ de $\mathbf{I}_n$ dans $O(n)$. Puisque cette dernière est contenue dans l'image inverse de 1 par l'application continue $det : O(n) \to \{1, -1\}$, on a bien $SO(n) = O(n)^+$, comme annoncé plus haut.

Espaces affines euclidiens

Un espace affine euclidien $\mathcal{E}$ est un espace affine réel dont l'espace vectoriel directeur est un espace euclidien (E, b). Pour $P, Q \in \mathcal{E}$, on appelle distance de P à Q le nombre $d(P,Q) = (q(\overrightarrow{PQ}))^{\frac{1}{2}}$. Une *isométrie affine* de $\mathcal{E}$ est une application $f : \mathcal{E} \to \mathcal{E}$ qui conserve les distances. Ainsi, les translations $t_{\overrightarrow{v}}$, $\overrightarrow{v} \in E$ sont des isométries affines. L'énoncé suivant montre que les isométries affines de $\mathcal{E}$ forment un sous-groupe, noté $OA(\mathcal{E})$, du groupe affine $GA(\mathcal{E})$, et que $OA(\mathcal{E}) \simeq E \rtimes_\tau O(E)$, pour l'action naturelle τ de $O(E)$ sur E (voir chapitre I, §3.2).

Proposition III.3.4. *Soit f une isométrie affine de $\mathcal{E}$. Alors,*
i) f est un automorphisme affine de $\mathcal{E}$; c-à-d. $f \in GA(\mathcal{E})$. On note $\overrightarrow{f} \in GL(E)$ l'application linéaire correspondante ;
ii) Il existe une unique couple $(\overrightarrow{v}, g) \in E \times OA(\mathcal{E})$ vérifiant les 3 propriétés suivantes :
(a) $f = t_{\overrightarrow{v}} \circ g$; donc $\overrightarrow{g} = \overrightarrow{f}$;
(b) $Fix(g) = \{P \in \mathcal{E}, g(P) = P\}$ est non vide ; donc c'est un sous-espace affine de $\mathcal{E}$, d'espace vectoriel directeur $Ker(\overrightarrow{f} - id_E)$;
(c) $\overrightarrow{v} \in Ker(\overrightarrow{f} - id_E)$.
De plus, on a alors automatiquement : $t_{\overrightarrow{v}} \circ g = g \circ t_{\overrightarrow{v}}$.

Preuve : i) Fixons une origine O de $\mathcal{E}$. Alors, $f' = t_{\overrightarrow{f(O)O}} \circ f$ est une isométrie affine de $\mathcal{E}$ qui fixe O. En identifiant le vectorialisé $\mathcal{E}_O$ de $\mathcal{E}$ à E, on peut donc voir f' comme une application de E dans E telle que $f'(\overrightarrow{0}) = \overrightarrow{0}$. Pour tout $\overrightarrow{x} \in E$, on a alors : $q(f'(\overrightarrow{x})) = q(f'(\overrightarrow{x}) - f'(\overrightarrow{0})) = q(\overrightarrow{x})$, et la Proposition 3.1 entraîne que f' est linéaire. Donc $f' \in GA(\mathcal{E})$, et de même pour f.

ii) Montrons tout d'abord que si une décomposition $f = t_{\overrightarrow{v}} \circ g$ vérifie (a) et (b), elle vérifie (c) si et seulement si $t_{\overrightarrow{v}} \circ g = g \circ t_{\overrightarrow{v}}$. En effet, $g t_{\overrightarrow{v}} g^{-1} = t_{\overrightarrow{g}(\overrightarrow{v})}$ est égal à $t_{\overrightarrow{v}}$ si et slt si $\overrightarrow{v} \in Ker(\overrightarrow{g} - id_E) = Ker(\overrightarrow{f} - id_E)$.

Montrons maintenant qu'une décomposition $f = t_{\overrightarrow{v}} \circ g = g \circ t_{\overrightarrow{v}}$ vérifiant (a), (b), (c) est unique. Soit B un point fixe de g, et $B' = f(B) = B + \overrightarrow{v}$ son image par f. Pour tout $P \in \mathcal{E}$, d'image $P' = f(P)$, on a $\overrightarrow{PP'} = \overrightarrow{BP'} - \overrightarrow{BP} = \overrightarrow{BB'} + (\overrightarrow{f} - id_E)(\overrightarrow{BP}) = \overrightarrow{v} + (\overrightarrow{f} - id_E)(\overrightarrow{BP})$, où $v \in Ker(\overrightarrow{f} - id_E)$. Mais d'après le lemme 3.2.i, une telle décomposition de $\overrightarrow{PP'}$ est unique, donc $\overrightarrow{v}$, et par conséquent $g = f \circ t_{-\overrightarrow{v}}$, sont entièrement déterminés par f.

Construisons enfin une décomposition, en fixant une origine O de $\mathcal{E}$, d'image $O' = f(O)$ sous f. Selon le lemme 3.2.i, $\overrightarrow{OO'} = \overrightarrow{v} + (\overrightarrow{f} - id_E)(\overrightarrow{w}) \in Ker(\overrightarrow{f} - id_E) \oplus Im(\overrightarrow{f} - id_E)$. Soit $B = O - \overrightarrow{w}$, et soit $g \in OA(\mathcal{E})$ l'(unique) isométrie affine fixant B et de partie linéaire $\overrightarrow{g} = \overrightarrow{f}$. Alors f et $t_{\overrightarrow{v}} \circ g$ ont mêmes parties linéaires, et valent au point B

$$f(B) = O' + \overrightarrow{f}(\overrightarrow{OB}) = O + \overrightarrow{OO'} + \overrightarrow{f}(\overrightarrow{OB}) = O + \overrightarrow{v} + (\overrightarrow{f} - id_E)(\overrightarrow{BO}) + \overrightarrow{f}(\overrightarrow{OB})$$

$$= B + \overrightarrow{v} = t_{\overrightarrow{v}} \circ g(B), \text{ donc coïncident sur tout } \mathcal{E}.$$

Parmi les isométries affines f, on distingue en particulier :

- les *déplacements* (resp. antidéplacements), dont la partie linéaire $\overrightarrow{f} \in O(n)^+$ est une rotation vectorielle (resp. $\overrightarrow{f} \in O(n)^-$). Le groupe $OA(n)^+ \simeq E \rtimes_\tau O(n)^+$ des déplacements contient les translations, et est encore connexe.

- les *rotations*, qui sont les déplacements f possédant au moins un point fixe. Si $n = 2$, $f \neq id_{\mathcal{E}}$ est une rotation affine si et slt si $\overrightarrow{f}$ est une rotation vectorielle $\neq id_E$: alors, $Ker(\overrightarrow{f} - id_E) = 0$, et $f \neq id_{\mathcal{E}}$ a un unique point fixe, appelé centre de la rotation.

- les *symétries orthogonales* $\sigma_{\mathcal{W}}$ par rapport à un sous-espace affine $\mathcal{W}$ de $\mathcal{E}$: $\sigma_{\mathcal{W}}$ est la symétrie affine par rapport à $\mathcal{W}$ parallèlement à l'orthogonal $W^\perp$ de la direction W de $\mathcal{W}$. Quand $\mathcal{W} = \mathcal{H}$ est un hyperplan de $\mathcal{E}$, on parle de la *réflexion* $\tau_{\mathcal{H}}$ par rapport à $\mathcal{H}$; c'est un antidéplacement, d'ordre 2 dans le groupe $OA(n)$. Le composé de deux réflexions par rapports à des hyperplans parallèles de direction H est une translation par un vecteur $\overrightarrow{v} \in H^\perp$.

Si $n = 2$ et si f est un antidéplacement, alors $\overrightarrow{f}$ est une réflexion vectorielle (Prop. 3.3.iii) par rapport à la droite $D = Ker(\overrightarrow{f} - id_E)$. Si $Fix(f) \neq \emptyset$, c-à-d. si, dans la décomposition canonique $f = t_{\overrightarrow{v}} \circ g$ donnée par la Prop. 3.4.ii, le vecteur $\overrightarrow{v}$ est nul, f est une réflexion par rapport une droite $\mathcal{D} = Fix(f)$ de direction D. En revanche, si $Fix(f) = \emptyset$, f n'est pas une réflexion, mais une symétrie glissée (voir chap. I, §3.2, exemple iii). Comme la translation $t_{\overrightarrow{v}}$ est le produit de deux reflexions par rapport à des droites perpendiculaires à D, on voit qu'une symétrie glissée est la composée de p réflexions, avec $p = 2 + 1$, et que p ne peut être choisi < 3. En fait :

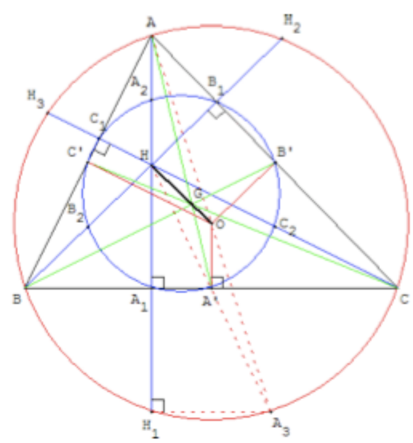
Proposition III.3.5. *Soit f une isométrie affine d'un espace affine euclidien $\mathcal{E}$ de dimension n. Il existe un entier $p \leq n + 1$ et des reflexions $\tau_{\mathcal{H}_1}, ..., \tau_{\mathcal{H}_p}$ telles que $f = \tau_{\mathcal{H}_p} \circ ... \circ \tau_{\mathcal{H}_1}$.*

Preuve : si f admet un point fixe O, f s'interprète comme une isométrie vectorielle du vectorialisé $\mathcal{E}_O$, et on peut lui appliquer la Proposition 3.3.ii. Sinon, $\overrightarrow{f} - id_E$ n'est *pas* inversible, de sorte que l'entier a apparaissant dans la décomposition de l'isométrie vectorielle $\overrightarrow{f}$ à la Proposition 3.3.i est nécessairement ≥ 1. Considérons alors la décomposition canonique $f = t_{\overrightarrow{v}} \circ g$ de f : comme $\overrightarrow{g} = \overrightarrow{f}$, son a est aussi ≥ 1, et comme g admet un point fixe, on peut, en passant au vectorialisé, écrire g comme un produit d'au plus $b + 2c < n$ réflexions de $\mathcal{E}_O$, d'où, puisque $t_{\overrightarrow{v}}$ est produit de deux réflexions, une décomposition de f en produits d'au plus $b + 2c + 2 \leq n + 1$ réflexions.

