Géométrie

Introduction.

La géométrie est un domaine parmi les plus anciens en mathématiques, qui était déjà très développé chez les grecs (avec à l'origine sans doute des questions pratiques de mesure des distances et des surfaces, comme indique le nom « géométrie »). Comme le discours porte sur des objets tels des points, des droites et des figures, qui contrairement aux nombres ne peuvent pas être complètement captés par une expression numérique ou une formule, le caractère de la géométrie est un peu abstrait dès le départ. (Il est cependant intéressant d'observer que pour les grecs la notion de nombre était *moins* fondamental que les notions géométriques ; par exemple quand Euclide montre que pour une collection de nombres premiers donnée (nous dirions finie) il existe toujours un nouveau nombre premier hors la collection, il formule et prouve l'énoncé en représentant les nombres comme les longueurs géométriques, multiples entiers d'une « unité » de longueur.) Il n'est donc pas étonnant que la géométrie ait fourni le premier exemple d'une théorie mathématique abstraite, basée sur des axiomes: les Éléments d'Euclide.

La géométrie élémentaire donne donc une occasion pour formuler des raisonnements mathématiques qui ne portent pas sur les nombres. Malgré cette abstraction, les énoncés peuvent être assez parlants: on peut « voir » le sens d'un énoncé dans une figure, sans avoir recours à des formules compliquées. C'est ainsi que beaucoup auront vu pas mal de géométrie dans l'enseignement secondaire. Pourtant, la géométrie n'est pas un sujet facile, car en contraste avec par exemple l'algèbre linéaire, il y a relativement peu de méthodes *systématiques* pour résoudre des problèmes géométriques. Le but de ce cours est de présenter néanmoins la géométrie comme une théorie structurée, en évitant de présenter (comme il serait certainement possible) une grande collection de théorèmes remarquables, assez jolis mais isolés. L'accent sera toujours sur la *déduction*: il est moins important de savoir tout ce qui est vrai (ce qui serait d'ailleurs impossible) que de savoir *pourquoi* ce que l'on affirme est vrai (c'est-à-dire découle des axiomes).

Il serait tout à fait dans l'esprit géométrique de commencer avec des notions primitives et abstraites telles que « point » et « droite », et avec des axiomes qui expriment la nature fondamentale de ces notions, et ensuite de développer une théorie de plus en plus élaborée par un raisonnement rigoureux, se basant soit sur les axiomes soit sur des propriétés déjà établies. Néanmoins va on adopter une approche un peu différente pour des raisons pratiques, car le chemin des axiomes purement géométriques aux énoncés intéressants est long. Sachant qu'en admettant un nombre suffisant d'axiomes, on pourra montrer que l'espace géométrique peut être muni d'un *système de coordonnés* avec valeurs dans un corps K (par exemple $K = \mathbf{R}$ pour la géométrie euclidienne), ce qui permet de décrire les points entièrement en termes d'éléments de K (c'est-à-dire de nombres), on peut raccourcir considérablement le chemin en supposant un corps K donné dès le départ. On admettra donc des axiomes qui décrivent l'espace géométrique en relation avec un espace vectoriel sur K. Ainsi les axiomes algébriques d'un corps et d'un espace vectoriel remplaceront certains axiomes géométriques, et on pourra appliquer des résultats connus de l'algèbre linéaire. C'est cette approche, qui donne un caractère plus algébrique à la théorie, qu'on va suivre.

La théorie géométrique traitée dans ce cours peut être divisée en deux parties principales: la géométrie affine et la géométrie euclidienne. La géométrie affine est plus limitée en notions (notamment il n'y a pas de notion de distance) et peut en conséquence être plus générale (on peut développer cette théorie pour n'importe quel corps K). En revanche la géométrie euclidienne, qui doit supposer que $K = \mathbf{R}$, peut exprimer beaucoup plus de choses (comme définition de cercles, notion d'angle) et est plus riche en théorèmes. Comme la géométrie euclidienne est une spécialisation de la géométrie affine, elle pourra s'appuyer sur les résultats de la géométrie affine, et il est donc naturel de commencer avec cette dernière.

L'auteur de ce cours est douloureusement conscient du fait que ce cours écrit ne contient aucune illustration, ce qui est un sérieux défaut pour un cours de géométrie. Cette circonstance s'explique par l'effort très important nécessaire pour créer de bonnes illustrations. J'espère que cela servira comme incitation à suivre plus activement le cours magistral, où il est plus facile de donner des illustrations (sans oublier bien sûr l'assiduité aux travaux dirigés, d'une importance fondamentale pour la réussite).

17 janvier 2012 *Marc van Leeuwen*

§1. Géométrie affine.

1.1. *Espaces affines.*

On suppose un corps K fixé; il interviendra dans les définitions des notions ci-dessous, mais sera en général pas mentionné explicitement (par exemple on parlera le plus souvent de «espace affine» et «espace vectoriel» au lieu de «K-espace affine» et «espace vectoriel sur K»).

Notre développement de la théorie des espaces affines sera fortement inspiré par la théorie des espaces vectoriels sur K; par exemple on va considérer pour les espaces affines des notions correspondant à «sous-espace», «combinaison linéaire» et «(in)dépendance linéaire». Mais malgré l'existence d'un grand nombre de telles analogies, il y a une différence fondamentale entre les deux théories: la théorie des espaces vectoriels est une théorie *algébrique*, avec accent sur les opérations définies sur les éléments telles que l'addition et le produit scalaire, pendant que la théorie des espaces affines est une théorie *géométrique*, avec accent sur les relations entre les objets (points, droites,...). Fondamentale pour cette dernière approche est que le caractère des objets *individuels* de la même classe (disons les points) est toujours le même, c'est-à-dire qu'il n'y a pas de point privilégié ni de droite privilégiée dans un espace affine. En contraste, dans un espace vectoriel il est inévitable d'avoir un vecteur privilégié: le vecteur nul, qui est (l'unique) élément neutre pour l'addition vectorielle.

Pour introduire les espaces affines *à partir* des espaces vectoriels, notre but principal est de faire oublier cet élément privilégié; cela nécessite l'abandon de la notion d'une addition entre les éléments de notre espace. Par contre on peut retenir dans une certaine mesure la soustraction: pour deux points donnés, leur déplacement relatif a un sens, et le rôle privilégié d'un déplacement nul est tout à fait souhaitable, car cela exprime l'égalité des deux points. Ainsi on est amené à distinguer deux type d'objets: les points de l'espace affine, et les déplacements (vecteurs) entre des paires de points; ces derniers vont en effet constituer un espace vectoriel, dit la *direction* de l'espace affine.

1.1.1. Définition. *Soit E un espace vectoriel sur K. Un espace affine $\mathcal{A}$ de direction E est une ensemble muni d'une application $\mathcal{A} \times \mathcal{A} \to E$ notée $(A,B) \mapsto \overrightarrow{AB}$, et d'une loi externe $\mathcal{A} \times E \to \mathcal{A}$ notée $(A,\vec{x}) \mapsto A + \vec{x}$, vérifiant les conditions suivantes:*
(1) Pour tout $A \in \mathcal{A}$ on a $\overrightarrow{AA} = \vec{0}$.
(2) Pour tout $A,B \in \mathcal{A}$ on a $A + \overrightarrow{AB} = B$.
(3) Pour tout $A \in \mathcal{A}$ et tout $\vec{x} \in E$ on a $\overrightarrow{A(A+\vec{x})} = \vec{x}$.
(4) Pour tout $A \in \mathcal{A}$ et tout $\vec{x},\vec{y} \in E$ on a $(A+\vec{x})+\vec{y} = A + (\vec{x}+\vec{y})$.
On note $E = \overrightarrow{\mathcal{A}}$ pour la direction de $\mathcal{A}$, et on définit la dimension de $\mathcal{A}$ comme celle de l'espace vectoriel $\overrightarrow{\mathcal{A}}$.

On écrira $A - \vec{x}$ pour $A + -\vec{x}$. Une conséquence immédiate de ces axiomes est l'équivalence

$$\overrightarrow{AB} = \vec{x} \iff A + \vec{x} = B, \tag{1}$$

car substitution de $\overrightarrow{AB} = \vec{x}$ dans l'axiome (2) donne $A + \vec{x} = B$, et substitution de $A + \vec{x} = B$ dans l'axiome (3) donne $\overrightarrow{AB} = \vec{x}$. Ceci montre un principe fondamental: si on fixe un point $A \in \mathcal{A}$, on obtient par l'équation (1) une correspondance entre les vecteurs $\vec{x} \in E$ de l'espace vectoriel et les points $B \in \mathcal{A}$ de l'espace affine, qui est clairement bijective. Ainsi les deux types d'espace sont très étroitement liés. Seulement, comme la correspondance dépend du choix de A et n'est donc pas unique, le rôle privilégié du vecteur nul ne se traduit pas en un rôle privilégié du point correspondant, qui n'est autre (d'après l'axiome (1)) que le point choisi A.

L'espace physique ambiant ou le plan d'un tableau forment des modèles intuitifs des espaces affines de dimension 3 respectivement 2 (ils ne sont pas munis d'une «origine» particulière, mais deux points déterminent un déplacement unique). Mais on ne peut pas les présenter comme un exemple mathématique d'un espace affine. Pour de tels exemples, il sera nécessaire de se baser sur une structure déjà connue, telle un espace vectoriel (qu'on saura construire comme $\mathbf{R}^n$ par exemple). Cela n'est pas très naturel du point de vue géométrique (on aimerait au contraire construire des espace vectoriels à partir des déplacements dans un espace affine), mais on a pas trop le choix. Donc voici quelques exemples.

- Un espace vectoriel E peut lui-même être vu comme espace affine (il figurera alors aussi bien à la place de $\mathcal{A}$ que celle de $\overrightarrow{\mathcal{A}}$) en définissant $\overrightarrow{AB} = B - A$ (à gauche A et B sont vus comme des points, à droite comme des vecteurs) et $A + \vec{x}$ comme la somme vectorielle. On vérifie facilement les conditions de la définition 1.1.1. Évidemment cet exemple n'est pas le meilleur pour comprendre l'utilité des espaces affines, mais il montre qu'on peut penser de $\overrightarrow{AB}$ comme une différence.
- Si E est un sous-espace d'un autre espace vectoriel F, et a un élément de F qui n'est pas dans E, on peut prendre $\mathcal{A} = a + E \stackrel{\text{déf}}{=} \{\, a + \vec{x} \mid \vec{x} \in E \,\}$. Cet exemple est déjà plus intéressant, car $a + E$ n'est pas un espace vectoriel, il n'est par exemple pas fermé par rapport á l'addition. Cependant, en définissant comme ci-dessus $\overrightarrow{AB} = B - A$, on a toujours $\overrightarrow{AB} \in E$ pour $A, B \in \mathcal{A}$, ainsi que $A + \vec{y} \in \mathcal{A}$ pour $A \in \mathcal{A}$ et $\vec{y} \in E$; on tombe dans les bons ensembles. Clairement, définition 1.1.1 reste vérifiée.
- L'exemple précédent arrive en particulier quand $\mathcal{A}$ est un ensemble non vide de solutions d'un système inhomogène d'équations linéaires, dont les solutions du système homogène associé forme l'espace vectoriel E. Ici $\mathcal{A} = a + E$ exprime le fait que chaque solution du système inhomogène est somme d'une solution particulière a et d'une solution du système homogène associé (qui est dans E).

Dans la pratique, on va toujours supposer *donné* au départ un espace affine $\mathcal{A}$ de direction $\overrightarrow{\mathcal{A}} = E$ (qui pour nous sera toujours de dimension finie). Par conséquent on ne pourra pas supposer que $\mathcal{A}$ soit une partie d'un espace vectoriel (même si c'est le cas dans les exemples ci-dessus), donc les expressions telles que $A + B$ et $3A$ sont *interdites* d'usage (cependant, à la place de $A - B$ on pourra écrire $\overrightarrow{BA}$). Un aspect important de la définition 1.1.1 est de nous forcer à distinguer dans notre langage entre points et vecteurs, et de nous limiter à écrire des expressions qui soit désignent un point, soit désignent un vecteur (on vérifiera cela dans la suite !).

Tirons quelques conséquences immédiates de la définition 1.1.1. Pour tous points $A, B \in \mathcal{A}$ on a

$$A + \vec{0} = A \qquad \text{et} \qquad \overrightarrow{BA} = -\overrightarrow{AB} \tag{2}$$

car $A + \vec{0} = A + \overrightarrow{AA} = A$, et $B + -\overrightarrow{AB} = (A + \overrightarrow{AB}) + -\overrightarrow{AB} = A + (\overrightarrow{AB} + -\overrightarrow{AB}) = A + \vec{0} = A$ d'où la seconde égalité. On a aussi les règles d'annulation suivantes

$$\overrightarrow{AB} = \overrightarrow{AC} \qquad \Rightarrow \qquad B = C, \tag{3}$$
$$\overrightarrow{AC} = \overrightarrow{BC} \qquad \Rightarrow \qquad A = B, \tag{4}$$
$$A + \vec{x} = B + \vec{x} \qquad \Rightarrow \qquad A = B, \tag{5}$$
$$A + \vec{x} = A + \vec{y} \qquad \Rightarrow \qquad \vec{x} = \vec{y}, \tag{6}$$

qui sont déduites facilement; par exemple pour la première on a $B = A + \overrightarrow{AB} = A + \overrightarrow{AC} = C$, et dans la seconde $\overrightarrow{AC} = \overrightarrow{BC}$ implique $-\overrightarrow{AC} = -\overrightarrow{BC}$, c'est-à-dire $\overrightarrow{CA} = \overrightarrow{CB}$, et on peut appliquer la première. L'axiome (4) de la définition 1.1.1 n'est que rarement invoqué explicitement, car il permet de tout simplement omettre les parenthèses dans une expression comme $A + \vec{x} + \vec{y}$. Une autre règle très utile est

$$\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC} \tag{7}$$

appelée «relation de Chasles» (malgré le fait qu'elle était connue bien avant les travaux géométriques de Michel Chasles (1793–1880)). Elle découle du simple calcul $A + \overrightarrow{AB} + \overrightarrow{BC} = B + \overrightarrow{BC} = C$ et l'équivalence (1). On termine notre petite collection de règles utiles par

$$\overrightarrow{(A + \vec{x})(B + \vec{y})} = \overrightarrow{AB} - \vec{x} + \vec{y} \tag{8}$$

qui montre que la notation $\overrightarrow{AB}$ a un caractère «négatif» par rapport à A et «positif» par rapport à B (et qui découle de $(A + \vec{x}) + (\overrightarrow{AB} - \vec{x} + \vec{y}) = B + \vec{y}$ comme on vérifie facilement), ainsi que par l'égalité

$$\overrightarrow{AB} - \overrightarrow{CD} = \overrightarrow{AC} - \overrightarrow{BD} \tag{9}$$

qui découle de $\overrightarrow{AB} + \overrightarrow{BD} = \overrightarrow{AD} = \overrightarrow{AC} + \overrightarrow{CD}$ (deux fois Chasles). Quand les deux membres de (9) sont nuls, c'est-à-dire quand $\overrightarrow{AB} = \overrightarrow{CD}$ et donc $\overrightarrow{AC} = \overrightarrow{BD}$, on dit que $ABDC$ forme un parallélogramme.

1.2. *Sous-espaces affines.*

Dans un espace affine tous les points sont semblables, donc pour pouvoir formuler des énoncés intéressants, il faut considérer des *relations* entre les points. Certaines relations centrales dans la géométrie affine, telles que la condition pour plusieurs points d'être alignés ou non, proviennent de la notion de sous-espace affine. Ce sont des sous-ensembles de points «plats» tels que des droites ou des plans. Ils peuvent être considérés en soi comme des espaces affines, mais plongés dans l'espace affine $\mathcal{A}$ de départ. La définition de sous-espace affine est dérivée de la notion de sous-espace vectoriel.

1.2.1. Définition. *Soit $\mathcal{A}$ un espace affine de direction E. Un sous-ensemble $\mathcal{V}$ de $\mathcal{A}$ est un sous-espace affine s'il peut être écrit sous la forme $\mathcal{V} = A_0 + F \overset{\text{déf}}{=} \{ A_0 + \vec{x} \mid \vec{x} \in F \}$ où A_0 est un point de $\mathcal{A}$ est F est un sous-espace vectoriel de E. On pose par définition $\overrightarrow{\mathcal{V}} = F$ (la direction de $\mathcal{V}$) et $\dim \mathcal{V} = \dim F$.*

Il faut remarquer que la dimension d'un sous-espace affine est un attribut fondamental. En effet, la notion générale de sous-espace affine n'existe même pas en géométrie classique, où on parle toujours de *points* (dimension 0), de *droites* (dimension 1), ou de *plans* (dimension 2), donc en spécifiant toujours la dimension. La notion générale de sous-espace affine permet de ne pas se limiter *a priori* en dimension, et d'éviter de faire certains raisonnements généraux séparément pour les points, droites et plans ; ceci dit, dans les configurations concrètes qu'on étudiera, on connaîtra toujours les dimensions des sous-espaces affines donnés. Il est utile de considérer pour les énoncés généraux ci-dessous ce qu'ils veulent dire dans les cas concrets de points, droites et plans d'un espace affine de dimension 2 ou 3.

On remarque que la définition demande l'existence d'une écriture sous la forme $A_0 + F$ mais pas son unicité. On verra qu'on peut choisir pour A_0 tout élément de $\mathcal{V}$ (l'utilisation d'un indice 0 veut suggérer le choix d'un point de base). Mais dans l'expression pour $\mathcal{V}$, le sous-espace vectoriel F sera toujours le même, car on peut le retrouver à partir de $\mathcal{V}$ comme $F = \{ \overrightarrow{AB} \mid A, B \in \mathcal{V} \}$; cela justifie la notation $\overrightarrow{\mathcal{V}} = F$. Pour montrer cette identité si $\mathcal{V} = A_0 + F$, on établit l'inclusion dans les deux sens : on a $A_0 \in \mathcal{V}$ (car $\vec{0} \in F$) et pour $\vec{x} \in F$ on a donc $\vec{x} = \overrightarrow{A_0(A_0 + \vec{x})} \in \{ \overrightarrow{AB} \mid A, B \in \mathcal{V} \}$, ce qui montre $F \subseteq \{ \overrightarrow{AB} \mid A, B \in \mathcal{V} \}$, et pour $A = A_0 + \vec{x}$ et $B = A_0 + \vec{y}$ on a $\overrightarrow{AB} = \vec{y} - \vec{x} \in F$ d'après (8) (car F est un espace vectoriel), ce qui montre $\{ \overrightarrow{AB} \mid A, B \in \mathcal{V} \} \subseteq F$.

En passant on a vu que toujours $A_0 \in \mathcal{V}$, donc *un sous-espace affine n'est jamais vide.* C'est une remarque importante, car on rencontrera des constructions donnant un ensemble qui est toujours un sous-espace affine, *sauf* qu'il peut aussi être vide; dans ces situations on sera obligé de distinguer explicitement les deux cas. Montrons maintenant que le choix de $A_0 \in \mathcal{V}$ est libre dans l'expression $\mathcal{V} = A_0 + F$.

1.2.2. Proposition. *Si $\mathcal{V}$ est un sous-espace affine de $\mathcal{A}$, alors pour tout $A_1 \in \mathcal{A}$ on a $\mathcal{V} = A_1 + \overrightarrow{\mathcal{V}}$.*

Preuve. Par définition d'un sous-espace affine, on a $\mathcal{V} = A_0 + F$ pour un certain point A_0. Alors pour tout $A_1 \in \mathcal{A}$ il existe $\vec{x} \in F$ avec $A_1 = A_0 + \vec{x}$; fixons un tel A_1 et donc $\vec{x} = \overrightarrow{A_0A_1}$, alors on aussi $A_0 = A_1 - \vec{x}$. Il s'agit de montrer que les ensembles $A_0 + F$ et $A_1 + F$ sont identiques, ce qu'on fera en montrant que chacun est contenu dans l'autre. Tous les points de $A_0 + F$ sont de la forme $A_0 + \vec{y}$ avec $\vec{y} \in F$. En récrivant $A_0 + \vec{y} = (A_1 - \vec{x}) + \vec{y} = A_1 + (\vec{y} - \vec{x})$ on voit qu'un tel point appartient à $A_1 + F$, car la différence $\vec{x} - \vec{y}$ de deux vecteurs du sous-espace vectoriel F de $\overrightarrow{\mathcal{A}}$ reste dans F. Ceci établit l'inclusion $A_0 + F \subseteq A_1 + F$, et la démonstration de l'inclusion opposée est similaire: tout point de $A_1 + F$ s'écrit $A_1 + \vec{y} = (A_0 + \vec{x}) + \vec{y} = A_0 + (\vec{x} + \vec{y}) \in A_0 + F$ car $\vec{x} + \vec{y} \in F$. □

Si $\dim(\mathcal{A}) = n$, la dimension des sous-espaces affines peut varier de 0 (les points) à n. Il n'y a qu'un seul sous-espace affine de dimension n (l'espace $\mathcal{A}$ tout entier), donc ce cas est rarement intéressant (sauf comme résultat possible d'une construction qui produit un sous-espace de dimension variable). Pour cette raison la notion de «plan» n'est utile que si l'espace entier $\mathcal{A}$ est de dimension au moins 3, et il n'y a pas de nom spécifique pour un sous-espace de dimension 3 (ni pour les dimensions supérieures), bien qu'ils existent si $\dim \mathcal{A}$ est suffisamment grande. Le nom *hyperplan* est réservé aux sous-espaces de A de la plus grande dimension intéressante dans $\mathcal{A}$, qui est $\dim \mathcal{A} - 1$ (si $\dim \mathcal{A} = 2$ ce terme est donc synonyme «droite», si $\dim \mathcal{A} = 3$ il l'est avec «plan»).

1.2.3. Proposition. *Soit $\mathcal{A}$ un espace affine, et $\mathcal{V}, \mathcal{W}$ deux sous-espaces affines. Alors précisément une des deux possibilités suivantes se produit:*
(1) $\mathcal{V} \cap \mathcal{W} = \emptyset$
(2) $\mathcal{V} \cap \mathcal{W}$ est un sous-espace affine de A, de direction $\overrightarrow{\mathcal{V}} \cap \overrightarrow{\mathcal{W}}$.
Dans le second cas on dit que $\mathcal{V}$ et $\mathcal{W}$ se coupent (en $\mathcal{V} \cap \mathcal{W}$).

Preuve. Par définition un sous-espace affine n'est pas vide, donc les deux possibilités ne peuvent pas se produire simultanément. Montrons que si la première ne se produit pas, c'est-à-dire s'il existe un point $A_0 \in \mathcal{V} \cap \mathcal{W}$, alors la seconde possibilité se produit. Cela revient à montrer que $\mathcal{V} \cap \mathcal{W} = A_0 + (\overrightarrow{\mathcal{V}} \cap \overrightarrow{\mathcal{W}})$, car $\overrightarrow{\mathcal{V}} \cap \overrightarrow{\mathcal{W}}$ est, comme toute intersection de sous-espaces vectoriels de E, un sous-espace vectoriel de E. Comme $A_0 \in \mathcal{V}$ et $A_0 \in \mathcal{W}$ on peut écrire $\mathcal{V} = A_0 + \overrightarrow{\mathcal{V}}$ et $\mathcal{W} = A_0 + \overrightarrow{\mathcal{W}}$, et les points $P \in \mathcal{V} \cap \mathcal{W}$ sont ceux qui s'écrivent à la fois comme $P = A_0 + \vec{x}$ avec $\vec{x} \in \overrightarrow{\mathcal{V}}$ et comme $P = A_0 + \vec{y}$ avec $\vec{y} \in \overrightarrow{\mathcal{W}}$. Mais pour de telles écritures on a forcément $\vec{x} = \vec{y}$ (règle (6)), et donc $\vec{x} \in \overrightarrow{\mathcal{V}} \cap \overrightarrow{\mathcal{W}}$, d'où $P \in A_0 + (\overrightarrow{\mathcal{V}} \cap \overrightarrow{\mathcal{W}})$. Le fait que réciproquement tout point de $A_0 + (\overrightarrow{\mathcal{V}} \cap \overrightarrow{\mathcal{W}})$ appartient à $(A_0 + \overrightarrow{\mathcal{V}}) \cap (A_0 + \overrightarrow{\mathcal{W}}) = \mathcal{V} \cap \mathcal{W}$ est évident, car il appartient à la fois à $A_0 + \overrightarrow{\mathcal{V}}$ et à $A_0 + \overrightarrow{\mathcal{W}}$. $\square$

On a donné cette démonstration en tout détail pour montrer qu'elle repose uniquement sur les définitions et des faits déjà établis. Mais une fois que le principe est clair, il serait fastidieux de continuer à faire ainsi dans toutes les démonstrations. On se permettra donc dans la suite des raccourcis du style "pour montrer $\mathcal{V} \cap \mathcal{W} = A_0 + (\overrightarrow{\mathcal{V}} \cap \overrightarrow{\mathcal{W}})$, il suffit d'écrire $\mathcal{V} \cap \mathcal{W} = (A_0 + \overrightarrow{\mathcal{V}}) \cap (A_0 + \overrightarrow{\mathcal{W}}) = A_0 + (\overrightarrow{\mathcal{V}} \cap \overrightarrow{\mathcal{W}})$", car les arguments nécessaires pour justifier ces deux égalités sont faciles à fournir une fois qu'on se rend compte du sens des expressions concernées (l'application la règle (6) doit être presque un automatisme).

