Université Pierre et Marie Curie

Master de Sciences et Technologies

Mention Mathématiques et applications

M1, Année 2016–2017

4M001

GÉOMÉTRIE AFFINE ET PROJECTIVE

Patrick Polo

Patrick Polo
Université P. & M. Curie
Institut de Mathématiques de Jussieu (UMR 7586 du CNRS)
URL : http://webusers.imj-prg.fr/~patrick.polo
E-mél : patrick.polo@upmc.fr

CHAPITRE 1

ACTIONS DE GROUPES, ESPACES ET APPLICATIONS AFFINES

[(1)] Ceci est le chap. 1 du polycopié 2015-16. En 2016-17, certains numéros ne seront pas traités en cours, afin de dégager du temps pour parler de l'algèbre extérieure d'un espace vectoriel et, si possible, des variétés grassmanniennes. Plutôt que de supprimer des numéros, on indiquera en tête (ou fin) de chapitre les numéros qui n'ont pas été traités. Leur présence dans le polycopié peut donc être considéré comme des compléments de cours.

Ainsi, les numéros 1.8, 1.9 et 1.12 à 1.15 n'ont pas été traités (mais ils le seront peut-être en TD). Le numéro 3.4 ne sera pas traité.

1. Actions de groupes

Rappels 1.0. — On rappelle qu'un *groupe* est un ensemble non vide G muni d'une « loi de composition » $G \times G \to G$, $(g,h) \mapsto gh$, qui vérifie les trois propriétés suivantes :

(1) (Associativité) Pour tout $x, y, z \in G$, $(xy)z = x(yz)$.

(2) (Neutre) Il existe un élément e de G tel que $eg = g = ge$ pour tout $g \in G$. Cet élément est appelé l'élément neutre de G. Il est unique (car si e' est aussi élément neutre, alors $e' = ee' = e$).

(3) (Inverse) Pour tout $g \in G$, il existe un élément $g' \in G$ tel que $gg' = e = g'g$. Cet élément est appelé l'inverse de g et est noté g^{-1}. Il est unique, car si h est un autre inverse, alors $h = h(gg') = (hg)g' = g'$.

Si de plus la loi de composition vérifie $gh = hg$ pour tout $g, h \in G$, on dit que G est un groupe *commutatif* ou *abélien*. Dans ce cas, la loi est souvent notée $+$ et le neutre est noté 0. Par exemple, $(\mathbb{Z}, +)$ est un groupe abélien, de même que $(\mathbb{Z}/n\mathbb{Z}, +)$.

Exemples. — Soit k un corps (par exemple $k = \mathbb{R}$).

(1) Si E est un k-espace vectoriel, alors E muni de l'addition est un groupe abélien.

(2) Le groupe $\mathrm{GL}(E)$ (isomorphe à $\mathrm{GL}_n(k)$ si $\dim(E) = n$) est non commutatif si $\dim(E) > 1$.

(3) Soit $n \in \mathbb{N}^*$. On note S_n le groupe de toutes les *permutations* de $\{1, \dots, n\}$, i.e. des bijections de l'ensemble $\{1, \dots, n\}$ dans lui-même. C'est un groupe non commutatif si $n > 2$. Par exemple, si on note (ij) la *transposition* qui échange i et j alors $(12)(23)$ est la permutation

$$c = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

qui est un 3-cycle et est notée (123). Mais $(23)(12)$ est le 3-cycle (132), égal à c^{-1}.

[(1)] Version du 4 octobre 2016 : une coquille corrigée dans la démo de 7.2.

***Définition 1.1* (Actions).** — Soient G un groupe et X un ensemble. Une *action à gauche* de G sur X est une application $G \times X \to X$, $(g, x) \mapsto g \cdot x$, qui vérifie les deux propriétés suivantes :

(1) Pour tout $g, h \in G$ et $x \in X$, $g \cdot (h \cdot x) = (gh) \cdot x$.

(2) Pour tout $x \in X$, $e \cdot x = x$.

De même, une *action à droite* de G sur X est une application $X \times G \to X$, $(x, g) \mapsto x \cdot g$, qui vérifie les deux propriétés suivantes :

(1) Pour tout $g, h \in G$ et $x \in X$, $(x \cdot h) \cdot g = x \cdot (hg)$.

(2) Pour tout $x \in X$, $x \cdot e = x$.

Souvent on omettra le $\cdot$ et l'on écrira simplement gx (resp. xg) au lieu de $g \cdot x$ (resp. $x \cdot g$).

Exemples 1.2. — 1) Soient G un groupe et H un sous-groupe. On a une action à gauche (resp. à droite) de H sur G définie par $h \cdot g = hg$ (resp. $g \cdot h = gh$). La vérification des propriétés (1) et (2) est immédiate. Cette action s'appelle l'action de H sur G par translations à gauche (resp. à droite).

2) S_n agit à gauche sur $\{1, \dots, n\}$.

3) On munit $\mathbb{R}^n$ du produit scalaire standard : $(x \mid y) = x_1 y_1 + \cdots + x_n y_n$ et l'on note $O(n)$ le sous-groupe de $\mathrm{GL}_n(\mathbb{R})$ formé des automorphismes qui préservent ce produit scalaire. Alors $O(n)$ agit à gauche sur $\mathbb{R}^n$ et aussi, pour tout $r \geq 0$, sur la sphère

$$S^{n-1}(r) = \{x \in \mathbb{R}^n \mid (x \mid x) = r^2\}.$$

Remarques 1.3. — Dans les exemples précédents, on voit à l'oeuvre deux principes généraux utiles. Soit donnée une action à gauche $\lambda : G \times X \to X$, $(g, x) \mapsto gx$.

(1) Pour tout sous-groupe H de G, la restriction de λ à $H \times X$, définie par $(h, x) \mapsto hx$, est une action à gauche de H sur X. Idem bien sûr pour les actions à droite. Ainsi, l'exemple 1) plus haut est la restriction à H de l'action de G sur lui-même par translations à gauche (ou à droite).

(2) Soit Y un sous-ensemble de X *stable par* G, i.e. tel que $gy \in Y$ pour tout $y \in Y$. Alors la restriction de λ à $G \times Y$, définie par $(g, y) \mapsto gy$, est une action à gauche de G sur Y. (Idem bien sûr pour les actions à droite.)

L'exemple 3) plus haut est une combinaison de ces deux points : l'action à gauche de $\mathrm{GL}_n(\mathbb{R})$ sur $\mathbb{R}^n$ induit, par restriction, une action à gauche de $O(n)$ sur $\mathbb{R}^n$. Chaque sphère $S^{n-1}(r)$ est stable pour cette action, donc on obtient une action à gauche de $O(n)$ sur $S^{n-1}(r)$.

***Exemple 1.4* (Important).** — Soient X un ensemble non vide et n un entier ≥ 2. Alors, « on voit bien » que le groupe S_n agit sur le produit $X^n = \{(x_1, \dots, x_n) \mid x_i \in X\}$, disons à gauche. Mais quelle est la bonne formule :

$$(*) \qquad \sigma \cdot (x_1, \dots, x_n) = \begin{cases} (x_{\sigma(1)}, \dots, x_{\sigma(n)})\,? \\ (x_{\sigma^{-1}(1)}, \dots, x_{\sigma^{-1}(n)})\,? \end{cases}$$

Comment se souvenir laquelle des deux est la bonne ? [2] Voici deux façons de répondre à la question.

[2] Ceci a longtemps troublé l'auteur de ces lignes.

(a) On prend un exemple pour $n = 3$: pour une transposition $\tau = (ij)$, on a $\tau^{-1} = \tau$ donc il n'y a aucun doute possible : $(12) \cdot (a, b, c) = (b, a, c)$ puis $(23) \cdot (b, a, c) = (b, c, a)$. Donc le 3-cycle $c = (132) = (23)(12)$ transforme (x_1, x_2, x_3) en (x_2, x_3, x_1) qui est $(x_{c^{-1}(1)}, x_{c^{-1}(2)}, x_{c^{-1}(3)})$, donc si l'une des formules de $(*)$ est correcte, c'est la 2ème.

(b) En réfléchissant un cran au-dessus, il faut comprendre que X^n est la même chose que l'ensemble des *applications* $f : \{1, \dots, n\} \to X$, une telle application correspondant au n-uplet $(f(1), \dots, f(n))$. D'autre part, $G = S_n$ agit à gauche sur l'ensemble de départ $Y = \{1, \dots, n\}$ donc il agit *à droite* sur l'ensemble $\mathrm{Applic}(Y, X)$ par la formule :

$$(\star) \qquad (f \cdot g)(y) = f(gy).$$

En effet, notant e le neutre de G, on a pour tout $g, h \in G$ et $y \in Y$: $(f \cdot e)(y) = f(ey) = f(y)$ et

$$\big((f \cdot g) \cdot h\big)(y) = (f \cdot g)(hy) = f(g(hy)) = f((gh)y) = \big(f \cdot (gh)\big)(y)$$

et donc $f \cdot e = f$ et $(f \cdot g) \cdot h = f \cdot (gh)$. Ceci montre que $(\star)$ est bien une action à droite, et on peut la transformer en une action à gauche en posant (exercice : vérifier ceci)

$$g \cdot f = f \cdot g^{-1} \qquad \text{i.e.} \qquad \forall y \in Y, \quad (g \cdot f)(y) = f(g^{-1}y).$$

On obtient ainsi que l'action à gauche de S_n sur X^n est donnée par la 2ème formule de $(*)$.

Remarque 1.5. — Dans l'exemple précédent, on voit (encore) à l'oeuvre deux principes généraux utiles.

(1) Si l'on a une action *à gauche* $G \times X \to X$, $(g, x) \mapsto gx$, on peut la transformer en une action *à droite* en posant $x \cdot g = g^{-1}x$. Et idem dans l'autre sens. Donc le choix d'avoir une action à gauche ou à droite est un peu une question de goût. La plupart du temps on considèrera des actions à gauche, mais dans certains cas il est naturel (ou usuel) de considérer une action à droite.

Attention, on prendra garde que si l'on a une action à gauche $G \times X \to X$, $(g, x) \mapsto gx$ alors en général la formule $x \cdot g = gx$ ne définit **pas** une action à droite, sauf si G est commutatif. En effet, on a alors par définition : $(x \cdot g) \cdot h = h(x \cdot g) = h(gx) = (hg)x$ tandis que $x \cdot (gh) = (gh)x$ et les deux peuvent être différents si $hg \neq gh$.

(2) Si l'on a une action à gauche $G \times X \to X$ et si Y est un second ensemble, l'ensemble $\mathrm{Applic}(X, Y)$ est muni d'une action de G à droite.

***Remarque* (Agrég).** — Un cas particulier important du point (2) plus haut est celui où $X = V$ est un k-espace vectoriel et où l'action de G est linéaire, i.e. pour tout $g \in G$ l'application $x \mapsto gx$ est linéaire (exercice : se donner une telle action linéaire équivaut à se donner un morphisme de groupes $\rho : G \to \mathrm{GL}(V)$) ; on dit alors que V (ou, plus précisément, ρ) est une représentation linéaire de G. Alors l'espace vectoriel dual V^* est un sous-espace vectoriel de l'espace vectoriel $\mathrm{Applic}(V, k)$ et est muni d'une action à droite de G définie, pour tout $f \in V^*$, $g \in G$ et $v \in V$ par

$$(f \cdot g)(v) = f(gv),$$

i.e. $f \cdot g$ n'est autre que ${}^t\rho(g)(f)$, où ${}^t\rho(g) : V^* \to V^*$ est la transposée de $\rho(g)$. En changeant ceci en une action à gauche, on obtient que l'action à gauche de G sur V^* est donnée par le morphisme $\rho^* : G \to \mathrm{GL}(V^*)$, $g \mapsto {}^t\rho(g^{-1})$. On dit que ρ^* est la représentation *duale* (ou, en langage ancien : *contragrédiente*) de ρ.

***Définitions 1.6* (Orbites, stabilisateurs, action transitive ou libre)**

Soit $G \times X \to X$ une action à gauche.

(1) Pour tout $x \in X$, l'ensemble $\mathscr{O}_x = \{gx \mid g \in G\}$ des transformés de x par les éléments de G s'appelle *l'orbite* de x sous (l'action de) G ou, plus simplement, la G-orbite de x.

(2) Deux orbites distinctes sont disjointes. En effet, si $z \in \mathscr{O}_x \cap \mathscr{O}_y$ alors il existe $g, h \in G$ tels que $z = gx = hy$, d'où $y = h^{-1}gx$ et donc $y \in \mathscr{O}_x$ et $\mathscr{O}_y = \mathscr{O}_x$. Par conséquent, les G-orbites forment une *partition* de X.

(3) L'ensemble des orbites est noté X/G et appelé l'ensemble quotient de X par G.

(4) Pour tout $x \in X$, l'ensemble $G_x = \{g \in G \mid gx = x\}$ est un *sous-groupe* de G appelé le *stabilisateur* (ou *fixateur*) de x dans G.

(5) Si x, y sont dans la même orbite, leurs stabilisateurs G_x et G_y sont *conjugués* dans G. En effet, soit $g \in G$ tel que $gx = y$. Alors, pour tout $h \in G$ on a :

$$h \in G_y \iff y = hy = hgx \iff x = g^{-1}y = (g^{-1}hg)x \iff g^{-1}hg \in G_x$$

et donc on a $G_y = gG_xg^{-1}$.

(6) On dit que l'action est *transitive* si X est une seule orbite, i.e. si pour tout $x, y \in X$ il existe $g \in G$ tel que $gx = y$.

(7) On dit que l'action est *libre* si tous les stabilisateurs sont triviaux, i.e. si $G_x = \{e\}$ pour tout $x \in X$.

(8) Si l'action est transitive, il résulte de (5) que tous les stabilisateurs sont conjugués, donc si l'un d'eux est trivial alors ils le sont tous. On obtient ainsi que :

L'action est libre et transitive ssi il existe $x \in X$ tel que l'application $G \to X$, $g \mapsto gx$ soit bijective. Dans ce cas, ceci a lieu pour tout $x \in X$.

Terminologie. — Reprenons l'exemple 1) de 1.2 : soit H un sous-groupe de G ; pour tout $g \in G$, son orbite sous l'action par translation *à droite* de H est notée gH. Elle est appelée la classe *à gauche* de g modulo H, le motif étant que « G agit à gauche sur l'ensemble des classes à gauche » via $g \cdot g'H = gg'H$, pour tout $g, g' \in G$.[3]

De même, l'orbite $Hg = \{hg \mid h \in H\}$ sous l'action par translation à gauche est appelée la classe à droite de g modulo H.

On peut aussi considérer l'action à gauche de $H \times H$ sur G définie par $(h', h) \cdot g = h'gh^{-1}$; dans ce cas, l'orbite HgH de g est appelée la « double classe » de g modulo H.

Rappel 1.7. — Un sous-groupe H de G est dit *distingué*[4] si pour tout $g \in G$ on a $gHg^{-1} = H$, i.e. si pour tout $g \in G$ et $h \in H$ on a $ghg^{-1} \in H$.[5] Dans ce cas, pour tout $g \in G$ on a $gH = (gHg^{-1})g = Hg$, donc les classes à gauche et à droite sont identiques. De plus, l'ensemble des classes G/H est muni d'une structure de groupe définie par $(gH)(g'H) = gg'H$. (Exercice : vérifier que ceci est bien défini.)

***Proposition 1.8* (Théorème de Lagrange)**. — *Soit H un sous-groupe d'un groupe fini G et soit G/H l'ensemble des classes (à gauche) modulo H. Alors on a $|G/H| = |G|/|H|$. En particulier, $|H|$ divise $|G|$.*

[3] Mais attention, en anglais gH est un *right coset*.

[4] En anglais *normal*.

[5] Faisant varier h, ceci donne $gHg^{-1} \subset H$. Appliquant ceci à g^{-1} on a aussi $g^{-1}Hg \subset H$ d'où $H = g(g^{-1}Hg)g^{-1} \subset gHg^{-1}$, d'où $gHg^{-1} = H$.