Voici une autre preuve, plus géométrique, de cet énoncé. Soient $\Phi \subset Fix(f)$ un ensemble de points fixes de f, et P un point de $\mathcal{E}$ non fixé par f. Montrons qu'il existe une réflexion $\tau_{\mathcal{H}}$ telle que l'isométrie $g_1 = \tau_{\mathcal{H}} \circ f$ admette P et les points de Φ parmi ses points fixes. Puisque $f(P) = P'$ n'est pas égal à P, il existe un unique hyperplan $\mathcal{H}$ orthogonal à (PP') et passant par le milieu M du segment $[PP']$. Alors, $\tau_{\mathcal{H}}(P') = P$, donc $P \in Fix(g_1)$. Par ailleurs, tout point Q de Φ vérifie $d(Q,P) = d(f(Q)f(P)) = d(Q,P')$, donc $(QM) \perp (PP')$, et $Q \in \mathcal{H}$ est fixé par $\tau_{\mathcal{H}}$, donc par $\tau_{\mathcal{H}} \circ f = g_1$.

Soit alors $\{P_0, ..., P_n\}$ un repère affine de $\mathcal{E}$. En itérant le procédé supra p fois, avec $p \leq n + 1$, on construit p hyperplans $\mathcal{H}_i$ tels que l'isométrie $g = \tau_{\mathcal{H}_p} \circ ... \circ \tau_{\mathcal{H}_1} \circ f$ admette $P_0, ..., P_n$ parmi ses points fixes. Comme une application affine est entièrement déterminée par ses valeurs sur un repère affine, c'est que $g = id_{\mathcal{E}}$, et $f = \tau_{\mathcal{H}_p} \circ ... \circ \tau_{\mathcal{H}_1}$.

Signalons pour conclure que sans être des isométries, les projections orthogonales et les similitudes (c'est-à-dire les composées d'isométries par des homothéties) jouent un rôle important en géométrie euclidienne. Le cercle des 9 points en fournit une illustration classique : voir http ://pagesperso-orange.fr/debart/geoplan/feuerbach.html , d'où est extraite la figure ci-dessous.



Chapitre IV

Géométrie multilinéaire

Université de Paris VI | Master M1 2010-2011
D. Bertrand | *Algèbre géométrique*

CHAPITRE IV

GÉOMÉTRIE MULTILINÉAIRE

IV.1 Produits tensoriels.

Dans ce paragraphe, on considère des espaces vectoriels, de dimension éventu-ellement infinie, sur un corps (commutatif) K quelconque. En première lecture, on pourra les supposer tous de dimension finie.

IV.1.1 Définition

Soient V et W deux espaces vectoriels. On cherche à construire un espace vectoriel $V \otimes W$ tel que la donnée d'une forme bilinéaire b sur $V \times W$ équivale à celle d'une forme *linéaire* f_b sur $V \otimes W$. Plus généralement, on étudie le "problème universel" suivant.

Existe-t-il un couple (E, β) formé d'un espace vectoriel E et d'une application bilinéaire $\beta : V \times W \to E$ vérifiant la propriété suivante : pour toute application bilinéaire b de $V \times W$ vers un espace vectoriel U quelconque, il existe une **unique application linéaire** *$f = f_b : E \to U$ telle que*

$$\textit{pour tout } (v, w) \in V \times W,\ b(v, w) = f_b\big(\beta(v, w)\big)\ ,$$

autrement dit, telle que $b = f_b \circ \beta$.

Du fait de la condition d'unicité, deux solutions $(E, \beta), (E', \beta')$ de ce problème sont liées par un isomorphisme K-linéaire $\phi : E \to E'$ tel que $\phi \circ \beta = \beta'$. S'il y a une solution, elle est donc unique à isomorphisme près. On va montrer qu'il existe une solution. On la désignera (à isomorphisme près) par $(E = V \otimes W, \beta)$. Pour tout $(v, w) \in V \times W$, on écrira

$$\beta(v, w) := v \otimes w \in V \otimes W,$$

de sorte que la bilinéarité de β se traduit par les relations :

$$\forall v, v' \in V, w, w' \in W, \lambda \in K\ ,\ (\lambda v) \otimes w = v \otimes (\lambda w) = \lambda(v \otimes w)\,;$$
$$v \otimes (w + w') = v \otimes w + v \otimes w'\ ,\ (v + v') \otimes w = v \otimes w + v' \otimes w.$$

Lorsqu'on veut préciser le corps des scalaires K, on écrit $V \otimes_K W$. On dit que $V \otimes_K W$ est “le” produit tensoriel de V par W sur K. Ses éléments s'appellent des tenseurs, ceux de la forme $v \otimes w$ des tenseurs purs (ou décomposés). Attention : une somme $v \otimes w + v' \otimes w' \in V \otimes W$ n'est en général pas un tenseur pur !

Existence du produit tensoriel.

Première preuve : soit $\mathcal{E}$ l'espace vectoriel formé par toutes les applications ensemblistes de $V \times W$ dans K nulles en dehors d'un sous-ensemble fini de $V \times W$. Il admet pour base $\{\delta_{(v,w)}, (v, w) \in V \times W\}$ où δ_x désigne la fonction de Dirac relative au point $x \in V \times W$. Considérons le sous-espace vectoriel $\mathcal{F}$ de $\mathcal{E}$ engendré par les éléments $\delta_{(v,w+w')} - \delta_{(v,w)} - \delta_{(v,w')}, \delta_{(v+v',w)} - \delta_{(v,w)} - \delta_{(v',w)}, \delta_{(\lambda v,w)} - \lambda\delta_{(v,w)}, \delta_{(v,\lambda w)} - \lambda\delta_{(v,w)}$, ainsi que l'espace vectoriel quotient $E = \mathcal{E}/\mathcal{F}$. Alors, $\beta : V \times W \to E : \beta(v, w) :=$ classe de $\delta_{(v,w)}$ dans E, est une application bilinéaire de $V \times W$ dans E, dont l'image ensembliste $\beta(V \times W)$ engendre l'espace vectoriel E. Pour toute application $b : V \times W \to U$, l'application $\tilde{f}_b : \mathcal{E} \to U$ qui attache à un élément $\Sigma_{(v,w)\in A\subset V\times W, A \text{ fini}}\,\lambda_{(v,w)}\delta_{(v,w)}$ de $\mathcal{E}$ l'élément $\Sigma_{(v,w)\in A}\lambda_{(v,w)}b(v, w)$ de U est bien définie (car les $\delta_{(v,w)}$ forment une base de $\mathcal{E}$), et est linéaire. Si b est de plus bilinéaire, $\tilde{f}_b$ s'annule sur $\mathcal{F}$, et définit donc par passage au quotient par $\mathcal{F}$ une application linéaire $f_b : E \to U$, qui vérifie bien $f_b \circ \beta = b$. Enfin, f_b est la seule application linéaire de E dans U vérifiant cette propriété, puisque l'image de β engendre E tout entier.

Deuxième preuve : soient $\{e_i; i \in I\}, \{\epsilon_j; j \in J\}$ des bases de V, W. À tout $(i, j) \in I \times J$, attachons un symbole $\vec{m}_{i,j}$ et considérons l'espace vectoriel $E := \oplus_{(i,j)\in I\times J} K.\vec{m}_{i,j}$, de base $\{\vec{m}_{i,j}, (i, j) \in I \times J\}$. Il existe alors une (unique) application bilinéaire $\beta : V \times W \to E$ telle que $\beta(e_i, \epsilon_j) = \vec{m}_{i,j}$ pour tout $(i, j) \in I \times J$. Soit maintenant $b : V \times W \to U$ une application bilinéaire. Puisque les $\vec{m}_{i,j}$ forment une base de E, il existe une unique application linéaire $f_b : E \to U$ telle que $f(\vec{m}_{i,j}) = b(e_i, \epsilon_j)$ pour tout $(i, j) \in I \times J$. De la bilinéarité de b et de β, et de la linéarité de f_b, on déduit alors que pour tout $v = \Sigma_{i\in A \text{ fini } \subset I}\lambda_i e_i \in V\ ,\ w = \Sigma_{j\in B \text{ fini } \subset J}\mu_j\epsilon_j \in W$, on a : $b(v, w) = \Sigma_{(i,j)\in A\times B}\lambda_i\mu_j b(e_i, \epsilon_j) = f_b(\Sigma_{(i,j)\in A\times B}\lambda_i\mu_j\vec{m}_{i,j}) = f_b(\beta(v, w))$. Ainsi, l'application linéaire f_b vérifie la propriété requise, et c'est la seule puisque que l'image de β contient la base $\vec{m}_{i,j}$ de E. On voit ainsi de plus que

Théoreme IV.1.1. *: si V, W sont des espaces vectoriels de dimension finies n, m, de bases $\{e_i, 1 \leq i \leq n\}, \{\epsilon_j, 1 \leq j \leq m\}$, alors $V \otimes W$ est un espace vectoriel de dimension nm, de base $\{e_i \otimes \epsilon_j, 1 \leq i \leq n, 1 \leq j \leq m\}$.*

Remarques :

i) pour $V = K$, l'application bilinéaire $K \times W \to W : (\lambda, w) \mapsto \lambda w$ établit un isomorphisme canonique $K \otimes W \simeq W$.

ii) Inversion : L'application bilinéaire $V \times W \to W \otimes V : (v, w) \to w \otimes v$ établit un isomorphisme canonique $V \otimes W \simeq W \otimes V$. Mais attention quand $V = W$: dans $V \otimes V$, $v \otimes v'$ n'est égal à $v' \otimes v$ que si v et v' sont colinéaires.