La distinction entre les deux cas dans la proposition sera d'un intérêt particulier: beaucoup d'énoncés géométriques disent que dans certaines configurations deux sous-espaces affines se coupent toujours (ou au contraire jamais). Par exemple dire que dans le plans trois droites sont concourantes (passent par un même point) veut dire que le deux premières se coupent, et que la troisième coupe leur intersection. Ceci est en contraste avec la situation pour les sous-espaces vectoriels d'un espace vectoriel, qui se coupent toujours car chacun doit au moins contenir le vecteur nul.

Lequel des deux cas de la proposition se produira le plus souvent dépend de la dimension de $\mathcal{A}$ et celles de $\mathcal{V}$ et $\mathcal{W}$. Par exemple dans le plan ($\dim \mathcal{A} = 2$) il est rare qu'un point ($\dim \mathcal{V} = 0$) et une droite ($\dim \mathcal{W} = 1$) se coupent, car pour cela il faut que le point se trouve sur la droite. Mais dans le cas de deux droites ($\dim \mathcal{V} = \dim \mathcal{W} = 1$), la situation la plus commune est qu'elles se coupent, et seulement dans le cas de deux droites distinctes et parallèles on aura une intersection vide. En général on verra que le cas d'une intersection vide devient exceptionnel dès que la somme des dimensions des sous-espaces atteint $\dim \mathcal{A}$ ou plus; ceci dit, une intersection vide reste toujours possible (sauf si l'un de $\mathcal{V}, \mathcal{W}$ est égal à l'espace $\mathcal{A}$ tout entier) à cause de la possibilité de sous-espace parallèles, quelle notion est définie ainsi.

1.2.4. Définition. *Deux sous-espaces $\mathcal{V}, \mathcal{W}$ de $\mathcal{A}$ sont dits parallèles s'ils ont la même direction: $\overrightarrow{\mathcal{V}} = \overrightarrow{\mathcal{W}}$.*

Parfois on notera la relation d'être parallèles par $\mathcal{V} \parallel \mathcal{W}$. Remarquons que selon cette définition, le fait que $\mathcal{V}$ et $\mathcal{W}$ sont parallèles implique $\dim \mathcal{V} = \dim \mathcal{W}$, et n'exclut pas la possibilité $\mathcal{V} = \mathcal{W}$, c'est-à-dire tout sous-espace affine est parallèle à lui-même. En fait c'est le seul cas où $\mathcal{V}, \mathcal{W}$ parallèles se coupent:

1.2.5. Proposition. *Soit $\mathcal{V}, \mathcal{W}$ deux sous-espaces parallèles de $\mathcal{A}$. Alors on a précisément un des deux cas suivants: (1) $\mathcal{V} = \mathcal{W}$, ou (2) $\mathcal{V} \cap \mathcal{W} = \emptyset$.*

Preuve. Si on avait les deux possibilités au même temps on aurait $\mathcal{V} = \emptyset$, mais cela n'est pas un sous-espace. Donc au plus une des possibilités se produit; montrons que si la seconde ne se produit pas, alors la première se produit. Le fait que $\mathcal{V} \cap \mathcal{W} \neq \emptyset$ permet de choisir un point $A_0 \in \mathcal{V} \cap \mathcal{W}$, et on aura $\mathcal{V} = A_0 + \overrightarrow{\mathcal{V}} = A_0 + \overrightarrow{\mathcal{W}} = \mathcal{W}$, car $\overrightarrow{\mathcal{V}} = \overrightarrow{\mathcal{W}}$ d'après l'hypothèse que $\mathcal{V}$ et $\mathcal{W}$ sont parallèles. $\square$

On rappelle que deux sous-espaces vectoriels V, W d'un espace vectoriel E sont «supplémentaires» dans E si $E = V \oplus W$, c'est-à-dire si tout vecteur $\vec{x} \in E$ peut être écrit de façon unique comme $\vec{x} = \vec{v} + \vec{w}$ avec $v \in V$ et $w \in W$. Une condition nécessaire pour cela est $\dim V + \dim W = \dim E$. Cette condition n'est certainement pas suffisante. Cependant, pour V, W avec $\dim V + \dim W = \dim E = n$, il est plutôt exceptionnel que les sous-espaces ne soient pas supplémentaires: supposons avoir choisi (au hasard) des bases des deux sous-espaces, alors ces sous-espaces seront supplémentaires si la famille à n vecteurs formée par ces deux bases est libre (et donc une base de E), ce qui sera le cas dès que le déterminant de cette famille est différent de 0. On transféra maintenant cette notion dans le monde des espaces affines.

1.2.6. Définition. *Deux sous-espaces affines $\mathcal{V}, \mathcal{W}$ de $\mathcal{A}$ sont dits supplémentaires dans $\mathcal{A}$ si leurs directions $\overrightarrow{\mathcal{V}}, \overrightarrow{\mathcal{W}}$ sont supplémentaires dans $\overrightarrow{\mathcal{A}}$.*

Par exemple, comme deux droites vectoriels dans un espace de dimension 2 sont supplémentaires sauf si elles sont identiques, deux droites affines dans le plan sont supplémentaires sauf si elles sont parallèles.

On va montrer que deux sous-espaces affines supplémentaires dans $\mathcal{A}$ se coupent toujours dans un point unique. En fait on va déduire cela de la caractérisation suivante de la condition de se couper.

1.2.7. Proposition. *Soient $\mathcal{V}, \mathcal{W}$ des sous-espaces affines de $\mathcal{A}$, et $A \in \mathcal{V}$ et $B \in \mathcal{W}$ des points quelconques. Alors $\mathcal{V}$ et $\mathcal{W}$ se coupent si et seulement si $\overrightarrow{AB} \in \overrightarrow{\mathcal{V}} + \overrightarrow{\mathcal{W}}$ $(= \{\, \vec{x} + \vec{y} \mid \vec{x} \in \overrightarrow{\mathcal{V}}, \vec{y} \in \overrightarrow{\mathcal{W}} \,\})$.*

Preuve. Considérons pour un point $P \in \mathcal{A}$ la condition $P \in \mathcal{V} \cap \mathcal{W}$. Comme $\mathcal{V} = A + \overrightarrow{\mathcal{V}}$ et $\mathcal{W} = B + \overrightarrow{\mathcal{W}}$ elle est vérifiée si et seulement si à la fois $\overrightarrow{AP} \in \overrightarrow{\mathcal{V}}$ et $\overrightarrow{BP} \in \overrightarrow{\mathcal{W}}$. Si un P vérifiant la condition existe, c'est-à-dire si $\mathcal{V}$ et $\mathcal{W}$ se coupent, on aura $\overrightarrow{AB} = \overrightarrow{AP} - \overrightarrow{BP} \in \overrightarrow{\mathcal{V}} + \overrightarrow{\mathcal{W}}$. Réciproquement si $\overrightarrow{AB} \in \overrightarrow{\mathcal{V}} + \overrightarrow{\mathcal{W}}$, cela veut dire qu'on peut écrire $\overrightarrow{AB} = \vec{x} + \vec{y}$ avec $\vec{x} \in \overrightarrow{\mathcal{V}}$ et $\vec{y} \in \overrightarrow{\mathcal{W}}$, et le point $P = A + \vec{x} = A + \overrightarrow{AB} - \vec{y} = B - \vec{y}$ appartient à $(A + \overrightarrow{\mathcal{V}}) \cap (B + \overrightarrow{\mathcal{W}}) = \mathcal{V} \cap \mathcal{W}$, donc $\mathcal{V}$ et $\mathcal{W}$ se coupent. □

1.2.8. Corollaire. *Deux sous-espaces affines $\mathcal{V}, \mathcal{W}$ de $\mathcal{A}$ dont les directions engendrent $\overrightarrow{\mathcal{A}}$ tout entier, en formule $\overrightarrow{\mathcal{V}} + \overrightarrow{\mathcal{W}} = \overrightarrow{\mathcal{A}}$, se coupent toujours. En particulier si $\mathcal{V}$ et $\mathcal{W}$ sont supplémentaires ils se coupent; dans ce cas leur intersection est réduit à un seul point.*

Preuve. On peut toujours choisir $A \in \mathcal{V}$ et $B \in \mathcal{W}$, et d'après la proposition précédente, $\mathcal{V}$ et $\mathcal{W}$ se coupent si $\overrightarrow{AB} \in \overrightarrow{\mathcal{V}} + \overrightarrow{\mathcal{W}}$; c'est toujours le cas, car $\overrightarrow{\mathcal{V}} + \overrightarrow{\mathcal{W}} = \overrightarrow{\mathcal{A}}$. Si $\mathcal{V}$ et $\mathcal{W}$ sont supplémentaires on a $\overrightarrow{\mathcal{V}} + \overrightarrow{\mathcal{W}} = \overrightarrow{\mathcal{A}}$, et en plus cette somme est directe, d'où $\overrightarrow{\mathcal{V}} \cap \overrightarrow{\mathcal{W}} = \{\vec{0}\}$. D'après la proposition 1.2.3 (2), $\overrightarrow{\mathcal{V}} \cap \overrightarrow{\mathcal{W}}$ est la direction de $\mathcal{V} \cap \mathcal{W}$, qui est donc de dimension 0, c'est-à-dire réduit à un seul point. □

Ainsi on voit que pour deux sous espaces $\mathcal{V}, \mathcal{W}$ affines de $\mathcal{A}$, le cas où $\mathcal{V} \cap \mathcal{W} = \emptyset$ devient exceptionnel dès que $\dim \mathcal{V} + \dim \mathcal{W} \geq \dim \mathcal{A}$: on a vu que si $\dim \mathcal{V} + \dim \mathcal{W} = \dim \mathcal{A}$, alors dans le plupart des cas $\mathcal{V}$ et $\mathcal{W}$ seront supplémentaires et auront donc un unique point d'intersection, et si $\dim \mathcal{V} + \dim \mathcal{W} > \dim \mathcal{A}$ cela ne fera qu'augmenter les chances que $\mathcal{V}$ et $\mathcal{W}$ se coupent. Pour les sous-espaces vectoriels $\overrightarrow{\mathcal{V}}$ et $\overrightarrow{\mathcal{W}}$ on connaît la relation $\dim \overrightarrow{\mathcal{V}} + \dim \overrightarrow{\mathcal{W}} = \dim(\overrightarrow{\mathcal{V}} \cap \overrightarrow{\mathcal{W}}) + \dim(\overrightarrow{\mathcal{V}} + \overrightarrow{\mathcal{W}})$ qui est valable quelle que soit la position de $\overrightarrow{\mathcal{V}}$ par rapport à $\overrightarrow{\mathcal{W}}$ (elle peut être déduite du théorème du rang). Comme $\dim(\overrightarrow{\mathcal{V}} + \overrightarrow{\mathcal{W}})$ ne peut pas excéder $\dim \mathcal{A}$ on en déduit que $\dim(\overrightarrow{\mathcal{V}} \cap \overrightarrow{\mathcal{W}}) \geq \dim \mathcal{V} + \dim \mathcal{W} - \dim \mathcal{A}$. Or si $\mathcal{V}$ et $\mathcal{W}$ se coupent, leur intersection sera de direction $\overrightarrow{\mathcal{V}} \cap \overrightarrow{\mathcal{W}}$, et dans le cas où $\dim \mathcal{V} + \dim \mathcal{W} \geq \dim \mathcal{A}$, cette dernière inégalité affirme une dimension minimale pour l'intersection:

$$\dim(\mathcal{V} \cap \mathcal{W}) \geq \dim \mathcal{V} + \dim \mathcal{W} - \dim \mathcal{A} \qquad \text{si } \mathcal{V} \cap \mathcal{W} \neq \emptyset. \tag{10}$$

Par exemple dans l'espace ($\dim \mathcal{A} = 3$), deux plans ($\dim \mathcal{V} = \dim \mathcal{W} = 2$) qui se coupent auront pour intersection au moins une droite ($\dim(\mathcal{V} \cap \mathcal{W}) \geq 2 + 2 - 3 = 1$). Malgré cela, il ne faut pas oublier qu'une intersection vide reste toujours possible (dans l'exemple donné, c'est le cas de deux plans parallèles).

Outre la borne inférieure donnée par (10) pour la dimension d'une intersection de deux sous-espaces affines qui se coupent, et la borne supérieure évidente $\min(\dim \mathcal{V}, \dim \mathcal{W})$ (car $\overrightarrow{\mathcal{V} \cap \mathcal{W}}$ est contenu dans $\overrightarrow{\mathcal{V}}$ et aussi dans $\overrightarrow{\mathcal{W}}$), on ne peut pas dire beaucoup en général sur de telles intersections. Cependant, dans le cas où $\mathcal{W}$ est un hyperplan, ces résultats peuvent être résumés dans une forme qui s'avérera utile. Pour le faire, on aura besoin d'une notion similaire à celle d'être parallèle, mais qui n'est pas limitée au cas de deux sous-espaces affines de la même dimension.

1.2.9. Définition. *Deux sous-espaces affines $\mathcal{V}, \mathcal{W}$ de $\mathcal{A}$ avec $\dim \mathcal{V} \leq \dim \mathcal{W}$ sont faiblement parallèles si $\vec{\mathcal{V}} \subseteq \vec{\mathcal{W}}$. Si le contexte fixe $\dim \mathcal{V}$ et $\dim \mathcal{W}$, on omet parfois "faiblement", et écrit même $\mathcal{V} \parallel \mathcal{W}$.*

On remarque que deux sous-espaces affines sont toujours faiblement parallèles si l'un est réduit à un point ($\vec{\mathcal{V}} = \{\vec{0}\}$) ou égal à l'espace entier ($\mathcal{W} = \mathcal{A}$), et que dans le cas $\dim \mathcal{V} = \dim \mathcal{W}$ ils sont faiblement parallèles si et seulement si ils sont parallèles. La notion est donc surtout utile dans les cas restants.

L'intersection de deux sous-espaces affines faiblement parallèles est soit vide, soit égale au plus petit des deux. Car avec $\vec{\mathcal{V}} \subseteq \vec{\mathcal{W}}$ et $A_0 \in \mathcal{V}$ on aura $A_0 + \vec{\mathcal{W}} \parallel \mathcal{W}$, et la proposition 1.2.5 s'applique; si $(A_0 + \vec{\mathcal{W}}) \cap \mathcal{W} = \emptyset$ alors aussi $\mathcal{V} \cap \mathcal{W} = \emptyset$, et si au contraire $A_0 + \vec{\mathcal{W}} = \mathcal{W}$, alors $\mathcal{V} \cap \mathcal{W} = A_0 + \vec{\mathcal{V}} = \mathcal{V}$.

1.2.10. Proposition. *Soit $\mathcal{V}, \mathcal{W}$ deux sous-espaces affines de $\mathcal{A}$ non faiblement parallèles, dont $\mathcal{W}$ est un hyperplan (avec donc $\dim \mathcal{W} = \dim \mathcal{A} - 1$). Alors $\mathcal{V}$ et $\mathcal{W}$ se coupent, et $\dim(\mathcal{V} \cap \mathcal{W}) = \dim \mathcal{V} - 1$.*

Preuve. D'après la remarque ci-dessus, l'hypothèse implique $0 < \dim \mathcal{V} < \dim \mathcal{A}$. On a $\vec{\mathcal{V}} \not\subseteq \vec{\mathcal{W}}$ donc $\vec{\mathcal{V}} + \vec{\mathcal{W}}$ est un sous-espace vectoriel de $\vec{\mathcal{A}}$ qui contient strictement $\vec{\mathcal{W}}$, et comme ce dernier a déjà dimension $\dim \mathcal{A} - 1$, la seule possibilité qui reste est $\vec{\mathcal{V}} + \vec{\mathcal{W}} = \vec{\mathcal{A}}$. D'après le corollaire ci-dessus, $\mathcal{V}$ et $\mathcal{W}$ se coupent donc. Or, comme $\dim \mathcal{A} - \dim \mathcal{W} = 1$, l'inégalité qu'on a établie ci-dessus devient $\dim(\mathcal{V} \cap \mathcal{W}) \geq \dim \mathcal{V} - 1$. D'autre part on a toujours $\dim(\vec{\mathcal{V}} \cap \vec{\mathcal{W}}) \leq \dim \mathcal{V}$, et si on avait $\dim(\vec{\mathcal{V}} \cap \vec{\mathcal{W}}) = \dim \mathcal{V}$ cela voudrait dire $\vec{\mathcal{V}} \cap \vec{\mathcal{W}} = \vec{\mathcal{V}}$ et donc $\vec{\mathcal{V}} \subseteq \vec{\mathcal{W}}$, contrairement à l'hypothèse; par conséquent il ne reste que $\dim(\mathcal{V} \cap \mathcal{W}) = \dim \mathcal{V} - 1$ (l'intersection est donc un hyperplan de $\mathcal{V}$). □

1.3. *Barycentres.*

En algèbre linéaire, on peut former des combinaisons linéaires de vecteurs, qui est un outil fondamental. Cette notion ne se traduit pas directement en géométrie affine, car on ne peut pas additionner des points. Cependant on peut donner un sens à certaines combinaisons linéaires, au moins dans l'exemple où $\mathcal{A}$ est un translaté d'un sous-espace affine qui ne passe pas par $\vec{0}$: par exemple la différence $B - A$ se laisse (par définition) interpréter comme le vecteur $\overrightarrow{AB}$. En fait, dans cet exemple toute combinaison linéaire de points dans $\mathcal{A}$ *dont la somme des coefficients est nulle* donne un vecteur de $E = \vec{\mathcal{A}}$, et toute combinaison linéaire de points dans $\mathcal{A}$ dont la somme des coefficients est 1 donne un nouveau point de $\mathcal{A}$. Par exemple $A + B - C$ pour $A, B, C \in \mathcal{A}$ est une combinaison linéaire de points dont la somme des coefficients est 1, et le résultat est égal à $A + \overrightarrow{CB} \in \mathcal{A}$, et aussi à $B + \overrightarrow{CA}$. Ces deux expressions désignent toujours le même point, indépendamment de la nature de l'espace affine $\mathcal{A}$, car $\overrightarrow{(A + \overrightarrow{CB})(B + \overrightarrow{CA})} = \overrightarrow{AB} - \overrightarrow{CB} + \overrightarrow{CA} = \vec{0}$. La notion de barycentre montrera qu'en général on peut donner un sens bien défini à des combinaisons linéaires de points, dans les conditions mentionnées. Pour une combinaison linéaire de points dont la somme des coefficients vaut ni 0, ni 1, on pourra multiplier tous les coefficients par les même facteur pour ramener leur somme à 1; ainsi chaque combinaison linéaire correspond soit à un vecteur, soit a un point.

La notion de barycentre donne un sens précis à tout ceci, mais sans utiliser la notion de combinaison linéaire de points. Cela évitera toute confusion entre vecteurs et points, entre autre celle liée au fait qu'une expression de la même forme (combinaison linéaire) désignerait un vecteur ou un point, dépendant uniquement d'une condition numérique (somme des coefficients).

Au lieu d'une combinaison linéaire de points, on parlera d'une *collection de points pondérés* (c.p.p.); c'est une famille finie $X = \big((A_1, \mu_1), \ldots, (A_n, \mu_n)\big)$ où pour tout i, $A_i \in \mathcal{A}$ est un point et $\mu_i \in K$ est un scalaire, appelé le poids du point $\mathcal{A}_i$. La *masse* de la collection, notée $\mu(X)$, est définie comme la somme $\sum_{i=1}^{n} \mu_i$ de ces poids. Pour faciliter la définition de barycentres, on associera d'abord un vecteur à la donnée d'un point de base P et une telle collection X, á savoir $\sum_{i=1}^{n} \mu_i \overrightarrow{PA_i}$, qu'on notera $v(P, X)$ (il est tentant d'utiliser la notation $\overrightarrow{PX}$, mais cela risque la confusion avec $\overrightarrow{PQ}$ où Q est un point).

1.3.1. Lemme. *Soit X une c.p.p., et $P, Q \in \mathcal{A}$ deux points, alors $v(P, X) - v(Q, X) = \mu(X)\overrightarrow{PQ}$.*

Preuve. $v(P, X) - v(Q, X) = \sum_{i=1}^{n} \mu_i \overrightarrow{PA_i} - \sum_{i=1}^{n} \mu_i \overrightarrow{QA_i} = \sum_{i=1}^{n} \mu_i (\overrightarrow{PA_i} + \overrightarrow{A_iQ}) = \mu(X)\overrightarrow{PQ}$. □

1.3.2. Proposition/définition. *Soit X une collection de points pondérés.*

(1) *Si $\mu(X) = 0$ alors le vecteur $v(P, X)$ est indépendant du point $P \in \mathcal{A}$, et sera noté $v(X)$.*

(2) *Si $\mu(X) \neq 0$, alors il existe un point $B \in \mathcal{A}$ unique, appelé le barycentre de X, tel que pour tout $P \in \mathcal{A}$ on ait $v(P, X) = \mu(X)\overrightarrow{PB}$. Le barycentre est noté $B = \operatorname{bar}(X)$, peut être calculé comme $B = A + \frac{1}{\mu(X)}v(A, X)$ pour $A \in \mathcal{A}$ quelconque, est est caractérisé par la propriété $v(B, X) = \vec{0}$.*

Preuve. Pour $\mu(X) = 0$, le lemme dit que $v(P, X) = v(Q, X)$ quels que soient $P, Q \in \mathcal{A}$, d'où le premier point. Pour le second point vérifions d'abord qu'un tel point B sera unique. La propriété générale $v(P, X) = \mu(X)\overrightarrow{PB}$ entraîne (avec $P = B$) que $v(B, X) = \vec{0}$, et pour tout autre point $A \neq B$ que $v(A, X) = \mu(X)\overrightarrow{AB} \neq \vec{0}$, d'où A ne partage pas ladite propriété. Pour terminer il suffira donc de montrer que le point B définie par $B = A + \frac{1}{\mu(X)}v(A, X)$ (avec $A \in \mathcal{A}$ quelconque) possède la propriété, établissant ainsi l'existence. D'après la définition de B on a $\mu(X)\overrightarrow{AB} = v(A, X)$, donc pour $P \in \mathcal{A}$ on calcule avec le lemme $v(P, X) = v(A, X) + \mu(X)\overrightarrow{PA} = \mu(X)(\overrightarrow{AB} + \overrightarrow{PA}) = \mu(X)\overrightarrow{PB}$, comme voulu. □

On note que $\operatorname{bar}(X)$ n'est défini que si $\mu(X) \neq 0$. Dans la notation pour les collections de points pondérés, on se permettra d'écrire A_i au lieu de (A_i, μ_i) quand $\mu_i = 1$; ainsi $\operatorname{bar}(A, (B, 2))$ est une abréviation pour $\operatorname{bar}((A, 1), (B, 2))$, et $\operatorname{bar}(A, B)$ en est une pour $\operatorname{bar}((A, 1), (B, 1))$. Dans les cas, comme ce dernier, où tous les poids μ_i sont 1, on parle de l'*isobarycentre* des points $A_1, \ldots, A_n$. L'isobarycentre de deux points s'appelle aussi simplement le milieu des deux points. Dans la formule $B = A + \frac{1}{\mu(X)}v(A, X)$ pour le barycentre on peut prendre A égal à l'un des points de la collection, et alors l'un des vecteurs dans l'expression $v(A, X)$ sera nul. Par exemple pour le milieu de A et B on obtient les expressions $\operatorname{bar}(A, B) = A + \frac{1}{2}(\overrightarrow{AA} + \overrightarrow{AB}) = A + \frac{1}{2}\overrightarrow{AB}$ et $\operatorname{bar}(A, B) = B + \frac{1}{2}\overrightarrow{BA}$, ce qui correspond bien à ce qu'on pourrait attendre pour le milieu de deux points.