Démonstration. — Pour tout $g \in G$, l'application $h \mapsto gh$ est une bijection de H sur la classe gH (la bijection inverse étant donnée par $x \mapsto g^{-1}x$), donc $|gH| = |H|$. De plus, G est la réunion disjointe des classes gH, dont le nombre est par définition $|G/H|$. On en déduit que $|G| = |G/H| \times |H|$, d'où la proposition. $\square$

Proposition 1.9. — *Soit $G \times X \to X$ une action à gauche et soit $x \in X$.*

(i) *L'application $\lambda_x : G \to X$, $g \mapsto gx$ induit une bijection $G/G_x \xrightarrow{\sim} \mathscr{O}_x$.*

(ii) *En particulier, si $|G| < \infty$, on a $|\mathscr{O}_x| = |G|/|G_x|$.*

Démonstration. — Posons $H = G_x$. L'application $\overline{\lambda}_x : G/H \to \mathscr{O}_x$, $gH \mapsto gx$ est bien définie, car si $g' = gh$ avec $h \in H$ on a bien $g'x = ghx = gx$. Elle est surjective, puisque tout élément de $\mathscr{O}_x$ égale gx pour un certain $g \in G$. Montrons qu'elle est injective : supposons que gH et $g'H$ aient même image, i.e. que $gx = g'x$. Alors $x = g^{-1}g'x$ donc $h = g^{-1}g'$ appartient à $G_x = H$. Comme $g' = gh$, ceci entraîne que $g'H = gH$. Ceci montre que $\overline{\lambda}_x$ est injective, d'où le point (i). Le point (ii) découle alors de la proposition précédente. $\square$

Exemple 1.10. — Soit $n \geq 2$. On munit $\mathbb{R}^n$ de la norme euclidienne standard. Soit $r \in \mathbb{R}_+^*$ et $x, y \in \mathbb{R}^n$ de norme r. Notons e_1 et f_1 les vecteurs unitaires associés, i.e. $e_1 = (1/r)x$ et $f_1 = (1/r)y$. On peut compléter e_1 (resp. f_1) en une base orthonormée $\mathscr{B} = (e_1, \ldots, e_n)$ et $\mathscr{C} = (f_1, \ldots, f_n)$. Soit P l'automorphisme de $\mathbb{R}^n$ qui envoie $\mathscr{B}$ sur $\mathscr{C}$; comme ce sont deux bases orthonormées, $P \in O(n)$ (et donc $\det(P) = \pm 1$). Si $\det(P) = -1$ alors, comme $n \geq 2$ on peut changer f_n en $-f_n$ tout en préservant la condition $P(e_1) = f_1$. On obtient ainsi un élément $g \in \mathrm{SO}(n)$ tel que $g(e_1) = f_1$ et donc $g(x) = y$. Ceci montre que $G = \mathrm{SO}(n)$ agit transitivement sur chaque sphère $S^{n-1}(r)$, donc celles-ci sont les orbites de G dans $\mathbb{R}^n - \{0\}$. Et $\{0\}$ est bien sûr une orbite.

On obtient ainsi que $\mathbb{R}^n$ est partitionné par les G-orbites $S^{n-1}(r)$, pour $r \in \mathbb{R}_+$, et l'application naturelle $\mathbb{R}^n \to \mathbb{R}_+$, $x \mapsto \|x\|$ induit une bijection de l'espace des orbites $\mathbb{R}^n/G$ avec $\mathbb{R}_+$.

D'autre part, supposons pour fixer les idées que $n = 3$ et notons (e_1, e_2, e_3) la base canonique de $\mathbb{R}^3$. Le stabilisateur dans $\mathrm{SO}(3)$ du vecteur e_3 est formé des rotations $\begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$, avec $\theta \in \mathbb{R}/2\pi\mathbb{Z}$. Elles forment un groupe isomorphe à $\mathrm{SO}(2)$, et l'on a donc une bijection[6] $\mathrm{SO}(3)/\mathrm{SO}(2) \xrightarrow{\sim} S^2(1)$.

Exemple 1.11. — Dans $\mathbb{R}^3$ muni du produit scalaire standard, soit C le cube de centre 0 dont les sommets sont les huits points $(\pm 1, \pm 1, \pm 1)$ et soit G le groupe des isométries directes[7] de $\mathbb{R}^3$ qui laissent C stable. Soit I le centre de la face supérieure, i.e. $I = (0, 0, 1)$ et soit H le stabilisateur de I dans G. On voit que H est le groupe isomorphe à $\mathbb{Z}/4\mathbb{Z}$ engendré par la rotation d'axe orienté par e_3 et d'angle $\pi/4$.

D'autre part, on se convainc facilement que la G-orbite de I est formée des centres des six faces de C, donc est de cardinal 6. D'après le point (i) de la proposition précédente, on en déduit que $|G| = 24$.

Terminologie. — Soit G un groupe agissant sur un ensemble fini X et soient $\mathscr{O}_1, \ldots, \mathscr{O}_n$ les orbites. Si, pour $i = 1, \ldots, n$ on choisit (arbitrairement) un élément $x_i \in \mathscr{O}_i$, on dit que $(x_1, \ldots, x_n)$ est un « système de représentants (de l'ensemble) des orbites » (i.e. on a choisi un élément et un seul dans chaque orbite).

---

[6] (Agrég) Par ailleurs, on peut montrer que c'est un homéomorphisme.

[7] i.e. de déterminant 1.

Proposition 1.12 **(Formule des classes)**. — *Soit G un groupe fini opérant sur un ensemble fini X et soit $x_1, \ldots, x_n$ un système de représentants des orbites. Pour tout i, notons G_i le stabilisateur dans G de x_i. Alors $|X| = \sum_{i=1}^n \dfrac{|G|}{|G_i|}$.*

Démonstration. — X est la réunion disjointe des orbites $\mathscr{O}_i$ donc $|X| = \sum_{i=1}^n |\mathscr{O}_i|$. D'autre part, pour tout i on a $|\mathscr{O}_i| = |G|/|G_i|$. □

Donnons deux applications (disons en vue de l'Agrégation). Soit p un nombre **premier**.

Rappel. — Soient G un groupe et $g \in G$. On dit que g est *d'ordre fini* s'il existe un entier $n > 0$ tel que $g^n = e$; dans ce cas, notant d le plus petit tel entier, on dit que G est **d'ordre** d. Alors, tout $n > 0$ tel que $g^n = e$ est un multiple de d : en effet, par division euclidienne on a $n = dq + r$ avec $0 \leq r < d$, d'où $e = g^n = g^{dq} g^r = g^r$ et comme $r < d$ ceci entraîne $r = 0$. Par conséquent, si p est un nombre premier et si $g^p = e$ alors $g = e$ ou bien g est d'ordre p.

Proposition 1.13 **(Théorème de Cauchy)**. — *Soit G un groupe fini de cardinal divisible par p. Alors G contient au moins un élément d'ordre p.*

Démonstration. — Soit H le sous-groupe de S_p engendré par le p-cycle $c = (12 \cdots p)$; il est isomorphe à $\mathbb{Z}/p\mathbb{Z}$. Comme S_p agit sur le produit G^p, on obtient par restriction une action de H sur G^p et, d'après 1.4, on a $c \cdot (g_1, \ldots, g_p) = (g_p, g_1, \ldots, g_{p-1})$ pour tout $(g_1, \ldots, g_p) \in G^p$. Montrons que H laisse stable le sous-ensemble

$$X = \{(g_1, \ldots, g_p) \in G^p \mid g_1 \cdots g_p = e\}.$$

Comme $H = \{c, \ldots, c^p\}$ il suffit de montrer que $c \cdot x \in X$ pour tout $x = (g_1, \ldots, g_p) \in X$. Or l'égalité $g_1 \cdots g_p = e$ entraîne que g_p est l'inverse de $h = g_1 \cdots g_{p-1}$, donc on a aussi $g_p h = e$ d'où $c \cdot x \in X$.

Comme H n'a pas de sous-groupe autre que $\{e\}$ et H (pourquoi ?), ses orbites dans X sont donc soit de cardinal p, soit des points fixes. Or $(g_1, \ldots, g_p)$ est un point fixe de c si et seulement si les g_i sont tous égaux à un élément g qui vérifie $g^p = e$. Donc il s'agit de montrer qu'il existe dans X un point fixe distinct de $x_0 = (e, \ldots, e)$.

Notant $\mathscr{O}_1, \ldots, \mathscr{O}_r$ les orbites de cardinal p et m le nombre de points fixes, on a donc

$$|X| = m + pn.$$

D'autre part, l'application $X \to G^{p-1}$, $(g_1, \ldots, g_p) \mapsto (g_1, \ldots, g_{p-1})$ est une bijection (car g_p est déterminé par $g_p = (g_1 \cdots g_{p-1})^{-1}$), donc on a

$$|X| = |G|^{p-1}$$

et comme p divise $|G|$ il divise aussi $|X|$ et donc m. Or $m > 0$ car $x_0 = (e, \ldots, e)$ est un point fixe. Donc m est un multiple non nul de p, donc il existe au moins $p - 1$ points fixes distincts de x_0. □

Définitions 1.14. — 1) Le *centre* $Z(G)$ d'un groupe G est l'ensemble des éléments de G qui commutent à tout les autres, i.e.

$$Z(G) = \{z \in G \mid \forall g \in G, \quad zg = gz\}.$$

On vérifie facilement que c'est un sous-groupe distingué de G (exercice).

2) Un p-groupe fini est un groupe fini dont le cardinal est une puissance de p.

Proposition 1.15. — *Le centre d'un p-groupe fini G est non trivial.*

Démonstration. — Posons $|G| = p^N$ et considérons l'action de G sur lui-même par conjugaison : $g \cdot g' = gg'g^{-1}$. Les orbites sont les classes de conjugaison, et les orbites formées d'un seul point correspondent aux éléments de $Z(G)$. Notons $x_1, \dots, x_n$ un système de représentants des *autres* classes de conjugaison et pour tout i notons G_i le stabilisateur de x_i, i.e.

$$G_i = \{g \in G \mid gx_ig^{-1} = x_i\} = \{g \in G \mid gx_i = x_ig\}.$$ [8]

D'après la formule des classes, on a :

$$(\dagger) \qquad p^N = |G| = |Z(G)| + \sum_{i=1}^{n} \frac{p^N}{|G_i|}.$$

De plus, chaque G_i est un sous-groupe de G distinct de G donc de cardinal p^{m_i} avec $0 \leq m_i < n$, donc $p^N/|G_i| = p^{n-m_i}$ est divisible par p.

Donc, d'après (†), $|Z(G)|$ est divisible par p, donc $Z(G)$ est non trivial (i.e. n'est pas égal à $\{e\}$). □

Terminons cette section avec les produits semi-directs de groupes. Même si cette notion ne figure plus explicitement au programme de l'Agrégation, elle est utile pour décrire la structure de certains groupes, dont le groupe affine et celui des homothéties-translations (voir plus loin) et le groupe symétrique S_4 (en lien avec le birapport).

***Définition 1.16* (Produits semi-directs).** — (1) Soient H un groupe, G et N deux sous-groupes, N étant *distingué* dans H. On dit que H est le *produit semi-direct* de N par G, et l'on note $H = N \rtimes G$, si tout $h \in H$ s'écrit de façon unique $h = ng$, avec $n \in N$ et $g \in G$, i.e. si l'application $N \times G \to H$, $(n, g) \mapsto ng$ est bijective. **Attention**, cette application n'est **pas** un morphisme de groupes en général : le produit de ng et de $n'g'$ est donné par : $(ng)(n'g') = n(gn'g^{-1})\, gg'$ (noter que $gn'g^{-1} \in N$ puisque N est distingué).[9]

(1 bis) Remarquons que l'application $G \times N \to N$, $(g, n) \mapsto gng^{-1}$ est une action à gauche de G sur N. De plus, cette action respecte la structure de groupe de N, i.e. pour tout $g \in G$ l'application $n \mapsto g\cdot = gng^{-1}$ est un morphisme de groupes. En effet, pour tout $g \in G$ et $n, n' \in N$, on a :

$$g(nn')g^{-1} = (gng^{-1})(gn'g^{-1}).$$

Notant $\mathscr{S}(N)$ le groupe de toutes les bijections $N \to N$, on peut exprimer ceci en disant que le morphisme de groupes $G \to \mathscr{S}(N)$ donné par l'action, est à valeurs dans le sous-groupe $\mathrm{Aut}(N)$ de $\mathscr{S}(N)$ formé des automorphismes de groupe.

(2) On peut alors reformuler la définition de façon plus abstraite (mais plus souple). Soient G, N deux groupes ; on dit que G agit sur N « par automorphismes de groupe » si l'on s'est donné un morphisme de groupes $\varphi : G \to \mathrm{Aut}(N)$. On peut alors former le *produit semi-direct* de N par G (relativement à l'action φ), noté $N \rtimes_\varphi G$ (ou simplement $N \rtimes G$), en déclarant que c'est l'ensemble produit $N \times G$ muni de la loi de groupe ainsi définie :

$$(n, g) \cdot (n', g') = \big(n\varphi(g)(n'), gg'\big).$$ [10]

[8] G_i s'appelle le *centralisateur* de x_i et se note $C_G(x_i)$.

[9] Tout $h \in H$ s'écrit aussi de façon unique gn', avec $g \in G$ et $n' \in N$, puisque $ng = g(g^{-1}ng)$.

[10] On laisse au lecteur le soin de vérifier que ceci est bien une loi de groupe.

Dans le cas particulier où l'action de G sur N est *triviale*[11] le groupe obtenu est noté $N \times G$ et appelé le *produit direct* de N et G : c'est l'ensemble produit $N \times G$ muni de la loi de groupe $(n, g) \cdot (n', g') = (nn', gg')$.

(3) Dans la situation de (1), lorsque N et G sont des sous-groupes d'un groupe H dans lequel N est distingué, l'action φ est la conjugaison : $\varphi(g)(n) = gng^{-1}$. En fait, la situation (2) n'est pas plus générale que (1), car posant $H = N \rtimes_\varphi G$ et notant e_N (resp. e_G) l'élément neutre de N (resp. G), le groupe N (resp. G) s'identifie au sous-groupe $\{(n, e_G) \mid n \in N\}$ (resp. $\{(e_N, g) \mid g \in G\}$) de H, et alors pour tout $n \in N$ et $g \in G$ on a :

$$gng^{-1} = (e_N, g)(n, e_G)(e_N, g^{-1}) = (e_N, g)(n, g^{-1}) = (\varphi(g)(n), e_G) = \varphi(g)(n),$$

i.e. l'action donnée φ de G sur N devient dans $H = N \rtimes_\varphi G$ l'action par conjugaison de G sur le sous-groupe distingué N.

(4) Si $H = N \rtimes G$, alors le groupe quotient H/N est isomorphe à G. (Exercice !)

2. Sous-espaces affines

Commençons par les exemples suivants.

***Exemple 2.1* (Droites affines dans $\mathbb{R}^2$).** — Soit $u = (u_1, u_2) \in \mathbb{R}^2$ un vecteur non nul et soit $A = (x_0, y_0)$ un « point » de $\mathbb{R}^2$. La droite affine passant par A et de vecteur directeur u est

$$\mathscr{D} = \{A + tu \mid t \in \mathbb{R}\};$$

c'est l'ensemble des « points » $B = (x, y)$ tels que le vecteur $\overrightarrow{AB} = (x - x_0, y - y_0)$ soit colinéaire à u. Une équation de $\mathscr{D}$ est donnée par

$$0 = \begin{vmatrix} x - x_0 & u_1 \\ y - y_0 & u_2 \end{vmatrix} = (x - x_0)u_2 - (y - y_0)u_1, \qquad \text{i.e.} \qquad u_2x - u_1y = u_2x_0 - u_1y_0.$$

Une telle droite affine apparaît par exemple en analyse : si $f : I \to \mathbb{R}$ une fonction dérivable, où I est un intervalle ouvert de $\mathbb{R}$, et si $x_0 \in I$ et $f(x_0) = y_0$, la tangente au graphe Γ_f de f en x_0 (i.e. au point (x_0, y_0)) est la droite affine d'équation $y - y_0 = f'(x_0)(x - x_0)$. Elle passe par $A = (x_0, y_0)$ et a pour vecteur directeur le vecteur $u = (1, f'(x_0))$.