iii) *Complexification* : soit V un espace vectoriel sur $\mathbb{R}$, de dimension n. Comme $\mathbb{C}$ est un ev de dimension 2 sur $\mathbb{R}$, on peut former $\mathbb{C} \otimes_{\mathbb{R}} V := V_{\mathbb{C}}$, qui est un $\mathbb{R}$- ev de dimension $2n$. De plus, pour tout $\lambda \in \mathbb{C}$, l'application $\mathbb{C} \times V \to V_{\mathbb{C}} : (\mu, v) \mapsto \lambda\mu \otimes v$ est $\mathbb{R}$-bilinéaire, donc induit un endomorphisme $\mathbb{R}$-linéaire m_λ de $V_{\mathbb{C}}$. On vérifie que $m_\lambda + m_{\lambda'} = m_{\lambda+\lambda'}, m_\lambda \circ m_{\lambda'} = m_{\lambda\lambda'}, m_1 = id_{V_{\mathbb{C}}}$, de sorte que la loi $\mathbb{C} \times V_{\mathbb{C}} \ni (\lambda, \mathbf{v}) \mapsto \lambda.\mathbf{v} := m_\lambda(\mathbf{v}) \in V_{\mathbb{C}}$ munit $V_{\mathbb{C}}$ d'une structure d'espace vectoriel sur $\mathbb{C}$, de dimension n sur $\mathbb{C}$. Un tel $\mathbb{C}$-ev est naturellement muni d'un involution $\mathbb{C}$-antilinéaire $\sigma = c : V_{\mathbb{C}} \to V_{\mathbb{C}}$, appelée conjugaison complexe, et définie à partir des relations $c(\lambda \otimes v) = \overline{\lambda} \otimes v$. On récupère la structure réelle en notant que $\{v \in V_{\mathbb{C}}, c(v) = v\}$ coïncide avec $1 \otimes V$, qu'on identifie à V. Ceci montre d'ailleurs qu'il n'y a pas de conjugaison complexe naturelle sur un $\mathbb{C}$-espace vectoriel général (les essais "évidents" dépendent du choix d'une base).

iv) *K-algèbres* : ce sont les K-espaces vectoriels A munis d'une application linéaire $\mu : A \otimes A \to A$. Traduire ce que cela signifie en terme de la loi de composition interne $(a, b) \mapsto a.b := \mu(a \otimes b)$. Pour l'associativité, qui n'est pas requise dans la définition d'une K-algèbre (penser aux algèbres de Lie), voir le §2.1 ci-dessous.

v) Soient $V_1 \subset V, W_1 \subset W$ deux sous-espaces vectoriels. Alors, l'application $V_1 \times W_1 \ni (v_1, w_1) \mapsto v_1 \otimes w_1 \in V \otimes W$ fournit une application linéaire injective $\iota : V_1 \otimes W_1 \to V \otimes W$, qui permet d'identifier $V_1 \otimes W_1$ à un sous-espace vectoriel de $V \otimes W$.

vi) (Hors programme) Attention toutefois : si une bonne partie des énoncés de ce chapitre s'étend sans difficulté au cas des modules sur un anneau commutatif, ce n'est pas le cas pour l'injectivité de ι au (v) précédent. Ainsi, le $\mathbb{Z}$-module $V = \mathbb{Z}$ admet $V_1 = 2\mathbb{Z} \simeq \mathbb{Z}$ comme sous-module, et en considérant le $\mathbb{Z}$-module $W = \mathbb{Z}/2\mathbb{Z}$, on a $V_1 \otimes_{\mathbb{Z}} W \simeq \mathbb{Z} \otimes_{\mathbb{Z}} W \simeq W$, tandis que le produit tensoriel *dans* $V \otimes_{\mathbb{Z}} W$ d'un élément $2n \in V_1$ par $\overline{1} \in W$ vaut $(2n) \otimes_{\mathbb{Z}} \overline{1} = 2(n \otimes_{\mathbb{Z}} \overline{1}) = n \otimes_{\mathbb{Z}} 2.\overline{1} = n \otimes_{\mathbb{Z}} \overline{0} = 0$. L'application $\iota : V_1 \otimes_{\mathbb{Z}} W \to V \otimes_{\mathbb{Z}} W$ est donc ici identiquement nulle.

IV.1.2 Propriétés fonctorielles

Théoreme IV.1.2. *Soient $g : V \to V', h : W \to W'$ deux applications linéaires. Alors, il existe une unique application linéaire, notée*

$$g \otimes h : V \otimes W \to V' \otimes W'$$

telle que pour tout $(v, w) \in V \times W$, on ait $(g \otimes h)(v \otimes w) = g(v) \otimes h(w)$.

Preuve : l'application $b : V \times W \to V' \otimes W' : (v, w) \mapsto g(v) \otimes h(w)$ est bilinéaire. L'application linéaire correspondante f_b est donc la seule vérifiant ces conditions, et on la note $f_b := g \otimes h$.

Mais cette notation impose de vérifier une compatibilité. Les applications g, h appartiennent aux espaces vectoriels $\mathcal{L}(V, V'), \mathcal{L}(W, W')$, et la notation $g \otimes h$ représente déjà un élément de leur produit tensoriel. Heureusement, l'application $\mathcal{L}(V, V') \times \mathcal{L}(W, W') \to \mathcal{L}(V \otimes W, V' \otimes W') : (g, h) \mapsto f_b$ fournit une application linéaire canonique

$$\kappa : \mathcal{L}(V, V') \otimes \mathcal{L}(W, W') \to \mathcal{L}(V \otimes W, V' \otimes W'),$$

envoyant bien $g \otimes h$ (au sens du membre de gauche) sur f_b. Elle est injective, mais pas surjective en général (voir le théorème 2).

Transcription matricielle : supposons $V' = V, W' = W$ de dimensions finies n, m, et soient A, B les matrices représentatives $(n \times n), (m \times m)$ de g, h dans des bases e_i, ϵ_j de V, W. Alors, la matrice représentative de l'endomorphisme $g \otimes h$ de $V \otimes W$ dans la base $e_i \otimes \epsilon_j$ est donnée par le *produit de Kronecker* $A \otimes B$ de A par B : c'est la matrice $(nm \times nm)$ obtenue en remplacant chaque coefficient $a_{i,i'}$ de A par le bloc $a_{i,i'}B$.

Pour tout espace vectoriel V, notons $V^* := \mathcal{L}(V, K)$ le dual de V. Si V est de dimension finie, et muni d'une base $\{e_1, ..., e_n\}$, on notera $\{e_1^*, ..., e_n^*\}$ la base duale de V^*, définie par $e_j^*(e_i) = \delta_{i,j}, 1 \leq i, j \leq n$. On rappelle que le rang d'une application lineaire est la dimension de son image.

Théoreme IV.1.3. *: soient V, W deux espaces vectoriels.*

i) Il existe une unique injection linéaire

$$\phi : V^* \otimes W \hookrightarrow \mathcal{L}(V, W),$$

telle que pour tout $\ell \in V^*, w \in W, v \in V$, on ait $\big(\phi(\ell \otimes w)\big)(v) = \ell(v)w$.

ii) L'image de ϕ est formée des applications linéaires de V dans W de rang fini. En particulier, si V ou W est de dimension finie, ϕ est un isomorphisme.

iii) Si $W = V$ est de dimension finie, et si $u \in End(V)$ est l'image par ϕ de $\Sigma_\alpha \ell_\alpha \otimes v_\alpha$, alors sa trace vaut $Tr(u) = \Sigma_\alpha \ell_\alpha(v_\alpha)$. Dans les bases décrites plus haut, l'application identité $u = id_V$ de V est l'image par ϕ de $\Sigma_{i=1,\dots n} e_i^* \otimes e_i$.

Preuve : i) pour définir ϕ (qui est un cas particulier de l'application κ supra), considérer l'application bilinéaire $V^* \times W \to \mathcal{L}(V, W) : (\ell, w) \mapsto \{v \mapsto \ell(v)w\}$. Pour l'injectivité, choisissons des bases $\{\ell_i, i \in I\}, \{\epsilon_j, j \in J\}$ de V^*, W. Alors, les $\{\ell_i \otimes \epsilon_j, (i,j)\}$ forment une base de $V^* \otimes W$, et il suffit de montrer que leurs images par ϕ sont des applications de V dans W linéairement indépendantes. Dans le cas contraire, celles-ci enverraient tout vecteur v de E sur des vecteurs $\ell_i(v)\epsilon_j$ de W liés. Les ϵ_j étant linéairement indépendants, on aurait donc $\ell_i(v) = 0$ pour tout $i \in I, v \in V$, et les ℓ_i seraient toutes nulles.

ii) Comme $V^* \otimes W$ est formé de combinaisons linéaires *finies* des $\ell_i \otimes \epsilon_j$, leurs images par ϕ sont bien de rang fini. Inversément, soit $u : V \to W$, d'image W' de dimension finie, et soit $\{\epsilon_1, .., \epsilon_m\}$ une base de W'. Le lecteur concluera en la complétant en une base $\{\epsilon_j, j \in J\}$ de W. Enfin, si V ou W est de dimension finie, tout u est de rang fini.

iii) En complétant un élément non nul ℓ, resp. w, de V^*, resp. W en une base de V^*, W, et en considérant la base duale correspondante dans V, on voit que $\phi(\ell \otimes w)$ est représentée par une matrice U dont le seul coefficient non nul est $U_{1,1} = \ell(w)$. Sa trace vaut donc $\ell(w)$, et la linéarité de ϕ et de Tr permet de conclure. Voici une façon plus intrinsèque de présenter les choses : la forme bilinéaire $V^* \times V \to K : (\ell, v) \mapsto \ell(v)$ définit une forme linéaire canonique (appelée : contraction) $T : V^* \otimes V \to K$. Alors, $T \circ \phi^{-1}$ est une forme linéaire sur $End(V)$, qui coincide avec Tr sur les endomorphismes de type $\phi(\ell \otimes v)$. Pour la denière assertion, noter que pour tout k, $\phi(\Sigma_i e_i^* \otimes e_i)(e_k) = \Sigma_i \delta_{i,k} e_i = e_k = id_V(e_k)$. On retrouve ainsi la formule $Tr(id_V) = T(\Sigma_i e_i^* \otimes e_i) = \Sigma_i e_i^*(e_i) = n = dimV$ (évidente, mais qui peut servir de définition de la dimension de V).