On peut distinguer un certain nombre de modifications qu'on peut apporter à une collection de points pondérés X sans changer $\operatorname{bar}(X)$; elles sont déduites de propriétés correspondantes pour la combinaison linéaire dans la définition de $v(P, X)$. D'abord une permutation des couples (A_i, μ_i) n'a pas d'effet sur le barycentre, donc par exemple $\operatorname{bar}(A, (B, 2)) = \operatorname{bar}((B, 2), A)$. Ensuite la présence d'un couple $(A, 0)$ dans la collection est également sans effet, donc on pourra librement supprimer (ou rajouter) des points avec poids nul. On pourra aussi remplacer un couple (A_i, μ_i) par deux couples $(A_i, \mu'), (A_i, \mu'')$ avec $\mu_i = \mu' + \mu''$. Finalement on pourra multiplier simultanément tous les poids μ_i par un même facteur non nul $\lambda \in K$, car cela a pour effet de multiplier $v(A, X)$ par λ, mais aussi la masse $\mu(X)$, et du coup la formule $A + \frac{1}{\mu(X)}v(A, X)$ pour le barycentre donne un point qui n'a pas changé.

Toutes ces opérations ne suffisent pas encore pour pouvoir manipuler les barycentres avec la même flexibilité qu'on connaît pour les combinaisons linéaires en algèbre linéaire. Notamment il manque une manière de regrouper un sous-ensemble de points pondérés dans la collection et de simplifier l'expression par un calcul sur ces points, comme la loi associative permet de calculer une partie d'une combinaison linéaire sans toucher aux termes restants. La règle qui permet ce type de regroupement est connue sous le nom «associativité de barycentres»; de manière informelle elle dit qu'on peut remplacer un sous ensemble de points par leur barycentre, muni de leur masse comme poids. Elle permettra par exemple de calculer $\operatorname{bar}(P, (Q, 2), R)$ comme $\operatorname{bar}((B, 3), R)$ où $B = \operatorname{bar}(P, (Q, 2))$. La règle s'avère plus difficile à formuler de façon formelle (et on ne le fera donc pas) que d'appliquer ou de prouver. Après permutation, on peut supposer que l'ensemble à regrouper est une suite de points pondérés consécutifs au début de la collection.

1.3.3. Proposition. *Le barycentre d'un c.p.p. $X = \big((A_1, \mu_1), \ldots, (A_n, \mu_n)\big)$ ne change pas si l'on remplace une sous-collection $Y = ((A_1, \mu_1), \ldots, (A_k, \mu_k))$, avec $k < n$ et $\mu(Y) \neq 0$, par $(\operatorname{bar}(Y), \mu(Y))$.*

Preuve. La proposition découle du simple calcul suivant, valable pour tout point $P \in \mathcal{A}$:

$$v(P, X) = \sum_{i=1}^{n} \mu_i \overrightarrow{PA_i} = \sum_{i=1}^{k} \mu_i \overrightarrow{PA_i} + \sum_{i=k+1}^{n} \mu_i \overrightarrow{PA_i} = \mu(Y)\overrightarrow{P\operatorname{bar}(Y)} + \sum_{i=k+1}^{n} \mu_i \overrightarrow{PA_i} = v(P, X'),$$

où $X' = ((\operatorname{bar}(Y), \mu(Y)), (A_{k+1}, \mu_{k+1}), \ldots, (A_n, \mu_n))$ est la collection X après le remplacement indiqué. □

Si A et B son deux points distincts, et μ, μ' deux poids de somme non nul, alors la formule $\operatorname{bar}((A, \mu), (B, \mu')) = A + \frac{\mu'}{\mu + \mu'}\overrightarrow{AB}$ montre que leur barycentre se trouve sur la droite affine $A + \langle\overrightarrow{AB}\rangle$, où $\langle\overrightarrow{AB}\rangle$ désigne la droite vectorielle engendré par le vecteur $\overrightarrow{AB}$.

1.3.4. Définition. *Soit $A, B \in \mathcal{A}$ deux points distincts. La droite passant par A et B, notée $\mathcal{D}_{A,B}$, est la droite affine $A + \langle \overrightarrow{AB} \rangle = B + \langle \overrightarrow{AB} \rangle$.*

Il est évident que $A \in \mathcal{D}_{A,B}$ et $B \in \mathcal{D}_{A,B}$, et $\mathcal{D}_{A,B}$ est la seule droite avec cette propriété, ce qui justifie de l'appeler *la* droite passant par A et B. Car si $\mathcal{D}$ est une droite quelconque passant par A et B, l'intersection $\mathcal{D}_{A,B} \cap \mathcal{D}$ contiendra au moins ces deux points distincts; n'étant pas vide, elle est un sous-espace affine d'après la proposition 1.2.3, et sa dimension n'est pas 0 (elle consisterait alors d'un seul point), donc elle est égale à $\mathcal{D}_{A,B}$, et $\mathcal{D} = \mathcal{D}_{A,B}$. Trois points qui se trouvent sur une même droite affine sont dits *alignés*, une notion qui inclut aussi le cas où au moins deux de ces points sont identiques (si les trois sont identiques, il n'y a pas de droite unique les contenant, mais les points sont appelés alignés tout de même). Pour une famille de plus de trois points on dira également qu'ils sont alignés si chaque triplet parmi eux sont alignés; si A, B sont deux points distincts de la famille, $\mathcal{D}_{A,B}$ passe par tous les points de la famille et est la seule droite à le faire.

On appelle deux vecteurs $\vec{x}, \vec{y}$ proportionnels s'il existe des scalaires $\alpha, \beta \in K$, pas tous les deux nuls, tels que $\alpha\vec{x} = \beta\vec{y}$, une relation qu'on notera $\vec{x} : \vec{y} = \beta : \alpha$ (attention à l'ordre!). Une famille de vecteurs contenus dans une droite vectorielle $\langle \vec{v} \rangle$ sont tous proportionnels (c'est-à-dire proportionnels deux à deux), car si $\vec{x} = \lambda\vec{v}$ et $\vec{y} = \mu\vec{v}$ on a $\mu\vec{x} = \lambda\vec{y}$ et donc $\vec{x} : \vec{y} = \lambda : \mu$. Réciproquement une famille de vecteurs qui sont tous proportionnels (et qui ne sont pas tous nuls) est contenue dans une droite vectorielle, car si $\vec{x}$ est un vecteur non nul de la famille, tout autre vecteur $\vec{y}$ de la famille vérifie $\vec{x} : \vec{y} = \beta : \alpha$ pour $\alpha, \beta \in K$, avec $\beta \neq 0$ (car $\beta = 0$ entraînerait $\vec{x} = \vec{0}$), ce qui permet d'écrire $\vec{y} = \frac{\alpha}{\beta}\vec{x} \in \langle \vec{x} \rangle$, et la droite vectorielle $\langle \vec{x} \rangle$ contient donc toute la famille de vecteurs.

Attention, s'il est vrai que pour deux vecteurs $\vec{x}, \vec{y}$ proportionnels on peut toujours écrire l'un comme un multiple de l'autre, on a pas toujours le choix d'écrire $\vec{y}$ comme multiple de $\vec{x}$ (car si par hasard $\vec{x} = \vec{0}$ et $\vec{y} \neq \vec{0}$, on sera obligé d'écrire $\vec{x} = 0\vec{y}$); c'est la raison pour laquelle dans l'argument ci-dessous on fait attention d'assurer que $\vec{x}$ soit non nul. Par contre si $\vec{x}$ et $\vec{y}$ sont des vecteurs proportionnels et *non nuls*, alors une relation $\vec{x} : \vec{y} = \beta : \alpha$ permettra d'écrire aussi bien $\vec{x} = \frac{\beta}{\alpha}\vec{y}$ que $\vec{y} = \frac{\alpha}{\beta}\vec{x}$, d'où $\frac{\beta}{\alpha}$ et $\frac{\alpha}{\beta}$ sont des scalaires non nuls déterminés (contrairement à α et β) par le couple $\vec{x}, \vec{y}$, et il conviendra dans ces cas de noter $\frac{\vec{x}}{\vec{y}}$ pour $\frac{\beta}{\alpha}$, et $\frac{\vec{y}}{\vec{x}}$ pour $\frac{\alpha}{\beta}$. Mais un quotient de vecteurs n'a un sens que dans ce cas précis.

Pour une famille de points alignés, tous les vecteurs $\overrightarrow{AB}$ formés par des points A, B de la famille sont proportionnels, car tous sont contenus dans la direction de la droite affine sur laquelle les points sont alignés. (Au cas où tous les points seraient identiques, on ne peut pas parler de cette droite, mais dans ce cas les vecteurs mentionnés seront tous nuls, et donc aussi tous proportionnels, car la définition donnée nous permet de dire $\vec{0} : \vec{0} = 1 : 1$.)

1.3.5. Proposition. *Pour des points $A \neq B$ et poids μ, ν de somme non nulle, $P = \operatorname{bar}((A, \mu), (B, \nu))$ est un point situé sur $\mathcal{D}_{A,B}$ de telle façon qu'on a la proportionnalité de vecteurs $\overrightarrow{AP} : \overrightarrow{PB} = \nu : \mu$.*

Preuve. La propriété du barycentre donne $\mu\overrightarrow{PA} + \nu\overrightarrow{PB} = \vec{0} = -\mu\overrightarrow{AP} + \nu\overrightarrow{PB}$, donc $\overrightarrow{AP} : \overrightarrow{PB} = \nu : \mu$.

On fera attention à l'ordre dans cette proportionnalité. Réciproquement si $C = A + \lambda\overrightarrow{AB}$ est un point quelconque, on pourra écrire C comme barycentre à partir de A et B: on vérifie par simple calcul que $C = \operatorname{bar}((A, 1 - \lambda), (B, \lambda))$. Ceci permet de caractériser $\mathcal{D}_{A,B}$ comme collection de barycentres:

1.3.6. Corollaire. *Pour des points $A \neq B$ on a $\mathcal{D}_{A,B} = \{ \operatorname{bar}((A, \mu), (B, \nu)) \mid \mu, \nu \in K, \mu + \nu \neq 0 \}$.*

Un triangle est une configuration géométrique déterminée par trois points non alignés, les sommets du triangle. (On fait exprès de garder un peu de flou sur ce qu'*est* une configuration, et un triangle en particulier, pour permettre d'y associer par la suite des attributs supplémentaires. Par exemple on pourra associer à une triangle ses trois cotés, qui sont des droites définies en termes des sommets du triangle, et considérer qu'elles font partie de la configuration.) Comme première application de barycentres, on a la caractérisation suivante du barycentre des sommets d'un triangle.

1.3.7. Proposition. *Soit A, B, C les sommets d'un triangle. Alors le barycentre $P = \operatorname{bar}(A, B, C)$ du triangle est situé sur chacune des droites $\mathcal{D}_{A,\operatorname{bar}(B,C)}$, $\mathcal{D}_{B,\operatorname{bar}(A,C)}$, et $\mathcal{D}_{C,\operatorname{bar}(A,B)}$, et il vérifie les proportionnalités $\overrightarrow{AP} : \overrightarrow{P\operatorname{bar}(B,C)} = 2 : 1$, $\overrightarrow{BP} : \overrightarrow{P\operatorname{bar}(A,C)} = 2 : 1$, et $\overrightarrow{CP} : \overrightarrow{P\operatorname{bar}(A,B)} = 2 : 1$.*

Preuve. C'est une simple application de l'associativité des barycentres (proposition 1.3.3), qui dit en particulier que $\mathrm{bar}(A, B, C) = \mathrm{bar}(A, (\mathrm{bar}(B, C), 2)) = \mathrm{bar}(B, (\mathrm{bar}(A, C), 2)) = \mathrm{bar}(C, (\mathrm{bar}(A, B), 2))$ □

On remarque que l'énoncé utilise implicitement des conditions telles que $A \neq \mathrm{bar}(B, C)$, nécessaires pour que les droites indiquées soient bien définies; ces conditions sont assurées par l'hypothèse que A, B, C ne sont pas alignés (car ils sont les sommets d'un triangle).

Ainsi la notion de barycentre permet de conclure que les droites $\mathcal{D}_{A,\mathrm{bar}(B,C)}$, $\mathcal{D}_{B,\mathrm{bar}(A,C)}$, $\mathcal{D}_{C,\mathrm{bar}(A,B)}$, dites *médianes* du triangle, sont *concourantes*, c'est-à-dire qu'elles ont un point en commun. Si cet énoncé est qualitatif, la notion de barycentre sert surtout à trouver de propriétés quantitatives. Voici un exemple.

1.3.8. Théorème [Ceva]. *Soit A, B, C les sommets d'un triangle, et $\lambda, \mu \in K$ des scalaires tels que $\{\lambda, \mu, 1-\lambda-\mu\} \cap \{0, 1\} = \emptyset$; on pose $P = \mathrm{bar}\big((A, 1-\lambda-\mu), (B, \lambda), (C, \mu)\big)$. Alors les droites $\mathcal{D}_{A,P}$, $\mathcal{D}_{B,P}$, $\mathcal{D}_{C,P}$ sont bien définies et coupent respectivement $\mathcal{D}_{B,C}$, $\mathcal{D}_{A,C}$, $\mathcal{D}_{A,B}$ en des points uniques, distincts des sommets du triangle, et qu'on appellera respectivement a, b, et c. On a les proportionnalités $\overrightarrow{Ba} : \overrightarrow{aC} = \mu : \lambda$, $\overrightarrow{Cb} : \overrightarrow{bA} = 1-\lambda-\mu : \mu$, et $\overrightarrow{Ac} : \overrightarrow{cB} = \lambda : 1-\lambda-\mu$, et par conséquent*

$$\frac{\overrightarrow{Ba}}{\overrightarrow{aC}} \times \frac{\overrightarrow{Cb}}{\overrightarrow{bA}} \times \frac{\overrightarrow{Ac}}{\overrightarrow{cB}} = 1.$$

Remarquons que l'hypothèse $\{\lambda, \mu, 1-\lambda-\mu\} \cap \{0, 1\} = \emptyset$ n'est qu'une forme compacte pour exprimer les 6 inégalités $\lambda \neq 0$, $\lambda \neq 1$, $\mu \neq 0$, ..., $1-\lambda-\mu \neq 1$. Le fait d'introduire P de cette façon un peu artificielle (au lieu de prendre un point quelconque du plan contentant le triangle) évite de devoir exclure explicitement des points sur les côtés du triangle, ou sur les droites parallèles à ces côtés et passant par le sommet restant, ce qui est nécessaire pour assurer que les points a, b, c soient bien définies. Le démonstration ci-dessous montre qu'avec l'approche choisie, le théorème relève d'un simple calcul.

Preuve. On montrera qu'en effet $a = \mathrm{bar}\big((B, \lambda), (C, \mu)\big)$, ainsi que $b = \mathrm{bar}\big((A, 1-\lambda-\mu), (C, \mu)\big)$ et $c = \mathrm{bar}\big((A, 1-\lambda-\mu), (B, \lambda)\big)$, après quoi le résultat découle de la proposition 1.3.5. Les barycentres figurant ici sont bien définis ($\lambda + \mu \neq 0$ est conséquence de $1-\lambda-\mu \neq 1$, et il est similaire pour les deux autres masses), ils sont situés respectivement sur $\mathcal{D}_{B,C}$, $\mathcal{D}_{A,C}$, et $\mathcal{D}_{A,B}$ d'après la proposition 1.3.5, et ils sont distincts des sommets car aucun des poids λ, μ, $1-\lambda-\mu$ n'est nul. Or, en prenant ces expressions barycentriques comme définitions de a, b, c, l'associativité des barycentres permet d'écrire

$$P = \mathrm{bar}\big((A, 1-\lambda-\mu), (a, \lambda+\mu)\big) = \mathrm{bar}\big((B, \lambda), (b, 1-\lambda)\big) = \mathrm{bar}\big((C, \mu), (c, 1-\mu)\big),$$

qui montre que P est situé sur chacune des droites $\mathcal{D}_{A,a}$, $\mathcal{D}_{B,b}$, et $\mathcal{D}_{C,c}$ (qui sont bien définies, car par exemple a étant aligné avec B et C ne peut pas être égal à A). Pour terminer, il suffira de montrer qu'on peut écrire ces droites aussi respectivement comme $\mathcal{D}_{A,P}$, $\mathcal{D}_{B,P}$, et $\mathcal{D}_{C,P}$, et que les points a, b, c, sont les *seuls* points respectivement de $\mathcal{D}_{A,P} \cap \mathcal{D}_{B,C}$, de $\mathcal{D}_{B,P} \cap \mathcal{D}_{A,C}$, et de $\mathcal{D}_{C,P} \cap \mathcal{D}_{A,B}$. Pour le premier point, le seul souci est de vérifier que P ne coïncide pas avec l'un des sommets du triangle, mais par exemple comme $\overrightarrow{AP} : \overrightarrow{Pa} = \lambda + \mu : 1-\lambda-\mu$ (toujours la proposition 1.3.5) et $\lambda + \mu \neq 0$, on a $P \neq A$; de façon similaire $P \neq B$ et $P \neq C$. Le second point est évident, car si par exemple $\mathcal{D}_{A,P} \cap \mathcal{D}_{B,C}$ contenait plus que le seul point a, il serait égal à $\mathcal{D}_{B,C}$ tout entier, contredisant le fait que A n'est pas aligné avec B et C. □

1.4. *Sous-espace affines et barycentres.*

On a vu que pour $A \neq B$ on peut caractériser $\mathcal{D}_{A,B}$ comme l'ensemble des barycentres formés à partir de A et B. En fait on peut caractériser l'ensemble des sous-espaces affines à l'aide des barycentres.

1.4.1. Proposition. *Une partie non vide S d'un espace affine $\mathcal{A}$ est un sous-espace affine si et seulement pour toute collection pondérée X de points de S, de masse non nulle, on a $\mathrm{bar}(X) \in S$.*

Preuve. On choisit $P \in S$. Supposons d'abord que S soit un espace affine, et montrons que S est fermé pour la formation de barycentres. Pour alors pour $X = \big((A_1, \mu_1), \ldots, (A_n, \mu_n)\big)$ on a $\mathrm{bar}(X) = P + \frac{1}{\mu(X)} \sum_{i=1}^{n} \mu_i \overrightarrow{PA_i}$, et comme $\overrightarrow{PA_i} \in \overrightarrow{S}$ pour tout i, on a $\mathrm{bar}(X) \in P + \overrightarrow{S} = S$. Réciproquement, supposons maintenant S une partie de $\mathcal{A}$ fermée pour la formation de barycentres. Avec

$F = \{\, \overrightarrow{PQ} \mid Q \in S \,\}$ on a $S = P + F$, et il suffit de montrer que F est sous-espace vectoriel de E pour conclure que S est sous-espace affine de $\mathcal{A}$; comme $\vec{0} \in F$, cela revient à vérifier que F est fermé pour les combinaisons linéaires. Soit donc $\vec{x}, \vec{y} \in F$ et $\lambda, \mu \in K$, et posons $Q = P+\vec{x}$, $R = P+\vec{y}$. On a $P+\lambda\vec{x}+\mu\vec{y} = P+\lambda\overrightarrow{PQ}+\mu\overrightarrow{PR} = \mathrm{bar}\big((P, 1-\lambda-\mu), (Q, \lambda), (R, \mu)\big) \in S$ par hypothèse, donc $\lambda\vec{x}+\mu\vec{y} \in F$ comme voulu. □

1.4.2. Définition. *Pour une partie non vide S de $\mathcal{A}$, l'ensemble des barycentres de toutes les collections pondérées de points de S de masse non nulle est appelé le sous-espace affine engendré par S, et noté* $\mathrm{Aff}(S)$.

Pour vérifier que $\mathrm{Aff}(S)$ est en effet un sous-espace affine de $\mathcal{A}$, on appliquera la proposition ci-dessus. Pour montrer qu'on a $\mathrm{bar}(X) \in \mathrm{Aff}(S)$ pour toute collection X de points pondérés A_i de $\mathrm{Aff}(S)$, de masse non nulle, il suffira (par définition de $\mathrm{Aff}(S)$) de récrire $\mathrm{bar}(X)$ comme barycentre d'une collection pondérée de points de S. Chaque A_i s'écrit $A_i = \mathrm{bar}\big((B_1, \nu_1), \ldots, (B_n, \nu_n)\big)$ avec les $B_i \in S$, où on peut supposer $\sum_{i=1}^{n} \nu_i = 1$, et en utilisant l'associativité des barycentres dans le sens opposé, on peut remplacer (A_i, μ_i) dans l'expression $\mathrm{bar}(X)$ par $(B_1, \mu_i\nu_1), \ldots, (B_n, \mu_i\nu_n)$. Faisant ainsi pour tout A_i, cette expression se trouve transformée en un barycentre d'une collection pondérée de points tous dans S.

Tout sous-espace affine de $\mathcal{A}$ qui contient S doit aussi contenir $\mathrm{Aff}(S)$, car il doit être fermé pour la formation de barycentres; ainsi on peut décrire $\mathrm{Aff}(S)$ comme le plus petit sous-espace affine de $\mathcal{A}$ contenant S, dans le sens que tout autre tel sous-espace contient $\mathrm{Aff}(S)$.

Exemples. Pour un singleton $S = \{P\}$ on a $\mathrm{Aff}(S) = S$. Pour $S = \{P, Q\}$ avec $P \neq Q$ on a $\mathrm{Aff}(S) = \mathcal{D}_{P,Q}$, d'après le corollaire 1.3.6. Pour une droite $\mathcal{D}$ on a $\mathrm{Aff}(\mathcal{D}) = \mathcal{D}$, et plus généralement chaque sous-espace affine $\mathcal{V}$ est égal à $\mathrm{Aff}(\mathcal{V})$. Pour trois points non-alignés P, Q, R, le sous-espace $\mathrm{Aff}(\{P, Q, R\})$ est l'unique plan qui contient le triangle P, Q, R. On simplifiera la notation à $\mathrm{Aff}(P, Q, R)$ dans ce cas.

Si $S = \{A_0, \ldots, A_n\}$, tout élément de $\mathrm{Aff}(S)$ s'écrit (en tant que barycentre) sous la forme $A_0 + \vec{x}$ où $\vec{x}$ est une combinaison linéaire des vecteurs $\overrightarrow{A_0A_1}, \ldots, \overrightarrow{A_0A_n}$, d'où ces vecteurs engendrent la direction $\overrightarrow{\mathrm{Aff}(S)}$. On en déduit que $\dim(\mathrm{Aff}(S)) \leq n = \#S - 1$ (où $\#S$ désigne le cardinal $n+1$ de S). Parmi les questions qui sont d'un intérêt particulier en géométrie, on trouve (outre la question déjà mentionnée si certains sous-espaces se coupent ou non) la question de la dimension de $\mathrm{Aff}(S)$ pour certaines collections S de points. Ainsi les points de S sont alignés si $\dim(\mathrm{Aff}(S)) \leq 1$, et ils sont *coplanaires* si $\dim(\mathrm{Aff}(S)) \leq 2$. On vient de donner une borne supérieure pour cette dimension, il est donc sans intérêt de dire que deux points sont alignés ou que trois point sont coplanaires. On va maintenant considérer la condition que cette dimension maximale est atteinte (c'est donc par exemple la condition pour trois points d'être *non-alignés*, pour quatre points d'être *non-coplanaires*, etc.).