Plus généralement, dans un cours de fonctions de plusieurs variables, vous avez peut-être vu que si $F : U \to \mathbb{R}$ est une application différentiable, où U est un ouvert de $\mathbb{R}^2$, si C désigne la courbe d'équation $F(x, y) = 0$ et si $p = (x_0, y_0)$ est un point de C en lequel les dérivées partielles $\alpha = \dfrac{\partial F}{\partial x}(p)$ et $\beta = \dfrac{\partial F}{\partial y}(p)$ ne sont pas toutes les deux nulles, alors la *tangente à* C en p est la droite affine d'équation

$$\beta(y - y_0) + \alpha(x - x_0) = 0.$$

Dans le cas précédent, prenant $U = I \times \mathbb{R}$ et $F(x, y) = y - f(x)$, on retrouve comme cas particulier la tangente à Γ_f au point $p = (x_0, f(x_0))$. Exercice : déterminer la tangente au cercle d'équation $x^2 + y^2 = 1$ au point $p = (1, 0)$.

[11] Si G est un groupe et X un ensemble, l'action *triviale* de G sur X est donnée par $g \cdot x = x$ pour tout $g \in G$ et $x \in X$.

***Exemple 2.2* (Plans affines dans $\mathbb{R}^3$).** — De même, si $f : U \to \mathbb{R}$ est une application différentiable, où U est un ouvert de $\mathbb{R}^2$, on peut considérer dans $\mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R}$ son graphe :

$$\Gamma_f = \{(x, y, z) \in U \times \mathbb{R} \mid z = f(x, y)\}.$$

Le *plan tangent* à Γ_f « en $p = (x_0, y_0)$ » (i.e. au point $A = (x_0, y_0, f(x_0, y_0))$) est le « plan affine » $\mathscr{P}$ d'équation

$$z - z_0 = \alpha(x - x_0) + \beta(y - y_0),$$

où $\alpha = \dfrac{\partial f}{\partial x}(p)$ et $\beta = \dfrac{\partial f}{\partial y}(p)$. On a $\mathscr{P} = A + P = \{A + u \mid u \in P\}$, où P désigne le sous-espace vectoriel de $\mathbb{R}^3$ d'équation

$$z = \alpha x + \beta y.$$

(Exercice : vérifier ceci).

Plus généralement, si $F : \Omega \to \mathbb{R}$ est une application différentiable, où Ω est un ouvert de $\mathbb{R}^3$, si S désigne la surface d'équation $F(x, y, z) = 0$ et si $p = (x_0, y_0, z_0)$ est un point de S en lequel les dérivées partielles $\alpha = \dfrac{\partial F}{\partial x}(p)$, $\beta = \dfrac{\partial F}{\partial y}(p)$ et $\gamma = \dfrac{\partial F}{\partial z}(p)$ ne sont pas toutes les trois nulles, alors le *plan tangent à* S en p est le plan affine d'équation

$$\gamma(z - z_0) + \beta(y - y_0) + \alpha(x - x_0) = 0.$$

Dans le cas précédent, prenant $\Omega = U \times \mathbb{R}$ et $F(x, y, z) = z - f(x, y)$, on retrouve comme cas particulier le plan tangent à Γ_f au point $p = (x_0, y_0, f(x_0, y_0))$. Exercice : déterminer le plan tangent à la sphère d'équation $x^2 + y^2 + z^2 = 1$ au point $p = (1, 0, 0)$.

Ces exemples conduisent à la définition suivante.

***Définition 2.3* (Sous-espaces affines de $\mathbb{R}^n$).** — Soit F un sous-espace vectoriel (sev) de $E = \mathbb{R}^n$ et $p = (x_1, \dots, x_n)$ un « point » de $\mathbb{R}^n$. Alors

$$p + F = \{p + u \mid u \in F\} = \{y \in \mathbb{R}^n \mid y - p \in F\}$$

est appelé le « sous-espace affine de direction F » passant par p. Remarquons tout de suite que si $q \in p + F$ alors $q = p + u_0$ pour un certain $u_0 \in F$ et donc, comme l'application $F \to F$, $u \mapsto u_0 + u$ est une bijection (d'inverse $v \mapsto v - u_0$), on a :

$$q + F = \{p + u_0 + u \mid u \in F\} = \{p + v \mid v \in F\} = p + F.$$

Donc, si on pose $\mathscr{F} = p + F$, alors pour tout $q \in \mathscr{F}$ on a $\mathscr{F} = q + F$ donc $\mathscr{F}$ est aussi le « sous-espace affine de direction F » passant par q.

Bien entendu, on peut remplacer ci-dessus $\mathbb{R}$ par un corps quelconque k et $\mathbb{R}^n$ par un k-espace vectoriel (en abrégé : k-ev) arbitraire E. On obtient ainsi la notion de « sous-espace affine[12] de E de direction F », qui est suffisante pour la plupart des applications mathématiques.

***Exemple 2.4* (fondamental).** — Soient E, V des k-espaces vectoriels, $f : E \to V$ une application linéaire et v un élément de V appartenant à $\mathrm{Im}(f)$. Alors

$$\mathscr{F} = f^{-1}(v) = \{x \in E \mid f(x) = v\}$$

est un sous-espace affine de E de direction $F = \mathrm{Ker}(f)$.

[12] Pour abréger on écrira souvent « sea » pour « sous-espace affine ».

Démonstration. — $\mathscr{F}$ est non vide puisque $v \in \mathrm{Im}(f)$. Fixons $p \in \mathscr{F}$ et soit u un élément arbitraire de E, on a :

$$p + u \in \mathscr{F} \iff v = f(p+u) = f(p) + f(u) = v + f(u) \iff f(u) = 0.$$

Ceci montre que $\mathscr{F} = \{p + u \mid u \in \mathrm{Ker}(f)\} = p + \mathrm{Ker}(f)$. □

Déclinons quelques exemples « concrets » de l'exemple fondamental ci-dessus.

***Exemple 2.5* (Systèmes linéaires).** — Considérons une matrice $A \in M_{p,n}(k)$, un vecteur fixé $Y \in k^p$ et le système linéaire d'inconnue $X \in k^n$ donné par $AX = Y$ (c'est un système linéaire de p équations à n inconnues, avec second membre Y). Soit $\mathscr{S}$ l'ensemble des solutions de ce système.

Si $\mathscr{S} \neq \varnothing$ (i.e. si $Y \in \mathrm{Im}(A)$) alors $\mathscr{S}$ est un sous-espace affine de k^n de direction le sev $\mathrm{Ker}(A)$ (= l'ensemble des solutions du système homogène $AX = 0$). En effet, si X_0 est une solution arbitraire, on sait que $\mathscr{S} = X_0 + \mathrm{Ker}(A)$.

***Exemple 2.6* (sea défini par des équations).** — Considérons l'exemple « géométrique » suivant : soit E un k-ev, soient $f_1, \dots, f_r$ des formes linéaires sur E et $c_1, \dots, c_r \in k$. On suppose que le sous-ensemble

$$\mathscr{F} = \{p \in E \mid f_1(p) = c_1, \dots, f_r(p) = c_r\}$$

est *non vide*. Alors c'est un sea de E de direction le sev suivant :

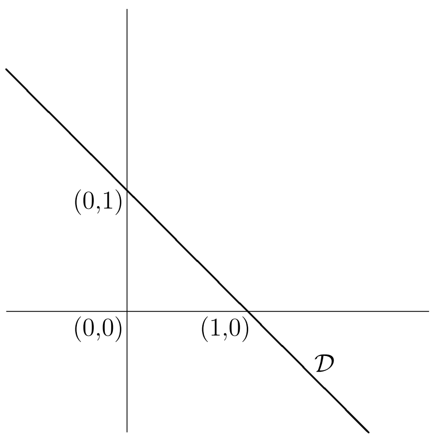
$$F = \{u \in E \mid f_1(u) = 0, \dots, f_r(u) = 0\} = \bigcap_{i=1}^{r} \mathrm{Ker}(f_i).$$

En effet, si l'on considère l'application linéaire $f : E \to k^r$ dont les composantes sont $f_1, \dots, f_r$, alors $\mathrm{Ker}(f) = F$ et $\mathscr{F} = f^{-1}(v)$, où $v = (c_1, \dots, c_r)$.

En particulier, prenons $k = \mathbb{R}$ et f la forme linéaire sur $\mathbb{R}^2$ définie par $f(x_1, x_2) = x_1 + x_2$. Alors

$$\mathscr{D} = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 + x_2 = 1\}$$

est une droite affine de direction la droite vectorielle D d'équation $x_1 + x_2 = 0$; celle-ci est engendrée, par exemple, par le vecteur $u = (1, -1)$ et l'on a $\mathscr{D} = (0,1) + \mathbb{R}u = (1,0) + \mathbb{R}u$:



Donnons encore un exemple, plus élaboré (Agrég.) Dans ce qui suit, EDL signifie Équation Différentielle Linéaire.

***Exemple 2.7* (EDL avec second membre).** — Soient I un intervalle ouvert de $\mathbb{R}$, $y : I \to \mathbb{R}$ une fonction continue, et $a \in \mathbb{R}$. On considère l'EDL avec second membre :

$$(*) \qquad x'(t) - ax(t) = y(t) \qquad \forall t \in I.$$

On sait que l'ensemble des solutions de l'équation homogène (i.e. sans second membre) est le $\mathbb{R}$-ev F de dimension 1 formé des fonctions $t \mapsto \lambda e^{at}$, pour $\lambda \in \mathbb{R}$. On peut résoudre l'équation $(*)$ par la méthode « de variation de la constante », i.e. on cherche $x(t)$ sous la forme $\lambda(t)e^{at}$; on obtient alors que

$$\lambda'(t)e^{at} = y(t) \qquad \text{i.e.} \qquad \lambda'(t) = e^{-at}y(t).$$

Fixant un point $t_0 \in I$, on obtient que $\lambda(t) = c + \int_{t_0}^{t} e^{-as}y(s)ds$, avec $c \in \mathbb{R}$ arbitraire. Posant

$$x_0(t) = e^{at}\int_{t_0}^{t} e^{-as}y(s)ds$$

on obtient que l'ensemble des solutions de $(*)$ est $\mathscr{F} = \{x_0(t) + ce^{at} \mid c \in \mathbb{R}\} = x_0 + F$. Notant E le $\mathbb{R}$-ev des fonctions $I \to \mathbb{R}$ de classe C^1, on voit donc que $\mathscr{F}$ est le sea de E de direction F passant par x_0.

3. Espaces affines et repères

Ce qui précède n'est pas suffisant pour faire de la géométrie, ou de la physique, par exemple car le monde dans lequel nous vivons est, en 1ère approximation, assimilable à un espace affine (euclidien) de dimension 3, mais ce n'est pas un espace vectoriel, car il n'y a pas d'origine privilégiée. Pour rendre compte de cette situation, on est amené à introduire le concept d'espace affine, donné par la définition abstraite suivante.

***Définition 3.1* (Espaces affines).** — Soient k un corps et E un k-ev. Un *espace affine de direction* E est un ensemble non vide $\mathscr{E}$ muni de l'une des structures équivalentes suivantes :

(i) On s'est donné une application $\phi : \mathscr{E} \times \mathscr{E} \to E$, $(A, B) \mapsto \overrightarrow{AB}$ vérifiant les deux propriétés suivantes :

$$(1) \qquad \boxed{\text{Relation de Chasles : } \quad \overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}} \qquad \forall A, B, C \in \mathscr{E}.$$

$$(2) \qquad \text{Pour tout } A \in \mathscr{E}, \text{ l'application } \phi_A : \mathscr{E} \to E, \quad B \mapsto \overrightarrow{AB} \quad \text{est } \textbf{bijective}.$$

On notera $\overrightarrow{u} \mapsto A + \overrightarrow{u}$ la **bijection inverse**, c.-à-d., pour tout $A \in \mathscr{E}$ et $\overrightarrow{u} \in E$, $A + \overrightarrow{u}$ désigne l'unique $B \in \mathscr{E}$ tel que $\overrightarrow{AB} = \overrightarrow{u}$.

(ii) On s'est donné une action à droite $\mathscr{E} \times E \to \mathscr{E}$, $(A, \overrightarrow{u}) \mapsto A + \overrightarrow{u}$ qui est libre et transitive.[13]

<u>Vocabulaire</u> : les éléments de $\mathscr{E}$ sont appelés « points », ceux de E sont appelés « vecteurs ». Si E est de dimension finie n, ce que nous supposerons par la suite, on pose $\dim \mathscr{E} = \dim E$. Si $\dim \mathscr{E} = 1$, resp. 2, on dira que $\mathscr{E}$ est une droite affine, resp. un plan affine.

<u>Notation</u> : pour abréger, on dira : « $(\mathscr{E}, E)$ est un espace affine ».

***Remarque 3.2*.** — Appliquant la relation de Chasles d'abord à $A = B = C$, on obtient $\overrightarrow{AA} = \overrightarrow{AA} + \overrightarrow{AA}$ d'où $\overrightarrow{AA} = \overrightarrow{0}$ pour tout A ; en l'appliquant ensuite à A, B arbitraires et $C = A$, on obtient $\overrightarrow{BA} = -\overrightarrow{AB}$.

[13] Ici, E est considéré juste comme un groupe abélien.

Remarque 3.3. — Supposons donnée ϕ vérifiant la relation de Chasles. Alors, pour montrer (2), il suffit de montrer qu'il existe $A_0 \in \mathscr{E}$ tel que ϕ_{A_0} est bijective. En effet, supposons que ce soit le cas et soit A un autre point, arbitraire mais fixé. Alors, pour B variant dans $\mathscr{E}$, on a $\phi_A(B) = \overrightarrow{AB} = \overrightarrow{A_0B} - \overrightarrow{A_0A} = \phi_{A_0}(B) - u_0$, où l'on a posé $u_0 = \overrightarrow{A_0A}$. Ceci montre que $\phi_A : \mathscr{E} \to E$ est la composée de ϕ_{A_0} et de l'application $E \to E$, $u \mapsto u - u_0$. Or cette dernière est bijective, car elle admet comme réciproque l'application $E \to E$, $v \mapsto v + u_0$. Comme on a supposé ϕ_{A_0} bijective, il en résulte que ϕ_A l'est aussi.

Avant de donner un exemple non trivial d'espace affine, démontrons l'équivalence entre (i) et (ii). [14]

Démonstration. — Supposons (i) vérifié. Alors, pour tout $I \in \mathscr{E}$ et $u \in E$, on note $I + u$ l'unique point J tel que $\overrightarrow{IJ} = u$. Fixons $A \in \mathscr{E}$. Comme $\overrightarrow{AA} = 0$, on a déjà que $A + 0 = A$.

Fixons maintenant $u, v \in E$ et posons $B = A + u$ et $C = B + v$. On a donc $u = \overrightarrow{AB}$ et $v = \overrightarrow{BC}$. Par définition, on a aussi $C = A + \overrightarrow{AC}$. D'autre part, d'après la relation de Chasles, on a $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC} = u + v$, d'où $C = A + (u + v)$. On a donc, pour tout $A \in \mathscr{E}$ et $u, v \in E$:

$$A + (u + v) = C = B + v = (A + u) + v.$$

Joint à l'égalité $A + 0 = A$, ceci montre que l'application $\mathscr{E} \times E \to \mathscr{E}$, $(A, u) \mapsto A + u$, définit une action à droite de E sur $\mathscr{E}$. De plus, d'après l'axiome (2) de (i), cette action est libre et transitive. Ceci montre que (i) implique (ii).

Réciproquement, supposons (ii) vérifié. Soit $A \in \mathscr{E}$ arbitraire mais fixé. Par hypothèse, l'application $E \to \mathscr{E}$, $u \mapsto A + u$ est bijective donc à tout $B \in \mathscr{E}$ on peut associer l'unique $u \in E$ tel que $B = A + u$, qu'on notera $\overrightarrow{AB}$. Ceci définit une application $\mathscr{E} \times \mathscr{E} \to E$, $(A, B) \mapsto \overrightarrow{AB}$ qui vérifie l'axiome (2) de (i).