Traduction matricielle : pour V, W de dimensions n, m finies, et munies, ainsi que leurs duaux, de bases, une application linéaire u de V dans W est représentée par une matrice U, dont le coefficient $U_{i,j}$ de la i-ème ligne et de la j-ième colonne vaut $\epsilon_i^*(u(e_j))$. En particulier, si $u = \phi(e_h^* \otimes \epsilon_k)$, on a $U_{i,j} = \epsilon_i^*\big(e_h^*(e_j)\epsilon_k\big) = e_h^*(e_j)\epsilon_i^*(\epsilon_k) = \delta_{h,j}\delta_{i,k}$. Donc $\phi(e_h^* \otimes \epsilon_k) = \phi_{k,h}$ est représentée par la matrice élémentaire usuelle E_{kh} (seul coefficient non nul, et égal à 1, en k-ième ligne, h-ième colonne). Et $u = \Sigma_{i,j} U_{i,j}\phi_{i,j}$ est l'image par ϕ de $\Sigma_{i,j} U_{i,j} e_j^* \otimes \epsilon_i$. Sous-entendant ϕ, on peut écrire l'image sous u d'un vecteur $v = \Sigma_k x_k e_k$ sous la forme

$$u(v) = \Sigma_{i,j} U_{i,j}(e_j^* \otimes \epsilon_i)(\Sigma_k x_k e_k) = \Sigma_i\big(\Sigma_j U_{i,j} x_j\big)\epsilon_i := \Sigma_i y_i \epsilon_i.$$

Traditionnellement, on pose $U_{i,j} = U^i_j$, $v = (x^1, ..., x^n), u(v) = (y^1, ..., y^n)$, où $y^i = \Sigma_{j=1,...,n} U^i_j x^j$, formule qu'on écrit sous la forme : $y^i = U^i_j x^j$. Il s'agit là de la convention d'Einstein, où un indice apparaissant en positions supérieure et inférieure dans un même terme représente l'effet d'une contraction, et signifie donc qu'on somme sur toutes les valeurs de cet indice. Un autre exemple est donné par $Tr(U) = T(\Sigma_{i,j} U^j_i e_j^* \otimes e_i) = U^i_i$.

Remarques :

i) Pour W de dimension finie, W est isomorphe à son bidual W^{**}, $\mathcal{L}(V, W) \simeq V^* \otimes W, \mathcal{L}(W^*, V^*) \simeq W^{**} \otimes V^* \simeq W \otimes V^*$, et l'opération d'inversion décrite à la remarque (ii) du §1.1 : $V^* \otimes W \simeq W \otimes V^*$ correspond à la transposition des applications linéaires : $\mathcal{L}(V, W) \ni u \mapsto {}^t u \in \mathcal{L}(W^*, V^*)$.

ii) D'après la définition-même de $V \otimes W$, son dual $(V \otimes W)^*$ *s'identifie à l'espace des formes bilinéaires sur* $V \times W$. Supposons maintenant V, W de dimensions finies. On a alors un isomorphisme canonique

$$\kappa_0 : V^* \otimes W^* \simeq (V \otimes W)^*,$$

tel que $(\kappa_0(\ell \otimes m))(v \otimes w) = \ell(v)m(w)$ pour toutes formes linéaires ℓ, m sur V, W, et tous v, w dans V, W. En particulier, pour $V = W$, toute forme bilinéaire sur V s'écrit sous la forme $b = \Sigma_{1 \le i,j \le n} b_{ij} e_i^* \otimes e_j^*$. Traduire en terme de $V^* \otimes V^*$ la condition que b soit symétrique ; qu'elle soit de rang r ; de rang 1. Pour $v = \Sigma_i x^i e_i, w = \Sigma_j y^j e_j$, on a $b(v, w) = \Sigma_{1 \le i,j \le n} b_{ij} x^i y^j$, ce qui s'écrit $b_{ij} x^i y^j$ avec la convention d'Einstein. La distinction entre les matrices U^j_i et b_{ij} est claire : les tenseurs qu'elles représentent vivent dans des espaces $V^* \otimes V, V^* \otimes V^*$ différents.

iii) Pour V, V', W, W' de dimensions finies, le morphisme $\kappa : \mathcal{L}(V, V') \otimes \mathcal{L}(W, W') \to \mathcal{L}(V \otimes W, V' \otimes W')$ équivaut, par le biais de ϕ, à l'isomorphisme

$(V^* \otimes V') \otimes (W^* \otimes W') \simeq (V \otimes W)^* \otimes (V' \otimes W')$ obtenu en composant κ_0, l'associativité et l'inversion des facteurs.

iv) (Produit tensoriel de mesures) Soient X, Y deux espaces topologiques compacts, munis de mesures μ, ν. Pour tout $f \in C(X, \mathbb{R}), g \in C(Y, \mathbb{R})$, on pose $(f \otimes g)\big((x, y)\big) = f(x)g(y)$. Justifier cette écriture et montrer que $C(X, \mathbb{R}) \otimes C(Y, \mathbb{R})$ s'identifie à un sous-espace de $C(X \times Y, \mathbb{R})$ dense pour la topologie de la convergence uniforme. On construit alors une mesure $\mu \otimes \nu$ sur $X \times Y$ vérifiant $(\mu \otimes \nu)(f \otimes g) = \mu(f)\nu(g)$. Justifier cette notation, et énoncer le théorème de Fubini. Pour $X = Y = [01] \subset \mathbb{R}$, munis de la mesure de Lebesgue $\mu = dx, \nu = dy$, on note $dx \otimes dy := dxdy$.

IV.2 Algèbre tensorielle.

En dehors des paragraphes 2.2 et 2.3, la caractéristique de K reste ici quelconque, mais on pourra la supposer nulle en première lecture.

IV.2.1 L'algèbre tensorielle $\otimes V$.

Soient p un entier naturel, et $V_1, ..., V_p$ des espaces vectoriels sur K. Une application de $V_1 \times ... \times V_p$ dans un espace vectoriel U est dite p-linéaire, si pour tout i et tout $(v_1, ..., v_{i-1}, v_{i+1}, ..., v_p)$, sa restriction à $v_1 \times ... \times v_{i-1} \times V_i \times v_{i+1} \times ... \times v_p$ fournit une application linéaire de V_i dans U. En remplaçant bi- par p- au paragraphe précédent, on pose un problème universel, dont la solution existe et est notée $V_1 \otimes ... \otimes V_p$ (à isomorphisme près).

Ainsi, pour $p = 3$, l'application trilinéaire $V_1 \times V_2 \times V_3 \to V_1 \otimes (V_2 \otimes V_3)$: $(v_1, v_2, v_3) \mapsto v_1 \otimes (v_2 \otimes v_3)$ définit une application linéaire canonique de $V_1 \otimes V_2 \otimes V_3$ dans $V_1 \otimes (V_2 \otimes V_3)$. On vérifie, par exemple en prenant des bases, que c'est un isomorphisme, et on identifie désormais $V_1 \otimes V_2 \otimes V_3$ avec $V_1 \otimes (V_2 \otimes V_3)$, et, de même, avec $(V_1 \otimes V_2) \otimes V_3$.

Quand tous les V_i sont égaux à un même V, sa p-ième puissance tensorielle est notée $\otimes^p V$. On pose $\otimes^0 V = K$. Du fait de l'associativité supra (étendue à un nombre quelconque de facteurs), la loi $\otimes^p V \times \otimes^q V \ni (t_1, t_2) \mapsto t_1 \otimes t_2 \in \otimes^{p+q} V$ permet de munir

$$\otimes V := \oplus_{p \geq 0} \otimes^p V$$

d'une structure de K-algèbre associative (et graduée), appelée *l'algèbre tensorielle de V*. Pour V de dimension 1, $\otimes V$ est isomorphe à l'algèbre $K[T]$

des polynômes en une variable. En dimension plus grande, $\otimes V$ n'est pas commutative.

IV.2.2 Tenseurs symétriques et antisymétriques.

Nous supposons ici que K *est de caractéristique nulle* (et en particulier, $\neq 2$).

Le groupe symétrique $\mathfrak{S}_p$ est le groupe, d'ordre $p!$, formé par les permutations de l'ensemble $\{1, ..., p\}$, avec la loi $\sigma\tau = \sigma \circ \tau$. Pour tout $\sigma \in \mathfrak{S}_p$, l'application $(v_1, ..., v_p) \mapsto v_{\sigma^{-1}(1)} \otimes ... \otimes v_{\sigma^{-1}(p)}$ définit un automorphisme $\rho(\sigma)$ de $\otimes^p V$ vérifiant $\rho(\sigma)(v_1 \otimes ... \otimes v_p) = v_{\sigma^{-1}(1)} \otimes ... \otimes v_{\sigma^{-1}(p)}$: en k-ième position, on met le vecteur qui était en $\sigma^{-1}(k)$-ième position, autrement dit, on place en $\sigma(k)$-ième position le vecteur qui était en k-ième position. Par conséquent, pour tout $\tau \in \mathfrak{S}_p$, $\rho(\tau)\big(\rho(\sigma)(v_1 \otimes ... \otimes v_p)\big) = \rho(\tau)(v_{\sigma^{-1}(1)} \otimes ... \otimes v_{\sigma^{-1}(p)}) = v_{\sigma^{-1}(\tau^{-1}(1))} \otimes ... \otimes v_{\sigma^{-1}(\tau^{-1}(p))} = v_{(\tau\sigma)^{-1}(1)} \otimes ... \otimes v_{(\tau\sigma)^{-1}(p)} = \rho(\tau\sigma)(v_1 \otimes ... \otimes v_p)$. On a ainsi défini une action à gauche $(\sigma, t) \mapsto \sigma.t := \rho(\sigma)(t)$ de $\mathfrak{S}_p$ sur $\otimes^p V$.

Un tenseur $t \in \otimes^p V$ est dit symétrique, resp. antisymétrique, si pour tout $\sigma \in \mathfrak{S}_p$, on a $\sigma.t = t$, resp. $\sigma.t = sgn(\sigma)t$, où $sgn(\sigma) = sgn(\sigma^{-1}) = \pm 1$ désigne la signature de la permutation σ. On note $Sym^p V$, resp. $Ant^p V$, le sous-ev de $\otimes^p V$ formé par les tenseurs symétriques, resp. antisymétriques. Clairement : $Sym^p V \cap Ant^p V = \{0\}$, mais dès que $p > 2$ et $dim(V) > 1$ (voir plus bas), $Sym^p V \oplus Ant^p V$ ne remplit pas tout $\otimes^p V$.