1.4.3. Définition. *Une famille $S = (A_0, \ldots, A_n)$ de est dit affinement libre si* $\dim(\mathrm{Aff}(S)) = n$.

Par exemple, (A, B) est une famille affinement libre si $\mathrm{Aff}(A, B)$ est une droite, c'est-à-dire si $A \neq B$, et (A, B, C) est une famille affinement libre si $\mathrm{Aff}(A, B, C)$ est un plan, c'est-à-dire quand A, B, C ne sont pas alignés. Une famille (finie) S de points qui n'est pas affinement libre, c'est-à-dire pour laquelle $\dim(\mathrm{Aff}(S)) < \#S - 1$, est appelée *affinement liée*; cette notion généralise celle d'être «alignés» pour trois points. On peut caractériser la relation d'être affinement libre de diverses autres manières.

1.4.4. Théorème. *Soit $S = (A_0, \ldots, A_n)$ une famille avec $n \geq 1$, et $\mathcal{V} = \mathrm{Aff}(S)$. Sont équivalents:*
(i) S est une famille affinement libre, c'est-à-dire $\dim \mathcal{V} = n$*;*
(ii) Pour tout $i \leq n$ la famille de vecteurs $\{\, \overrightarrow{A_iA_j} \mid 0 \leq j \leq n, j \neq i \,\}$ est libre.
(iii) Il existe $i \leq n$ tel que la famille de vecteurs $\{\, \overrightarrow{A_iA_j} \mid 0 \leq j \leq n, j \neq i \,\}$ soit libre.
(iv) Pour tout $P \in \mathcal{V}$ il existe $\mu_0, \ldots, \mu_n$ uniques avec $\sum_{i=0}^{n} \mu_i = 1$ et $P = \mathrm{bar}\big((A_0, \mu_0), \ldots, (A_n, \mu_n)\big)$.
(v) Pour tout $i \leq n$ on a $A_i \notin \mathrm{Aff}(A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_n)$.
(vi) Pour tout i avec $0 < i \leq n$ on a $A_i \notin \mathrm{Aff}(A_1, \ldots, A_{i-1})$.

Preuve. (i)⇒(ii): La famille de vecteurs indiquée engendre $\overrightarrow{\mathrm{Aff}(S)}$ qui est de dimension n; comme ils sont n vecteurs, ils doivent former une famille libre. (ii)⇒(iii): Évident. (iii)⇒(iv): L'existence des μ_i est assurée par la définition de $\mathrm{Aff}(S)$; or pour une telle écriture on a $P = A_i + \sum_{j \neq i} \mu_j \overrightarrow{A_iA_j}$, une identité qui détermine les μ_j avec $j \neq i$ par l'hypothèse de liberté, et le dernier coefficient μ_i est fixé par la relation $\sum_{k=0}^{n} \mu_k = 1$. (iv)⇒(v): On a toujours $A_i = \mathrm{bar}\big((A_0, \mu_0), \ldots, (A_n, \mu_n)\big)$ avec $\mu_i = 1$ et les autres $\mu_j = 0$, et si on avait $A_i \in \mathrm{Aff}(A_1, \ldots, A_{i-1}, A_{i+1,}, \ldots, A_n)$ cela donnerait (après normalisation) une autre écriture, celle-là avec $\mu_i = 0$, ce qui contredit l'unicité dans l'hypothèse. (v)⇒(vi): Évident car $\mathrm{Aff}(A_1, \ldots, A_{i-1}) \subseteq \mathrm{Aff}(A_1, \ldots, A_{i-1}, A_{i+1,}, \ldots, A_n)$. (vi)⇒(i): En considérant la chaîne d'espaces vectoriels $\mathrm{Aff}(A_0) \subseteq \mathrm{Aff}(A_0, A_1) \subseteq \cdots \subseteq \mathrm{Aff}(A_0, \ldots, A_n) = \mathcal{V}$, l'hypothèse dit que toutes les n inclusions sont strictes. Cela implique $\dim \mathcal{V} \geq n$, et on a déjà vu que toujours $\dim \mathcal{V} \leq n$, donc $\dim \mathcal{V} = n$. □

1.5. *Repères et coordonnées cartésiens et barycentriques.*

La caractérisation (iv) du théorème précédent montre que pour une famille affinement libre S de m points on peut associer à tout point $P \in \mathrm{Aff}(S)$ un m-uplet particulier de poids μ_i pour lequel P s'écrit comme barycentre à partir des points de S; bien que d'autres telles écritures existent, ce m-uplet est distingué par la condition $\sum_i \mu_i = 1$. Il est naturel de considérer les μ_i comme des coordonnées qui décrivent P, et on les appellera des *coordonnées barycentriques*. Mais il existe une autre façon de décrire des points en coordonnées, en utilisant les coordonnées cartésiens pour les vecteurs. On décrira maintenant les deux systèmes de coordonnées, les «repères» nécessaires pour les introduire, et les relations entre eux.

1.5.1. Définition. *Un repère cartésien dans un espace affine $\mathcal{A}$ est un couple $\mathcal{R} = (\mathcal{O}, \mathcal{E})$, où $\mathcal{O} \in \mathcal{A}$ est un point appelé l'origine du repère, et $\mathcal{E}$ est une base de l'espace vectoriel $\overrightarrow{\mathcal{A}}$. Un repère affine $\mathcal{S}$ dans $\mathcal{A}$ est une famille affinement libre maximale, c'est-à-dire avec $\mathrm{Aff}(\mathcal{S}) = \mathcal{A}$, ou encore avec $\#S = \dim \mathcal{A} + 1$. Le repère affine associé à un repère cartésien $\mathcal{R} = (\mathcal{O}, (\vec{b}_1, \ldots, \vec{b}_n))$ est $\mathcal{S} = (\mathcal{O}, \mathcal{O} + \vec{b}_1, \ldots, \mathcal{O} + \vec{b}_n)$.*

Si $\mathcal{R} = (\mathcal{O}, (\vec{b}_1, \ldots, \vec{b}_n))$ est un repère cartésien, on peut écrire tout point $P \in \mathcal{A}$ de façon unique comme $P = \mathcal{O} + p_1\vec{b}_1 + \cdots + p_n\vec{b}_n$; on écrira $P = (p_1, \ldots, p_n)_{\mathcal{R}}$ qu'on appellera les coordonnées (cartésiennes) de P par rapport au repère $\mathcal{R}$. Si $\mathcal{S} = (A_0, \ldots, A_n)$ est le repère affine associé à $\mathcal{R}$, où donc $A_0 = \mathcal{O}$ et $A_i = \mathcal{O} + \vec{b}_i$ pour $i > 0$, l'égalité $P = A_0 + p_1\overrightarrow{A_0A_1} + \cdots + p_n\overrightarrow{A_0A_n}$ est équivalente à $P = \mathrm{bar}\big((A_0, p_0), \ldots, (A_n, p_n)\big)$, où l'on définit $p_0 = 1 - \sum_{i=1}^{n} p_i$ pour assurer que la somme des poids soit 1. On notera $P = (p_0, \ldots, p_n)_{\mathcal{S}}$, le point coordonnées barycentriques $p_0, \ldots, p_n$ par rapport à $\mathcal{S}$.

Les deux types de coordonnées ont chacun leurs avantages. Les coordonnées barycentriques ont besoin d'un coefficient de plus, mais qui est contraint par la condition que la somme des coordonnées doit toujours être 1. En contrepartie les coordonnées barycentriques mettent en évidence une symétrie plus grande que les coordonnées cartésiennes: pour toute permutation des $n + 1$ points du repère affine, les coordonnées barycentriques s'adaptent en appliquant la même permutation au $n + 1$-uplet; pour les coordonnées cartésiennes une telle permutation, qui peut changer l'origine $\mathcal{O}$, changera les coordonnées d'une façon beaucoup moins simple à décrire. Mais en tout cas, on a vu que les deux systèmes de coordonnées pour des repères associés ne diffèrent que peu, donc on peut facilement passer de l'un à l'autre si cela facilite les opérations qu'on veut effectuer.

Pour les coordonnées cartésiennes, on a aussi des coordonnées pour décrire les vecteurs, en notant $(v_1, \ldots, v_n)_{\mathcal{E}}$ le vecteur $v_1\vec{b}_1 + \cdots + v_n\vec{b}_n$. Le passage en coordonnées entre points et vecteurs est naturel:

$$\begin{aligned}(p_1, \ldots, p_n)_{\mathcal{R}} + (x_1, \ldots, x_n)_{\mathcal{E}} &= (p_1 + x_1, \ldots, p_n + x_n)_{\mathcal{R}},\\ \overrightarrow{(p_1, \ldots, p_n)_{\mathcal{R}}(q_1, \ldots, q_n)_{\mathcal{R}}} &= (q_1 - p_1, \ldots, q_n - p_n)_{\mathcal{E}}.\end{aligned}$$

Pour calculer de la même façon en coordonnées barycentriques, il nous faudra une notation barycentrique pour les vecteurs. Dans cette notation la somme des coordonnées sera toujours 0; nous mettrons aussi un 'v' dans la notation pour indiquer qu'il s'agit d'un vecteur. Si $\mathcal{S}$ est le repère affine associé à $\mathcal{R}$ on pose

$$v(p_0, \ldots, p_n)_{\mathcal{S}} = (p_1, \ldots, p_n)_{\mathcal{E}} \in E \qquad \text{où obligatoirement } \textstyle\sum_{i=0}^{n} p_i = 0.$$

Cette notation est liée à celle qui permet d'écrire, pour une c.p.p. X de masse nulle, $v(X)$ pour le vecteur $v(P, X)$ qui dans ce cas ne dépend pas de P : on a

$$v(p_0, \ldots, p_n)_{\mathcal{S}} = v\big(((A_0, p_0), \ldots, (A_n, p_n))\big) \qquad \text{si } \mathcal{S} = (A_0, \ldots, A_n) \text{ et } \textstyle\sum_{i=0}^{n} p_i = 0.$$

On déduira facilement les règles de calcul similaires à celles pour les coordonnées cartésiennes :

$$\begin{aligned}(p_0, \ldots, p_n)_{\mathcal{S}} + v(x_0, \ldots, x_n)_{\mathcal{S}} &= (p_0 + v_0, \ldots, p_n + v_n)_{\mathcal{S}},\\ \overrightarrow{(p_0, \ldots, p_n)_{\mathcal{S}}(q_0, \ldots, q_n)_{\mathcal{S}}} &= v(q_0 - p_0, \ldots, q_n - p_n)_{\mathcal{S}}.\end{aligned}$$

1.6. *Calcul de dimension d'un sous-espace affine en coordonnées.*

Comme il a été dit ci-dessus, beaucoup de questions étudiées en géométrie affine reviennent à déterminer $\dim(\mathrm{Aff}(S))$ pour une collection finie S de points. Dans cette section on va montrer comment cette question peut être traduit en termes de coordonnées dans un repère cartésien ou affine. Commençons par rappeler que la question correspondante en algèbre linéaire, à savoir déterminer la dimension d'un sous-espace engendré par une famille finie de vecteurs, donne lieu à la notion de *rang* d'une matrice: si on exprime les vecteurs dans une base quelconque, la dimension cherchée est égale au rang de la matrice M dont les colonnes sont données par les coordonnées des vecteurs. Ce rang peut être caractérisé de plusieurs façons: il est clairement égal à la dimension de l'espace engendré par les colonnes de M, autrement dit au nombre maximum de colonnes qui forment une famille libre, et il est aussi égal au nombre maximum de lignes de M qui forment une famille libre, et à la taille maximale d'une matrice carrée *inversible* extraite de M (ce que veut dire : obtenue en supprimant éventuellement certaines lignes et colonnes), c'est-à-dire dont le déterminant est non nul. Cette dernière caractérisation affirme en particulier que le rang d'une matrice $n \times n$ est $< n$ si et seulement si le déterminant de la matrice est nul (sinon son rang est n).

1.6.1. Théorème. *Soit $A_1, \ldots, A_k$ des points d'un espace affine $\mathcal{A}$ de dimension n.*

(1) *Si $\mathcal{R}$ est un repère cartésien, et M est la matrice de taille $(n+1) \times k$ dont toutes ses entrées de la première ligne sont égales à 1, et dont les entrées dans les autres lignes de la colonne j sont données par les coordonnées cartésiennes de A_j dans le repère $\mathcal{R}$, alors* $\dim(\mathrm{Aff}(A_1, \ldots, A_k)) = \mathrm{rg}(M) - 1$.

(2) *Si $\mathcal{S}$ est un repère affine, et M est la matrice de taille $(n+1) \times k$ dont les entrées de la colonne j sont données par les coordonnées barycentriques de A_j dans le repère $\mathcal{S}$, alors*

$$\dim(\mathrm{Aff}(A_1, \ldots, A_k)) = \mathrm{rg}(M) - 1.$$

Par exemple, si $\mathcal{A}$ est un plan affine, $\mathcal{R}$ et $\mathcal{S}$ sont des repères cartésien et affine associés, et on considère les trois points $A = (a_1, a_2)_{\mathcal{R}} = (a_0, a_1, a_2)_{\mathcal{S}}$, $B = (b_1, b_2)_{\mathcal{R}} = (b_0, b_1, b_2)_{\mathcal{S}}$ et $C = (c_1, c_2)_{\mathcal{R}} = (c_0, c_1, c_2)_{\mathcal{S}}$, la proposition dit que ces point seront alignés si et seulement si

$$\begin{vmatrix} 1 & 1 & 1 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix} = 0, \quad \text{ou de façon équivalente} \quad \begin{vmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix} = 0.$$

En plus elle dit que les trois points sont confondus si et seulement si les matrices en question sont de rang 1 (pour la matrice à gauche cela implique clairement $a_1 = b_1 = c_1$ et $a_2 = b_2 = c_2$; pour la matrice à droite il peut surprendre que la seule condition d'avoir rang 1 force les colonnes de la matrice à être identiques, mais cela est conséquence de la condition implicite que les coordonnées barycentriques d'un point sont toujours de somme 1). Pour quatre points dans un espace de dimension 3 on aura de façon similaire qu'ils sont coplanaires si les matrices 4×4 correspondantes sont de déterminant 0 (et alignés si elles sont de rang ≤ 2).

Cette proposition est surtout utile pour décider si certaines familles de points sont liées ou non, ce qui revient simplement à tester si le déterminant de la matrices donnée est nul. La version avec les coordonnées barycentriques sera à préférer dans les cas où parmi a_0, b_0, c_0 certains sont nuls, ou quand on a une symétrie supplémentaire dans les expressions dans le cas de coordonnées barycentriques.

Preuve. On se servira du fait que les opérations habituelles avec les lignes et les colonnes d'une matrice ne changent pas son rang, pour ramener les deux cas à la situation en algèbre linéaire qu'on vient de mentionner. Pour le cas (1) des coordonnées cartésiennes, on soustrait d'abord la première colonne de M de toutes les autres, ce qui fait apparaître des zéros dans la première ligne et les coordonnées des vecteurs $\overrightarrow{A_1A_2}, \ldots, \overrightarrow{A_1A_k}$ dans les autre lignes, et puis par des opérations avec la première ligne on fait en sorte que l'entrée 1 en position 1 soit la seule non nulle dans sa colonne (elle l'était déjà dans sa ligne). Alors toute matrice carrée extraite du résultat M' qui est de déterminant non nul et qui contient la première ligne ou la première colonne doit aussi contenir l'autre des deux (pour éviter une ligne ou colonne nulle); en plus si elle est de taille maximale, elle contient effectivement cette ligne et cette colonne. Par conséquent le rang de M' est un de plus que le rang de la matrice obtenue en supprimant sa première ligne et colonne.

Le rang de cette dernière est $\dim(\mathrm{Vect}(\overrightarrow{A_1A_2},\dots,\overrightarrow{A_1A_k})) = \dim(\overrightarrow{\mathrm{Aff}(A_1,\dots,A_k)})$, d'où l'énoncé. Pour le cas (2), il suffit d'ajouter à la première ligne de la matrice toutes les autres lignes, pour le ramener au cas précédent, parce que les sommes de coordonnées barycentriques des points $A_1,\dots,A_k$ sont toutes 1. □

Pour illustrer l'utilisation de ce théorème, nous démontrons le théorème dit de Ménélaüs.

1.6.2. Théorème. *Soit A,B,C un triangle du plan affine, et a,b,c des points sur ses côtés, donnés par $a = \mathrm{bar}((B,\lambda),(C,1-\lambda))$, $b = \mathrm{bar}((C,\mu),(A,1-\mu))$ et $c = \mathrm{bar}((A,\nu),(B,1-\nu))$. Les points a,b,c sont alignés si soit $\{0,1\} \subseteq \{\lambda,\mu,\nu\}$ (il sont alors alignés sur l'un des côtés), soit $\frac{1-\lambda}{\lambda} \times \frac{1-\mu}{\mu} \times \frac{1-\nu}{\nu} = -1$.*

Preuve. On utilise $\mathcal{S} = (A,B,C)$ comme repère affine. Le condition d'alignement s'exprime comme $\dim(\mathrm{Aff}(a,b,c)) \leq 1$, ce qui se traduit d'après le théorème 1.6.1 par

$$\begin{vmatrix} 0 & 1-\mu & \nu \\ \lambda & 0 & 1-\nu \\ 1-\lambda & \mu & 0 \end{vmatrix} = 0, \quad \text{ce qui donne} \quad (1-\lambda)(1-\mu)(1-\nu) = -\lambda\mu\nu.$$

Si un paramètre est nul un autre doit être 1, donc $\{0,1\} \subseteq \{\lambda,\mu,\nu\}$; sinon on a $\frac{1-\lambda}{\lambda} \times \frac{1-\mu}{\mu} \times \frac{1-\nu}{\nu} = -1$. □

On peut observer que les trois fractions dans le produit sont les mêmes qu'on a vues dans le théorème de Ceva, c'est juste le résultat qui est de signe opposé. On aurait pu utiliser des coordonnées cartésiennes, ce qui mène a remplacer les entrées de la première ligne de la matrice par des 1; cela ne change pas le déterminant, mais la manque de symétrie aurait rendu la factorisation utilisée moins facile à trouver.

1.7. *Description de sous-espaces affines en coordonnées.*

Les repères permettent d'écrire les sous-espaces affines en coordonnées. Si $\mathcal{V} = P + \overrightarrow{\mathcal{V}}$ est un sous-espace affine et $(\vec{v}_1,\dots,\vec{v}_k)$ une base de $\overrightarrow{\mathcal{V}}$, on a la description dite paramétrique de $\mathcal{V}$:

$$\mathcal{V} = \{\, P + \lambda_1\vec{v}_1 + \cdots + \lambda_k\vec{v}_k \mid \lambda_1,\dots,\lambda_k \in K \,\},$$

et si $P = (p_1,\dots,p_n)_{\mathcal{R}}$ et $\vec{v}_j = (a_{1,j},\dots,a_{n,j})_{\mathcal{E}}$ pour $j = 1,\dots,k$ dans un repère cartésien $\mathcal{R} = (\mathcal{O},\mathcal{E})$:

$$\mathcal{V} = \{\, \bigl(p_1 + \sum_{j=1}^{k}\lambda_j a_{1,j},\dots,p_n + \sum_{j=1}^{k}\lambda_j a_{n,j}\bigr)_{\mathcal{R}} \mid \lambda_1,\dots,\lambda_k \in K \,\}.$$

Cette forme est surtout utile quand $\mathcal{V}$ est une droite $\mathcal{D}$, c'est-à-dire avec $k = 1$:

$$\mathcal{D} = \{\, \bigl(p_1 + \lambda a_1,\dots,p_n + \lambda a_n\bigr)_{\mathcal{R}} \mid \lambda \in K \,\}.$$

Les expressions en coordonnées barycentriques sont similaires, mais évidemment avec une coordonnée de plus. Pour les sous-espaces affines de dimensions élevée, notamment pour les hyperplans, il est plus pratique de les décrire par des *équations* en coordonnées, ce qui donne le suivant.

1.7.1. Proposition. *Soit $\mathcal{H}$ un hyperplan de $\mathcal{A}$, soit $\mathcal{R}$ un repère cartésien, et $\mathcal{S}$ le repère affine associé.*

(1) Il existe $\alpha_1,\dots,\alpha_n \in K$, pas tous nuls, et $b \in K$, tels que $\mathcal{H}$ soit donné en coordonnées cartésiennes par

$$\mathcal{H} = \{\, (x_1,\dots,x_n)_{\mathcal{R}} \mid \alpha_1x_1 + \cdots + \alpha_nx_n = b \,\}.$$

Or le $n+1$-uplet $(\alpha_1,\dots,\alpha_n,b)$ est unique à multiplication simultanée par un scalaire $\lambda \in K^*$ près.

(2) Il existe $\beta_0,\dots,\beta_n \in K$, pas tous égaux, tels que $\mathcal{H}$ soit donné en coordonnées barycentriques par

$$\mathcal{H} = \{\, (y_0,\dots,y_n)_{\mathcal{S}} \mid \textstyle\sum_{i=0}^{n} y_i = 1, \quad \beta_0y_0 + \cdots + \beta_ny_n = 0 \,\}.$$

Or le $n+1$-uplet $(\beta_0,\dots,\beta_n)$ est unique à multiplication simultanée par un scalaire $\lambda \in K^*$ près.

(3) Pour un $n+1$-uplet d'un de deux types, on peut en trouver un de l'autre type en se servant des relations $\beta_0 = -b$, et $\beta_i = \alpha_i - b$ pour $i = 1,\dots,n$.

Preuve. Posons $\mathcal{R} = (\mathcal{O}, \mathcal{E})$. Comme $\overrightarrow{\mathcal{H}}$ est un hyperplan vectoriel de $E = \overrightarrow{\mathcal{A}}$, il existe une forme linéaire non nulle $\alpha \in E^* = \mathcal{L}(E, K)$ dont le noyau est $\overrightarrow{\mathcal{H}}$, et elle est unique à un facteur scalaire près. En coordonnées une telle forme est donnée par $\alpha : (x_1, \ldots, x_n)_{\mathcal{E}} \mapsto \alpha_1 x_1 + \cdots + \alpha_n x_n$ pour $\alpha_1, \ldots, \alpha_n \in K$. Si $P = (p_1, \ldots, p_n)_{\mathcal{R}}$ est un point quelconque de $\mathcal{H}$, on pose $b = \alpha(\overrightarrow{\mathcal{O}P}) = \alpha_1 p_1 + \cdots + \alpha_n p_n$, et $Q = (x_1, \ldots, x_n)_{\mathcal{R}} \in \mathcal{H}$ donne $\alpha(\overrightarrow{PQ}) = 0 = \alpha_1(x_1 - p_1) + \cdots + \alpha_n(x_n - p_n)$ et donc $\alpha_1 x_1 + \cdots + \alpha_n x_n = b$. Pour en déduire les autres points on remarque que $(x_1, \ldots, x_n)_{\mathcal{R}} = (y_0, \ldots, y_n)_{\mathcal{S}}$ signifie que $x_i = y_i$ pour $i > 0$ et $y_0 = 1 - \sum_{i=1}^n x_i$; dans l'équation on remplace b par $b \sum_{i=0}^n y_i$, et après regroupement des termes on obtient $-by_0 + (\alpha_1 - b)y_1 + \cdots + (\alpha_n - b)y_n = 0$, d'où on trouve les énoncés (2) et (3). $\square$

La forme barycentrique de l'équation pour un hyperplan (point (2)) est particulièrement utile, car elle rajoute une équation *homogène* à la condition $\sum_{i=0}^n y_i = 1$ à laquelle le coordonnées barycentriques sont soumises d'office (bien sûr celle-ci n'est pas homogène). On utilisera la notation

$$\mathcal{H} = [\beta_0, \ldots, \beta_n]_{\mathcal{S}}$$

pour l'hyperplan ainsi détermine par les coefficients $\beta_0, \ldots, \beta_n$. Les crochets servent pour distinguer cette notation de celle d'un *point* en coordonnées barycentriques. On rappelle que (contrairement aux coordonnées barycentriques d'un point) les β_i ne sont déterminés par $\mathcal{H}$ qu'à un scalaire près, et qu'ils ne doivent pas être tous égaux (s'ils l'étaient, les équations dans le point (2) seraient contradictoires).