Il ne reste qu'à vérifier la relation de Chasles. Soit $A, B, C \in \mathscr{E}$ et soient u, v les éléments de E (uniques) tels que $B = A + u$ et $C = B + v$. Alors $C = (A + u) + v = A + (u + v)$ et donc $\overrightarrow{AC} = u + v = \overrightarrow{AB} + \overrightarrow{BC}$. Ceci montre que (ii) implique (i). □

***Exemple 3.4* (fondamental)**. — [15] On prend $k = \mathbb{R}$. Historiquement, les notions de droite affine, de plan affine et d'espace affine ont été introduites dans l'Antiquité grecque, sur des bases axiomatiques suggérés par l'observation physique du monde dans lequel nous vivons.

On « sait bien » ce qu'est une droite affine réelle $\mathscr{D}$:

$\mathscr{D}$

Elle est formée de « points », chacun d'épaisseur nulle, et si l'on se donne deux points A, B de $\mathscr{D}$, le segment $[A, B]$ possède une certaine « longueur » AB. Le *rapport* de deux longueurs est un nombre réel ; plus précisément, si l'on fixe un segment orienté $[O, I]$, d'origine O et d'extrémité $I \neq O$, pris comme « segment unité » alors à tout point B de $\mathscr{D}$ on peut associer le réel $\varepsilon\, OB/OI$, où $\varepsilon = +1$ si B est dans la demi-droite fermée de sommet O contenant I, et $\varepsilon = -1$ sinon. On pose les deux axiomes suivants, qui sont conformes à l'intuition physique de mesure des longueurs.

[14] Pour alléger l'écriture, un élément arbitraire de E (resp. le vecteur nul) sera noté u (resp. 0) au lieu de $\overrightarrow{u}$ (resp. $\overrightarrow{0}$).

[15] Cet exemple a pour but d'expliquer la *raison d'être* de la définition 3.1, mais ne prétend pas être de lecture facile. Le lecteur qui voudra bien accepter 3.1 comme point de départ de la théorie peut en omettre la lecture.

• Axiome 1. Ce qui précède définit une *bijection* $B \mapsto x_B$ entre $\mathscr{D}$ et $\mathbb{R}$, et x_B est appelé l'abscisse de B dans le repère (O, I). Noter que $x_I = +1$.

• Axiome 2. Ces bijections sont « compatibles » entre elles au sens suivant.

(1) (Changement d'origine) Fixons un point A, d'abscisse x_A dans le repère (O, I), et soit J l'unique point de $\mathscr{D}$ d'abscisse $x_A + 1$. Alors, pour tout point B, ses abscisses x_B et x'_B dans les repères (O, I) et (A, J) respectivement, sont reliées par : $x'_B = x_B - x_A$.

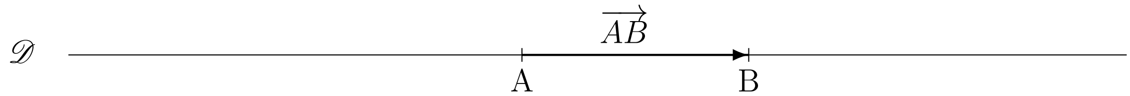
(2) (Changement d'unité [16]) Fixons un point K, d'abscisse $t \neq 0$ dans le repère (O, I). Alors, pour tout point B, ses abscisses x_B et x'_B dans les repères (O, I) et (O, K) respectivement, sont reliées par : $x'_B = x_B/t$.

(3) Bien sûr, on peut combiner les deux : si dans le repère (O, I) les abscisses de A et K sont x_A et $x_A + t$ avec $t \neq 0$, alors pour tout point B, ses abscisses x_B et x'_B dans les repères (O, I) et (A, K) respectivement, sont reliées par : $x'_B = (x_B - x_A)/t$.

On peut alors construire un $\mathbb{R}$-ev V de dimension 1 comme suit. On appelle *bipoint* un couple (A, B) de points de $\mathscr{D}$. Fixant un repère $\mathscr{R}_0 = (O, I)$, on dit que deux bipoints (A, B) et (C, D) sont *équipollents* si les abscisses $x_A, \ldots, x_D$ vérifient

$$x_B - x_A = x_D - x_C.$$

Ceci est clairement une relation d'équivalence. De plus, il résulte des formules de changement de repère données plus haut que cette relation d'équivalence ne dépend pas du choix du repère $\mathscr{R}_0$. La classe d'équivalence du bipoint (A, B) sera notée $\overrightarrow{AB}$ et appelée un *vecteur* ; on la représente par une flèche allant de A vers B :



Le vecteur $\overrightarrow{AA}$ est noté $\overrightarrow{0}$ et appelé vecteur nul. Alors deux vecteurs sont égaux ssi ils sont tous deux nuls ou bien si les flèches correspondantes ont la même longueur $\neq 0$ et le même sens.

On voit de plus que pour tout point A et tout vecteur $\overrightarrow{BC}$, il existe un unique point D tel que $\overrightarrow{AD} = \overrightarrow{BC}$: ceci équivaut à la condition $x_D - x_A = x_C - x_B$ (qui, encore une fois, ne dépend pas du repère $\mathscr{R}_0$). L'ensemble V des vecteurs est donc en bijection avec $\mathscr{D}$ via $A \mapsto \overrightarrow{OA}$, et l'on a des applications

$$\mathbb{R} \times V \to V, \quad (\lambda, \overrightarrow{OA}) \mapsto \lambda \cdot \overrightarrow{OA} = \overrightarrow{OA_\lambda}, \qquad V \times V \to V, \quad (\overrightarrow{OA}, \overrightarrow{OB}) \mapsto \overrightarrow{OA} + \overrightarrow{OB} = \overrightarrow{OC},$$

où A_λ (resp. C) est le point d'abscisse λx_A (resp. $x_A + x_B$) dans le repère $\mathscr{R}_0$. D'une part, ceci munit V d'une structure de $\mathbb{R}$-espace vectoriel de dimension 1, engendré par le vecteur $\overrightarrow{OI}$: la vérification des axiomes découle immédiatement des propriétés de l'addition et la multiplication dans $\mathbb{R}$. D'autre part, ces applications ne dépendent pas, en fait, du choix de $\mathscr{R}_0$, donc la structure de $\mathbb{R}$-ev sur V ainsi obtenue est « canonique ».

Enfin, il résulte de la construction que l'application $\mathscr{D} \times \mathscr{D} \to V$, $(A, B) \mapsto \overrightarrow{AB}$ vérifie les conditions (1) et (2) de 3.1 (i). Donc la « droite affine réelle » $\mathscr{D}$ est bien un espace affine de direction V.

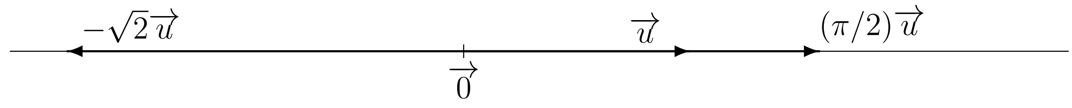
---

[16] et aussi de sens si $t < 0$.

Ayant fait tout ce travail (non trivial !), on voit alors que les axiomes 1 et 2 sont des conséquences de la structure d'espace affine, et lui sont donc équivalents. Ceci justifie de définir la droite affine réelle $\mathscr{D}$:

$\mathscr{D}$

comme un espace affine sous le $\mathbb{R}$-ev V (de dimension 1) des vecteurs « flèches » :



(À la différence de $\mathbb{R}$, qui possède 1 comme générateur canonique, V n'a pas de générateur canonique.)

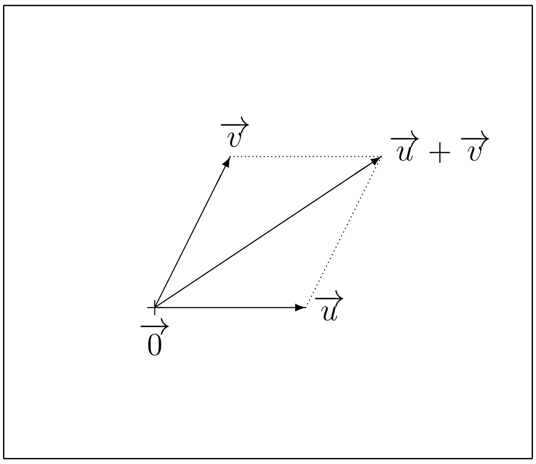
De même, en se basant sur l'intuition physique donnée par une feuille de papier ou un tableau noir, tous deux supposés de largeur et hauteur infinies, on peut définir le « plan affine réel » en se donnant :

(1) Un ensemble non vide $\mathscr{P}$, formé de *points*.

(2) Une famille de parties de $\mathscr{P}$, appelées les *droites*.

(3) Un certain nombre d'axiomes, plusieurs choix étant possibles pour arriver au même résultat. Par exemple, pour tout triplet (O, I, J) de points non alignés, une bijection entre $\mathscr{P}$ et $\mathbb{R}^2$, ces bijections étant soumises à des conditions de compatibilité comme plus haut. Ou bien, des axiomes plus géométriques : l'axiome que toute « droite » est une droite affine réelle, des axiomes d'incidence et un axiome assurant la validité du théorème de Thalès.

Dans ce cas aussi, en partant d'axiomes géométriques suggérés par l'intuition physique, on arrive à montrer (avec pas mal de travail !) que le plan affine réel $\mathscr{P}$ est un espace affine au sens de 3.1 sous le $\mathbb{R}$-ev (de dimension 2) des vecteurs « flèches » ci-dessous :



la somme de deux vecteurs non colinéaires étant définie par la règle du parallélogramme. Ceci étant obtenu, on voit alors que les axiomes de départ découlent facilement de la structure affine, donc lui sont équivalents, et que finalement, la définition 3.1 est la plus commode en pratique.

***Remarque 3.5* (importante).** — Soit E un k-espace vectoriel.

a) E lui-même est un espace affine $\mathscr{E}$ de direction E. En effet, l'addition définit une action à droite $E \times E \to E$, $(x, u) \mapsto x + u$. En termes du point (i) de 3.1, ceci revient à définir $\phi : E \times E \to E$ par $(x, y) \mapsto y - x$. Alors (1) est vérifiée car $\phi(x, z) = z - x = (z - y) + (y - x) = \phi(x, y) + \phi(y, z)$

et (2) est vérifiée car pour tout x fixé dans E, l'application $y \mapsto y - x$ est une bijection de E sur lui-même (sa réciproque étant l'application $u \mapsto x + u$).

b) Pour tout sev F de E et tout $p \in E$, on laisse au lecteur le soin de vérifier que le sea $\mathscr{F} = p + F$ défini dans la section 2 est bien un *espace affine de direction* F. (Ce qui justifie la terminologie employée dans la section 2.) Ceci est généralisé dans la proposition suivante.

Définition et proposition 3.6. — *Soit $\mathscr{E}$ un espace affine de direction E. Soit $A_0 \in \mathscr{E}$ et soit F un sous-espace vectoriel (en abrégé, sev) de E. On pose*

$$\mathscr{F} = A_0 + F = \{A_0 + u \mid u \in F\}.$$

Alors $\mathscr{F}$ est un espace affine de direction F. On dira que c'est un sous-espace affine *(en abrégé, sea) de $\mathscr{E}$.*

Démonstration. — $\mathscr{E}$ est muni d'une action à droite de E, qui est libre (et transitive). Donc, par restriction (cf. 1.3), $\mathscr{E}$ est muni d'une action de F, qui est libre. Cette action laisse stable $\mathscr{F}$, qui est l'orbite de A_0 sous l'action de F, donc on obtient une action de F sur $\mathscr{F}$, qui est libre et transitive. Ceci prouve que $\mathscr{F}$ est un espace affine de direction F. □

***Définition 3.7* (Repères affines)**. — Soit $(\mathscr{E}, E)$ un espace affine de dimension n. Un <u>*repère affine*</u> de $\mathscr{E}$ est un $(n+1)$-uplet $(A_0, \dots, A_n)$ de points de $\mathscr{E}$ tel que les n vecteurs $\overrightarrow{A_0A_1}, \dots, \overrightarrow{A_0A_n}$ engendrent E et donc en forment une base.[17]

Dans ce cas, pour tout $M \in \mathscr{E}$ le vecteur $\overrightarrow{A_0M}$ s'écrit de façon unique $\sum_{i=1}^n x_i\overrightarrow{A_0A_i}$ et l'on dit que $(x_1, \dots, x_n)$ sont les *coordonnées* de M dans le repère $\mathscr{R} = (A_0, \dots, A_n)$.

De façon équivalente, se donner un repère équivaut à se donner un point O pris comme origine, et une base $\mathscr{B} = (e_1, \dots, e_n)$. Ceci équivaut à se donner le repère $(O, A_1, \dots, A_n)$, où $A_i = O + e_i$, et donc un point M a pour coordonnées $(x_1, \dots, x_n)$ ssi $\overrightarrow{OM} = x_1e_1 + \cdots + x_ne_n$.

***Théorème 3.8* (Changement de repère)**. — *Soit $(\mathscr{E}, E)$ un espace affine de dimension n, $\mathscr{B} = (e_1, \dots, e_n)$ et $\mathscr{C} = (f_1, \dots, f_n)$ deux bases de E, soit $P = \mathrm{Mat}_{\mathscr{B}}(\mathscr{C})$ la matrice de passage, et soient $O, O' \in \mathscr{E}$. Posons $\overrightarrow{OO'} = t_1e_1 + \cdots + t_ne_n$. Pour $M \in \mathscr{E}$ arbitraire, notons $(x_1, \dots, x_n)$ (resp. $(x'_1, \dots, x'_n)$) ses coordonnées dans le repère $\mathscr{R} = (O, \mathscr{B})$ (resp. $\mathscr{R}' = (O', \mathscr{C})$). Alors on a*

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} + \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} \qquad \textit{donc aussi} \qquad \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = P^{-1} \begin{pmatrix} x_1 - t_1 \\ \vdots \\ x_n - t_n \end{pmatrix}.$$

De façon abrégée, si l'on note X, X' et T les vecteurs colonnes ci-dessus, on a

$$\boxed{X = PX' + T} \qquad \boxed{X' = P^{-1}(X - T).}$$

Démonstration. — Soient X, X' comme ci-dessus et notons Y le vecteur colonne PX'. Alors, d'après la formule de changement de base dans E, on a :

$$\overrightarrow{O'M} = x'_1f_1 + \cdots + x'_nf_n = y_1e_1 + \cdots + y_ne_n,$$

[17] Ceci est une propriété du $(n+1)$-uplet lui-même, car pour tout j fixé on a $\overrightarrow{A_jA_i} = \overrightarrow{A_0A_i} - \overrightarrow{A_0A_j}$ donc l'espace vectoriel $\mathrm{Vect}(\overrightarrow{A_jA_i} \mid i = 0, \dots, n\}$ engendré par les $\overrightarrow{A_jA_i}$ est le même que $\mathrm{Vect}(\overrightarrow{A_0A_i} \mid i = 0, \dots, n\}$.

et comme $\overrightarrow{OM} = \overrightarrow{OO'} + \overrightarrow{O'M}$ on obtient

$$X = \overrightarrow{OM} = \sum_{i=1}^{n}(t_i + y_i)e_i = T + PX'$$

et l'on a donc aussi $X' = P^{-1}(X - T)$. □

Remarque. — Le théorème précédent donne un exemple d'*application affine* : si $(\mathscr{E}', E')$ et $(\mathscr{E}, E)$ sont deux espaces affines, de dimension n et p respectivement, et si l'on s'est fixé des repères $\mathscr{R}'$ et $\mathscr{R}$, d'où des coordonnées X' et X sur $\mathscr{E}'$ et $\mathscr{E}$, vues comme des vecteurs colonnes, on dit qu'une application $f : \mathscr{E}' \to \mathscr{E}$ est *affine* s'il existe une matrice $A \in M_{p,n}(k)$ et un vecteur colonne $T \in \mathbb{R}^p$ tels que $f(X') = AX' + T$.

En utilisant le théorème précédent[18], on peut montrer que si f vérifie cette propriété pour *un* couple de repères $(\mathscr{R}', \mathscr{R})$ alors elle la vérifie pour *tout* tel couple, donc la notion d'application affine est une notion *intrinsèque*, qui ne dépend pas du choix des repères. On va reprendre ce point, de façon plus abstraite, dans la section suivante.