L'endomorphisme de $\otimes^p V$ de symétrisation $s_p : t \mapsto s_p(t) = \frac{1}{p!}\Sigma_{\sigma \in \mathfrak{S}_p} \sigma.t$ (resp. d'antisymétrisation $a_p : t \mapsto a_p(t) = \frac{1}{p!}\Sigma_{\sigma \in \mathfrak{S}_p} sgn(\sigma)\sigma.t$) a une image contenue dans $Sym^p V$ (resp. $Ant^p V$), où il induit l'identité. On en déduit que $Im(s_p) = Sym^p V$, resp. $Im(a_p) = Ant^p V$.

Le noyau de s_p est le sous-espace vectoriel N_p engendré par les tenseurs de la forme $\{\sigma.t - t, \sigma \in \mathfrak{S}_p, t \in \otimes^p V\}$. En effet, N_p est clairement tué par s_p, tandis que tout élément t de $\otimes^p V$ est somme de $s_p(t) \in Sym^p(V)$ et de $(id - s_p)(t) = \frac{1}{p!}\Sigma_{\sigma \in \mathfrak{S}_p}(t - \sigma t) \in N_p$. Donc $Ker s_p = N_p$, $\otimes^p V = Sym^p V \oplus N_p$ et s_p est la projection de $\otimes^p V$ sur $Sym^p V$ parallèlement à N_p.

On montre de même que le noyau de a_p est le sous-espace vectoriel J_p engendré par les tenseurs de la forme $\{\sigma.t - sgn(\sigma)t, \sigma \in \mathfrak{S}_p, t \in \otimes^p V\}$, et que a_p est la projection de $\otimes^p V$ sur $Ant^p V$ parallèlement à J_p.

Notons ici que pour $p = 2$, N_2 coïncide avec $Im(a_2)$, J_2 avec $Im(s_2)$, de sorte que $\otimes^2 V = Sym^2 V \oplus Ant^2 V$. Ainsi, $\otimes^2(V^*) = Sym^2(V^*) \oplus Ant^2(V^*)$, soit, pour V de dimension finie (voir la Remarque (ii) du §1.2) : toute forme bilinéaire sur V s'écrit de façon unique comme somme d'une forme bilinéaire symétrique et d'une forme bilinéaire alternée.

IV.2.3 Bases de Sym^pV et de Ant^pV.

Ici, K est encore de caractéristique nulle, et nous supposons que V est un espace vectoriel de dimension finie n, muni d'une base $(e_1, ..., e_n)$.

Dans ces conditions, $\otimes^p V$ admet pour base les n^p tenseurs $e_{i_1} \otimes ... \otimes e_{i_p}$, où $(i_1, ..., i_p)$ parcourt l'ensemble des p-uplets, éventuellement répétés, d'éléments de $\{1, ..., n\} := [1, n]$, c'est-à-dire l'ensemble des applications $I : [1, p] \to [1, n]$, avec $I(1) = i_1, ..., I(p) = i_p$; on écrira alors $e_I = e_{i_1} \otimes ... \otimes e_{i_p}$. Pour n'importe quelle application $J : [1, p] \mapsto [1, n]$, il existe une permutation $\sigma \in \mathfrak{S}_p$ telle que $(j_{\sigma(1)}, ..., j_{\sigma(p)})$ soit une réécriture de $(j_1, ..., j_p)$ en termes croissants au sens large, c.à.d. telle que $I = J \circ \sigma$ soit une application croissante au sens large de même image que J, multiplicités comprises. Les classes de e_I et de e_J modulo N_p sont alors égales. Par conséquent, les images des e_I dans Sym^pV, où I parcourt l'ensemble des applications croissantes $(1 \leq i_1 \leq i_2 \leq ... \leq i_p \leq n)$, engendrent Sym^pV. Je dis maintenant que ces images sont linéairement indépendantes. En effet, le symétrisé d'un tel e_I est la somme d'un multiple non nul de e_I (correspondant aux σ stabilisant I), et d'éléments $e_{j_1} \otimes ... \otimes e_{j_p}$ de la base de $\otimes^p V$ ne correspondant pas à des suites croissantes. Ainsi, *Sym^pV admet pour base $\{s_p(e_{i_1} \otimes ... \otimes e_{i_p}), 1 \leq i_1 \leq ... \leq i_p \leq n\}$, et*

$$dim(Sym^pV) = \binom{n+p-1}{p}.$$

(Une façon de vérifier cette formule consiste à noter que les suites ci-dessus s'identifient aux monômes $T_{i_1}...T_{i_p} = T_1^{d_1}...T_n^{d_n}$ en n variables de degré $p = d_1 + ... + d_n$. Ou encore, qu'elles s'identifient, via $(i_1, ..., i_p) \mapsto (i_1, i_2+1, ..., i_p+p-1)$, aux suites strictement croissantes $[1, p] \mapsto [1, n+p-1]$ c.à.d. aux parties à p éléments d'un ensemble à $n + p - 1$ éléments.)

Passons à Ant^pV, et notons que comme $\mathrm{car}(K) \neq 2$, J_p contient les tenseurs $v_1 \otimes ... \otimes v_p$ pour lesquels $v_i = v_j$ pour deux indices $i \neq j$. En fait, $J_p = J'_p$, où J'_p désigne le sous-espace de $\otimes^p V$ engendré par ces tenseurs et par les $\{\sigma.e_I - sgn(\sigma)e_I, \sigma \in \mathfrak{S}_p,\ I : [1, p] \mapsto [1, n]$ strictement croissante$\}$. On peut donc reprendre le raisonnement précédent, en se limitant aux applications

$I : [1,p] \to [1,n]$ strictement croissantes. On en déduit que Ant^pV *admet pour base* $\{a_p(e_{i_1} \otimes ... \otimes e_{i_p}), 1 \leq i_1 < ... < i_p \leq n\}$. En particulier,

$$dim(Ant^pV) = \binom{n}{p}.$$

Pour $p = 2$, on a bien $dim(Sym^2V) + dim(Ant^2V) = \frac{(n+1)n}{2} + \frac{n(n-1)}{2} = n^2$, mais la somme des dimensions sera $< n^p$ pour $p > 2$ (et $n > 1$).

IV.2.4 Une présentation plus intrinsèque.

Nous ne faisons plus d'hypothèse sur la caractéristique de K , et montrons maintenant que Sym^pV et Ant^p sont solutions de problèmes universels, qui fournissent à leurs classes d'isomorphismes des définitions meilleures, car : (i) elles sont valables en toute caractéristique ; (ii) elles sont mieux adaptées à l'étude de leurs propriétés fonctorielles ; (iii) elles conduisent naturellement aux analogues symétrique et antisymétrique (mieux, alterné) de l'algèbre $\otimes V$. On fixe un espace vectoriel V sur le corps K et un entier p.

Pour toute application $b : V^p \to U$, et tout $\sigma \in \mathfrak{S}_p$, on pose $(\sigma.b)(v_1, ..., v_p)$ $= b(v_{\sigma(1)}, ..., v_{\sigma(p)})$. On dit que b est dite symétrique (resp. antisymétrique) si $\sigma.b = b$ (resp. $sgn(\sigma)b$) pour tout $\sigma \in \mathfrak{S}_p$; et que b est alternée si $b(v_1, ..., v_p) = 0$ dès qu'il existe deux indices $i \neq j$ tels que $v_i = v_j$, ce qui entraîne ($\Leftrightarrow$ en car. $\neq 2$) que b est antisymétrique, puisque $\mathfrak{S}_p$ est engendré par les transpositions.

Existe-t-il un couple (S, s) *(resp.* (Λ, λ) *) formé d'un espace vectoriel* S *(resp.* Λ*) et d'une application p-linéaire symétrique* s *(resp. alternée* λ*) :* $V^p \to S$ *(resp.* Λ*) vérifiant la propriété suivante : pour toute application p-linéaire symétrique (resp. alternée)* b *de* V^p *vers un espace vectoriel* U *quelconque, il existe une* **unique application linéaire** $f = f_b$ *telle que* $b = f_b \circ s$ *(resp.* $b = f_b \circ \lambda$*).*

La réponse est positive. La définition même des sous-espaces N_p (resp. J'_p) du §2.2 (voir aussi la fin de ce paragraphe) montre qu'elle est donnée par $S = (\otimes^pV)/N_p$ (resp. $\Lambda = (\otimes^pV)/J'_p$), muni de l'application canonique s (resp. λ) : $(v_1, ..., v_p) \mapsto$ classe de $v_1 \otimes ... \otimes v_p$ modulo N_p (resp. J'_p). On posera désormais :

$$S^pV := \otimes^pV/N_p \qquad , \qquad \wedge^pV = \otimes^pV/J'_p\,;$$

. . $s(v_1, v_2, .., v_p) = v_1 v_2 ... v_p$, $\lambda(v_1, v_2, ..., v_p) = v_1 \wedge v_2 \wedge ... \wedge v_p$.

Par exemple, les symboles $v.w \in S^2V$ (resp. $v \wedge w \in \wedge^2 V$) sont caractérisés par les relations de bilinéarité que vérifie $v \otimes w$, jointes aux relations $v.w = w.v$ (resp. $v \wedge v = 0$).

On dit que S^pV (resp. $\wedge^p V$) est la p-ième puissance symétrique (resp. extérieure) de V. Les éléments de $\wedge^p V$ sont appelés des *p-vecteurs*, ceux de $\wedge^p(V^*)$ des *p-formes*.

En caractéristique nulle, le §2.2 montre qu'on peut relever les quotients S^pV (resp. $\wedge^p V$) de $\otimes^p V$ en ses sous-espaces Sym^pV (resp. Ant^pV), mais comme l'a déjà fait deviner le calcul de leurs dimensions au §2.3, de tels relèvements sont inutiles, et les mêmes arguments montrent qu'en toute caractéristique :

Théoreme IV.2.1. *pour V de dimension n, de base $\{e_1, ..., e_n\}$:*
- S^pV, de dim. $\binom{n+p-1}{p}$, admet pour base les $\{e_{i_1} ... e_{i_p}, 1 \leq i_1 \leq ... \leq i_p \leq n\}$;
- $\wedge^p V$, de dim. $\binom{n}{p}$, admet pour base les $\{e_{i_1} \wedge ... \wedge e_{i_p}, 1 \leq i_1 < ... < i_p \leq n\}$.