1.7.2. Proposition. *Soit $\mathcal{A}$ un plan affine ($\dim \mathcal{A} = 2$), muni d'un repère affine $\mathcal{S} = (A_0, A_1, A_2)$.*

(1) La droite passant par deux points distincts $P = (p_0, p_1, p_2)_{\mathcal{S}}$, $Q = (q_0, q_1, q_2)_{\mathcal{S}}$, est donnée par

$$\mathcal{D}_{P,Q} = [p_1 q_2 - p_2 q_1, p_2 q_0 - p_0 q_2, p_0 q_1 - p_1 q_0]_{\mathcal{S}}.$$

(2) Deux droites $\mathcal{D} = [c_0, c_1, c_2]_{\mathcal{S}}$ et $\mathcal{D}' = [d_0, d_1, d_2]_{\mathcal{S}}$ sont parallèles si et seulement si

$$\begin{vmatrix} 1 & 1 & 1 \\ c_0 & c_1 & c_2 \\ d_0 & d_1 & d_2 \end{vmatrix} = 0.$$

(3) Le point d'intersection de deux droites non parallèles $\mathcal{D} = [c_0, c_1, c_2]$, $\mathcal{D}' = [d_0, d_1, d_2]$ est donné par

$$\mathcal{D} \cap \mathcal{D}' = \{ \frac{1}{f}(c_1 d_2 - c_2 d_1, c_2 d_0 - c_0 d_2, c_0 d_1 - c_1 d_0)_{\mathcal{S}} \},$$

où le facteur $f = c_1 d_2 - c_2 d_+ c_2 d_0 - c_0 d_2 + c_0 d_1 - c_1 d_0$ rend la somme des coordonnées égale à 1.

Preuve. D'après le théorème 1.6.1, un point $A = (a_0, a_1, a_2)_{\mathcal{S}}$ est aligné avec P et Q si et seulement si

$$\begin{vmatrix} a_0 & p_0 & q_0 \\ a_1 & p_1 & q_1 \\ a_2 & p_2 & q_2 \end{vmatrix} = 0,$$

et comme le déterminant est 3-linéaire par rapport à ces colonnes, ceci donne, pour P, Q fixes, une condition linéaire homogène sur les coordonnées de A. Cette condition s'écrit $\beta_0 a_0 + \beta_1 a_1 + \beta_2 a_2 = 0$ avec $(\beta_0, \beta_1, \beta_2) = (p_1 q_2 - p_2 q_1, p_2 q_0 - p_0 q_2, p_0 q_1 - p_1 q_0)$, d'où le point (1). Pour le point (2), on observe qu'un point $A = (a_0, a_1, a_2)_{\mathcal{S}}$ vérifie $A \in \mathcal{D} \cap \mathcal{D}'$ si et seulement si le vecteur $(a_0, a_1, a_2) \in K^3$ est envoyé sur $(1, 0, 0)$ par la matrice écrite. Dire que le déterminant de cette matrice est nul veut dire que ce système d'équations pour (a_0, a_1, a_2) *n'est pas* de Cramer ; il possède alors ou bien aucune solution, ou bien plusieurs solutions, et c'est précisément le cas si $\mathcal{D}$ et $\mathcal{D}'$ sont parallèles. Dans le cas contraire, le déterminant est égale à f du point (3), et la formule donnée est juste la règle de Cramer pour ce système. $\square$

1.7.3. Proposition. *Trois droites $\mathcal{D}_i = [\beta_{i,0}, \beta_{i,1}, \beta_{i,2}]_{\mathcal{S}}$ pour $i = 1, 2, 3$ sont concourantes ou parallèles si et seulement si*

$$\begin{vmatrix} \beta_{1,0} & \beta_{1,1} & \beta_{1,2} \\ \beta_{2,0} & \beta_{2,1} & \beta_{2,2} \\ \beta_{3,0} & \beta_{3,1} & \beta_{3,2} \end{vmatrix} = 0.$$

Preuve. Pour $\mathcal{D}_i$, l'équation homogène $\beta_{i,0} y_0 + \beta_{i,1} y_1 + \beta_{i,2} y_2$ définit un plan vectoriel V_i dans K^3, qui n'est pas parallèle au plan affine B donné par l'équation $y_0 + y_1 + y_2 = 1$ qui caractérise les triplets de coordonnées barycentriques valables (et l'intersection de $V_i \cap B$ donne les coordonnées barycentriques des points de $\mathcal{D}_i$). Le déterminant est nul si et seulement si $\dim(V_1 \cap V_2 \cap V_3) > 0$. Si c'est le cas, soit $D = V_1 \cap V_2 \cap V_3$ coupe B, et $\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3$ sont donc concourantes ; soit D est parallèle à B, et les droites sont parallèles : $D \subseteq \overrightarrow{\mathcal{D}}_1 \cap \overrightarrow{\mathcal{D}}_2 \cap \overrightarrow{\mathcal{D}}_3$. Dans le cas contraire $\mathcal{D}_1 \cap \mathcal{D}_2 \cap \mathcal{D}_3 = \emptyset$ et $\overrightarrow{\mathcal{D}}_1 \cap \overrightarrow{\mathcal{D}}_2 \cap \overrightarrow{\mathcal{D}}_3 = \{\vec{0}\}$. $\square$

1.8. *Applications affines.*

Dans tous les domaines de l'algèbre et de la géométrie, il est d'une importance fondamentale de connaître les applications d'une structure (espace) vers une autre du même type qui "respectent" ses relations pertinentes ; ces applications sont appelées de façon générique des morphismes (et ceux qui possèdent un morphisme réciproque sont des isomorphismes). Pour les espaces affines, ce sont les applications affines.

1.8.1. Définition. *Soit $\mathcal{A}, \mathcal{A}'$ deux espaces affines sur K. Une application $f : \mathcal{A} \to \mathcal{A}'$ est appelée une application affine (ou morphisme d'espaces affines, ou encore affinité) s'il existe une application K-linéaire $\varphi : \overrightarrow{\mathcal{A}} \to \overrightarrow{\mathcal{A}'}$ telle que pour tout $A, B \in \mathcal{A}$ on ait $\overrightarrow{f(A)f(B)} = \varphi(\overrightarrow{AB})$. On notera $\varphi = \overrightarrow{f}$.*

Pour justifier la notation $\overrightarrow{f}$ (qu'on appelle l'application linéaire associée à f), il faut l'unicité de φ, mais c'est clair car tout vecteur de la direction $\overrightarrow{\mathcal{A}}$ s'écrit (de multiples façons) sous la forme $\overrightarrow{AB}$, dont la définition prescrit l'image par φ en termes de f. Par contre l'existence de φ n'est pas évidente, car différentes écritures du même vecteur $\vec{x} \in \overrightarrow{\mathcal{A}}$ pourraient exiger des valeurs contradictoires pour $\varphi(\vec{x})$. Mais si pour $A_0 \in \mathcal{A}$ on *définit* $\varphi(\vec{x}) = \overrightarrow{f(A_0)f(A_0 + \vec{x})}$ (de sorte que $\overrightarrow{f(A_0)f(B)} = \varphi(\overrightarrow{A_0B})$ pour tout $B \in \mathcal{A}$) et *si ce φ est linéaire*, alors ces exigences seront concordantes: pour un autre point A on aura $\overrightarrow{f(A)f(A+\vec{x})} = \overrightarrow{f(A)f(A_0)} + \overrightarrow{f(A_0)f(A+\vec{x})} = -\varphi(\overrightarrow{A_0A}) + \varphi(\overrightarrow{A_0A} + \vec{x}) = \varphi(\vec{x})$ par linéarité, donc l'écriture $\vec{x} = \overrightarrow{A(A+\vec{x})}$ donne lieu à la même valeur pour $\varphi(\vec{x})$. Ainsi on voit que dans la définition, on pourrait se limiter à exiger la relation entre f et φ pour un seul point A (mais toujours pour tout point B) sans changer le sens de la définition. La formulation actuelle a cependant certains avantages, notamment elle rend évident le fait que pour deux applications affines $f : \mathcal{A} \to \mathcal{A}'$ et $g : \mathcal{A}' \to \mathcal{A}''$, la composée $g \circ f : \mathcal{A} \to \mathcal{A}''$ est aussi affine, avec $\overrightarrow{g \circ f} = \overrightarrow{g} \circ \overrightarrow{f}$.

1.8.2. Proposition.
(1) Si $f : \mathcal{A} \to \mathcal{A}'$ est affine, alors $f(P + \vec{x}) = f(P) + \overrightarrow{f}(\vec{x})$ pour tout $P \in \mathcal{A}$ et $\vec{x} \in \overrightarrow{\mathcal{A}}$.
(2) Une application $f : \mathcal{A} \to \mathcal{A}'$ est affine si et seulement si elle conserve les barycentres:

$$f(\mathrm{bar}((A_1, \mu_1), \ldots, (A_n, \mu_n))) = \mathrm{bar}((f(A_1), \mu_1), \ldots, (f(A_n), \mu_n)).$$

(3) Une application affine $f : \mathcal{A} \to \mathcal{A}'$ est déterminée par la donnée de son application linéaire associée $\overrightarrow{f}$ et l'image $f(A)$ d'un point $A \in \mathcal{A}$ quelconque.
(4) Réciproquement, si après avoir choisi un point $A \in \mathcal{A}$ on donne une application linéaire $\varphi : \overrightarrow{\mathcal{A}} \to \overrightarrow{\mathcal{A}'}$ ainsi qu'un point $A' \in \mathcal{A}'$, il existe une application affine $f : \mathcal{A} \to \mathcal{A}'$ telle que $\overrightarrow{f} = \varphi$ et $f(A) = A'$.
(5) Une application affine $f : \mathcal{A} \to \mathcal{A}'$ est déterminée par la donnée des images $f(A_i) \in \mathcal{A}'$ d'un repère affine $A_0, \ldots, A_n$ dans $\mathcal{A}$. Ces images peuvent être librement choisies dans $\mathcal{A}'$ (f existera toujours).

Preuve. Le point (1) est une conséquence directe de la définition, et il entraîne (3). Pour le point (4) on peut définir $f(A + \vec{x}) = A' + \varphi(\vec{x})$ pour tout $\vec{x} \in \overrightarrow{\mathcal{A}}$, et on remarqué que d'avoir ceci avec φ linéaire est suffisant pour que f soit affine avec $\overrightarrow{f} = \varphi$. En appliquant (1) et la formule pour le barycentre, on trouve que f conserve les barycentres: le premier membre est (avec barycentre calculé à l'aide de P)

$$f\big(P + \frac{1}{\mu(X)} \sum_{i=1}^{n} \mu_i \overrightarrow{PA_i}\big) = f(P) + \frac{1}{\mu(X)} \sum_{i=1}^{n} \mu_i \overrightarrow{f}(\overrightarrow{PA_i}) = f(P) + \frac{1}{\mu(X)} \sum_{i=1}^{n} \mu_i \overrightarrow{f(P)f(A_i)},$$

ce qui donne le barycentre du second membre, calculé à l'aide de $f(P)$. Réciproquement, si $f : \mathcal{A} \to \mathcal{A}'$ conserve le barycentres et $\mathcal{S} = (A_0, \ldots, A_n)$ est un repère affine de $\mathcal{A}$, avec repère cartésien associée $\mathcal{R} = (A_0, \mathcal{E})$, et si l'on pose $\varphi : \overrightarrow{\mathcal{A}} \to \overrightarrow{\mathcal{A}'}$ l'application linéaire qui envoie $\overrightarrow{A_0A_i} \mapsto \overrightarrow{f(A_0)f(A_i)}$, on vérifie pour $\vec{x} = (x_1, \ldots, x_n)_{\mathcal{E}}$ que $f(A_0 + \vec{x}) = f((x_0, \ldots, x_n)_{\mathcal{S}}) = f(A_0) + \sum_{i=1}^{n} x_i \varphi(\overrightarrow{A_0A_i}) = f(A_0) + \varphi(\vec{x})$ (où x_0 a complété les coordonnées barycentriques), donc f est l'application affine du point (4) pour $A = A_0$, φ, et $A' = f(A_0)$. L'application $f : (x_0, \ldots, x_n)_{\mathcal{S}} \mapsto \mathrm{bar}((A'_0, x_0), \ldots, (A'_n, x_n))$ convient pour le point (5). □

On donne maintenant quelques exemples d'applications affines, avec l'accent sur le cas $\mathcal{A} = \mathcal{A}'$ où l'application est une transformation de l'espace affine en lui-même (un endomorphisme affine).

- Toute application constante f est affine, avec $\overrightarrow{f} = 0$, l'application linéaire nulle.

- Pour un repère affine $\mathcal{S}$ de $\mathcal{A}$ donné, chaque coordonnée barycentrique définit une application affine $x_i : \mathcal{A} \to K$ (en considérant K comme espace affine); $\overrightarrow{x_i} : E \to K$ est égal à la i-ème fonction coordonnée de $E = \overrightarrow{\mathcal{A}}$ par rapport à la base du repère cartésien associé à $\mathcal{S}$, pour $i > 0$.
- Pour tout $\vec{x} \in E$, la *translation* $t_{\vec{x}} : A \mapsto A + \vec{x}$ est une application affine $\mathcal{A} \to \mathcal{A}$, avec $\overrightarrow{t_{\vec{x}}} = \mathrm{id}_E$.
- Pour tout point $P \in \mathcal{A}$, la *symétrie centrale* $S_P : A \mapsto P - \overrightarrow{PA}$ est une application affine $\mathcal{A} \to \mathcal{A}$, avec $\overrightarrow{S_P} = -\mathrm{id}_E$ (on appelle P le centre de la symétrie centrale S_P; c'est son unique point fixe).
- Pour tout point $P \in \mathcal{A}$ et scalaire $\lambda \in K^*$ l' *homothétie* $h_{P,\lambda} : A \mapsto A \mapsto P + \lambda\overrightarrow{PA}$ est une application affine $\mathcal{A} \to \mathcal{A}$, avec $\overrightarrow{h_{P,\lambda}} = \lambda\,\mathrm{id}_E$. (On interdit le cas $\lambda = 0$, qui donnerait l'application constante de valeur P; ainsi toute homothétie $h_{P,\lambda}$ possède une application inverse, à savoir $h_{P,\lambda^{-1}}$. D'ailleurs, pour $\lambda = 1$ on retrouve l'identité (quel que soit P), et pour $\lambda = -1$ la symétrie centrale de centre P.) Si $\lambda \neq 1$, l'homothétie $h_{P,\lambda}$ a P comme unique point fixe; on appelle P le centre de cette homothétie.
- Pour tout sous-espace affine $\mathcal{V}$ de $\mathcal{A}$, et tout supplémentaire (vectoriel) W dans $E = \overrightarrow{\mathcal{A}}$ de la direction $\overrightarrow{\mathcal{V}}$, le projecteur $p_{\mathcal{V},W}$ sur $\mathcal{V}$ parallèle à W est application qui envoie $A \in \mathcal{A}$ vers l'unique point B avec $B \in \mathcal{V}$ et $\overrightarrow{AB} \in W$. C'est une application affine, et $\overrightarrow{p_{\mathcal{V},W}}$ est la projection vectorielle de E sur $\overrightarrow{\mathcal{V}}$ parallèle à W (c'est-à-dire la projection sur le premier facteur dans $E = \overrightarrow{\mathcal{V}} \oplus W$).

1.8.3. Proposition. *L'image $f(\mathcal{V})$ par une application affine $f : \mathcal{A} \to \mathcal{A}'$ d'un sous-espace affine $\mathcal{V}$ de $\mathcal{A}$ est un sous-espace affine de $\mathcal{A}'$ de direction $\overrightarrow{f}(\overrightarrow{\mathcal{V}})$. Si $\mathcal{V}'$ est un sous-espace affine de $\mathcal{A}'$ qui coupe l'image $f(\mathcal{A})$, alors $f^{-1}(\mathcal{V}') = \{\, a \in \mathcal{A} \mid f(A) \in \mathcal{V}' \,\}$ est un sous-espace affine de $\mathcal{A}$, de direction $\overrightarrow{f}^{-1}(\overrightarrow{\mathcal{V}'})$.*

Preuve. Si $P \in \mathcal{V}$ on a $f(\mathcal{V}) = f(P + \overrightarrow{\mathcal{V}}) = f(P) + \overrightarrow{f}(\overrightarrow{\mathcal{V}})$ qui est un sous-espace affine de $\mathcal{A}'$ de direction $\overrightarrow{f}(\overrightarrow{\mathcal{V}})$. Si $P' \in f(\mathcal{V}) \cap f(\mathcal{A})$ il existe $Q \in \mathcal{A}$ tel que $f(Q) = P'$, et on vérifie que $f^{-1}(\mathcal{V}') = Q + \overrightarrow{f}^{-1}(\overrightarrow{\mathcal{V}'})$. □

En particulier, si en $\mathcal{A} = \mathcal{A}'$ et f est une translation ou une homothétie, on a $f(\mathcal{V}) \parallel \mathcal{V}$ car $\overrightarrow{f(\mathcal{V})} = \mathrm{id}_E(\overrightarrow{\mathcal{V}}) = \overrightarrow{\mathcal{V}}$. Le fait $\overrightarrow{f(\mathcal{V})} = \overrightarrow{f}(\overrightarrow{\mathcal{V}})$ implique aussi $\dim(f(\mathcal{V})) \leq \dim(\mathcal{V})$.

1.8.4. Proposition. *Une application affine f est injective/surjective/bijective si et seulement si son application linéaire associée $\overrightarrow{f}$ est injective/surjective/bijective. Si $(A_0, \ldots, A_n)$ est un repère affine dans $\mathcal{A}$, ces conditions sont équivalentes à celles que l'image $(f(A_0), \cdots, f(A_n))$ de ce repère soit affinement libre/génératrice de $\mathcal{A}'$/un repère affine dans $\mathcal{A}'$.*

Preuve. La première partie est une conséquence directe de la relation $f(P = \vec{x}) = f(P) + \overrightarrow{f}(\vec{x})$. La seconde partie découle alors de la propriété correspondante pour les applications linéaires, car $\overrightarrow{f}$ envoie la base $(\overrightarrow{A_0A_1}, \ldots, \overrightarrow{A_0A_n})$ de $\overrightarrow{\mathcal{A}}$ sur la famille $(\overrightarrow{f(A_0)f(A_1)}, \ldots, \overrightarrow{f(A_0)f(A_n)})$, qui est libre/génératrice de $\overrightarrow{\mathcal{A}'}$/une base de $\overrightarrow{\mathcal{A}'}$ si $(f(A_0), \cdots, f(A_n))$ est affinement libre/génératrice de $\mathcal{A}'$/un repère affine. □

Dans le cas d'une application affine bijective, il existe aussi une application affine $g : \mathcal{A}' \to \mathcal{A}$ qui, pour chaque point du repère affine $(f(A_0), \cdots, f(A_n))$ dans $\mathcal{A}'$, envoie $f(A_i) \mapsto A_i$. Comme $g \circ f$ et $f \circ g$ sont des applications affines $\mathcal{A} \to \mathcal{A}$ respectivement $\mathcal{A}' \to \mathcal{A}'$ qui fixent un repère, elle sont l'identité, et g est la réciproque de f. En conclusion : *la réciproque d'une application affine bijective est affine.*

Les applications affines sont souvent utiles pour démontrer des théorèmes; en voici un exemple simple.

1.8.5. Théorème [dit «de Thalès»]. *Soit $\mathcal{D}, \mathcal{D}'$ deux droites qui se coupent en un point P, et $\mathcal{H}_1, \mathcal{H}_2$ deux hyperplans parallèles, ne contentant pas P, et dont la direction commune ne contient ni $\overrightarrow{\mathcal{D}}$ ni $\overrightarrow{\mathcal{D}'}$. Alors $\mathcal{D}$ et $\mathcal{D}'$ coupent chacun $\mathcal{H}_i$ en un point, qu'on appelle A_i respectivement B_i, pour $i = 1, 2$. Il existe $\lambda \in K^*$ tel que $\overrightarrow{PA_1} : \overrightarrow{PA_2} = 1 : \lambda = \overrightarrow{PB_1} : \overrightarrow{PB_2}$, et on a aussi $\overrightarrow{A_1B_1} : \overrightarrow{A_2B_2} = 1 : \lambda$. L'homothétie $h_{P,\lambda}$ vérifie $h_{P,\lambda}(\mathcal{H}_1) = \mathcal{H}_2$, $h_{P,\lambda}(A_1) = A_2$ et $h_{P,\lambda}(B_1) = B_2$.*

Preuve. D'après la proposition 1.2.10, les conditions impliquent que les intersections d'une des droites et l'un des hyperplans sont toujours formées d'un seul point. Or P, A_1, A_2 étant alignés, les vecteurs non nuls $\overrightarrow{PA_1}$ et $\overrightarrow{PA_2}$ sont proportionnels, et on peut définir $\lambda = \frac{\overrightarrow{PA_1}}{\overrightarrow{PA_2}}$. Alors l'homothétie $h_{P,\lambda}$ envoie A_1 sur A_1, et comme $h_{P,\lambda}(\mathcal{H}_1)$ est un hyperplan parallèle à $\mathcal{H}_1$ passant par A_2, c'est forcément $\mathcal{H}_2$. Comme $h_{P,\lambda}(\mathcal{D}') = \mathcal{D}'$, l'image par $h_{P,\lambda}$ du point B_1 d'intersection de D' et de $\mathcal{H}_1$ est le point B_1 d'intersection de D' et de $\mathcal{H}_1$; le reste en découle. □

Pour une application affine $f : \mathcal{A} \to \mathcal{A}$, on définit l'ensemble $\mathrm{Fix}(f) = \{ A \in \mathcal{A} \mid f(A) = A \}$ de ces points fixes. Cette définition peut s'appliquer aussi aux applications linéaires $\varphi : E \to E$, en considérant E comme espace affine; dans ce cas $\mathrm{Fix}(\varphi) = \ker(\varphi - \mathrm{id}_E)$ est l'espace propre de φ pour valeur propre 1 s'il y en a, ou $\{\vec{0}\}$ si 1 n'est pas valeur propre de φ. Contrairement au cas linéaire, $\mathrm{Fix}(f)$ peut très bien être vide pour une application affine f; c'est la cas notamment pour toutes les translations (sauf l'identité, qui est l'unique application affine avec $\mathrm{Fix}(f) = \mathcal{A}$). Mais dans le cas contraire on a:

1.8.6. Proposition. *Si $f : \mathcal{A} \to \mathcal{A}$ est une application affine et $\mathrm{Fix}(f)$ n'est pas vide, alors $\mathrm{Fix}(f)$ est un sous-espace affine dont la direction est $\mathrm{Fix}(\overrightarrow{f})$.*

Preuve. Soit $A \in \mathrm{Fix}(f)$. Alors pour $B \in \mathcal{A}$, la condition $B \in \mathrm{Fix}(f)$ équivaut à $\overrightarrow{f(A)f(B)} = \overrightarrow{AB}$, c'est-à-dire $\overrightarrow{f}(\overrightarrow{AB}) = \overrightarrow{AB}$, et $\overrightarrow{AB}$ est dans l'espace propre de $\overrightarrow{f}$ pour la valeur propre 1, d'où le résultat. □

Une conséquence de cette proposition est que si $\mathrm{Fix}(f)$ contient précisément un point, alors on a $\mathrm{Fix}(\overrightarrow{f}) = \{\vec{0}\}$ (c'est-à-dire, 1 n'est pas valeur propre de $\overrightarrow{f}$). La réciproque est aussi vraie.