4. Applications affines : définition, exemples, th. de Thalès

Définition et proposition 4.1 **(Applications affines)**. — *Soient $(\mathscr{E}, E)$ et $(\mathscr{E}', E')$ deux espaces affines. Une application $f : \mathscr{E} \to \mathscr{E}'$ est une* **application affine** *s'il existe un point $O \in \mathscr{E}$ tel que l'application $\phi : E \to E'$ définie par :*

$$\forall \overrightarrow{u} \in E, \quad f(O + \overrightarrow{u}) = f(O) + \phi(\overrightarrow{u}) \qquad \text{c.-à-d.} \qquad \forall P \in \mathscr{E}, \quad \overrightarrow{f(O)f(P)} = \phi(\overrightarrow{OP})$$

est **linéaire**. *Dans ce cas, ϕ est notée $\overrightarrow{f}$ et appelée la* **partie linéaire** *de f ; de plus les égalités ci-dessus sont vraies en remplaçant O par un point I arbitraire, i.e. on a :*

$$\boxed{\forall I, P \in \mathscr{E}, \qquad \overrightarrow{f(I)f(P)} = \overrightarrow{f}(\overrightarrow{IP})}$$

$$\text{et} \qquad \boxed{\forall I \in \mathscr{E},\ \forall \overrightarrow{u} \in E, \qquad f(I + \overrightarrow{u}) = f(I) + \overrightarrow{f}(\overrightarrow{u}).}$$

Démonstration. — Fixons un point $O \in \mathscr{E}$, posons $O' = f(O)$ et notons $\theta : E \to \mathscr{E}$ et $\theta' : E' \to \mathscr{E}'$ les bijections $u \mapsto O + u$ et $u' \mapsto O' + u'$. Alors l'application ϕ de l'énoncé est $\theta'^{-1} \circ f \circ \theta$: en effet, la 1ère égalité définissant ϕ équivaut à l'égalité $f \circ \theta = \theta' \circ \phi$.

On suppose que, pour ce choix de O, l'application $\phi : E \to E'$ est **linéaire**. Soit $I \in \mathscr{E}$ arbitraire. D'après la relation de Chasles, et la définition de ϕ, on a, pour tout $P \in \mathscr{E}$:

$$\overrightarrow{f(I)f(P)} = \overrightarrow{f(O)f(P)} - \overrightarrow{f(O)f(I)} = \phi(\overrightarrow{OP}) - \phi(\overrightarrow{OI})$$

et comme ϕ est supposée linéaire, ceci égale $\phi(\overrightarrow{OP} - \overrightarrow{OI})$. D'après la relation de Chasles, à nouveau, on a $\overrightarrow{OP} - \overrightarrow{OI} = \overrightarrow{IP}$. Posant $\overrightarrow{f} = \phi$, on a donc obtenu :

$$\forall I, P \in \mathscr{E}, \qquad \overrightarrow{f(I)f(P)} = \overrightarrow{f}(\overrightarrow{IP}).$$

La proposition est démontrée. □

Remarque. — Pour $f : \mathscr{E} \to \mathscr{E}$ arbitraire et $O \in \mathscr{E}$, l'application $\phi : E \to E$ définie plus haut n'a aucune raison d'être linéaire : prendre par exemple $\mathscr{E} = E = \mathbb{R}$, $O = 0$ et $f(t) = t^2$ pour tout $t \in \mathbb{R}$, alors $\phi(t) = f(t)$ n'est pas linéaire.

Commençons par un exemple important d'applications affines $\mathscr{E} \to \mathscr{E}$.

[18] qui est le cas où $(\mathscr{E}', E') = (\mathscr{E}, E)$ et f est l'identité, mais où les repères $\mathscr{R}'$ et $\mathscr{R}$ peuvent être différents.

***Définition 4.2* (Translations).** — Pour tout $u \in E$, on note t_u la « translation de vecteur u », définie par $t_u(A) = A + u$ pour tout $A \in \mathscr{E}$. Posant $A' = t_u(A)$ et $B' = t_u(B)$, on a $\overrightarrow{AA'} = u = \overrightarrow{BB'}$ et donc :

$$\overrightarrow{A'B'} = \overrightarrow{A'A} + \overrightarrow{AB} + \overrightarrow{BB'} = -u + \overrightarrow{AB} + u = \overrightarrow{AB}.$$

Ceci prouve que t_u est affine, de partie vectorielle id_E, l'application identique de E. Noter au passage que la translation de vecteur nul t_0 est $\mathrm{id}_{\mathscr{E}}$, l'application identique de $\mathscr{E}$.

Exercice 4.3. — Réciproquement, montrer que si $f : \mathscr{E} \to \mathscr{E}$ est affine et vérifie $\overrightarrow{f} = \mathrm{id}_E$ alors $f = t_u$ pour un certain $u \in E$.

Proposition 4.4. — *Pour tout $u, v \in E$, on a $\boxed{t_v \circ t_u = t_{u+v} = t_u \circ t_v}$. Par conséquent, l'ensemble T des translations forme un groupe commutatif (isomorphe à E) : l'élément neutre est $t_0 = \mathrm{id}_{\mathscr{E}}$ et l'inverse de t_u est t_{-u}.*

Démonstration. — Soit $A \in \mathscr{E}$, posons $B = t_u(A)$ et $C = t_v(B) = (t_v \circ t_u)(A)$. Alors $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC} = u + v$ d'où l'égalité $t_{u+v}(A) = C = (t_v \circ t_u)(A)$. Ceci prouve que $t_v \circ t_u = t_{u+v}$, et comme $u + v = v + u$, ceci égale aussi $t_u \circ t_v$. Il est clair que $t_0 = \mathrm{id}_{\mathscr{E}}$ est élément neutre, et l'égalité $t_u \circ t_{-u} = t_0$ montre que t_{-u} est l'inverse de t_u (ce qui est aussi évident géométriquement).

L'application $E \to T$, $u \mapsto t_u$ est donc un morphisme de groupes, surjectif par définition. Il est injectif car si $t_u = \mathrm{id}_{\mathscr{E}}$ alors $u = \overrightarrow{AA} = \overrightarrow{0}$. Donc c'est un isomorphisme de E, considéré comme groupe abélien, sur le groupe des translations. □

Donnons maintenant quelques propriétés générales (et utiles) des applications affines.

Proposition 4.5. — *Soient $(\mathscr{E}, E)$, $(\mathscr{E}', E')$, $(\mathscr{E}'', E'')$ trois espaces affines, $f : \mathscr{E} \to \mathscr{E}'$ et $g : \mathscr{E}' \to \mathscr{E}''$ des applications affines.*

(i) *Alors l'application $g \circ f$ est affine, de partie linéaire $\overrightarrow{g \circ f} = \overrightarrow{g} \circ \overrightarrow{f}$.*

(ii) *D'autre part, f est bijective ssi $\overrightarrow{f}$ l'est. Dans ce cas, f^{-1} est affine, de partie linéaire $(\overrightarrow{f})^{-1}$.*

Démonstration. — (i) Soient $A, B \in \mathscr{E}$, posons $A' = f(A)$ et $A'' = g(A') = (g \circ f)(A)$ et définissons de même B' et B''. On a :

$$\begin{aligned} \overrightarrow{A''B''} &= \overrightarrow{g}(\overrightarrow{A'B'}) && \text{(car } g \text{ est affine)} \\ &= \overrightarrow{g}\big(\overrightarrow{f}(\overrightarrow{AB})\big) && \text{(car } f \text{ est affine)} \\ &= (\overrightarrow{g} \circ \overrightarrow{f})(\overrightarrow{AB}). \end{aligned}$$

Ceci prouve que $g \circ f$ est affine, de partie linéaire égale à $\overrightarrow{g} \circ \overrightarrow{f}$.

(ii) Fixons $O \in \mathscr{E}$, posons $O' = f(O)$ et notons $\theta : E \to \mathscr{E}$ et $\theta' : E' \to \mathscr{E}'$ les bijections $u \mapsto O + u$ et $u' \mapsto O' + u'$. Comme $f \circ \theta = \theta' \circ \overrightarrow{f}$, on voit que f est bijective ssi $\overrightarrow{f}$ l'est. Dans ce cas, on a $f^{-1}(O') = O$ et $\theta \circ (\overrightarrow{f})^{-1} = f^{-1} \circ \theta'$, i.e. le diagramme :

$$\begin{tikzcd} \mathscr{E} & \mathscr{E}' \arrow[l, "f^{-1}"'] \\ E \arrow[u, "\theta"] & E' \arrow[l, "(\overrightarrow{f})^{-1}"'] \arrow[u, "\theta'"'] \end{tikzcd}$$

est commutatif, et ceci montre que f^{-1} est affine de partie linéaire $(\overrightarrow{f})^{-1}$ (cf. la preuve de 4.1). $\square$

***Proposition 4.6* (Écriture d'une application affine dans des repères)**
Soient $(\mathscr{E}, E)$ et $(\mathscr{E}', E')$ deux espaces affines, de dimensions respectives m et n, et munis de repères $\mathscr{R} = (O, \mathscr{B})$ et $\mathscr{R}' = (O', \mathscr{C})$.

(i) *Soit $f : \mathscr{E} \to \mathscr{E}'$ une application affine. Soit $A = \mathrm{Mat}_{\mathscr{C},\mathscr{B}}(\overrightarrow{f}) \in M_{n,m}(\mathbb{R})$ et soient $(b'_1, \dots, b'_n)$ les coordonnées de $f(O)$ dans $\mathscr{R}'$. Alors, pour tout $P \in \mathscr{P}$, de coordonnées $(x_1, \dots, x_m)$ dans $\mathscr{R}$, les coordonnées $(y_1, \dots, y_n)$ de $f(P)$ dans $\mathscr{R}'$ sont données par :*

$$\boxed{\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} + \begin{pmatrix} b'_1 \\ \vdots \\ b'_n \end{pmatrix}.}$$

(ii) *Réciproquement, si f est donnée par la formule ci-dessus, alors f est affine et A est la matrice de $\overrightarrow{f}$ dans les bases $\mathscr{B}$ (au départ) et $\mathscr{C}$ (à l'arrivée).*

Démonstration. — (i) En effet, on a $\begin{pmatrix} y_1 - b'_1 \\ \vdots \\ y_n - b'_n \end{pmatrix} = \overrightarrow{f(O)f(P)} = \overrightarrow{f}(\overrightarrow{OP}) = A \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$.

(ii) Notons $\phi : E \to E'$ l'application définie par $f(O + u) = f(O) + \phi(u)$ et notons X, resp. Z, le vecteur colonne représentant u dans la base $\mathscr{B}$ (resp. $\phi(u)$ dans la base $\mathscr{C}$). La formule donnée entraîne que $Z = AX$: ceci montre que ϕ est linéaire, donc f est affine et $\phi = \overrightarrow{f}$, et que $\mathrm{Mat}_{\mathscr{C},\mathscr{B}}(\overrightarrow{f}) = A$. $\square$

Remarques 4.7. — (1) On dit que deux espaces affines $(\mathscr{E}, E)$ et $(\mathscr{E}', E')$ sont *isomorphes* s'il existe une application affine bijective $f : \mathscr{E} \to \mathscr{E}'$. Dans ce cas, $\overrightarrow{f}$ est un isomorphisme de E sur E', donc si $\mathscr{E}$ est de dimension finie n, il en est de même de $\mathscr{E}'$.

(2) Exercice. Soit $(\mathscr{E}, E)$ un espace affine de dimension n et $\mathscr{R} = (O, \mathscr{B})$ un repère de $\mathscr{E}$. Montrer que l'application $f : \mathscr{E} \to k^n$, $p \mapsto (x_1, \dots, x_n)$, où $(x_1, \dots, x_n)$ sont les coordonnées de p dans $\mathscr{R}$, est affine et bijective.

Donnons un autre exemple important d'applications affines $\mathscr{E} \to \mathscr{E}$.

***Définition 4.8* (Homothéties).** — Soit $\lambda \in k^\times = k - \{0\}$. Pour A fixé dans $\mathscr{E}$, on note $h = h(A, \lambda)$ l'application qui à tout M de $\mathscr{E}$ associe l'unique point M' tel que $\overrightarrow{AM'} = \lambda\overrightarrow{AM}$. Remarquons que :

a) Pour $M = A$, on a $\overrightarrow{AA'} = \overrightarrow{0}$ d'où $h(A) = A$, donc A est un *point fixe* de h.

b) Pour tout M, on a donc : $\overrightarrow{A'M'} = \overrightarrow{AM'} = \lambda\overrightarrow{AM}$. Ceci montre que h est affine, de partie linéaire l'homothétie vectorielle $h_\lambda = \lambda\,\mathrm{id}_E$ (définie par $h_\lambda(u) = \lambda u$ pour tout $u \in E$).

c) Si $\lambda = 1$ alors h est l'application identique $\mathrm{id}_{\mathscr{E}}$ de $\mathscr{E}$.

d) Si $\lambda \neq 1$ alors A est l'unique point fixe de h. En effet si $M = h(M)$ alors $\overrightarrow{AM} = \overrightarrow{Ah(M)} = \lambda\overrightarrow{AM}$ d'où $(1 - \lambda)\overrightarrow{AM} = \overrightarrow{0}$ et si $\lambda \neq 1$ ceci entraîne $\overrightarrow{AM} = \overrightarrow{0}$ d'où $M = A$. Donc, si $\lambda \neq 1$, on dira que $h(A, \lambda)$ est « *l'homothétie de rapport λ et de centre A* ».

e) On voit facilement que $h(A, \lambda) \circ h(A, \mu) = h(A, \lambda\mu)$ donc les homothéties fixant A (i.e. l'identité et celles de centre A et rapport $\neq 1$) forment un groupe isomorphe au groupe multiplicatif $k^\times$.

Proposition 4.9. — *Soit $h : \mathscr{E} \to \mathscr{E}$ une application affine telle que $\overrightarrow{h} = \lambda\,\mathrm{id}_E$ avec $\lambda \neq 1$. Alors h possède un unique point fixe A, et si $\lambda \neq 0$ alors $h = h(A, \lambda)$.*

Démonstration. — Fixons un point $O \in \mathscr{E}$. Alors, un point $A \in \mathscr{E}$ arbitraire vérifie $h(A) = A$ si et seulement si l'on a $\overrightarrow{Oh(A)} = \overrightarrow{OA}$. Or on a

$$\overrightarrow{Oh(A)} = \overrightarrow{Oh(O)} + \overrightarrow{h(O)h(A)} = \overrightarrow{Oh(O)} + \overrightarrow{h}(\overrightarrow{OA}) = \overrightarrow{Oh(O)} + \lambda\overrightarrow{OA}$$

donc on voit que $h(A) = A$ équivaut à $(1-\lambda)\overrightarrow{OA} = \overrightarrow{Oh(O)}$ et comme $\lambda \neq 1$ ceci détermine A de façon unique, i.e. on a $\overrightarrow{OA} = (1-\lambda)^{-1}\,\overrightarrow{Oh(O)}$.

Alors, pour tout $M \in \mathscr{E}$, on a $\overrightarrow{Ah(M)} = \overrightarrow{h(A)h(M)} = \lambda\overrightarrow{AM}$ donc, si $\lambda \neq 0$ alors h est bien l'homothétie $h(A, \lambda)$ de centre A et de rapport λ. (Si $\lambda = 0$ alors $\overrightarrow{Ah(M)} = \overrightarrow{0}$ donc $h(M) = A$ pour tout M i.e. h est l'application « constante » qui envoie tout M sur A). □

Exercice 4.10. — Soient $A, B \in \mathscr{E}$ et $\lambda, \mu \in k^\times$, on considère la composée $h(A, \lambda) \circ h(B, \mu)$. Montrer que si $\lambda\mu \neq 1$ (resp. $= 1$), c'est $h(C, \lambda\mu)$ pour un point C à déterminer (resp. une translation, de vecteur à déterminer).

Avant d'introduire d'autres exemples : les projections et, si $\mathrm{car}(k) \neq 2$, les symétries, démontrons la proposition suivante.