(Il est d'ailleurs clair sur sa propriété universelle que $\wedge^p V = \{0\}$ pour $p > dimV$, toute forme p-linéaire alternée sur V^p étant alors identiquement nulle.)

Ainsi, pour $p = n$, $\wedge^n V$ est une droite, engendrée par le n-vecteur $e_1 \wedge ... \wedge e_n$. Pour tout n-uplet $\mathbf{v} = (v_1, ..., v_n)$ de vecteurs de V, on appelle *déterminant de la famille $\{v_1, ..., v_n\}$ relativement à la base* $\mathbf{e} = (e_1, ..., e_n)$ l'unique scalaire $det_{\mathbf{e}}(\{\mathbf{v}\}) \in K$ tel que $v_1 \wedge ... \wedge v_n = det_{\mathbf{e}}(\{\mathbf{v}\})\, e_1 \wedge ... \wedge e_n$. Si $v_j = \Sigma_{i=1,...,n} x_{i,j} e_i$, on a : $v_1 \wedge ... \wedge v_n = \Sigma_{i_1,...,i_n} x_{i_1,1} x_{i_2,2} ... x_{i_n,n}\; e_{i_1} \wedge ... \wedge e_{i_n}$, où $(i_1, ..., i_n)$ parcourt l'ensemble de tous les n-uplets ordonnés d'éléments distincts de $\{1, .., n\}$, d'où :

$$det_{\mathbf{e}}(v_1, ..., v_n) = \Sigma_{\sigma \in \mathfrak{S}_n}\; sgn(\sigma)\; x_{1,\sigma(1)} ... x_{n,\sigma(n)}$$

Mais $det_{\mathbf{e}}(\{\mathbf{v}\})$ ne dépend que de $v_1 \wedge ... \wedge v_n$ et de $e_1 \wedge ... \wedge e_n$. Déduire de cette remarque que $det_{\mathbf{e}}$ est une forme n-linéaire alternée sur V, et retrouver toutes les propriétés élémentaires du déterminant.

Remarquons enfin que puisque $\mathfrak{S}_p$ est engendré par les transpositions $(i, i+1), 1 \leq i < p$, une forme p-linéaire b sur V (c'est-à-dire sur V^p) est symétrique si et seulement si elle s'annule sur le sous-espace N''_p engendré par les tenseurs de la forme $v_1 \otimes .. \otimes v_{i-1} \otimes (v_i \otimes v_{i+1} - v_{i+1} \otimes v_i) \otimes v_{i+2} \otimes ... \otimes v_p$. On en déduit (on aurait pu le montrer directement) que $N''_p = N_p$. De même, une forme p-linéaire b sur V est alternée si et seulement si elle s'annule sur le sous-espace J''_p engendré par les tenseurs de la forme $v_1 \otimes .. \otimes v_{i-1} \otimes v_i \otimes v_i \otimes v_{i+2} \otimes ... \otimes v_p$. On en déduit (ou on montre directement) que $J'_p = J''_p$ (qui, en car. $\neq 2$, coïncide avec J_p).

IV.2.5 Algèbres symétrique et extérieure

Soit maintenant $q \geq 0$ un second entier. Comme $N_p \otimes (\otimes^q V)$ et $(\otimes^p V) \otimes N_q$ sont contenus dans N_{p+q}, la loi de produit $(\otimes^p V) \times (\otimes^q V) \to \otimes^{p+q} V$ définie au §2.1 induit par passage au quotient une application bilinéaire canonique $S^p V \times S^q V \to S^{p+q} V : (t_1, t_2) \mapsto t_1.t_2 :=$ classe de $\tilde{t}_1 \otimes \tilde{t}_2$ modulo N_{p+q} (indépendante du choix des relevés $\tilde{t}_1, \tilde{t}_2$ de t_1, t_2 dans $\otimes^p V, \otimes^q V$). On a ainsi muni

$$S(V) := \oplus_{p \geq 0} S^p V$$

(avec $S^0 V = K$) d'une structure de K-algèbre graduée, associative et commutative, appelée *l'algèbre symétrique de V*. Pour V de dimension finie n, le théorème 2.1 montre que $S(V)$ s'identifie à l'algèbre $K[T_1, ..., T_n]$ des polynômes en n variables à coefficients dans K. Pour V de dimension 1, $\otimes V = S(V) \simeq K[T]$, et $K[T] \otimes K[T] \simeq K[T_1, T_2]$. De façon générale, $K[X_1, ..., X_n] \otimes K[Y_1, ..., Y_m] \simeq K[X_1, ..., X_n, Y_1, ..., Y_m]$.

Remarque : en fait, c'est à $S(V^*)$ qu'il vaut mieux identifier $K[T_1, ..., T_n]$. En effet, $T_1, ..., T_n$ peuvent être vus comme des formes linéaires sur V formant une base de V^*, et les monômes $\{T_1^{d_1}, ...T_n^{d_n}, d_1 + ... + d_n = p\}$, qui s'interprètent naturellement comme des formes p-linéaires symétriques sur V, forment une base de $S^p(V^*)$, .

De même, $J'_p \otimes (\otimes^q V)$ et $(\otimes^p V) \otimes J'_q$ sont contenus dans J'_{p+q}, et la loi de produit $(\otimes^p V) \times (\otimes^q V) \to \otimes^{p+q} V$ induit par passage au quotient une application bilinéaire canonique $\wedge^p V \times \wedge^q V \to \wedge^{p+q} V$, notée : $(t_1, t_2) \mapsto t_1 \wedge t_2$; ce $(p+q)$-vecteur s'appelle le produit extérieur du p-vecteur t_1 par le q-vecteur t_2. Pour V de dimension n, on a ainsi muni

$$\wedge(V) := \oplus_{p=0,...,n} \wedge^p V$$

(avec $\wedge^0 V = K$) d'une structure de K-algèbre graduée de dimension 2^n, associative et, si $n > 1$, non commutative, appelée *l'algèbre extérieure de V*. On a néanmoins la propriété d'anticommutativité suivante :

$$\forall t_1 \in \wedge^p V, t_2 \in \wedge^q V, \; t_2 \wedge t_1 = (-1)^{pq} t_1 \wedge t_2 \text{ dans } \wedge^{p+q} V,$$

car pour passer de $(1, ..., p, p+1, ..., p+q)$ à $(p+1, ..., p+q, 1, ..., p)$, il suffit de faire pq chevauchements. En particulier, au moins en car. $\neq 2$, si t est un p-vecteur avec p impair, $t \wedge t = 0$ dans $\wedge^{2p} V$ (quand $p = 1$, on le savait déjà, par définition) ; mais si p est pair, $t \wedge t$ n'a aucune raison d'être nul : par exemple, pour $dimV \geq 4$, $t = e_1 \wedge e_2 + e_3 \wedge e_4 \in \wedge^2 V$ a pour carré extérieur $t \wedge t = 2e_1 \wedge e_2 \wedge e_3 \wedge e_4 \neq 0$ dans $\wedge^4 V$.

IV.2.6 Propriétés fonctorielles

Du fait de leurs applications géométriques, nous allons nous concentrer sur le cas des puissances extérieures, mais la théorie peut être développée de façon parallèle pour les puissances symétriques. Pour simplifier, nous supposons désormais que notre espace vectoriel V est de dimension n finie, muni d'une base $\mathbf{e} = (e_1, ..., e_n)$. Nous désignons par $(e_1^*, ..., e_n^*)$ la base de V^* duale, et pour tout $p \geq 0$:

• par $S(p, n)$ l'ensemble, à $\binom{n}{p}$ éléments, des suites strictement croissantes $J : [1, p] \to [1, n]$, qu'on identifie aux parties à p éléments $\{j_1 < j_2 < ... < j_p\}$ de $\{1, ..., n\}$, avec $J(i) = j_i$;

• par $\{e_J = e_{j_1}, \wedge ... \wedge e_{j_n}, J \in S(p, n)\}$ la base de $\wedge^p V$ donné par le théorème 2.1 ;

• par $\{e_J^* = e_{j_1}^*, \wedge ... \wedge e_{j_n}^* J \in S(p, n)\}$ la base de $\wedge^p(V^*)$ définie de la même façon.

Le caractère "fonctoriel" des constructions tensorielles conduit aux propositions-définitions suivantes.

Soit $u : V \to W$ une application linéaire. Il existe une unique application linéaire, notée $\wedge^p u : \wedge^p V \to \wedge^p W$ telle que pour tout $(v_1, ..., v_p) \in V^p$, on ait $(\wedge^p u)(v_1 \wedge ... \wedge v_p) = u(v_1) \wedge ... \wedge u(v_p)$.