1.8.7. Proposition. *Si $f : \mathcal{A} \to \mathcal{A}$ est une application affine (avec $\mathcal{A}$ un espace affine de dimension finie), telle que $\mathrm{Fix}(\overrightarrow{f}) = \{\vec{0}\}$, alors $\mathrm{Fix}(f)$ contient précisément un point.*

Preuve. Choisissons d'abord un point $A \in \mathcal{A}$ et posons $B = f(A)$. On veut montrer qu'il existe un vecteur $\vec{x}$ tel que $A + \vec{x} \in \mathrm{Fix}(f)$, en sorte que $\mathrm{Fix}(f)$ ne soit pas vide (après on saura appliquer la proposition précédente). La condition $A + \vec{x} = f(A + \vec{x})$ donne $A + \vec{x} = B + \overrightarrow{f}(\vec{x})$, c'est-à-dire $\vec{x} - \overrightarrow{f}(\vec{x}) = \overrightarrow{AB}$, ou encore $(\mathrm{id}_E - \overrightarrow{f})(\vec{x}) = \overrightarrow{AB}$. Or $\ker(\mathrm{id}_E - \overrightarrow{f}) = \mathrm{Fix}(\overrightarrow{f}) = \{\vec{0}\}$, donc $\mathrm{id}_E - \overrightarrow{f} : \overrightarrow{\mathcal{A}} \to \overrightarrow{\mathcal{A}}$ est injectif, et comme $\overrightarrow{\mathcal{A}}$ est de dimension finie, il est aussi surjectif (théorème du rang). Ceci montre qu'une solution pour $\vec{x}$ existe (et est unique, comme le confirmera la proposition précédente). □

On remarque que dans la liste d'exemples donnée ci-dessus, cette proposition s'applique aux applications constantes et aux homothéties de facteur $\lambda \neq 1$ (dont les symétries centrales), de façon assez triviale (car $\overrightarrow{f}$ n'a qu'une seule valeur propre dans ces cas, qui n'est pas 1). La conclusion de la proposition est d'ailleurs également évidente dans ces cas (la valeur constante de l'application, respectivement le centre de l'homothétie, étant l'unique point fixe); l'intérêt de la proposition est surtout pour les situations plus compliquées, où on peut néanmoins déterminer les valeurs propres de $\overrightarrow{f}$.

1.9. *Le groupe affine de $\mathcal{A}$.*

Le fait que la composée de deux applications affines est une application affine implique que l'ensemble des applications affines inversibles $\mathcal{A} \to \mathcal{A}$ forme un groupe (avec composition comme loi interne). Ce groupe est appelé le groupe affine de $\mathcal{A}$, et noté $\mathbf{GA}(\mathcal{A})$. La règle $\overrightarrow{g \circ f} = \overrightarrow{g} \circ \overrightarrow{f}$ montre que $f \mapsto \overrightarrow{f}$ définit un *homomorphisme* de groupes $\mathbf{GA}(\mathcal{A}) \to \mathbf{GL}(E)$ où $E = \overrightarrow{\mathcal{A}}$.

1.9.1. Proposition. *Le noyau de cet homomorphisme $\mathbf{GA}(\mathcal{A}) \to \mathbf{GL}(E)$ est l'ensemble $\{ t_{\vec{x}} \mid \vec{x} \in E \}$ des translations. Celui-ci forme donc un sous-groupe distingué de $\mathbf{GA}(\mathcal{A})$.*

Preuve. Les applications f dans le noyau sont celles avec $\overrightarrow{f} = \mathrm{id}_E$, c'est-à-dire tel que $\overrightarrow{f(A)f(B)} = \overrightarrow{AB}$ pour tout $A, B \in \mathcal{A}$, autrement dit $\overrightarrow{Af(A)} = \overrightarrow{Bf(B)}$. En appelant $\vec{x}$ ce vecteur $\overrightarrow{Af(A)}$ qui est indépendant du point A, on voit que $f = t_{\vec{x}}$, et on avait déjà constaté que réciproquement $\overrightarrow{t_{\vec{x}}} = \mathrm{id}_E$. □

Plus généralement, l'image réciproque dans $\mathbf{GA}(\mathcal{A})$ de tout sous-groupe distingué de $\mathbf{GL}(E)$ est un sous-groupe distingué de $\mathbf{GA}(\mathcal{A})$. En particulier, pour les sous-groupes $\{\mathrm{id}_E, -\mathrm{id}_E\}$ et $\{ \lambda\,\mathrm{id}_E \mid \lambda \in K^* \}$:

- L'ensemble des translations et des symétries centrales forme un sous-groupe distingué de $\mathbf{GA}(\mathcal{A})$;
- L'ensemble des translations et des homothéties de facteur $\lambda \neq 1$ forme un sous-groupe distingué de $\mathbf{GA}(\mathcal{A})$, appelé le sous-groupe des homothéties-translations.

Le fait qu'il s'agit de sous-groupes veut dire qu'en composant leur membres, on ne sort pas de l'ensemble. Pour les translations et symétries centrales on voit facilement que la composée de deux symétries centrales donne une translation ($S_Q \circ S_P = t_{2\overrightarrow{PQ}}$ où on rappelle que c'est le facteur à droite S_P

qui est appliqué en premier), et la composée d'une translation et une symétrie centrale donne une autre symétrie centrale. Donc dans ce cas les deux parties forment deux «moitiés» du sous-groupe.

Pour les homothéties-translations la situation est similaire, mais, les homothéties sont beaucoup plus «présentes» dans le group que les translations : le facteur de l'homothétie vectorielle $\overrightarrow{f}$ est multiplicatif, et f est une translation seulement lorsque ce facteur vaut 1. Pour toute autre valeur de ce facteur il s'agit d'une homothétie, dont on trouve le centre (unique) facilement quand le facteur λ et l'image d'un seul point P sont connus : c'est le point $P + \frac{1}{1-\lambda}\overrightarrow{Pf(P)}$. Voici un exemple où on exploite la structure de groupe et la multiplicativité du facteur, pour prouver une formulation alternative du théorème 1.6.2.

1.9.2. Théorème de Ménélaüs. *Soit A, B, C un triangle, et $a \in \mathcal{D}_{B,C} \setminus \{B, C\}$, $b \in \mathcal{D}_{C,A} \setminus \{C, A\}$, $c \in \mathcal{D}_{A,B} \setminus \{A, B\}$. Alors a, b, et c sont alignés si et seulement si*

$$\frac{\overrightarrow{aC}}{\overrightarrow{aB}} \times \frac{\overrightarrow{bA}}{\overrightarrow{bC}} \times \frac{\overrightarrow{cB}}{\overrightarrow{cA}} = 1.$$

Preuve. On considère la composition d'homothéties $f = h_{c,\nu} \circ h_{b,\mu} \circ h_{a,\lambda}$, où λ, μ et ν (dans cet ordre) sont les facteurs (tous $\neq 1$) du produit dans l'énoncé. Alors f est une homothétie-translation, et $\overrightarrow{f} = \lambda\mu\nu\,\mathrm{id}$. Or par construction $B \in \mathrm{Fix}(f)$: les homothéties successives envoient $B \mapsto C \mapsto A \mapsto B$; on a donc $f = h_{B,\lambda\mu\nu}$. Une homothétie $h_{P,\alpha}$ envoie chaque droite $\mathcal{D}$ vers une droite parallèle à $\mathcal{D}$, qui est égale à $\mathcal{D}$ si et seulement si, soit $\alpha = 1$ (l'homothétie est alors l'identité), soit $P \in \mathcal{D}$ (c'est clairement suffisant pour $h_{P,\alpha}(\mathcal{D}) = \mathcal{D}$; réciproquement si $h_{P,\alpha}(\mathcal{D}) = \mathcal{D}$ avec $\alpha \neq 1$, la restriction de $h_{P,\alpha}$ à $\mathcal{D}$ contient (d'après la proposition 1.8.7) un point fixe, qui ne peut être que le centre P). Ainsi les homothéties $h_{a,\lambda}$ et $h_{b,\mu}$ envoient la droite $\mathcal{D}_{a,b}$ sur elle-même; on obtient les équivalences $c \in \mathcal{D}_{a,b} \iff h_{c,\nu}(\mathcal{D}_{a,b}) = \mathcal{D}_{a,b} \iff f(\mathcal{D}_{a,b}) = \mathcal{D}_{a,b} \iff \lambda\mu\nu = 1$ (car $B \notin \mathcal{D}_{a,b}$). □

§2. Géométrie euclidienne.

On avance maintenant vers la géométrie euclidienne. Elle est basée sur la géométrie affine, c'est-à-dire un espace euclidien est un cas particulier d'un espace affine, muni d'une structure additionnelle. Pour introduire la structure euclidienne, il est nécessaire de se restreindre au cas où le corps de base K est égal au corps $\mathbf{R}$ des nombres réels, car cette structure ne peut pas être définie pour (par exemple) les corps $\mathbf{C}$, $\mathbf{Q}$, ou pour des corps finis. La structure additionnelle peut prendre la forme d'une notion de distance entre chaque paire de points, mais pour caractériser les propriétés que cette structure doit avoir, il sera plus commode d'introduire une structure de *produit scalaire* sur la direction de l'espace affine; la distance sera facile à exprimer en termes de ce produit. Une conséquence de cette approche est que pour les débuts de la géométrie euclidienne on va considérer juste les espace vectoriels, pour revenir aux espaces affines plus tard.

2.1. *Espaces vectoriels euclidiens.*

2.1.1. Définition. *Un produit scalaire dans un espace vectoriel E sur $\mathbf{R}$ est une application $E \times E \to \mathbf{R}$, noté $(\vec{x}, \vec{y}) \mapsto \langle \vec{x}, \vec{y} \rangle$, telle que les conditions suivantes sont vérifiées:*

(1) L'application est bilinéaire: pour tout vecteur $\vec{y}$ fixé, l'application $\vec{x} \mapsto \langle \vec{x}, \vec{y} \rangle$ est linéaire $E \to \mathbf{R}$, et pour tout vecteur $\vec{x}$ fixé, l'application $\vec{y} \mapsto \langle \vec{x}, \vec{y} \rangle$ est linéaire $E \to \mathbf{R}$.
(2) L'application est symétrique: pour tout $\vec{x}, \vec{y}$ on a $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$.
(3) L'application est non-dégénérée: si pour un certain $\vec{x}$ on a que $\vec{y} \mapsto \langle \vec{x}, \vec{y} \rangle$ est l'application nulle, alors $\vec{x} = \vec{0}$.
(4) L'application est positive: pour tout $\vec{x}$ on a $\langle \vec{x}, \vec{x} \rangle \geq 0$.

Un espace vectoriel euclidien est un espace vectoriel E sur $\mathbf{R}$ muni d'un produit scalaire.

Les deux premières conditions caractérisent les «formes bilinéaires symétriques», les deux autres conditions disent que cette forme est «définie positive»; elle peuvent être combinées en disant que $\vec{x}$ on a $\langle \vec{x}, \vec{x} \rangle \geq 0$ avec égalité uniquement si $\vec{x} = \vec{0}$. Il est clair que cela implique les deux dernières conditions (car $\langle \vec{x}, \vec{x} \rangle > 0$ dit en particulier que $\vec{y} \mapsto \langle \vec{x}, \vec{y} \rangle$ n'est pas l'application nulle), mais la réciproque est moins évidente, et sa démonstration illustre une technique de preuve dans ce domaine:

2.1.2. Proposition. *Si $\langle\cdot,\cdot\rangle$ est un produit scalaire, alors pour tout $\vec{x} \neq \vec{0}$ on a $\langle\vec{x},\vec{x}\rangle > 0$.*

Preuve. Par l'absurde: supposons $\vec{x} \neq \vec{0}$ avec $\langle\vec{x},\vec{x}\rangle = 0$. Alors d'après la propriété (3), il existe $\vec{y}$ avec $\langle\vec{x},\vec{y}\rangle \neq 0$. Pour $\lambda \in \mathbf{R}$ on calcule $\langle\lambda\vec{x}+\vec{y},\lambda\vec{x}+\vec{y}\rangle = 2\langle\vec{x},\vec{y}\rangle\lambda + \langle\vec{y},\vec{y}\rangle$ (en utilisant bilinéarité et symétrie; le terme $\langle\vec{x},\vec{x}\rangle\lambda^2$ est nul par hypothèse). D'un côté, cette valeur doit être positive pour tout λ (propriété (4)), de l'autre côté c'est une fonction polynomiale en λ de degré 1, qui s'annule donc pour une certaine valeur de λ, et prend des signes opposés de part et d'autre de cette valeur; contradiction. □

Cette proposition permet notamment de prouver que $\vec{x} = \vec{0}$ en montrant que $\langle\vec{x},\vec{x}\rangle = 0$; ainsi on aura « une égalité vectorielle pour le prix d'une égalité scalaire », ce qui normalement ne se produit pas (une équation vectorielle *linéaire* aura en général pour solution tout un hyperplan affine au lieu dún seul point). Ce qui distingue l'équation $\langle\vec{x},\vec{x}\rangle = 0$ (outre le fait qu'elle est quadratique) est qu'elle est au bord de l'impossible: en remplaçant 0 par un nombre négatif aussi proche qu'on veut, l'équation n'aura plus de solutions du tout.

Le fait que la propriété « définie positive » du produit scalaire peut être formulée comme « $\langle\vec{x},\vec{x}\rangle \geq 0$ avec égalité uniquement si $\vec{x} = \vec{0}$ » a une conséquence importante: la restriction d'un produit scalaire à une sous-espace vectoriel est encore un produit scalaire, et donc tout sous-espace vectoriel d'un espace vectoriel euclidien E est automatiquement un espace euclidien, dit sous-espace euclidien de E. Par contre, la propriété (3) *en isolation* ne se transfère pas toujours d'un espace à tous ses sous-espaces.

Le produit scalaire permet d'introduire des notions dérivées de lui. Notamment on dira que deux vecteurs $\vec{x},\vec{y}$ sont *orthogonaux*, noté $\vec{x} \perp \vec{y}$, si $\langle\vec{x},\vec{y}\rangle = 0$, et que $\sqrt{\langle\vec{x},\vec{x}\rangle}$ (qui est bien défini grâce à la propriété (4)) est la *norme* de $\vec{x}$, notée $\|\vec{x}\|$. Une propriété importante est l'inégalité de Cauchy-Schwarz:

$$|\langle\vec{x},\vec{y}\rangle| \leq \|\vec{x}\| \cdot \|\vec{y}\| \quad \text{pour tout } \vec{x},\vec{y}, \text{ avec égalité si et seulement si } \dim(\mathrm{Vect}(\vec{x},\vec{y})) \leq 1. \tag{11}$$

La dernière formule dit simplement que $\vec{x}$ et $\vec{y}$ sont linéairement dépendants. De nouveau on peut prouver cet énoncé en introduisant un paramètre λ. D'abord si $\vec{x} = \vec{0}$ les deux membres de l'inégalité sont nuls, et l'énoncé vrai; on suppose donc $\vec{x} \neq \vec{0}$. Alors $0 \leq \langle\lambda\vec{x}+\vec{y},\lambda\vec{x}+\vec{y}\rangle = \langle\vec{x},\vec{x}\rangle\lambda^2 + 2\langle\vec{x},\vec{y}\rangle\lambda + \langle\vec{y},\vec{y}\rangle$; le second membre est une fonction quadratique (car le coefficient $\langle\vec{x},\vec{x}\rangle$ est non nul) dont les valeurs sont toutes positives, et la valeur zéro est atteinte si et seulement si un λ avec $\lambda\vec{x} + \vec{y} = \vec{0}$ existe, c'est-à-dire si $\dim(\mathrm{Vect}(\vec{x},\vec{y})) \leq 1$. La théorie des fonction quadratiques nous dit que le discriminant $4\langle\vec{x},\vec{y}\rangle^2 - 4\langle\vec{x},\vec{x}\rangle\langle\vec{y},\vec{y}\rangle$ est négative, c'est-à-dire $\langle\vec{x},\vec{y}\rangle^2 \leq \langle\vec{x},\vec{x}\rangle\langle\vec{y},\vec{y}\rangle$, et égal à 0 si $\dim(\mathrm{Vect}(\vec{x},\vec{y})) \leq 1$. Les deux membres de $\langle\vec{x},\vec{y}\rangle^2 \leq \langle\vec{x},\vec{x}\rangle\langle\vec{y},\vec{y}\rangle$ sont positifs, donc on pourra appliquer la racine carrée aux deux membres, ce qui donne l'inégalité cherchée.

On en déduit une autre inégalité qui justifie le nom « norme ». Pour tout $\vec{x},\vec{y}$ on a $\|\vec{x}+\vec{y}\|^2 = \langle\vec{x}+\vec{y},\vec{x}+\vec{y}\rangle = \|\vec{x}\|^2 + 2\langle\vec{x},\vec{y}\rangle + \|\vec{y}\|^2 \leq \|\vec{x}\| + 2\|\vec{x}\|\,\|\vec{y}\| + \|\vec{y}\|^2 = (\|\vec{x}\| + \|\vec{y}\|)^2$, et en prenant encore des racines carrées des deux membres on obtient $\|\vec{x}+\vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$, l'inégalité triangulaire, qui est exigée pour toute « norme » en analyse.

Si dans ce dernier calcul on suppose simplement que $\langle\vec{x},\vec{y}\rangle = 0$, on obtient le théorème de Pythagore:

$$\|\vec{x}+\vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2 \quad \text{pour tout } \vec{x},\vec{y} \text{ avec } \vec{x} \perp \vec{y}.$$

L'orthogonalité a une importance pour les sous-espaces. On appelle deux sous-espaces F_1, F_2 orthogonaux, noté $F_1 \perp F_2$, si on a $\vec{x} \perp \vec{y}$ pour tout $\vec{x} \in F_1$ et $\vec{y} \in F_2$. Pour un sous-espace euclidien F de E on définit son *complément orthogonal* comme le sous-espace $F^\perp = \{\, \vec{x} \in E \mid \forall\vec{y} \in E : \vec{x} \perp \vec{y} \,\}$.

2.1.3. Proposition. *Pour tout sous-espace F d'un espace vectoriel euclidien E (de dimension finie) on a $E = F \oplus F^\perp$, ainsi que $F^{\perp\perp} = F$.*

Preuve. La proposition 2.1.2 implique que $F \cap F^\perp = \{\vec{0}\}$ (tout vecteur de $F \cap F^\perp$ est orthogonal à lui-même), donc F et $F^\perp$ sont en somme directe. Par linéarité, un vecteur $\vec{x}$ vérifiera $\vec{x} \in F^\perp$ dès qu'on a $\vec{x} \perp \vec{y}$ pour des vecteurs $\vec{y}$ parcourant une base de F; cela donne $\dim F$ équations linéaires décrivant $F^\perp$, d'où $\dim F^\perp \geq \dim E - \dim F$. Mais avec $\dim(F \oplus F^\perp) = \dim F + \dim F^\perp$, cela permet de conclure que $\dim F + \dim F^\perp = \dim E$, et donc $E = F \oplus F^\perp$. Par définition du complément orthogonal on a $F \subseteq F^{\perp\perp}$, et comme on vient de montrer que $\dim F = \dim E - \dim F^\perp = \dim F^{\perp\perp}$, on conclut que $F = F^{\perp\perp}$. □

Tout sous-espace d'un espace vectoriel euclidien admet donc un choix privilégié pour un sous-espace supplémentaire, à savoir $F^\perp$; cela fait une distinction importante avec la situation dans un espace vectoriel sans structure additionnelle, où des sous-espaces supplémentaires existent toujours, mais en général sans qu'il y ait un candidat distingué parmi eux. Les compléments orthogonaux permettent une méthode de preuve par récurrence sur la dimension, qu'on illustrera pour l'existence des base orthonormales.

2.1.4. Définition. *Une base $\left(\vec{b}_1, \ldots, \vec{b}_n\right)$ d'un espace vectoriel euclidien est dit orthogonal si $\vec{b}_i \perp \vec{b}_j$ dès que $i \neq j$. Il est dit orthonormale (ou orthonormée) si en plus $\|\vec{b}_i\| = 1$ pour tout i.*

Ainsi pour une base orthonormale on a $\langle \vec{b}_i, \vec{b}_j \rangle = \delta_{i,j}$, où $\delta_{i,j}$ (appelé symbole de Kronecker) désigne 0 si $i \neq j$, et 1 si $i = j$ (autrement dit, il donne le coefficient i, j d'une matrice identité). Cela permet de calculer le produit scalaire de deux vecteurs exprimés sur la même base orthonormale: si l'on a $\vec{x} = x_1\vec{b}_1 + \cdots + x_n\vec{b}_n$ et $\vec{y} = y_1\vec{b}_1 + \cdots + y_n\vec{b}_n$, on trouve par bilinéarité

$$\langle \vec{x}, \vec{y} \rangle = \sum_{i,j=1}^{n} x_i y_j \langle \vec{b}_i, \vec{b}_j \rangle = \sum_{i,j=1}^{n} x_i y_j \delta_{i,j} = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + \cdots + x_n y_n.$$

C'est précisément la définition du produit scalaire «canonique» sur $\mathbf{R}^n$, ce qui donne donc un exemple d'un espace vectoriel euclidien, avec sa base canonique comme base orthonormale. La proposition suivante montre que tous les espaces vectoriels euclidiens de dimension n sont de la même nature, car ils admettent une base similaire.

2.1.5. Proposition. *Tout espace vectoriel euclidien E de dimension finie a une base orthonormale.*

Preuve. Par récurrence sur la dimension de l'espace, avec pour un espace de dimension 0 la base vide. On montre d'abord l'existence d'une base orthogonale. Si $E = F \oplus F^\perp$ pour un sous-espace F quelconque, il est clair que pour des bases orthogonales de F et de $F^\perp$, leur réunion sera une base orthogonale de E. Il suffit donc de prendre un vecteur non nul $\vec{x}$, et $F = \langle \vec{x} \rangle$ la droite vectorielle qu'il engendre (qui a évidemment $\{\vec{x}\}$ pour base); alors $\dim F^\perp = \dim E - 1$, donc on peut supposer par récurrence que $F^\perp$ possède une base orthogonale, et en y rajoutant $\vec{x}$ on obtient une base orthogonale de E. Ensuite on transforme la base orthogonale en une base orthonormale en divisant chaque vecteur de base $\vec{b}$ par $\|\vec{b}\|$. □

2.1.6. Définition. *Si E_1, E_2 sont deux espaces vectoriels euclidiens, et $f : E_1 \to E_2$ une application linéaire, alors f est une isométrie si pour tout $\vec{x}, \vec{y} \in E_1$ on a $\langle \vec{x}, \vec{y} \rangle_{E_1} = \langle f(\vec{x}), f(\vec{y}) \rangle_{E_2}$. Si en plus f est inversible, on l'appelle un isomorphisme d'espaces vectoriels euclidiens, ou isomorphisme orthogonal.*

On remarque qu'une isométrie conserve les normes ($\|f(\vec{x})\| = \sqrt{\langle f(\vec{x}), f(\vec{x}) \rangle_{E_2}} = \sqrt{\langle \vec{x}, \vec{x} \rangle_{E_1}} = \|\vec{x}\|$), et est donc toujours injective: $f(\vec{x}) = \vec{0} \iff \|f(\vec{x})\| = 0 \iff \|\vec{x}\| = 0 \iff \vec{x} = \vec{0}$. Un isomorphisme d'espaces vectoriels euclidiens est donc la même chose qu'une isométrie surjective. Réciproquement on a:

2.1.7. Proposition. *Une application linéaire $f : E_1 \to E_2$ qui conserve les normes est une isométrie.*

Preuve. Pour tout $\vec{x}, \vec{y} \in E_1$ on a par linéarité $\|f(\vec{x}) - f(\vec{y})\|_{E_2} = \|f(\vec{x} - \vec{y})\|_{E_2} = \|\vec{x} - \vec{y}\|_{E_1}$; en identifiant les carrés des deux membres on obtient $\|f(\vec{x})\|_{E_2}^2 + 2\langle f(\vec{x}), f(\vec{y}) \rangle_{E_2} + \|f(\vec{y})\|_{E_2}^2 = \|\vec{x}\|_{E_1}^2 + 2\langle \vec{x}, \vec{y} \rangle_{E_1} + \|\vec{y}\|_{E_1}^2$, et donc, en utilisant le fait que f conserve les normes, $\langle f(\vec{x}), f(\vec{y}) \rangle_{E_2} = \langle \vec{x}, \vec{y} \rangle_{E_1}$, comme voulu. □

2.1.8. Proposition. *Si $\left(\vec{b}_1, \ldots, \vec{b}_n\right)$ est une base orthonormale de E_1, alors une application linéaire $f : E_1 \to E_2$ est une isométrie si et seulement si pour tout i, j on a $\langle f(\vec{b}_i), f(\vec{b}_j) \rangle_{E_2} = \delta_{i,j}$, et f est un isomorphisme orthogonal si et seulement si $\left(f(\vec{b}_1), \ldots, f(\vec{b}_n)\right)$ est une base orthonormale de E_2.*

Preuve. La condition $\langle f(\vec{b}_i), f(\vec{b}_j) \rangle_{E_2} = \delta_{i,j}$ est juste la définition d'une isométrie appliquée au éléments de la base orthonormale de E_1; elle est donc certainement nécessaire. Pour voir qu'elle est aussi suffisante, on peut reprendre $\vec{x} = x_1\vec{b}_1 + \cdots + x_n\vec{b}_n$ et $\vec{y} = y_1\vec{b}_1 + \cdots + y_n\vec{b}_n$, et calculer, comme ci-dessus

$$\langle f(\vec{x}), f(\vec{y}) \rangle_{E_2} = \sum_{i,j=1}^{n} x_i y_j \langle f(\vec{b}_i), f(\vec{b}_j) \rangle_{E_2} = \sum_{i,j=1}^{n} x_i y_j \delta_{i,j} = \sum_{i=1}^{n} x_i y_i = \langle \vec{x}, \vec{y} \rangle_{E_1}.$$

Pour que f soit un isomorphisme, il faut en plus que l'image de la base de E_1 soit une *base* de E_2, c'est-à-dire que les conditions pour que $\left(f(\vec{b}_1), \ldots, f(\vec{b}_n)\right)$ soit une base orthonormale de E_2 soient réunies. □

En résumant de qui précède, tout espace vectoriel euclidien E de dimension n possède (au moins) une base orthonormale, ce qui permet d'établir (au moins) un isomorphisme avec $\mathbf{R}^n$ avec son produit scalaire canonique, en envoyant cette base de E de façon bijective vers la base canonique de $\mathbf{R}^n$.