Proposition 4.11. — *Soient $(\mathscr{E}, E)$ un espace affine, $P, Q \in \mathscr{E}$, F, G deux sev de E. On considère les deux sous-espaces affines $\mathscr{F} = P + F$ et $\mathscr{G} = Q + G$.*

a) *Si $\mathscr{F} \cap \mathscr{G}$ est non vide, c'est un sous-espace affine de direction $F \cap G$.*
b) *Ceci est le cas ssi $\overrightarrow{PQ}$ appartient au sev $F + G$.*
c) *Par conséquent, si $F \oplus G = E$ alors $\mathscr{F} \cap \mathscr{G}$ est un singleton $\{I\}$.*

Démonstration. — (a) Supposons que $\mathscr{F} \cap \mathscr{G}$ contienne un point R. Pour un point $M \in \mathscr{E}$ arbitraire, on a les équivalences : $M \in \mathscr{F} \cap \mathscr{G} \Leftrightarrow (M \in \mathscr{F}$ et $M \in \mathscr{G}) \Leftrightarrow (\overrightarrow{RM} \in F$ et $\overrightarrow{RM} \in G) \Leftrightarrow \overrightarrow{RM} \in F \cap G \Leftrightarrow M \in R + (F \cap G)$. Ceci montre que $\mathscr{F} \cap \mathscr{G} = R + (F \cap G)$.

(b) $\mathscr{F} \cap \mathscr{G}$ est non vide ssi il existe $u \in F$ et $v \in G$ tels que $P + u = R = Q + v$, ce qui équivaut à $Q = P + u - v$ ou encore à $\overrightarrow{PQ} = u - v$. Ceci équivaut à dire que $\overrightarrow{PQ}$ appartient au sous-espace vectoriel $F + G = \{u + w \mid u \in F,\ w \in G\} = \{u - v \mid u \in F,\ v \in G\}$.

(c) Supposons que $F + G = E$. Alors $\overrightarrow{PQ}$ appartient à $F + G$ donc, d'après (b) et (a), $\mathscr{F} \cap \mathscr{G}$ est un sous-espace affine $R + (F \cap G)$ de direction $(F \cap G)$. Si de plus F et G sont en somme directe, c.-à-d. si $F \cap G = \{0\}$ alors $\mathscr{F} \cap \mathscr{G}$ est le singleton $\{R\}$. (Noter que, pour tout point $I \in \mathscr{E}$, le sous-espace affine de direction $\{0\}$ passant par I est le singleton $\{I\}$.) □

Définition et proposition 4.12 **(Projections et symétries)**

Soient $(\mathscr{E}, E)$ un espace affine, F, G deux sev de E qui sont supplémentaires, et $\mathscr{F}$ un sous-espace affine de $\mathscr{E}$ de direction F.

(i) *Pour tout $M \in \mathscr{E}$, il existe un unique point $p(M) \in \mathscr{F}$ tel que $\overrightarrow{Mp(M)} \in G$. L'application $p : M \mapsto p(M)$ est* affine, *de partie linéaire la projection π de E sur F de noyau G. On dit que p est la* projection de $\mathscr{E}$ sur $\mathscr{F}$ de direction G. [19]

[19] On dit aussi : « *parallèlement à G* ».

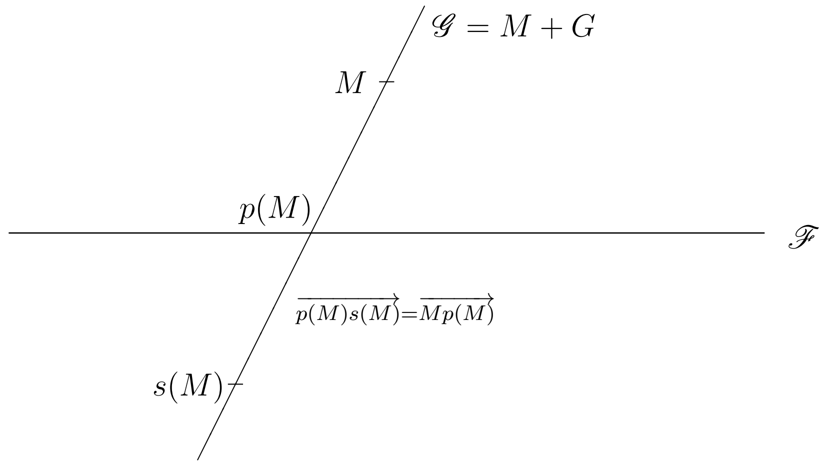
(ii) *Supposons* car$(k) \neq 2$. *Alors, pour tout* $M \in \mathscr{E}$, *on pose* $s(M) = M + 2\overrightarrow{Mp(M)}$. *L'application* $s : M \mapsto s(M)$ *est* affine, *de partie linéaire la symétrie par rapport à* F *de direction* G. *On dit que* s *est la* symétrie par rapport à $\mathscr{F}$ de direction G.[19]

Démonstration. — (i) Comme $F \oplus G = E$, les sous-espaces affines $\mathscr{F}$ et $M + G$ se coupent en un unique point, qu'on note $M' = p(M)$. Pour $P \in \mathscr{E}$, le vecteur $\overrightarrow{MP}$ s'écrit de façon unique $\overrightarrow{MP} = u + v$ avec $u \in F$ et $v \in G$, et par définition l'on a $\pi(\overrightarrow{MP}) = u$. D'autre part, posant $P' = p(P)$ on a :

$$\overrightarrow{MP} = \overrightarrow{MM'} + \overrightarrow{M'P'} + \overrightarrow{P'P}$$

et comme $\overrightarrow{MM'}, \overrightarrow{PP'} \in G$ et $\overrightarrow{M'P'} \in F$ (puisque $M', P' \in \mathscr{F}$), il en résulte que $\overrightarrow{M'P'} = \pi(\overrightarrow{MP})$. Ceci montre que p est affine, de partie linéaire π.

La démonstration de (ii) est laissée au lecteur. (On n'en a pas besoin pour le moment.) $\square$



Projection p sur $\mathscr{F}$ et symétrie s par rapport à $\mathscr{F}$, de direction G.

Une conséquence du fait que les projections sont des applications affines est le théorème de Thalès.[20] Avant de l'énoncer, on a besoin de la notation suivante.

Notation 4.13. — Soient u, v deux vecteurs non nuls d'un k-espace vectoriel V. S'ils sont colinéaires, il existe un unique $\lambda \in k^\times$ tel que $u = \lambda v$. Dans ce cas, on désigne par $\dfrac{u}{v}$ le scalaire λ.

***Théorème 4.14* (de Thalès).** — *Soit* $(\mathscr{E}, E)$ *un espace affine de dimension* ≥ 2, *soient* H *un hyperplan de* E, $\mathscr{D}, \mathscr{D}'$ *deux droites distinctes dont les directions* D *et* D' *ne sont pas contenues dans* H, *et* $\mathscr{H}_1, \mathscr{H}_2, \mathscr{H}_3$ *trois hyperplans de* $\mathscr{E}$ *de direction* H, *deux à deux distincts. Alors :*

(i) $\mathscr{D}$ *(resp.* $\mathscr{D}'$*) coupe chaque* $\mathscr{H}_i$ *en un unique point* A_i *(resp.* B_i*).*

(ii) *On a* $\dfrac{\overrightarrow{A_3A_2}}{\overrightarrow{A_3A_1}} = \dfrac{\overrightarrow{B_3B_2}}{\overrightarrow{B_3B_1}}$. *De plus,* A_3 *et* A_1 *étant fixés, ce scalaire* λ *détermine* A_2, *i.e. si un point* M *de* $\mathscr{D}$ *vérifie* $\dfrac{\overrightarrow{A_3M}}{\overrightarrow{A_3A_1}} = \lambda$ *alors* $M = A_2$.

(iii) *Si de plus* $\mathscr{D}$ *et* $\mathscr{D}'$ *sont concourantes en* $A_3 = B_3$, *on a aussi* $\dfrac{\overrightarrow{A_2B_2}}{\overrightarrow{A_1B_1}} = \lambda$.

[20] Thalès de Milet, mathématicien grec qui vécut d'environ -624 à -547 de notre ère.

Démonstration. — (i) Comme $D \not\subset H$ alors $E = D \oplus H$ donc $\mathscr{D}$ coupe chaque $\mathscr{H}_i$ en un unique point A_i, et de même pour $\mathscr{D}'$.

(ii) Soit p la projection de $\mathscr{E}$ sur $\mathscr{D}'$ parallèlement à H. Comme $\overrightarrow{A_iB_i} \in H$, alors $p(A_i) = B_i$ pour $i = 1, 2, 3$. Posons $\overrightarrow{A_3A_2} = \lambda\,\overrightarrow{A_3A_1}$. Alors on a

$$\overrightarrow{B_3B_2} = \overrightarrow{p(A_3)p(A_2)} = \pi(\overrightarrow{A_3A_2}) = \lambda\pi(\overrightarrow{A_3A_1}) = \lambda\,\overrightarrow{p(A_3)p(A_1)} = \lambda\,\overrightarrow{B_3B_1}.$$

Ceci prouve la première assertion de (ii). De plus, si un point $M \in \mathscr{D}$ vérifie $\dfrac{\overrightarrow{A_3M}}{\overrightarrow{A_3A_1}} = \lambda$ alors on a $\overrightarrow{A_3M} = \lambda\,\overrightarrow{A_3A_1} = \overrightarrow{A_3A_2}$ et donc $M = A_2$.

(iii) Dans ce cas, posons $O = A_3 = B_3$. Alors l'homothétie de centre O et de rapport λ envoie A_1 sur A_2 et B_1 sur B_2, donc on a $\overrightarrow{A_2B_2} = \lambda\,\overrightarrow{A_1B_1}$, ce qui prouve (iii). □

Remarque. — L'assertion (iii) admet une réciproque partielle : soient $\mathscr{D}$ et $\mathscr{D}'$ deux droites concourantes en un point O et A_1, A_2 (resp. B_1, B_2) deux points de $\mathscr{D}$ (resp. $\mathscr{D}'$) distincts de O. Si $\dfrac{\overrightarrow{OA_2}}{\overrightarrow{OA_1}}$ et $\dfrac{\overrightarrow{OB_2}}{\overrightarrow{OB_1}}$ sont égaux au même scalaire λ, alors l'homothétie de centre O et de rapport λ envoie A_1 sur A_2 et B_1 sur B_2, donc on a $\overrightarrow{A_2B_2} = \lambda\,\overrightarrow{A_1B_1}$. Par conséquent, les droites (A_1B_1) et (A_2B_2) sont parallèles.

Si de plus $\dim(\mathscr{E}) = 2$ ces droites sont des hyperplans $\mathscr{H}_1, \mathscr{H}_2$ et l'on obtient donc l'énoncé suivant.

***Théorème 4.15* (réciproque de Thalès dans le plan)**. — *Dans un plan affine $\mathscr{P}$, soient $\mathscr{D}, \mathscr{D}'$ deux droites distinctes, sécantes en un point O. Soient A_1, A_2 (resp. B_1, B_2) deux points de $\mathscr{D}$ (resp. $\mathscr{D}'$), distincts de O. Si $\dfrac{\overrightarrow{OA_2}}{\overrightarrow{OA_1}} = \dfrac{\overrightarrow{OB_2}}{\overrightarrow{OB_1}}$ alors les droites $\mathscr{H}_1 = (A_1B_1)$ et $\mathscr{H}_2 = (A_2B_2)$ sont parallèles.*

5. Groupe affine et produits semi-directs

Soit $(\mathscr{E}, E)$ un espace affine. Dans cette section, on étudie plus en détail la structure du groupe affine $\mathrm{GA}(\mathscr{E})$, formé des applications affines *bijectives* $\mathscr{E} \to \mathscr{E}$, et de certains de ses sous-groupes. Commençons par quelques propriétés valables pour toute application affine $f : \mathscr{E} \to \mathscr{E}$ (pas nécessairement bijective).

Notation. — On note $\mathrm{Taff}(\mathscr{E})$ l'ensemble des transformations affines de $\mathscr{E}$, i.e. applications affines $\mathscr{E} \to \mathscr{E}$.

Proposition 5.1. — *Soit $(\mathscr{E}, E)$ un espace affine.*

(i) $\mathrm{Taff}(\mathscr{E})$ *est stable par composition, c.-à-d., si $f, g \in \mathrm{Taff}(\mathscr{E})$ alors $g \circ f \in \mathrm{Taff}(\mathscr{E})$.*

(ii) *Soit $f \in \mathrm{Taff}(\mathscr{E})$, alors f bijective $\Leftrightarrow$ $\overrightarrow{f}$ bijective ; dans ce cas $f^{-1} \in \mathrm{Taff}(\mathscr{E})$ et sa partie linéaire est $(\overrightarrow{f})^{-1}$.*

(iii) *On note $\mathrm{GA}(\mathscr{E}) = \{f \in \mathrm{Taff}(\mathscr{E}) \mid f \text{ bijective}\}$; c'est un* **groupe** *et l'application $\mathrm{GA}(\mathscr{E}) \to \mathrm{GL}(E)$, $f \mapsto \overrightarrow{f}$ est un morphisme de groupes.*

(iv) *Pour tout $f \in \mathrm{Taff}(\mathscr{E})$ et $u \in E$, on a $f \circ t_u = t_{\overrightarrow{f}(u)} \circ f$; en particulier si $f \in \mathrm{GA}(\mathscr{E})$ alors $f \circ t_u \circ f^{-1} = t_{\overrightarrow{f}(u)}$.*

(v) *Pour tout* $O \in \mathscr{E}$, *tout* $f \in \mathrm{Taff}(\mathscr{E})$ *s'écrit* $f = t_{\overrightarrow{Of(O)}} \circ \phi_O$, *où* ϕ_O *est l'application affine définie par* $\phi_O(P) = O + \overrightarrow{f}(\overrightarrow{OP})$. *(i.e. via la bijection* $E \xrightarrow{\sim} \mathscr{E}$, $u \mapsto O + u$, ϕ_O *correspond à* $\overrightarrow{f}$ *).*

Démonstration. — (i), (ii) et (iii) découlent de 4.5. Prouvons (iv) : pour tout $P \in \mathscr{E}$, on a

$$(f \circ t_u)(P) = f(P+u) = f(P) + \overrightarrow{f}(u) = (t_{\overrightarrow{f}(u)} \circ f)(P)$$

ce qui prouve le premier point de (iv), et le second en découle.

Enfin, soient $O \in \mathscr{E}$ et $f \in \mathrm{Taff}(\mathscr{E})$, posons $u = \overrightarrow{Of(O)}$ et $g = t_{-u} \circ f$. Alors g est affine, sa partie linéaire est $\mathrm{id}_E \circ \overrightarrow{f} = \overrightarrow{f}$, et l'on a $g(O) = O$. Donc, pour tout $P \in \mathscr{E}$, on a $\overrightarrow{Og(P)} = \overrightarrow{g(O)g(P)} = \overrightarrow{f}(\overrightarrow{OP})$ donc g égale ϕ_O, et l'on a $f = t_u \circ \phi_O$. Ceci prouve (v). □

Terminologie 5.2. — Le groupe $\mathrm{GA}(\mathscr{E})$ s'appelle le *groupe affine* de $\mathscr{E}$.

Remarques 5.3. — Il résulte du point (iv) de la proposition précédente que le sous-groupe T des translations est un sous-groupe *distingué* de $\mathrm{GA}(\mathscr{E})$. D'autre part, il résulte de l'exercice 4.3 que T est le noyau du morphisme de groupes $\mathrm{GA}(\mathscr{E}) \to \mathrm{GL}(E)$, $f \mapsto \overrightarrow{f}$. On a donc un isomorphisme canonique $\mathrm{GA}(\mathscr{E})/T \xrightarrow{\sim} \mathrm{GL}(E)$.

De plus, on a le théorème suivant.

Théorème 5.4. — *Soient* $(\mathscr{E}, E)$ *un espace affine,* $\mathrm{GA}(\mathscr{E})$ *le groupe affine de* $\mathscr{E}$ *et* T *le sous-groupe des translations (isomorphe à* E*). Pour tout* $O \in \mathscr{E}$ *on a :*

(i) *Le sous-groupe* $G_O = \{g \in \mathrm{GA}(\mathscr{E}) \mid g(O) = O\}$ *est isomorphe à* $\mathrm{GL}(E)$ *via le morphisme* $f \mapsto \overrightarrow{f}$.