En effet, l'application $b : V^p \to \wedge^p W : (v_1, ..., v_p) \mapsto u(v_1) \wedge ... \wedge u(v_p)$ est une application p-linéaire alternée, et on note $f_b := \wedge^p u$ l'application linéaire qu'elle induit sur $\wedge^p V$. On prendra garde au fait que la notation $\wedge^p u$ ne désigne *pas* un élément de $\wedge^p(\mathcal{L}(V, W))$; il n'y a pas de confusion possible, $\wedge^p u$ ne pouvant être que nul dans ce deuxième sens. Remarquons également que si $i : V_1 \hookrightarrow V$ est une injection, l'application $\wedge^p i$ est encore injective, et permet d'identifier $\wedge^p V_1$ à un sous-espace vectoriel de $\wedge^p V$. Pour $V = W$, et $p = n$, on obtient :

Soient V un ev de dimension n, et $u \in End(V)$. Alors, $\wedge^n u$ s'identifie à un scalaire, noté $det(u) \in K$ et appelé le déterminant de l'endomorphisme u *de V. Pour tout $u, u' \in End(V)$, on a : $det(u'u) = det(u')det(u)$.*

En effet, pour tout ev D de dimension 1, la contraction $D^* \otimes D \to K$ montre que $End(D)$ s'identifie de façon canonique à l'anneau K. Donc $\wedge^n u \in End(\wedge^n V)$ est un scalaire $det(u)$ *indépendant* du choix d'une base. Par ailleurs, $u' \circ u$ vérifie $\wedge^n(u' \circ u)(v_1 \wedge ... \wedge v_n) = u'(u(v_1)) \wedge ... \wedge u'(u(v_n)) = \wedge^n u'\big(\wedge^n u(v_1 \wedge ... \wedge v_n)\big)$, et $\wedge^n(u' \circ u) = \wedge^n u' \circ \wedge^n u$. Si $\mathbf{e} = (e_1, ..., e_n)$ est une base de V, on a $\wedge^n u(e_1, ..., e_n) = u(e_1) \wedge ... \wedge u(e_n)$, donc

$$det(u) = det_{\mathbf{e}}(u(e_1), ..., u(e_n)),$$

comme espéré. En développant l'expression $Car_u(\lambda) = \wedge^n(u - \lambda id_V)$ définissant le polynôme caractéristique de u, on retrouve pour la trace de u la formule $Tr(u) = \Sigma_{i=1,...,n} e_i^*(u(e_i))$.

Théoreme IV.2.2. *: Il existe un isomorphisme canonique $\wedge^p(V^*) \simeq (\wedge^p V)^*$ tel que pour tout $(\ell_1, ..., \ell_p) \in (V^*)^p$ et tout p-vecteur $v_1 \wedge .... \wedge v_p \in \wedge^p V$, la p-forme $\ell_1 \wedge ... \wedge \ell_p \in \wedge^p(V^*)$, vue dans $(\wedge^p V)^*$, vérifie*

$$(\ell_1 \wedge ... \wedge \ell_p)(v_1 \wedge .... \wedge v_p) = det\big(\ell_i(v_j)\big)_{1 \leq i,j \leq p}.$$

Dans cette identification de $\wedge^p(V^*)$ avec le dual de $\wedge^p V$, la base $\{e_J^*, J \in S(p,n)\}$ est la base duale de la base $\{e_J, J \in S(p,n)\}$ de $\wedge^p V$.

Preuve : pour tout p-uplet $\ell = (\ell_1, ..., \ell_p) \in (V^*)^p$, la forme p-linéaire sur V définie par $b_\ell : (v_1, ..., v_p) \mapsto det\big(\ell_i(v_j)\big)$ est alternée, donc définit une forme linéaire f_{b_ℓ} sur $\wedge^p V$, c'est-à-dire un élément f_ℓ de $(\wedge^p V)^*$ vérifiant $f_\ell(v_1 \wedge ... \wedge v_p) = det\big(\ell_i(v_j)\big)$, et l'application $\ell \mapsto f_\ell$ de $(V^*)^p$ dans $(\wedge^p V)^*$ est p-linéaire alternée, donc définit une application linéaire canonique : $\phi : \wedge^p(V^*) \to (\wedge^p V)^*$ telle que $\phi(\ell_1 \wedge ... \wedge \ell_p) = f_\ell$. Pour tout $I, J \in S(p,n)$, l'image de e_J^* par ϕ vérifie : $\phi(e_J^*)(e_I) = det(e_{j_h}^*(e_{i_k}), 1 \leq h, k \leq p)$, qui vaut 1 si $I = J$, et 0 sinon. Donc ϕ est un isomorphisme de $\wedge^p(V^*)$ sur $(\wedge^p V)^*$ envoyant la base $\{e_J^*\}$ sur la base duale de $\{e_J\}$.

Soient $q \leq p$ deux entiers compris entre 1 et n, et $v_j = \Sigma_{i=1,...,n} x_{i,j} e_i$, $j = 1, ..., p$ un p-uplet d'éléments de V. On note $X = (x_{ij}) \in Mat_{n,p}$ la matrice à n lignes et p colonnes représentant ce p-uplet dans la base $\mathbf{e}$ de V. Pour toute partie $J \in S(q,p)$, considérons le q-vecteur $v_J = v_{j_1} \wedge ... \wedge v_{j_q}$. Ses coordonnées relativement à la base $\{e_I, I \in S(q,n)\}$ de $\wedge^q V$ s'écrivent :

$$v_J = \Sigma_{I \in S(q,n)} X_{I,J} e_I, \text{ avec } X_{I,J} = det(x_{i,j})_{i \in I, j \in J},$$

qui s'appelle le (determinant) *mineur* d'ordre q, d'indices lignes I, d'indices colonnes J, de la matrice X. Pour vérifier cette formule, il suffit d'évaluer sur v_J la base duale de $\wedge^q V^*$ formée par $\{e_I^*, I \in S(q,n)\}$. On obtient : $e_I^*(v_J) = det(e_{i_h}^*(v_{j_k}))_{1\leq h,k\leq q} = det(x_{i_h,j_k})_{1\leq h,k\leq q} = det(x_{i,j})_{i\in I, j\in J}$. Citons comme corollaire la

Formule de Lagrange : *soient $k \leq p, q \leq n$ quatre entiers, et $Y \in Mat_{p,n}, Z \in Mat_{n,q}$ deux matrices rectangulaires, de sorte que $X = YZ \in Mat_{p,q}$. Soient $I \in S(k,p), J \in S(k,q)$. Alors le mineur d'ordre k, d'indices (I,J), de X est donné par*

$$X_{I,J} = \Sigma_{H\in S(k,n)} Y_{I,H} Z_{H,J}.$$

En effet, soient $(\ell_i, i \in I)$ les k formes linéaires sur $V = K^n$ apparaissant sur les lignes correspondantes de Y, et $(v_j, j \in J)$ les k vecteurs de V apparaissant sur les colonnes correspondantes de Z. Alors, $X_{I,J} = (\ell_{i_1} \wedge ... \wedge \ell_{i_k})(v_{j_1}, ..., v_{j_k})$. Mais $v_J = \Sigma_{H\in S(k,r)} Z_{H,J} e_H$, et comme $\{e_i\}$ est la base duale de $\{e_i^*\}$ dans la bidualité $V^{**} \simeq V$, on a de même : $\ell_I = \Sigma_{H'\in S(k,r)} Y_{I,H'} e_{H'}^*$. Alors, $\ell_I(v_J) = \Sigma_{H,H'} Y_{I,H'} Z_{H,J} e_{H'}^*(e_H) = \Sigma_H Y_{I,H} Z_{H,J}$.

IV.2.7 Groupes algébriques linéaires

Soit V un K-espace vectoriel de dimension finie. On appelle construction tensorielle de V tout espace vectoriel $C(V)$ construit à partir de V en itérant les procédés suivants : si W est une construction de V, il en est de même de son dual W^*, de ses puissances tensorielles, symétriques et alternées $\otimes^p W, S^p W, \wedge^p W$; si W_1, W_2 sont deux constructions de V, il en est de même de leur somme directe $W_1 \oplus W_2$ et de leur produit tensoriel $W_1 \otimes W_2$.

Soit alors $u \in GL(V)$ un automorphisme de V. Comme on l'a vu ci-dessus dans plusieurs cas, u définit naturellement sur toute construction $C(V)$ de V un automorphisme $C(u) \in GL(C(V))$ de $C(V)$. Par exemple, si $C(V) = V^*$ est le dual de V, $C(u) = {}^t u^{-1} \in GL(V^*)$, où ${}^t u$ est le transposé de u ; pour $C(V) = \otimes^p V$, $C(u) := \otimes^p u$ vérifie : $C(u)(v_1 \otimes ... \otimes p) = u(v_1) \otimes ... \otimes u(v_p)$; propriété similaire pour $C(V) = S^p V, \wedge^p V$. Et $C_1(u) \oplus C_2(u), C_1(u) \otimes C_2(u)$ sont bien des automorphismes de $C_1(V) \oplus C_2(V), C_1(V) \otimes C_2(V)$.

Si u, u' sont deux automorphismes de V, l'automorphisme $C(u \circ u')$ de $C(V)$ est égal à $C(u) \circ C(u')$. De plus, $C(id_V) = id_{C(V)}$. Par conséquent, toute construction tensorielle $C(V)$ de V est naturellement munie d'une action à gauche du groupe $GL(V)$.

Étant donnés un ensemble fini de constructions $C_1(V), ..., C_N(V)$ de V, et dans chacune d'elle, un sous-espace vectoriel $W_i \subset C_i(V)$, on peut donc attacher au sous-espace vectoriel $W = W_1 \oplus ... \oplus W_N$ de la construction $C(V) = C_1(V) \oplus ... \oplus C_N(V)$ son stabilisateur

$$Stab(W) = \{u \in GL(V), C(u)(W) = W\},$$

qui est un sous-groupe de $GL(V)$. Dans une base de V où u est représentée par une matrice $U \in GL_n$, la condition $C(u)(W) = W$, ou de façon équivalente (puisque $C(u)$ est un automorphisme) $C(u)(W) \subset W$, se traduit par des équations algébriques à coefficients dans K liant les coefficients U_{ij} de la matrice U. De tels sous-groupes de $GL(V)$ sont dit algébriques, et on montre inversement que tout sous-groupe algébrique G de $GL(V)$ est le stabilisateur d'un sous-espace vectoriel d'une construction de V. Par exemple, le sous-groupe G de $GL_n(K)$ formé par les matrices triangulaires supérieures est le stabilisateur de $W = W_1 \oplus ... \oplus W_{n-1} \subset V \oplus .... \oplus V$, où pour tout i, W_i est le sous-espace $Ke_1 \oplus .. \oplus Ke_i$ de $V = K^n$.

Plutot qu'une collection $\{W_i\}$ de sous-espaces des constructions $C_1(V), ...,$ $C_N(V)$, on peut également se donner pour chaque i un tenseur $t_i \in C_i(V)$, ou de façon équivalente, un tenseur $t = t_1 \oplus ... \oplus t_N \in C(V)$, et lui associer son fixateur

$$Fix(t) = \{u \in GL(V), C(u)(t) = t\},$$

qui est encore un sous-groupe algébrique de $GL(V)$. Les sous-groupes algébriques de $GL(V)$ ne sont pas tous des fixateurs, mais c'est le cas de la plupart des "groupes classiques" rencontrés au chapitre III. Ainsi,

$SL_n(K) = Fix(t)$, où $t = e_1 \wedge ... \wedge e_n \in \wedge^n(V)$.