2.2. *Espaces affines euclidiens.*

Un espace affine euclidien $\mathcal{A}$ est un espace affine sur $\mathbf{R}$ dont la direction $E = \overrightarrow{\mathcal{A}}$ est muni d'une structure d'espace vectoriel euclidien (c'est-à-dire d'un produit scalaire). D'après la propriété correspondante pour les espaces vectoriels, tout sous-espace affine d'un espace affine euclidien est (par restriction) un espace affine euclidien. On commence avec des notions "affines" qui dépendent du fait que $K = \mathbf{R}$ est ordonné.

2.2.1. Définition. *Pour $A, B \in \mathcal{A}$, l'intervalle $[A, B]$ est $\{\, A + \lambda\overrightarrow{AB} \mid 0 \leq \lambda \leq 1 \,\}$; c'est une partie de $\mathcal{D}_{A,B}$ si $A \neq B$. Une partie $\mathcal{P}$ de $\mathcal{A}$ est convexe si pour tout $A, B \in \mathcal{P}$ on a $[A, B] \subseteq \mathcal{P}$. Pour une partie quelconque $\mathcal{P}$, l'ensemble des barycentres de collection de points pondérés de $\mathcal{P}$ dans lesquels tous les poids sont positifs est l'enveloppe convexe de $\mathcal{P}$.*

En utilisant l'associativité des barycentres (proposition 1.3.3) et le fait que l'intervalle réel $[0, 1]$ est fermé pour la multiplication, on montre facilement qu'une enveloppe convexe est effectivement convexe, et que toute partie convexe de $\mathcal{A}$ contenant $\mathcal{P}$ contient tout point de son enveloppe convexe; elle est donc la plus petite partie convexe de $\mathcal{A}$ contenant $\mathcal{P}$.

2.2.2. Proposition. *Une partie convexe d'une espace affine euclidien est connexe. L'image par une application affine, ou l'image réciproque par une application affine, d'une partie convexe est convexe.*

Preuve. Un segment est un arc, dont convexe implique connexe par arcs et donc connexe. Pour tout $f : \mathcal{A} \to \mathcal{A}'$ affine et $A, B \in \mathcal{A}$ on a $f([A, B]) = [f(A), f(B)]$; le reste de l'énoncé en découle directement. □

2.2.3. Proposition/Définition [régionnement de l'espace par un hyperplan]. *Si $\mathcal{H}$ est un hyperplan d'un espace affine euclidien $\mathcal{A}$ et $P \in \mathcal{A} \setminus \mathcal{H}$, alors les deux ensembles des $A \in \mathcal{A} \setminus \mathcal{H}$ pour lesquels $[P, A] \cap \mathcal{H}$ est non-vide respectivement vide forment une partition de $\mathcal{A} \setminus \mathcal{H}$, qui ne dépend pas du choix de P. La partition de $\mathcal{A}$ en ces deux parties et $\mathcal{H}$ lui-même est appelé le régionnement de $\mathcal{A}$ défini par $\mathcal{H}$. Les trois parties sont convexes, les deux parties autres que $\mathcal{H}$ sont les demi-espaces délimités par $\mathcal{H}$.*

Preuve. Si on choisit un repère affine de $\mathcal{H}$, et on le complète par P pour former un repère affine de $\mathcal{A}$, alors la coordonnée barycentrique correspondant à P détermine une application affine $f : \mathcal{A} \to \mathbf{R}$ qui s'annule précisément en $\mathcal{H}$, et qui envoie P vers 1. Pour $A \in \mathcal{A} \setminus \mathcal{H}$ on a $[P, A] \cap \mathcal{H} = \emptyset$ si et seulement si $0 \notin [1, f(A)]$, c'est-à-dire si $f(A) > 0$. Les deux parties indiquées de $\mathcal{A} \setminus \mathcal{H}$ sont donc $f^{-1}(\mathbf{R}_{>0})$ et $f^{-1}(\mathbf{R}_{<0})$, en particulier elles sont convexes (et $\mathcal{H}$ l'est évidemment aussi). Deux points $A, B \in \mathcal{A} \setminus \mathcal{H}$ appartiennent à la même des deux parties si et seulement si $f(A)$ et $f(B)$ sont de même signe, donc si $0 \notin [f(A), f(B)]$ ou encore $[A, B] \cap \mathcal{H} = \emptyset$; cette dernière description ne mentionne pas P ou f, et est donc indépendant du choix de P (bien que la façon utilisée pour distinguer les deux parties en dépend). □

Le produit scalaire sur $\overrightarrow{\mathcal{A}}$ permet de définir plusieurs nouvelles notions au niveau de l'espace affine.

2.2.4. Définition. *La distance entre deux points A, B d'un espace affine euclidien est $d(A, B) = \|\overrightarrow{AB}\|$.*

2.2.5. Proposition [inégalité triangulaire]. *Pour tout $A, B, C \in \mathcal{A}$ on a $d(A, C) \leq d(A, B) + d(B, C)$.*

Preuve. C'est une traduction directe de l'inégalité triangulaire $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$ pour les vecteurs. □

2.2.6. Définition. *Une application affine $f : \mathcal{A} \to \mathcal{A}'$ entre espaces affines euclidiens est appelée une isométrie si elle conserve la distance: $d(f(A), f(B)) = d(A, B)$ pour tout $A, B \in \mathcal{A}$. Si en plus elle est bijective, on l'appelle un isomorphisme d'espaces affine euclidiens.*

On peut montrer qu'une application entre espaces affines euclidiens dont on exige uniquement qu'elle conserve la distance est automatiquement affine, et donc une isométrie. On pourrait donc simplifier la définition ci-dessous. En appliquant la proposition 2.1.7 et la définition de distance, on voit qu'une application affine f est une isométrie si et seulement si c'est le cas de l'application linéaire associée $\overrightarrow{f}$.

Il est évident que la composée de deux isométries est aussi une isométrie. Par conséquent, l'ensemble des automorphismes de l'espace affine euclidien $\mathcal{A}$ forme un sous-groupe, qu'on notera $\mathbf{Isom}(\mathcal{A})$, du groupe $\mathbf{GA}(\mathcal{A})$ des automorphismes de $\mathcal{A}$ en tant qu'espace affine. (Il ne s'agit pas d'un sous-groupe distingué, car la définition de $\mathbf{Isom}(\mathcal{A})$ dépend du choix d'un produit scalaire dans $\mathcal{A}$, qui n'est pas unique.) Les applications linéaires associées à ces automorphismes sont les automorphismes orthogonaux de $E = \overrightarrow{\mathcal{A}}$, et forment le *groupe orthogonal* $\mathbf{O}(E)$, qui est un sous-groupe du groupe $\mathbf{GL}(E)$.

2.2.7. Définition. *Deux sous-espaces affines $\mathcal{V}, \mathcal{W}$ sont dits orthogonaux, noté $\mathcal{V} \perp \mathcal{W}$, si $\overrightarrow{\mathcal{V}} \perp \overrightarrow{\mathcal{W}}$. Si en plus $\overrightarrow{\mathcal{W}} = \overrightarrow{\mathcal{V}}^{\perp}$, on appelle $\mathcal{W}$ un complément orthogonal de $\mathcal{V}$. La projection affine sur $\mathcal{V}$ parallèle à $\overrightarrow{\mathcal{V}}^{\perp}$ est appelée la projection orthogonale sur $\mathcal{V}$, et notée $p_{\mathcal{V}}$.*

Il est à noter que contrairement à la situation pour les sous-espaces vectoriels, un supplément orthogonal d'un sous-espace affine n'est pas unique: tout sous-espace parallèle à un complément orthogonal de $\mathcal{V}$ est aussi un complément orthogonal de $\mathcal{V}$. Par contre la projection orthogonale sur $\mathcal{V}$ est unique, car elle est définie en termes de $\mathcal{V}$ et du complément orthogonale $\overrightarrow{\mathcal{V}}^{\perp}$ de *sa direction*. Une projection (orthogonale) n'est en général pas injective, et ce n'est donc pas une isométrie (elle réduit la distance de certaines paires de points distincts à 0). Mais en "prolongeant" l'image de chaque point au delà de sa projection par un facteur 2, on obtiendra une isométrie, et c'est une classe importante d'isométries qui est ainsi obtenue.

2.2.8. Définition. *Pour un sous-espace affine $\mathcal{V}$ de $\mathcal{A}$, la symétrie orthogonale $s_{\mathcal{V}} : \mathcal{A} \to \mathcal{A}$ par rapport à $\mathcal{V}$ est définie par $s_{\mathcal{V}}(A) = A + 2\overrightarrow{Ap_{\mathcal{V}}(A)}$, de sorte qu'on ait $p_{\mathcal{V}}(A) = \operatorname{bar}(A, s_{\mathcal{V}}(A))$ pour tout $A \in \mathcal{A}$.*

2.2.9. Proposition. *Toute symétrie orthogonale par rapport à un sous-espace affine est une isométrie.*

Preuve. On voit facilement que $s_{\mathcal{V}}$ est une application affine dont l'application linéaire associée est donnée par $\overrightarrow{s_{\mathcal{V}}}(\vec{x}) = 2p_{\overrightarrow{\mathcal{V}}}(\vec{x}) - \vec{x}$, où $p_{\overrightarrow{\mathcal{V}}}$ est la projection linéaire sur $\overrightarrow{\mathcal{V}}$ parallèle à $\overrightarrow{\mathcal{V}}^{\perp}$. En effet, il suffit de montrer que $s_{\mathcal{V}}(P + \vec{x}) = s_{\mathcal{V}}(P) + 2p_{\overrightarrow{\mathcal{V}}}(\vec{x}) - \vec{x}$ pour tout vecteur $\vec{x}$ et un point particulier P, qu'on peut choisir dans $\mathcal{V}$ de sorte que $s_{\mathcal{V}}(P) = P$. Avec $A = P + \vec{x}$ on a alors $s_{\mathcal{V}}(A) = A + 2\overrightarrow{Ap_{\mathcal{V}}(A)} = P + \vec{x} + 2(-\vec{x} + p_{\overrightarrow{\mathcal{V}}}(\vec{x})) = P - \vec{x} + 2p_{\overrightarrow{\mathcal{V}}}(\vec{x})$, ce qui donne le résultat voulu. Pour montrer que ce $\overrightarrow{s_{\mathcal{V}}}$ est une isométrie vectorielle, on choisit une base orthonormale dans $\overrightarrow{\mathcal{V}}$, et une autre dans $\overrightarrow{\mathcal{V}}^{\perp}$, dont la réunion sera un base orthonormale de E; alors $\overrightarrow{s_{\mathcal{V}}}$ fixe les éléments des la première base, et envoie les éléments de la seconde base vers le vecteur opposé. L'ensemble des images forme donc une base orthonormale de E, et d'après la proposition 2.1.8 $\overrightarrow{s_{\mathcal{V}}}$ on peut conclure. □

Les symétries orthogonales forment une classe importante d'isométries $\mathcal{A} \to \mathcal{A}$, mais elles ne forment pas un sous-groupe du groupe $\mathbf{Isom}(\mathcal{A})$; au contraire, leur importance est surtout due au fait qu'elles *engendrent* $\mathbf{Isom}(\mathcal{A})$ (c'est-à-dire qu'il n'y a pas de sous-groupe propre de $\mathbf{Isom}(\mathcal{A})$ qui contienne tous les symétries orthogonales). On prouvera cela plus tard, et on n'aura même pas besoin de toutes les symétries orthogonales pour engendrer $\mathbf{Isom}(\mathcal{A})$: il suffit de considérer les symétries orthogonales par rapport à des hyperplans $\mathcal{A}$, quelles symétries on appellera des *réflexions*.

2.3. *Sphères et hyperplans médiateurs.*

2.3.1. Définition. *Soit $\Omega \in \mathcal{A}$ et $r \in \mathbf{R}_{\geq 0}$. La sphère de centre Ω et rayon r, notée $\mathcal{S}(\Omega, r)$, et l'ensemble $\{ A \in \mathcal{A} \mid d(A, \Omega) = r \}$. Dans le cas où $\dim \mathcal{A} = 2$, on l'appelle plutôt cercle que sphère.*

2.3.2. Proposition/Définition. *Soit $\Gamma = \mathcal{S}(\Omega, r)$ et $\mathcal{V}$ une sous-espace affine, $P = p_{\mathcal{V}}(\Omega)$ et $d = d(\Omega, P)$. Si $d > r$ on a $\Gamma \cap \mathcal{V} = \emptyset$, sinon $\Gamma \cap \mathcal{V} = \mathcal{S}(P, \sqrt{r^2 - d^2}) \cap \mathcal{V}$. Dans le cas $r = d$ on a $\Gamma \cap \mathcal{V} = \{P\}$ et $\mathcal{V}$ est dit tangent à Γ en P. Pour tout $A \in \Gamma$, il y a un unique hyperplan $A + \langle \overrightarrow{\Omega A} \rangle^{\perp}$ tangent à Γ en A.*

Preuve. La formule pour $\Gamma \cap \mathcal{V}$ découle du théorème de Pythagore. Pour $\mathcal{V}$ tangent en A on a $\mathcal{V} \perp \overrightarrow{\Omega A}$. □

2.3.3. Proposition. *Soit $\mathcal{H}$ un hyperplan d'un espace affine euclidien $\mathcal{A}$, et A un point de $\mathcal{A}$ non situé dans $\mathcal{H}$. Avec $B = s_{\mathcal{H}}(A)$, le lieu des points à distance égale de A et de B est $\mathcal{H}$:*

$$\mathcal{H} = \{\, P \in \mathcal{A} \mid d(P,A) = d(P,B) \,\}.$$

Preuve. Soit $M = p_{\mathcal{H}}(A)$; alors $M \in \mathcal{H}$, et comme par définition $B = A + 2\overrightarrow{AM}$, on a $M = \mathrm{bar}(A,B)$. Posons $\vec{y} = \overrightarrow{AM} = \overrightarrow{MB}$, et pour $P \in \mathcal{A}$ posons $\vec{x} = \overrightarrow{MP}$. Alors $\overrightarrow{AP} = \vec{x} + \vec{y}$ et $\overrightarrow{BP} = \vec{x} - \vec{y}$, et on a $d(P,A)^2 - d(P,B)^2 = \langle \vec{x}+\vec{y}, \vec{x}+\vec{y}\rangle - \langle \vec{x}-\vec{y}, \vec{x}-\vec{y}\rangle = 4\langle \vec{x}, \vec{y}\rangle$. Comme $d(P,A)$ et $d(P,B)$ sont des nombres positifs, on aura $d(P,A) = d(P,B)$ si et seulement si cette différence de leurs carrés est nulle, c'est-à-dire si $\vec{x} \perp \vec{y}$. Par hypothèse $A \notin \mathcal{H}$ et donc $\vec{y} \neq \vec{0}$; alors $\langle \vec{y}\rangle$ est une droite vectorielle orthogonale à $\overrightarrow{\mathcal{H}}$, et comme ce dernier est un hyperplan vectoriel, $\langle \vec{y}\rangle$ est a en effet le complément orthogonal de $\overrightarrow{\mathcal{H}}$. Ainsi $\overrightarrow{\mathcal{H}} = \langle \vec{y}\rangle^{\perp}$ et on a vu que $\vec{x}$ est dans cet espace si et seulement si $d(P,A) = d(P,B)$; en traduisant par la définition $\vec{x} = \overrightarrow{PM}$ cette condition devient $P \in M + \overrightarrow{\mathcal{H}} = \mathcal{H}$. □

Si on commence avec deux points distincts A, B donnés, on peut toujours trouver un hyperplan $\mathcal{H}$ tel qu'on soit dans la situation de la proposition: il suffit de prendre $\mathcal{H} = M + \langle \overrightarrow{AM}\rangle^{\perp}$ avec $M = \mathrm{bar}(A,B)$ (la condition $A \neq B$ garantit que $\overrightarrow{AM} \neq \vec{0}$, et donc que l'expression donnée décrit bien un hyperplan). Comme la proposition caractérise $\mathcal{H}$, c'est aussi le seul hyperplan qui convient. Ainsi on peut formuler:

2.3.4. Définition/Proposition. *Soit $A, B \in \mathcal{A}$ deux points distincts. Alors l'hyperplan médiateur $\mathcal{H}$ de A et B est donné par $\mathcal{H} = \{\, P \in \mathcal{A} \mid d(P,A) = d(P,B) \,\}$. C'est l'unique hyperplan tel que $s_{\mathcal{H}}(A) = B$.*

On a vu dans le calcul ci-dessus que l'expression $d(P,A)^2 - d(P,B)^2$ se simplifiait à une expression linéaire (en le vecteur $\overrightarrow{MP}$). On peut former des expressions similaires dans lesquelles se produisent d'autres formes d'annulation, et qui ont pour résultat une description simple de certains lieux géométriques. On considère toujours la situation où deux points A, B sont fixés, et P est un point variable; comme ci-dessus on pose $M = \mathrm{bar}(A,B)$ ainsi que $\vec{x} = \overrightarrow{MP}$ (variable) et $\vec{y} = \overrightarrow{AM}$ (fixe), et on pose également $r = \|\vec{y}\|$. Si on calcule la somme $d(P,A)^2 + d(P,B)^2$ au lieu de la différence, on obtient $\|\vec{x}+\vec{y}\|^2 + \|\vec{x}-\vec{y}\|^2 = 2\|\vec{x}\|^2 + 2\|\vec{y}\|^2 = 2(\|\vec{x}\|^2 + r^2)$. Cette quantité est minimale ($2r^2$) quand $\vec{x} = \vec{0}$ c'est-à-dire $P = M$, et le lieu où elle prend une valeur constante $2c \geq 2r^2$ est une sphère de centre M:

$$\{\, p \in \mathcal{A} \mid d(P,A)^2 + d(P,B)^2 = 2c \,\} = \mathcal{S}\left((\mathrm{bar}(A,B), \sqrt{c - r^2}\right). \tag{12}$$

Ensuite on a le produit remarquable $\langle \overrightarrow{AP}, \overrightarrow{BP}\rangle = \langle \vec{x}+\vec{y}, \vec{x}-\vec{y}\rangle = \|\vec{x}\|^2 - \|\vec{y}\|^2 = \|\vec{x}\|^2 - r^2$, quelle quantité est aussi minimale ($-r^2$) quand $P = M$, et le lieu où elle prend une valeur constante $c \geq -r^2$

$$\{\, p \in \mathcal{A} \mid \langle \overrightarrow{AP}, \overrightarrow{BP}\rangle = c \,\} = \mathcal{S}\left(\mathrm{bar}(A,B), \sqrt{c + r^2}\right) \tag{13}$$

est encore une sphère de centre M. En particulier on a pour $c = 0$:

$$\{\, p \in \mathcal{A} \mid \overrightarrow{AP} \perp \overrightarrow{BP} \,\} = \mathcal{S}\left(\mathrm{bar}(A,B), r\,\right), \tag{14}$$

le cercle (unique) dont le segment $[A,B]$ est une diamètre.

Terminons cette section avec un pendant de la dernière partie de la proposition 2.3.4, c'est-à-dire le fait que pour deux points distincts il existe une réflexion unique qui envoie l'un vers l'autre, pour les réflexions vectorielles (symétries par rapport à un hyperplan vectoriel de E). Une réflexion *vectorielle* préserve les normes, donc dans l'énoncé suivant on doit supposer les normes des vecteurs égales.

2.3.5. Corollaire. *Soit E un espace vectoriel euclidien, et $\vec{x}, \vec{y} \in E$ deux vecteurs distincts tels que $\|\vec{x}\| = \|\vec{y}\|$. Alors il existe une réflexion vectorielle unique qui envoie $\vec{x}$ vers $\vec{y}$ (elle envoie aussi $\vec{y}$ vers $\vec{x}$).*

Preuve. On peut considérer E comme espace affine euclidien; dans cette identification les réflexions vectorielles sont celles parmi les réflexions (affines) qui fixent le vecteur $\vec{0}$. La proposition 2.3.4 fournit une réflexion unique qui envoie $\vec{x}$ vers $\vec{y}$ (et aussi $\vec{y}$ vers $\vec{x}$), et elle fixe $\vec{0}$, car $\|\vec{x}\| = \|\vec{y}\|$ veut dire en termes affines que $d(\vec{x}, \vec{0}) = d(\vec{y}, \vec{0})$, c'est-à-dire que $\vec{0}$ se trouve sur l'hyperplan médiateur de $\vec{x}$ et $\vec{y}$. □

2.4. *Le groupe orthogonal en dimension* 2*, et les angles orientés.*

La notion d'angle est liée à celle des rotations. Le cas le plus simple se pose dans $\mathbf{R}^2$, considéré comme espace vectoriel euclidien avec le produit scalaire canonique $\langle (x_1, y_1), (x_2, y_2)\rangle = x_1x_2 + y_1y_2$, où la rotation d'angle $\varphi \in \mathbf{R}$ est par définition l'application linéaire $\mathbf{R}^2 \to \mathbf{R}^2$ donnée par matrice orthogonale

$$R_\varphi = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}. \tag{15}$$

On a $R_\varphi \cdot R_\psi = R_{\varphi+\psi}$ pour tout φ, ψ, et on a $R_\varphi = R_\psi$ si et seulement si $\varphi \equiv \psi \pmod{2\pi}$. En fait pour deux vecteurs $\vec{x}, \vec{y}$ unitaires (c'est-à-dire de norme 1) il existe φ, unique à un multiple entier de 2π près, tel que $R_\varphi(\vec{x}) = \vec{y}$; on peut définir l'angle défini par les vecteurs $\vec{x}$ et $\vec{y}$ (dans cet ordre) comme la valeur φ (ou plus précisément sa classe de dans $\mathbf{R}/2\pi\mathbf{Z}$). Mais pour transférer cette notion d'angle à un espace euclidien abstrait nécessite de la définir indépendamment d'un choix particulier d'une base (orthonormale). On peut se servir du produit scalaire pour obtenir un renseignement important sur φ, car pour tout vecteur $\vec{x}$ on a $\langle \vec{x}, R_\varphi(\vec{x})\rangle = \cos\varphi \|\vec{x}\|^2$; ainsi si pour deux vecteurs non nuls $\vec{x}, \vec{y}$ on pose

$$\alpha = \operatorname{Arccos}\left(\frac{\langle \vec{x}, \vec{y}\rangle}{\|\vec{x}\|\|\vec{y}\|}\right), \tag{16}$$

on a un bon candidat pour l'angle φ entre les vecteurs normalisés $\frac{\vec{x}}{\|\vec{x}\|}$ et $\frac{\vec{y}}{\|\vec{y}\|}$ (l'expression donnée est conçue pour dépendre uniquement de ces vecteurs normalisés). En fait, grâce à l'inégalité (11) de Cauchy-Schwarz on sait que l'argument de la fonction Arccos est dans l'intervalle $[-1, 1]$ (ceci est vrai même dans les espaces de dimension > 2), d'où α est un nombre réel bien défini dans l'intervalle $[0, \pi]$, et on l'appelle l'*angle géométrique* entre $\vec{x}$ et $\vec{y}$. Le problème est que l'angle φ cherché peut être aussi bien α que $-\alpha$, et le seul produit scalaire $\langle \vec{x}, \vec{y}\rangle$ ne permet pas de distinguer les deux cas (seulement pour $\alpha = 0$ ou $\alpha = \pi$, c'est-à-dire quand $\vec{x}$ et $\vec{y}$ sont proportionnels, les possibilités α et $-\alpha$ sont équivalentes). Ceci est lié au fait que si l'on intervertit $\vec{x}$ et $\vec{y}$, l'angle φ est remplacé par $-\varphi$ (d'où on appelle φ un angle orienté); par contre l'expression pour α est symétrique en $\vec{x}$ et $\vec{y}$. Ce problème n'est pas dû à une maladresse de notre part, car le corollaire 2.3.5 montre que l'échange de $\vec{x}$ et $\vec{y}$ peut être effectué par une réflexion vectorielle, qui est un isomorphisme d'espaces vectoriels euclidiens, et conserve donc *toute* la structure considérée.