(ii) $\mathrm{GA}(\mathscr{E})$ *est le produit semi-direct de* T *par* G_O.

Démonstration. — Fixons $O \in \mathscr{E}$. Alors l'application $G_O \to \mathrm{GL}(E)$, $f \mapsto \overrightarrow{f}$ est un isomorphisme de groupes, sa réciproque étant le morphisme qui à tout $\phi \in \mathrm{GL}(E)$ associe l'application affine $M \mapsto O + \phi(\overrightarrow{OM})$. Ceci prouve (i).

De plus, pour tout $h \in \mathrm{GA}(\mathscr{E})$, l'application affine $t_{\overrightarrow{h(O)O}} \circ h$ appartient à G_O et l'on obtient ainsi des bijections réciproques :

$$\begin{array}{ccccccc} T \times G_O & \longrightarrow & \mathrm{GA}(\mathscr{E}) & \quad \text{et} \quad & \mathrm{GA}(\mathscr{E}) & \longrightarrow & T \times G_O \\ (t_u, g) & \mapsto & t_u \circ g & & h & \mapsto & \big(t_{\overrightarrow{Oh(O)}},\, t_{\overrightarrow{h(O)O}} \circ h\big). \end{array}$$

Ceci prouve (ii). □

Remarque 5.5. — Le théorème dit que l'on a des isomorphismes $\mathrm{GA}(\mathscr{E}) \simeq T \rtimes \mathrm{GL}(E) \simeq E \rtimes \mathrm{GL}(E)$. Mais attention, le 1er isomorphisme n'est pas canonique, car $\mathrm{GL}(E)$ n'est pas de façon naturelle un sous-groupe de $\mathrm{GL}(\mathscr{E})$: il faut choisir une origine O pour avoir le sous-groupe G_O !

Théorème 5.6. — *Soient* $(\mathscr{E}, E)$ *un espace affine,* $\mathrm{GA}(\mathscr{E})$ *son groupe affine et* T *le sous-groupe des translations.*

(i) *L'ensemble* G *des translations et des homothéties forme un groupe (non commutatif), appelé le groupes des homothéties et translations et noté* $\mathrm{HT}(\mathscr{E})$.

(ii) T *en est un sous-groupe distingué et le quotient* $\mathrm{HT}(\mathscr{E})/T$ *est isomorphe au groupe des homothéties de* E*, lui-même isomorphe au groupe multiplicatif* $k^\times$.

(iii) *Pour tout* $O \in \mathscr{E}$, *le sous-groupe* $H_O = \{h \in \mathrm{HT}(\mathscr{E}) \mid h(O) = O\}$ *est formé des homothéties fixant O, donc isomorphe à* $k^\times$, *et* $\mathrm{HT}(\mathscr{E})$ *est le produit semi-direct de* T *par* H_O.

(iv) $\mathrm{HT}(\mathscr{E})$ *est un sous-groupe distingué de* $\mathrm{GA}(\mathscr{E})$.

Démonstration. — (i), (ii) et (iii) sont laissés en exercice. Prouvons (iv). Soient $f \in \mathrm{GA}(\mathscr{E})$ et $h \in \mathrm{HT}(\mathscr{E})$. Si $h = t_u$ on sait déjà que $f \circ t_u \circ f^{-1} = t_{\overrightarrow{f}(u)}$. Donc on peut supposer que $h = h(A, \lambda)$ avec $\lambda \neq 1$. Notant $L(\cdot)$ la partie linéaire, comme $L(h) = \lambda\,\mathrm{id}_E$ est dans le centre de $\mathrm{GL}(E)$, on a $L(f \circ h \circ f^{-1}) = L(f) \circ (\lambda\,\mathrm{id}_E) \circ L(f)^{-1} = \lambda\,\mathrm{id}_E$.

Donc, d'après la proposition 4.9, $f \circ h \circ f^{-1}$ est une homothétie de rapport λ et de centre B à déterminer. Or on voit que $f(A)$ est laissé fixe, donc c'est le centre B cherché.

A posteriori (i.e. connaissant le résultat), on peut aussi faire le calcul suivant : pour tout $u \in E$ on a

$$(f \circ h \circ f^{-1})(f(A) + u) = (f \circ h)\big(A + (\overrightarrow{f})^{-1}(u)\big) = f\big(A + \lambda(\overrightarrow{f})^{-1}(u)\big) = f(A) + \lambda u$$

et ceci montre que $f \circ h \circ f^{-1} = h(f(A), \lambda)$. $\square$

6. Barycentres

On fixe un k-espace vectoriel E et un espace affine $\mathscr{E}$ de direction E.

Définition et proposition 6.1. — *Soient* $A_1, \dots, A_n \in \mathscr{E}$ *et* $\lambda_1, \dots, \lambda_n \in k$.

(1) *On suppose* $\lambda_1 + \cdots + \lambda_n = 1$. *Alors, si on choisit un point auxiliaire* O *et qu'on définit le point* G *par l'égalité*

$$\overrightarrow{OG} = \sum_{i=1}^{n} \lambda_i \overrightarrow{OA_i}$$

ce point G ne dépend pas *en fait du point* O *choisi, i.e.* **pour tout** $I \in \mathscr{E}$, *on aura*

$$(*) \qquad \overrightarrow{IG} = \sum_{i=1}^{n} \lambda_i \overrightarrow{IA_i}.$$

Ce point G *est appelé «* **barycentre** *des points pondérés* (A_i, λ_i) *» et l'on écrira :*

$$\boxed{G = \lambda_1 A_1 + \cdots + \lambda_n A_n.}$$

Remarquons encore qu'en prenant $I = G$ *dans* $(*)$, *le point* G *est caractérisé par l'égalité* $\sum_{i=1}^{n} \lambda_i \overrightarrow{GA_i} = \overrightarrow{0}$.

(1′) *Plus généralement, si* $S = \lambda_1 + \cdots + \lambda_n \neq 0$, *il existe un unique point* $G \in \mathscr{E}$ *tel que,* **pour tout** $O \in \mathscr{E}$, *on ait* $(\lambda_1 + \cdots + \lambda_n)\overrightarrow{OG} = \sum_{i=1}^{n} \lambda_i \overrightarrow{OA_i}$. *Par abus de langage, on dit encore parfois que* G *est le barycentre des points pondérés* (A_i, λ_i). [21]

(2) *Au contraire, si* $\lambda_1 + \cdots + \lambda_n = 0$ *alors le vecteur* $\overrightarrow{u} = \sum_{i=1}^{n} \lambda_i \overrightarrow{OA_i}$ *est* **indépendant** *du choix de* O. *On écrit alors que* $\overrightarrow{u} = \lambda_1 A_1 + \cdots + \lambda_n A_n$.

[21] En sous-entendant qu'on remplace chaque λ_i par λ_i/S pour avoir une somme égale à 1.

Démonstration. — (1) Fixons un point $O \in \mathscr{E}$ et définissons G par l'égalité $\overrightarrow{OG} = \sum_{i=1}^n \lambda_i \overrightarrow{OA_i}$. Alors, pour tout point O' on a, d'après la relation de Chasles :

$$\sum_{i=1}^n \lambda_i \overrightarrow{O'A_i} = \sum_{i=1}^n \lambda_i (\overrightarrow{O'O} + \overrightarrow{OA_i}) = \underbrace{(\sum_{i=1}^n \lambda_i)}_{=1} \overrightarrow{O'O} + \sum_{i=1}^n \lambda_i \overrightarrow{OA_i} = \overrightarrow{O'O} + \overrightarrow{OG} = \overrightarrow{O'G}.$$

Ceci montre que $\overrightarrow{O'G} = \sum_{i=1}^n \lambda_i \overrightarrow{O'A_i}$ pour **tout** $O' \in \mathscr{E}$, et G est bien sûr unique, car cette égalité pour **un** O' fixé suffit à déterminer G.

(1′) Posons $S = \sum_{i=1}^n \lambda_i$. On voit que (1′) se ramène au cas (1) en remplaçant chaque λ_i par $\lambda_i' = \lambda_i/S$, car alors la somme des λ_i' vaut 1 donc $G = \lambda_1' A_1 + \cdots + \lambda_n' A_n$ est bien défini et pour tout point O, on a

$$\overrightarrow{OG} = \sum_{i=1}^n \frac{\lambda_i}{S} \overrightarrow{OA_i} \qquad \text{i.e.} \qquad (\lambda_1 + \cdots + \lambda_n)\overrightarrow{OG} = \sum_{i=1}^n \lambda_i \overrightarrow{OA_i}.$$

(2) Fixons un point $O \in \mathscr{E}$ et posons $\overrightarrow{u} = \sum_{i=1}^n \lambda_i \overrightarrow{OA_i}$. Alors, pour tout point O' on a, d'après la relation de Chasles :

$$\sum_{i=1}^n \lambda_i \overrightarrow{O'A_i} = \sum_{i=1}^n \lambda_i (\overrightarrow{O'O} + \overrightarrow{OA_i}) = \underbrace{(\sum_{i=1}^n \lambda_i)}_{=0} \overrightarrow{O'O} + \sum_{i=1}^n \lambda_i \overrightarrow{OA_i} = \overrightarrow{u}.$$

Ceci montre que $\overrightarrow{u}$ est indépendant du choix de O. □

***Terminologie 6.2* (Centre de gravité).** — Soient $A_1, \ldots, A_n \in \mathscr{E}$. Si k est un corps *de caractéristique nulle*, par exemple si $k = \mathbb{R}$, alors on peut prendre $\lambda_i = 1/n$ pour $i = 1, \ldots, n$. Dans ce cas, G s'appelle *l'isobarycentre* ou *centre de gravité* [22] des points $A_1, \ldots, A_n$.

***Exemple 6.3* (Médianes d'un triangle).** — Par exemple, soient A, B, C trois points non alignés du plan affine **réel** et soit G leur isobarycentre, appelé le « centre de gravité du triangle ABC » ; on a donc :

$$(*) \qquad \overrightarrow{0} = \overrightarrow{GA} + \overrightarrow{GB} + \overrightarrow{GC}.$$

D'autre part, notons A' le milieu du segment $[B, C]$ et définissons de même B' et C'. Alors A' est l'isobarycentre de B et C donc pour tout point P on a $2\overrightarrow{PA'} = \overrightarrow{PB} + \overrightarrow{PC}$. Appliquant ceci à $P = G$, on déduit de $(*)$ que :

$$-\overrightarrow{GA} = \overrightarrow{GB} + \overrightarrow{GC} = 2\overrightarrow{GA'}.$$

Ceci montre que le point G appartient au segment $[AA']$ et que le vecteur $\overrightarrow{GA}$ est deux fois plus long (et de sens contraire) que le vecteur $\overrightarrow{GA'}$. On a le même résultat pour $[BB']$ et $[CC']$. Ceci montre que les médianes du triangle, i.e. les droites (AA'), (BB') et (CC'), sont concourantes et que G est situé sur chaque segment $[AA']$, etc. aux deux-tiers de la longueur, en partant du sommet.

[22] On dit aussi *centre d'inertie*, notamment en physique.

Remarque 6.4. — Dans l'exemple précédent, l'hypothèse que $\mathrm{car}(k)$ soit distincte de 2 et 3 est essentielle. En effet, si $\mathrm{car}(k) = 2$ et $A \neq B$, le milieu du segment $[A, B]$ n'existe pas : il n'existe aucun point M tel que $\overrightarrow{MA} + \overrightarrow{MB} = \overrightarrow{0}$ car comme $1 + 1 = 0$, le vecteur $\overrightarrow{MA} + \overrightarrow{MB}$ est indépendant de M et égale $\overrightarrow{AB} \neq \overrightarrow{0}$.

Si $\mathrm{car}(k) = 3$, les milieux A', B', C' existent, mais les médianes (AA'), (BB') et (CC') ne sont pas concourantes ; en fait elles sont parallèles ! (cf. [Du, 3.8]). En effet, prenons A comme origine et les vecteurs $e_1 = \overrightarrow{AB}$ et $e_2 = \overrightarrow{AC}$ comme base de E (supposé de dimension 2), de sorte que B et C ont pour coordonnées $(1, 0)$ et $(0, 1)$. Alors C' a pour coordonnées $(1/2, 0) = (2, 0)$ (car $2^2 \equiv 1 \bmod. 3$). De même, B' et A' ont pour coordonnées $(0, 2)$ et $(2, 2)$. Tenant compte de l'égalité $-1 \equiv 2 \bmod. 3$, on a donc : $\overrightarrow{AA'} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$, $\overrightarrow{BB'} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ et $\overrightarrow{CC'} = \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$.

Remarque 6.5. — Dans l'exemple précédent, on a vu que G est le barycentre de $(A, 1)$ et de $(A', 2)$, où A' est le milieu de $[B, C]$, i.e. le barycentre de $(B, 1)$ et $(C, 1)$. Ceci est généralisé dans la proposition suivante, qui facilite dans certains cas le calcul du barycentre.

***Proposition 6.6* (Associativité des barycentres).** — *Soient $t_1, \dots, t_n \in k$ de somme égale à 1, $A_1, \dots, A_n \in \mathscr{E}$, et G le barycentre des points pondérés $(A_1, t_1), \dots, (A_n, t_n)$. Soit I un sous-ensemble de $\{1, \dots, n\}$ tel que $\mu = \sum_{i \in I} t_i$ et $1 - \mu$ soient non nuls et soit H (resp. H') le barycentre des points pondérés (A_i, t_i) pour $i \in I$ (resp. $i \notin I$). Alors G est le barycentre des points pondérés (H, μ) et $(H', 1 - \mu)$.*

Démonstration. — Fixons $O \in \mathscr{E}$. Alors H et H' sont définis par $\overrightarrow{OH} = \dfrac{1}{\mu} \sum_{i \in I} t_i \overrightarrow{OA_i}$ et $\overrightarrow{OH'} = \dfrac{1}{1 - \mu} \sum_{i \notin I} t_i \overrightarrow{OA_i}$. Notant G' le barycentre de (H, μ) et de $(H', 1 - \mu)$, on a donc

$$\overrightarrow{OG'} = \mu \overrightarrow{OH} + (1 - \mu) \overrightarrow{OH'} = \sum_{i=1}^{n} t_i \overrightarrow{OA_i} = \overrightarrow{OG},$$

d'où $G' = G$. □

Proposition 6.7. — *Soit $f : \mathscr{E} \to \mathscr{E}'$ une application affine. Alors f « préserve les barycentres », c.-à-d. pour tous $A_1, \dots, A_n \in \mathscr{E}$ et $\lambda_1, \dots, \lambda_n \in k$ de somme égale à 1, si $G = \sum_{i=1}^{n} \lambda_i A_i$ alors on a $f(G) = \sum_{i=1}^{n} \lambda_i f(A_i)$.*

Démonstration. — On a $\overrightarrow{0} = \sum_{i=1}^{n} \lambda_i \overrightarrow{GA_i}$. Comme $\overrightarrow{f}$ est linéaire, on a :

$$\overrightarrow{0} = \sum_{i=1}^{n} \lambda_i \overrightarrow{f}(\overrightarrow{GA_i}) = \sum_{i=1}^{n} \lambda_i \overrightarrow{f(G) f(A_i)}$$

et ceci montre que $f(G)$ est le barycentre (dans $\mathscr{E}'$) des points pondérés $(f(A_i), \lambda_i)$. □

***Définition et proposition 6.8* (Sea engendré).** — *Soient k un corps, E un k-ev, $\mathscr{E}$ un espace affine de direction E, X une partie non vide de E et F le sev de E engendré par les vecteurs $\overrightarrow{xy}$, pour $x, y \in X$.*

(i) *Pour tout $x_0 \in X$, le sea $\mathscr{F} = x_0 + F$ est le plus petit sea de $\mathscr{E}$ contenant X ; en particulier il ne dépend pas du choix de $x_0 \in X$. On dira que c'est le sea* engendré *par X et on le notera $\mathrm{Aff}\langle X \rangle$.*

(ii) *On peut aussi le décrire comme suit : c'est l'ensemble des combinaisons barycentriques de points de X, i.e. un point p de $\mathscr{E}$ appartient à $\mathrm{Aff}\langle X \rangle$ ssi il existe un entier $n \geq 1$, des points $x_1, \dots, x_n$ de X et des scalaires $t_i \in k$ de somme 1 tels que $p = t_1 x_1 + \dots + t_n x_n$.*

Démonstration. — En effet, $\mathscr{F}$ contient X et si $\mathscr{F}'$ est un sea contenant X alors sa direction F' contient tous les vecteurs $\overrightarrow{xy}$, pour $x, y \in X$, donc $\mathscr{F}'$ contient $\mathscr{F}$. Ceci prouve (i)

Pour prouver (ii), notons provisoirement $\mathscr{F}'$ l'ensemble introduit en (ii). Si p est comme en (ii), on a $\overrightarrow{x_0p} = \sum_{i=1}^n t_i\, \overrightarrow{x_0x_i}$ et donc $p \in \mathscr{F}$.