$O_n(K) = Fix(t)$, où $t = e_1^* \otimes e_1^* + ... + e_n^* \otimes e_n^* \in V^* \otimes V^*$.

On notera que le groupe unitaire $U(n)$ n'est pas un sous-groupe algébrique de $GL_n(\mathbb{C})$.

IV.3 Applications géométriques

IV.3.1 Grassmaniennes

Soient V un ev de dimension n sur K, et $p \in [0, n]$ un entier. On appelle grassmanienne d'ordre p de V l'ensemble des sous-ev de dimension p de V. On le note $Gr_p(V)$, ou $Gr(p,n)$ quand $V = K^n$. Par exemple, $Gr_1(V)$ est l'espace projectif $\mathbb{P}(V) \simeq \mathbb{P}_{n-1}$ des droites de V. Pour tout élément W de $Gr_p(V)$, le sous-espace $\wedge^p W$ de $\wedge^p V$ est une droite, donc un élément de l'espace projectif $\mathbb{P}(\wedge^p V) \simeq \mathbb{P}_N$, de dimension $N = \binom{n}{p} - 1$. On appelle plongement de Plücker l'application

$$\gamma : Gr_p(V) \to \mathbb{P}(\wedge^p V) : W \mapsto \wedge^p W.$$

L'énoncé suivant justifie (partiellement) cette dénomination.

Proposition IV.3.1. *: i) Soient $v_1, ..., v_p$ (resp. $w_1, ..., w_p$) des vecteurs de V linéairement indépendants. Alors, les $v_1 \wedge ... \wedge v_p$ et $w_1 \wedge ... \wedge w_p$ sont linéairement dépendants dans $\wedge^p V$ si et seulement si les sous-espaces vectoriels (de dimension p) de V engendrés par $\{v_1, ..., v_p\}$ et par $\{w_1, ..., w_p\}$ coïncident.*

ii) Le plongement de Plücker est injectif, et son image est formée des droites de $\wedge^p V$ engendrées par des p-vecteurs purs (c'est-à-dire : décomposables, c.à.d. de la forme $v_1 \wedge ... \wedge v_p$).

Preuve : i) Soit W' (resp. W) le sous-ev engendré par les v_i (resp. les w_i). Si $W \subset W'$, $w_1 \wedge ... \wedge w_p$ est une combinaison linéaire des $v_{\sigma(1)} \wedge ... \wedge v_{\sigma(p)}, \sigma \in \mathfrak{S}_p$, donc un multiple de $v_1 \wedge ... \wedge v_p$. Inversément, si $w_1 \wedge ... \wedge w_p$ est un multiple (automatiquement non nul) de $v_1 \wedge ... \wedge v_p$, et si l'un des w_i, soit w_1, n'appartient pas à W', il existe une base de V complétant la famille libre $\{v_1, ..., v_p, w_1\}$. Les coordonées non nulles de $w_1 \wedge ... \wedge w_p$ dans la base correspondante de $\wedge^p V$ font toutes intervenir w_1, alors que par hypothèse, sa seule coordonnée non nulle est portée par $v_1 \wedge ... \wedge v_p$, où w_1 n'intervient pas. Contradiction.

ii) On vient de voir que si $W \neq W' \in Gr_p(V)$, les p-vecteurs qui portent les points $\gamma(W)$ et $\gamma(W')$ ne peuvent être proportionnels, donc $\gamma(W) \neq \gamma(W')$. La deuxième assertion est un réécriture de la définition des $\gamma(W)$.

Fixons une base $\mathbf{e} = \{e_1, ..., e_n\}$ de V, et soit $X = (x_{i,j})_{1 \leq i \leq n, 1 \leq j \leq p} \in Mat_{n,p}$ la matrice représentant le p-uplet $(v_1, .., v_p)$ relativement à $\mathbf{e}$. Alors, $v_1 \wedge ... \wedge v_p = \Sigma_{I \in S(p,n)} X_{I,[1,..,p]} e_I$. À homothétie près, la collection de ces mineurs ne dépend que de W et de l'identification de $\mathbb{P}(V)$ à $\mathbb{P}_{n-1}$ que fournit $\mathbf{e}$. On appelle $(\{\lambda X_{I,[1,..,p]}, I \in S(p,n), \lambda \in K^*\}$ les coordonnés de Plücker de W. Elles sont liées par des relations de dépendance algébriques (indépendantes de W), qui montrent que $\gamma(Gr(p,n))$ est une intersection de quadriques de $\mathbb{P}_N$.

IV.3.2 Aires p-dimensionnelles

Soit V un espace vectoriel sur $\mathbb{R}$, de dimension n, de sorte que $\wedge^n V$ est une droite réelle. On appelle *orientation* de V le choix d'une direction $\mathbb{R}^{>0}.\omega$ sur cette droite.

Supposons maintenant V muni d'un produit scalaire b. Si $\mathbf{e}$ et $\mathbf{e}'$ sont deux bases orthonormées de V, alors $det_{\mathbf{e}} = \pm det_{\mathbf{e}'}$ dans $\wedge^n(V^*)$, et on peut attacher à tout n-uplet $(v_1, ..., v_n)$ de vecteurs de V le nombre

$$Vol_{b,n}(v_1, ..., v_n) = |det_{\mathbf{e}}(v_1, ..., v_n)|^{\frac{1}{2}} = (det(b(v_i, v_j))_{1 \le i,j \le n})^{\frac{1}{2}}.$$

La formule de changement de variables dans les intégrales multiples montre que c'est la mesure du parallélépipède porté par $(v_1, ..., v_n)$ relativement à la mesure de Lebesgue sur $\mathbb{R}^n$ (identifié à V par le choix de la base $\mathbf{e}$). S'il est non nul, et si V est muni d'une orientation, on peut affecter ce nombre du signe de $v_1 \wedge ... \wedge v_n$ relativement à ω. On peut donc définir un volume algébrique n-dimensionnel pour les espaces vectoriels euclidens et orientés de dimension n.

Soit en revanche p un entier compris entre 1 et $n-1$. Les sous-espaces vectoriels W de dimension p de V sont encore munis d'un produit scalaire (la restriction de b à W restant définie positive), mais ils n'ont *plus* d'orientation. Si $(v_1, ..., v_p)$ sont p éléments de V linéairement indépendants, on peut donc parler seulement de l'aire p-dimensionnelle

$$Vol_{b,p}(v_1, ..., v_p) = (det(b(v_i, v_j))_{1 \le i,j \le p})^{\frac{1}{2}}$$

(sans signe) de $(v_1, ..., v_p)$. Elle vérifie :

Proposition IV.3.2. (théorème de Pythagore) *: le carré de l'aire p-dimensionnelle d'un parallélépipède de dimension p de $\mathbb{R}^n$ est égal à la somme des carrés des aires (p-dimensionnelles) de ses projections sur les $\binom{n}{p}$ faces de dimension p de $\mathbb{R}^n$.*

Preuve : soit $X = (x_{i,j})_{1 \le i \le n, 1 \le j \le p} \in Mat_{n,p}$ la matrice représentant le p-uplet $(v_1, .., v_p)$ relativement à la base orthonormée usuelle $\mathbf{e}$ de $\mathbb{R}^n = V$. On a $(b(v_i, v_j))_{1 \le i,j \le p} = {}^tX.X$, et la formule de Lagrange, qu'on appelle alors formule de Cauchy-Binet, donne :

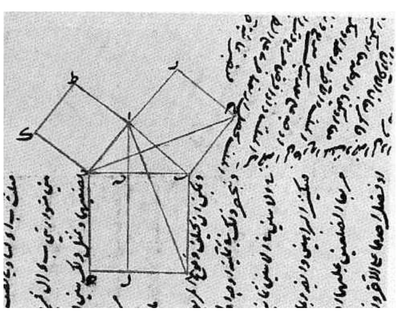
$$det({}^tX.X) = \Sigma_{H \in S(p,n)} \; ({}^tX)_{[1,...,p],H} X_{H,[1,..,p]} = \Sigma_{H \in S(p,n)} (X_{H,[1,..,p]})^2.$$

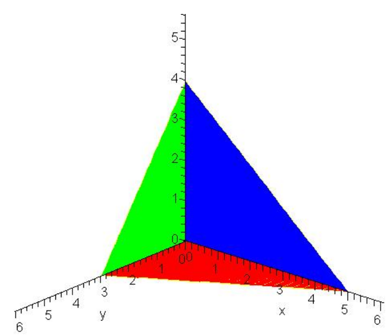
On conclut en notant que $X_{[h_1,...,h_p],[1,...,p]}$ est la matrice représentative des projections de $v_1, ..., v_p$ sur la face $\{0\} \times \mathbb{R}_{h_1} \times \{0\} \times ... \times \{0\} \times \mathbb{R}_{h_p} \times \{0\}$ de $\mathbb{R}^n$.

On peut présenter cette preuve de la façon équivalente suivante. Pour toute forme bilinéaire b sur E, il existe sur $\wedge^p V$ une forme bilinéaire canonique $\wedge^p b$ telle que

$$(\wedge^p b)(v_1 \wedge ... \wedge v_p, w_1 \wedge ... \wedge w_p) = det(b(v_i, w_j))_{1 \leq i,j \leq p}$$

pour tous v_i, w_j dans V. Si b est un produit scalaire, il en est de même pour $\wedge^p b$, et si $\mathbf{e} = (e_1, ..., e_n)$ est orthormée pour ce produit scalaire, la base $\{e_I, I \in S(p,n)\}$ de $\wedge^p V$ est orthormée pour $\wedge^p b$. Les coordonnées de $v_1 \wedge ... \wedge v_p$ dans cette base étant données par les mineurs $X_{H,[1,...,p]}$, le théorème de Pythagore usuel (avec "p" $= 1$, "n" $= \binom{n}{p}$) permet de conclure.





. Le cas $p = 1, n = 2$
. [Justifier, puis traduire...]

Un corollaire du cas $p = 2, n = 3$
[Calculer l'aire du triangle oblique]