Réciproquement, notons que F est engendré par les vecteurs $\overrightarrow{x_0x}$ pour $x \in X$, puisque $\overrightarrow{xy} = \overrightarrow{x_0y} - \overrightarrow{x_0x}$. Donc un élément arbitraire de $\mathscr{F}$ s'écrit sous la forme :

$$p = x_0 + \sum_{i=1}^n t_i\, \overrightarrow{x_0x_i}$$

pour certains $x_i \in X$ et $t_i \in k$. Posons $t_0 = 1 - \sum_{i=1}^n t_i$; d'après la définition du barycentre (cf. 6.1), ceci signifie que p est le barycentre des points pondérés (x_i, t_i) pour $i = 0, \dots, n$, donc $p \in \mathscr{F}'$. Ceci montre l'égalité $\mathscr{F} = \mathscr{F}'$. □

***Définition 6.9* (Points affinement indépendants ou liés)**

Soit $(\mathscr{E}, E)$ un espace affine et soit $p \in \mathbb{N}^*$. On dit que $(p+1)$ points $A_0, \dots, A_p \in \mathscr{E}$ sont **affinement indépendants** si $\mathrm{Aff}\langle A_0, \dots, A_p\rangle$ est de dimension p. Dans le cas contraire, on dit que $A_0, \dots, A_p \in \mathscr{E}$ sont **affinement liés**.

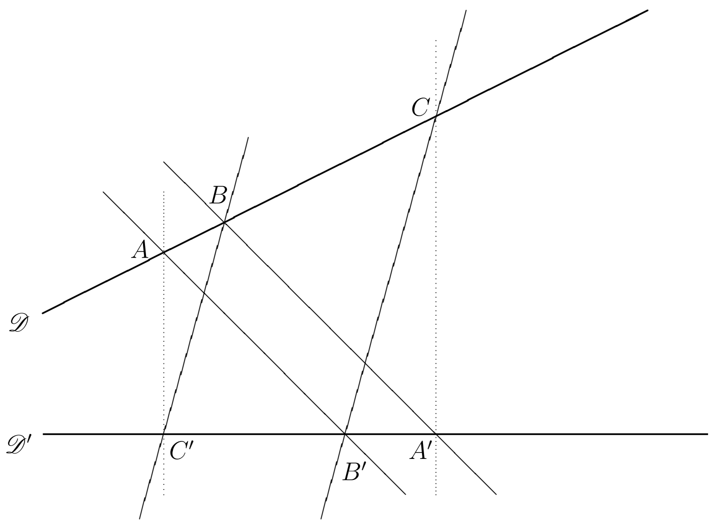
Terminologie. Si $p = 2$, trois points A_0, A_1, A_2 sont affinement liés $\Leftrightarrow$ ils sont alignés. Donc A_0, A_1, A_2 sont affinement indépendants $\Leftrightarrow$ ils sont **non alignés**.

On dit que des points $A_0, \dots, A_p \in \mathscr{E}$ sont **coplanaires** s'ils sont contenus dans un sea $\mathscr{P}$ de dimension 2 (un plan affine), i.e. si $\dim \mathrm{Aff}\langle A_0, \dots, A_p\rangle \leq 2$. Donc quatre points $A_0, \dots, A_3$ sont affinement indépendants $\Leftrightarrow$ ils sont **non coplanaires**.

7. Supplément : théorèmes de Pappus et de Desargues

On énonce dès maintenant la version affine des théorèmes de Pappus et de Desargues. Cette section sera traitée en cours plus tard, en même temps que la version projective de ces théorèmes.

***Théorème 7.1* (de Pappus).** — [23] *Dans un plan affine $\mathscr{P}$, soient $\mathscr{D}, \mathscr{D}'$ deux droites distinctes et A, B, C (resp. A', B', C') trois points distincts situés sur $\mathscr{D}$ (resp. $\mathscr{D}'$) et distincts de l'éventuel point de concours de $\mathscr{D}$ et $\mathscr{D}'$. On suppose que (AB') est parallèle à (BA') et que (BC') est parallèle à (CB'). Alors (AC') est parallèle à (CA').*



[23] Pappus d'Alexandrie, mathématicien grec du IVème siècle (dates approximatives : 290-350).

Démonstration. — Notons P la direction de $\mathscr{P}$. Distinguons deux cas, selon que $\mathscr{D}$ et $\mathscr{D}'$ sont parallèles ou concourantes.

(1) Si $\mathscr{D}$ et $\mathscr{D}'$ sont parallèles, les hypothèses entraînent que $(ABA'B')$ et $(CBC'B')$ sont des parallélogrammes, d'où $\overrightarrow{AB} = \overrightarrow{B'A'}$ et $\overrightarrow{BC'} = \overrightarrow{CB'}$. Alors on a :

$$\overrightarrow{AC'} = \overrightarrow{AB} + \overrightarrow{BC'} = \overrightarrow{B'A'} + \overrightarrow{CB'} = \overrightarrow{CA'}$$

et donc (AC') et (CA') sont parallèles.

(2) Supposons $\mathscr{D}$ et $\mathscr{D}'$ concourantes en un point O, qu'on prend comme origine de $\mathscr{P}$. Prenant une base (e_1, e_2) de P où e_1 (resp. e_2) est un vecteur directeur de $\mathscr{D}$ (resp. $\mathscr{D}'$), nos points ont les coordonnées suivantes : $A(a,0)$, $B(b,0)$, $C(c,0)$, $A'(0,a')$, $B'(0,b')$, $C'(0,c')$, où a,b,c sont distincts et non nuls, et de même pour a',b',c'. Par hypothèse, les vecteurs $\overrightarrow{AB'} = (-a,b')$ et $\overrightarrow{BA'} = (-b,a')$ d'une part, et $\overrightarrow{BC'} = (-b,c')$ et $\overrightarrow{CB'} = (-c,b')$ d'autre part, sont non nuls et liés, donc il existe $\lambda,\mu \in k^\times$ tels que

$$\begin{cases} b' = \lambda a' \\ a = \lambda b \end{cases} \qquad \text{et} \qquad \begin{cases} c' = \mu b' \\ b = \mu c \end{cases} \qquad \text{d'où} \qquad \begin{cases} c' = \mu\lambda a' \\ a = \lambda\mu c \end{cases}$$

Comme $\mu\lambda = \lambda\mu$,[24] on obtient que : $\overrightarrow{AC'} = (-a,c') = \lambda\mu(-c,a') = \lambda\mu\overrightarrow{CA'}$, et donc (AC') est parallèle à (CA').

□

Il existe plusieurs versions du théorème de Desargues.[25] Celle donnée ci-dessous est appelée « affine », pour la distinguer d'une version « projective » qui sera donnée dans un chapitre ultérieur.

***Théorème 7.2* (de Desargues affine).** — *Dans un plan affine, soient A,B,C,A',B',C' six points distincts, avec A,B,C (resp. A',B',C')* non alignés. *On suppose que :*

a) *les droites (AA'), (BB') et (CC') sont deux à deux distinctes,*

b) *les droites (AB) et $(A'B')$ sont parallèles, ainsi que les droites (AC) et $(A'C')$.*

Alors les droites (BC) et $(B'C')$ sont parallèles si et seulement si les droites (AA'), (BB') et (CC') sont concourantes ou parallèles.

Démonstration. — L'hypothèse (b) entraîne qu'il existe $\lambda,\mu \in k^\times$ tels que $\overrightarrow{A'B'} = \lambda\overrightarrow{AB}$ et $\overrightarrow{A'C'} = \mu\overrightarrow{AC}$. On a donc

$$\overrightarrow{B'C'} = \overrightarrow{A'C'} - \overrightarrow{A'B'} = \mu\overrightarrow{AC} - \lambda\overrightarrow{AB},$$

et comme $\overrightarrow{AC}, \overrightarrow{AB}$ sont linéairement indépendants, $\overrightarrow{B'C'}$ est colinéaire à $\overrightarrow{BC} = \overrightarrow{AC} - \overrightarrow{AB}$ si et seulement si $\mu = \lambda$. On a donc :

$$(BC) \text{ et } (B'C') \text{ parallèles} \iff \lambda = \mu. \tag{1}$$

D'autre part, remarquons que les droites (AB) et $(A'B')$ sont *distinctes*, car sinon $ABA'B'$ seraient alignés et l'on aurait $(AA') = (BB')$, contrairement à l'hypothèse (a).

[24] On fait cette remarque car on peut définir la notion d'espaces vectoriels ou affines sur un corps k éventuellement non commutatif, et alors la validité du théorème de Pappus équivaut à la commutativité de k. Le lecteur intéressé pourra consulter [LF, §IV.11] ou [Bo, §I.6.7].

[25] Girard Desargues (1591-1661) mathématicien et architecte français, contemporain de Fermat (1601-1665) et de Descartes (1596-1650). Par ailleurs, il eut pour élève vers 1640 le jeune Blaise Pascal (1623-1662).

De même, $(AC) \neq (A'C')$. Alors on a les équivalences suivantes : $\lambda = 1 \Leftrightarrow ABB'A'$ est un parallélogramme $\Leftrightarrow$ (AA') est parallèle à (BB'), et de même : $\mu = 1 \Leftrightarrow ACC'A'$ est un parallélogramme $\Leftrightarrow$ (AA') est parallèle à (CC'). Par conséquent, on a :

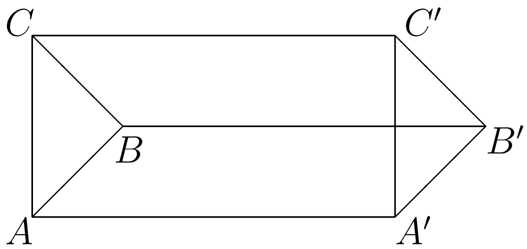
$$\lambda = 1 = \mu \Longleftrightarrow (AA'), (BB') \text{ et } (CC') \text{ sont parallèles.} \tag{2}$$

Et si $\lambda \neq 1$, alors les droites (AA') et (BB') se coupent en un point O, et d'après le théorème de Thalès, on a $\overrightarrow{OA'} = \lambda\overrightarrow{OA}$, d'où $\boxed{\overrightarrow{AO} = (1-\lambda)^{-1}\overrightarrow{AA'}.}$ De même, si $\mu \neq 1$, alors les droites (AA') et (CC') se coupent en un point T, et d'après le théorème de Thalès, on a $\overrightarrow{TA'} = \mu\overrightarrow{TA}$, d'où $\boxed{\overrightarrow{AT} = (1-\mu)^{-1}\overrightarrow{AA'}.}$ On en déduit que $(AA'), (BB')$ et (CC') sont concourantes ssi $\lambda \neq 1$, $\mu \neq 1$, et $T = O$. Or $T = O$ équivaut à $\overrightarrow{AT} = \overrightarrow{AO}$ et donc à $\mu = \lambda$. Donc finalement :

$$(AA'), (BB') \text{ et } (CC') \text{ sont concourantes} \Longleftrightarrow \lambda = \mu \neq 1. \tag{3}$$

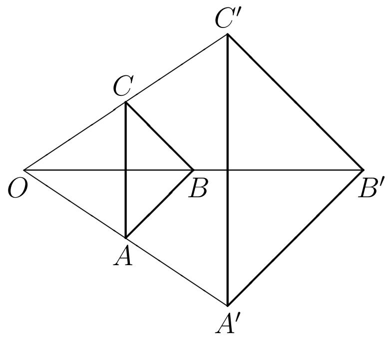
En combinant $(1), (2)$ et (3), on obtient bien que (BC) et $(B'C')$ sont parallèles si et seulement si $(AA'), (BB')$ et (CC') sont concourantes ou parallèles. $\square$

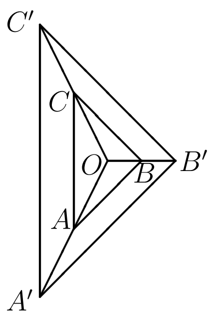
Lorsque $k = \mathbb{R}$, on peut illustrer les deux cas par les figures suivantes. Si (AA'), (BB') et (CC') sont parallèles, il y a essentiellement une figure :



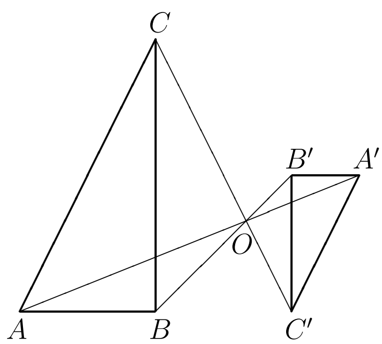
Si (AA'), (BB') et (CC') se coupent en un point O, le triangle $A'B'C'$ est l'image de ABC par une homothétie de centre O et de rapport λ, et la figure est différente selon le signe de λ :

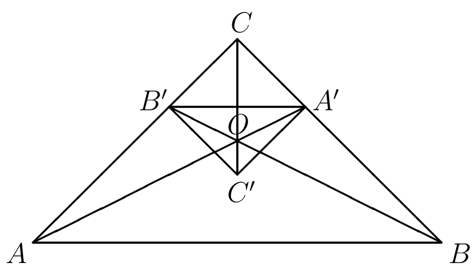
$\lambda > 0$





$\lambda < 0$





Prérequis pour ce cours : connaissances de Licence en algèbre linéaire et bilinéaire (produit scalaire, groupes orthogonaux, formes quadratiques, th. de réduction simultanée). Voir [Du] ou [Po] ci-dessous ou le livre : François Cottet-Emard, Algèbre linéaire et bilinéaire, de boeck, 2005 (disponible à la MIE).

Références pour ce chapitre :

∗ Polycopiés :

[Du] Antoine Ducros, Géométrie affine et euclidienne, Cours de L3 à l'UPMC 2009-2012 (sections 1 et 2), disponible sur la page de l'auteur : www.imj-prg.fr/~antoine.ducros

[It] Ilia Itenberg, Algèbre et géométrie, Cours de M1 à l'UPMC 2013-2014 (chap. 1), disponible sur la page de l'auteur : www.imj-prg.fr/~ilia.itenberg

[Po] Patrick Polo, Algèbre et géométrie, Cours de L2 à l'UPMC 2009-2013 (chap. 7), disponible sur la page de l'auteur : www.imj-prg.fr/~patrick.polo/L2

En complément du polycopié 2013-2014 d'Ilia Itenberg, on pourra aussi consulter les polycopiés faits par les enseignants précédents :

[Be] Daniel Bertrand, Algèbre et géométrie, Cours de M1 à l'UPMC 2009-2013, disponible sur la page de l'auteur : www.imj-prg.fr/~daniel.bertrand

[Ne] Jan Nekovar, Algèbre et géométrie, Cours de M1 à l'UPMC 2005-2009, disponible sur la page de l'auteur : www.imj-prg.fr/~jan.nekovar

---

∗ Et pour les (futurs) agrégatifs, qui ont besoin de références avec un No. d'ISBN, citons les ouvrages suivants :

[Au] Michèle Audin, Géométrie, EDP Sciences, 2006.
[Bo] Pascal Boyer, Algèbre et géométries, Calvage & Mounet, 2015.
[LF] Jacqueline Lelong-Ferrand, Les fondements de la géométrie, P.U.F., 1985.
[Ta] Patrice Tauvel, Géométrie, Dunod, 2005.
[Ti] Claude Tisseron, Géométries affine, projective et euclidienne, Hermann, 1988.