Par conséquent, il est impossible de définir, dans un espace vectoriel euclidien abstrait, une notion d'angle orienté à valeurs dans $\mathbf{R}/2\pi\mathbf{Z}$. Par contre, on saura définir une notion d'angle orienté dont les valeurs sont exprimés en termes de l'espace E lui-même (et notamment de son groupe d'automorphismes); l'effet d'une réflexion de E sur ce type de valeurs peut être non trivial. Si on veut néanmoins associer aux angles orientés une valeur numérique (dans $\mathbf{R}/2\pi\mathbf{Z}$; on parlera d'une *mesure d'angle*), il sera nécessaire d'ajouter à la structure euclidienne un élément qui *ne soit pas* conservé par les réflexions vectorielles. On pourrait ajouter le choix d'une base orthonormée, mais cela détruirait toute la symétrie: aucune isométrie non triviale peut conserver un choix de base. Ce qu'on veut est d'ajouter un petit élément à la structure, tel que un assez grand nombre d'isométries le respectent, mais pas les réflexions; cet élément sera appelé une *orientation* de l'espace.

Même avec une orientation, on ne saura définir la mesure d'angles orientés que si l'espace est de dimension 2. En fait, on peut définir l'orientation en toute dimension, mais pour fixer un choix pour l'angle orienté entre deux vecteurs non proportionnels $\vec{x}, \vec{y}$, il faut une orientation du *plan* vectoriel engendré par ces vecteurs, et la notion d'orientation ne permettra pas de déduire une orientation pour un tel sous-espace propre à partir d'une orientation donnée de espace ambiant.

Il peut sembler que ce qui suit représente un bien grand effort pour introduire une notion (d'angle orienté) bien limitée. Mais la géométrie euclidienne plane est un sujet très riche, et la notion d'angle orienté figure de façon proéminente dans beaucoup de ces énoncés et démonstrations.

Pour introduire l'orientation, on commence par observer que le déterminant d'un isomorphisme d'espaces vectoriels euclidiens (de dimension finie n) est toujours soit 1 soit -1. En effet, l'isomorphisme s'exprime dans une base orthonormée par une matrice M, dont les colonnes, qui décrivent les images par l'isomorphisme des éléments de cette base, forment une base orthonormée de $\mathbf{R}^n$ pour le produit scalaire canonique. On appelle M une *matrice orthogonale*, dont la condition s'écrit $M^\top M = \mathrm{id}_n$ où $M^\top$ désigne la transposée de M. On prend le déterminant dans cette identité, et comme la transposition ne change pas le déterminant, on obtient $(\det M)^2 = 1$ et donc $\det M \in \{+1, -1\}$. Ainsi on peut formuler:

2.4.1. Définition/Proposition. *Dans le groupe* $\mathbf{O}(E)$ *on distingue les deux parties:*

$\mathbf{O}^+(E) = \{ f \in \mathbf{O}(E) \mid \det f = +1 \}$ *dont les éléments les automorphismes orthogonaux directs,*

$\mathbf{O}^-(E) = \{ f \in \mathbf{O}(E) \mid \det f = -1 \}$ *dont les éléments les automorphismes orthogonaux indirects.*

Le sous-ensemble $\mathbf{O}^+(E)$ *forme un sous-groupe, appelé groupe spécial orthogonal de* E*, aussi noté* $\mathbf{SO}(E)$*.*

Si $(e_1, \ldots, e_n)$ est une base orthonormée de E, la proposition 2.1.8 nous dit que pour tout $f \in \mathbf{O}(E)$ on obtient une autre base orthonormée $(f(e_1), \ldots, f(e_n))$, et que la correspondance est surjective (toutes les bases orthonormées sont ainsi obtenues). Elle est d'ailleurs aussi injective (car une application linéaire est déterminée par ses valeurs sur une base), donc les bases orthonormées de E sont en bijection avec $\mathbf{O}(E)$. Par conséquent la partition de $\mathbf{O}(E)$ qu'on vient de décrire induit une partition de l'ensemble des bases orthonormées. La correspondance entre les parties $\mathbf{O}(E)$ et celles de l'ensemble de bases dépend du choix de la base $(e_1, \ldots, e_n)$ (car celle-ci appartient toujours à la partie correspondant à $\mathbf{O}^+(E)$), mais la *partition* de l'ensemble de bases est toujours la même (c'est un cas particulier de la partition d'un ensemble en *orbites* pour l'action d'un groupe, ici $\mathbf{O}^+(E)$). On remarque la ressemblance à la situation de la proposition 2.2.3, où on a aussi une partition naturelle d'un ensemble en deux parts, mais ayant besoin d'un choix supplémentaire pour les étiqueter.

2.4.2. Corollaire/Définition. *Il existe une partition canonique de l'ensemble des bases orthonormées d'un espace vectoriel euclidien* E *en deux parties, telle que pour une telle base* $(e_1, \ldots, e_n)$ *et* $f \in \mathbf{O}(E)$*, la base* $(f(e_1), \ldots, f(e_n))$ *appartient à la même partie que* $(e_1, \ldots, e_n)$ *si et seulement si* $f \in \mathbf{O}^+(E)$*. Choisir une orientation de* E *consiste à choisir l'une des deux parties de l'ensemble des bases orthonormées, qu'on appellera ensuite les bases orthonormées directes (les autres bases orthonormées étant indirectes). Un espace vectoriel (respectivement affine) euclidien orienté est un espace vectoriel euclidien muni (respectivement espace affine vectoriel dont la direction est munie) d'un choix d'orientation.*

Il est à noter que l'adjectif "direct" pour les automorphismes orthogonaux à un sens indépendant de l'orientation, mais que pour les bases orthonormées il n'a un sens que dans le cas d'un espace orienté. On remarque également que pour la distinction des bases directes et indirectes, l'ordre des vecteurs est crucial: comme la matrice d'une transposition est de déterminant -1, chaque permutation impaire des vecteurs change la base en une base de la classe opposée.

Finalement cette structure additionnelle a une particularité qui la distingue de la structure euclidienne: *dans un espace affine euclidien orienté, les sous-espaces ne sont pas orientés.* Autrement dit, il ne faut pas croire que parce qu'on annonce de travailler dans un espace orienté, on peut se servir de cette orientation dans les sous-espaces. En fait les choix éventuels d'orientations dans les sous-espaces (vectoriels) sont totalement indépendants les uns des autres, même lorsque les sous-espaces ont la même dimension. L'exemple des droites montre la difficulté : choisir une orientation pour une droite revient à choisir une direction "positive" le long de la droite, mais il n'y a pas de façon "cohérente" à le faire pour toutes les droites, car on peut déformer de manière continue une droite orientée pour revenir à la même droite, mais avec l'orientation opposée.

Pour considérer les angles, on focalise l'attention sur le plan vectoriel euclidien: dans le reste de cette section on suppose $\dim E = 2$. Mais commençons avec une remarque sur le cas de dimension 1: dans une droite vectoriel euclidien il existe précisément deux vecteurs unitaires, qui sont opposés (car en normalisant un vecteur non nul on en trouve un tel vecteur $\vec{u}$, et $\lambda\vec{u}$ est unitaire si et seulement si $\lambda^2 = 1$). Appelons maintenant $U = \mathrm{U}(E) = \{ \vec{u} \in E \mid \|\vec{u}\| = 1 \}$ l'ensemble des vecteurs unitaires; comme on est en dimension 2 il s'agit du cercle unité dans E. Tout élément d'une base orthonormée doit appartenir à U, et quand on choisit $\vec{u} \in U$ et cherche à le compléter à une base orthonormé, il y a deux possibilités opposées (car $\dim \langle \vec{u} \rangle^\perp = 1$). Or, pour une orientation fixée, l'une des deux possibilités donnera une base directe, et l'autre une base indirecte (car remplacer un vecteur pas son vecteur opposé change la classe de la base). Par conséquent, on a la propriété importante suivante:

2.4.3. Proposition. *Soit* $\dim E = 2$ *et* $\vec{u}, \vec{v} \in \mathrm{U}(E)$*; il existe un* $f \in \mathbf{O}^+(E)$ *unique tel que* $f(\vec{u}) = \vec{v}$*.*

Preuve. Choisissant une orientation pour E, alors chacun des vecteurs $\vec{u}, \vec{v}$ est le premier vecteur d'une base orthonormée directe, et l'application linéaire f qui envoie la première base vers la seconde appartient à $\mathbf{O}^+(E)$. Réciproquement tout $f \in \mathbf{O}^+(E)$ avec $f(\vec{u}) = \vec{v}$ envoie la première base vers la seconde, et est donc égal à l'application linéaire indiquée. On peut construire f ainsi: d'après le corollaire 2.3.5 il existe un réflection vectorielle unique σ tel que $\sigma(\vec{u}) = \vec{v}$ (si $\vec{u} = \vec{v}$, c'est $s_{\langle\vec{u}\rangle}$); comme toute réflection, σ appartient à $\mathbf{O}^-(E)$, mais en formant la composée $f = s_{\langle\vec{v}\rangle} \circ \sigma$ on a $f \in \mathbf{O}^+(E)$ et $f(\vec{u}) = s_{\langle\vec{v}\rangle}(\vec{v}) = \vec{v}$. $\square$

On remarque que la proposition est aussi vraie avec $f \in \mathbf{O}^-(E)$ au lieu de $f \in \mathbf{O}^+(E)$; en fait, en dimension 2 le corollaire 2.3.5 est valable sans l'hypothèse $\vec{x} \neq \vec{y}$, mais à sa place juste $\vec{x} \neq \vec{0}$.

Les éléments de $\mathbf{O}^+(E)$ sont appelés *rotations*, et on vient de montrer que pour deux vecteurs sur le cercle unité il existe une rotation unique qui envoie l'un vers l'autre. Cette situation sera utilisé pour introduire les rotations à partir de paires d'éléments de U, en analogie avec l'introduction de vecteurs à partir de paires de points; le groupe $\mathbf{O}^+(E)$ joue un rôle analogue au groupe de translations. Les renseignements suivants sur ce groupe, et le groupe $\mathbf{O}(E)$ qui le contient, sont fondamentales.

2.4.4. Proposition. *Soit E un espace vectoriel euclidien avec* $\dim E = 2$.
(1) Le groupe $\mathbf{O}^+(E)$ est commutatif.
(2) Pour tout $\rho \in \mathbf{O}^+(E)$ et $\sigma \in \mathbf{O}^-(E)$ on a $\sigma^2 = \mathrm{id}_E$ et $\sigma \circ \rho \circ \sigma = \rho^{-1}$.

Preuve. On propose deux preuves différentes, l'une en coordonnées, qui fournira en passant un modèle concret des groupes concernés, et une preuve plus géométrique. En choisissant une base orthonormée de E, on obtient un isomorphisme du groupe $\mathbf{O}(E)$ avec le groupe $\mathbf{O}(2, \mathbf{R})$ des matrices 2×2 orthogonales sur $\mathbf{R}$. Un vecteur $\binom{a}{b} \in \mathbf{R}^2$ est unitaire si $a^2 + b^2 = 1$, et dans ce cas les vecteurs unitaires qui lui sont orthogonaux sont $\pm\binom{b}{-a}$; on en déduit que

$$\mathbf{O}^+(2, \mathbf{R}) = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \,\middle|\, a, b \in \mathbf{R}; a^2 + b^2 = 1 \right\},$$

et

$$\mathbf{O}^-(2, \mathbf{R}) = \left\{ \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \,\middle|\, a, b \in \mathbf{R}; a^2 + b^2 = 1 \right\}.$$

La vérification de la proposition pour ces matrices relève d'un simple calcul; par exemple si on associe la matrice $\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$ au nombre complexe $a + b\mathbf{i}$ (de module 1), alors le produit de deux matrices de $\mathbf{O}^+(2, \mathbf{R})$ donne la matrice associée au produit des nombres complexes correspondants, et la partie (1) correspond donc à la commutativité de la multiplication complexe.

Pour une démonstration géométrique, on commence par la partie (2). On a vu que pour tout $\vec{u}, \vec{v} \in U$ il existe un $\sigma \in \mathbf{O}^-(E)$ *unique* avec $\sigma(\vec{u}) = \vec{v}$, et que la réflexion par la droite médiatrice de $\vec{u}$ et $\vec{v}$ est un tel σ; par conséquent tout $\sigma \in \mathbf{O}^-(E)$ est une réflexion en une droite vectorielle. Alors $\sigma^2 = \mathrm{id}_E$ exprime une propriété générale des réflexions (et même des symétries orthogonales). Pour $\rho \in \mathbf{O}^+(E)$ et $\sigma \in \mathbf{O}^-(E)$, la composée $\sigma \circ \rho$ appartient à $\mathbf{O}^-(E)$, et vérifie donc $(\sigma \circ \rho)^2 = \mathrm{id}_E$, ce qui après développement et multiplication des deux cotés à droite par ρ^{-1} donne $\sigma \circ \rho \circ \sigma = \rho^{-1}$. Finalement, pour la partie (1), il suffit pour $\rho_1, \rho_2 \in \mathbf{O}^+(E)$ de choisir $\sigma \in \mathbf{O}^-(E)$ quelconque, et de calculer $\rho_1 \circ \rho_2 = \sigma^2 \circ \rho_1 \circ \rho_2 \circ \sigma^2 = \sigma \circ (\rho_1 \circ \rho_2)^{-1} \circ \sigma = \sigma \circ \rho_2^{-1} \circ \rho_1^{-1} \circ \sigma = \sigma^2 \circ \rho_2 \circ \sigma^2 \circ \rho_1 \circ \sigma^2 = \rho_2 \circ \rho_1$. $\square$

Le fait que $\mathbf{O}^+(E)$ est commutatif permet de prendre ce groupe comme modèle pour les angles orientés dans E, et de les noter ces derniers additivement; c'est analogue aux translations d'un espace affine A, qui forment un groupe commutatif qui sert de modèle pour les vecteurs (dans la direction $\overrightarrow{\mathcal{A}}$) qui sont notés additivement.

2.4.5. Définition. *Le groupe des angles orientés d'un plan euclidien E, désigné par $\widehat{\mathrm{U}}(E)$, est une copie du groupe $\mathbf{O}^+(E)$ (donc canoniquement isomorphe à ce groupe), mais écrit additivement. Un angle dans $\widehat{\mathrm{U}}(E)$ peut être noté $\widehat{\vec{u}\vec{v}}$ pour toute paire $\vec{u}, \vec{v} \in \mathrm{U}(E)$ telle que $\rho(\vec{u}) = \vec{v}$ pour la rotation $\rho \in \mathbf{O}^+(E)$ correspondant à cet angle ($\vec{u}$ et $\vec{v}$ déterminent ρ, et donc l'angle, d'après la proposition 2.4.3).*

Pour ceux qui n'aiment pas l'idée de former une copie du groupe $\mathbf{O}^+(E)$ notée différemment, on peut interpréter la notation $\widehat{\vec{u}\vec{v}}$ comme désignant une classe d'équivalence de paires de vecteurs unitaires (à savoir celle qui contient $(\vec{u}, \vec{v})$), où sont équivalent par définition toutes les paires $(\vec{x}, \rho(\vec{x}))$ pour un même $\rho \in \mathbf{O}^+(E)$; c'est cette définition qu'on trouve souvent dans les livres de géométrie. En tout cas on peut appeler la $\rho \in \mathbf{O}^+(E)$ correspondant à $(\vec{u}, \vec{v})$ la rotation d'angle $\widehat{\vec{u}\vec{v}}$; l'angle $\widehat{\vec{u}\vec{u}}$ auquel correspond l'identité id_E sera noté $\hat{0}$, qui est l'élément neutre de $\widehat{\mathrm{U}}(E)$. La définition implique la "relation de Chasles"

$$\widehat{\vec{u}\vec{v}} + \widehat{\vec{v}\vec{w}} = \widehat{\vec{u}\vec{w}} \tag{17}$$

pour les angles car si le rotations $\rho, \rho' \in \mathbf{O}^+(E)$ sont telles que $\rho(\vec{u}) = \vec{v}$ et $\rho(\vec{v}) = \vec{w}$, la somme $\widehat{\vec{u}\vec{v}} + \widehat{\vec{v}\vec{w}}$ correspond à la rotation composée $\rho \circ \rho' = \rho' \circ \rho$ qui envoie $\vec{u}$ vers $\vec{w}$. Tout vecteur non nul $\vec{x}$ détermine un vecteur normalisé $\frac{\vec{x}}{\|\vec{x}\|} \in U$, et on étend la définition des angles aux vecteurs non nuls en décrétant que l'angle orienté entre deux tels vecteurs est celui entre leurs vecteurs orientés. Une classe de vecteurs avec le même vecteur normalisé (ils sont donc deux à deux proportionnels avec un rapport positif) forment une *demi-droite* vectorielle, et la notion d'angle se définit de la même manière entre ces demi-droites.

Comme les éléments du groupe des angles orientés sont définis en termes de l'espace E lui-même (notamment, ils ne sont pas donnés par des simples nombres), les automorphismes orthogonaux peuvent avoir un effet non trivial sur le angles. En fait, si ρ est la rotation d'angle $\widehat{\vec{u}\vec{v}}$, alors pour $f \in \mathbf{O}(E)$ la rotation d'angle $\widehat{f(\vec{u})f(\vec{v})}$ est $f \circ \rho \circ f^{-1}$, car cette composée appartient à $\mathbf{O}^+(E)$, et elle envoie $f(\vec{u})$ vers $f(\vec{v})$. D'après la proposition 2.4.4 cette composée est égale à ρ si $f \in \mathbf{O}^+(E)$ (les rotations conservent les angles orientés), et à ρ^{-1} si $f \in \mathbf{O}^-(E)$ (les réflexions changent le sens des angles orientés). Donc :

$$\widehat{\rho(\vec{u})\rho(\vec{v})} = \widehat{\vec{u}\vec{v}} \qquad \text{et} \qquad \widehat{\sigma(\vec{u})\sigma(\vec{v})} = -\widehat{\vec{u}\vec{v}} \qquad \text{pour } \rho \in \mathbf{O}^+(E) \text{ et } \sigma \in \mathbf{O}^-(E). \tag{18a, b}$$

En fait (18a) donne toutes les façons de présenter le même angle par des paires de vecteurs unitaires différentes: si $\widehat{\vec{u}\vec{v}} = \widehat{\vec{u}'\vec{v}'}$ avec $\vec{u}, \vec{v}, \vec{u}', \vec{v}' \in U$, alors il existe une rotation $\rho \in \mathbf{O}^+(E)$ avec $\rho(\vec{u}) = \vec{u}'$, et pour cette rotation on a $\vec{v}' = R_{\widehat{\vec{u}'\vec{v}'}}(\vec{u}') = R_{\widehat{\vec{u}\vec{v}}}(\rho(\vec{u})) = \rho(R_{\widehat{\vec{u}\vec{v}}}(\vec{u})) = \rho(\vec{v})$, car les rotations commutent entre elles. Le fait que les angles sont invariants sous $\mathbf{O}^+(E)$ nous permet d'associer une "mesure d'angle" indépendante de E aux angles orientés *si une orientation de E est fixée.*

2.4.6. Définition. *Soit E un plan vectoriel euclidien orienté. Alors pour $\vec{u}, \vec{v} \in \mathrm{U}(E)$, une mesure de l'angle $\widehat{\vec{u}\vec{v}}$ est un nombre $\varphi \in \mathbf{R}$ tel que, dans une base orthonomée directe de E, la rotation d'angle $\widehat{\vec{u}\vec{v}}$ ait la matrice $R_\varphi = \left(\begin{smallmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{smallmatrix}\right)$. Les mesures d'un angle orienté donné forment une classe dans $\mathbf{R}/2\pi\mathbf{Z}$.*

D'après ce qu'on vient d'observer la matrice pour la rotation ne dépend pas du choix d'une base orthonomée directe, mair sur une base orthonomée indirecte la matrice sera $R_{-\varphi} = \left(\begin{smallmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{smallmatrix}\right)$, ce qui confirme encore une fois qu'une mesure d'angle ne peut pas être définie sans orientation de l'espace.

2.5. *Doublement et bissection d'angles, Angles de droites.*

La notion de mesure d'angle nous permet de comprendre la structure du groupe $\widehat{\mathrm{U}}(E)$ des angles orientés dans le plan: il est isomorphe au groupe additif $\mathbf{R}/2\pi\mathbf{Z}$ (de deux manières, le choix d'une orientation de E en fixe une). Comme dans tout groupe additif, on peut définir dans $\widehat{\mathrm{U}}(E)$ la multiplication par un entier n par "addition répétée" de n copies d'un angle $\hat{\alpha}$ (ou $|n|$ copies de $-\hat{\alpha}$ si $n < 0$), et cette multiplication est un homomorphisme de groupes. L'isomorphisme $\widehat{\mathrm{U}}(E) \cong \mathbf{R}/2\pi\mathbf{Z}$ montre qu'on ne peut pas espérer d'étendre cette définition à d'autres nombres que les entiers à la place de n: si on voulait associer à un angle de mesure φ un angle de mesure $\lambda\varphi$ pour $\lambda \in \mathbf{R}$ non entier, le choix d'une autre mesure $\varphi + 2\pi$ pour le même angle donnerait pour résultat un angle de mesure $\lambda\varphi + 2\lambda\pi$, mais ce n'est pas dans la classe $\lambda\varphi + 2\pi\mathbf{Z}$, et donc pas une mesure pour l'angle proposé comme résultat auparavant. On pourrait résoudre l'ambiguïté en exigéant un choix particulier pour φ, par exemple avec $-\pi < \varphi \leq \pi$, mais dans ce cas application définie ne sera jamais un homomorphisme de groupes. Il est donc en particulier impossible de diviser un angle de façon univoque par un nomber entier $n > 1$; pour ce cas on peut décrire de manière précise le degré d'ambiguïté qui forme l'obstruction à une telle opération.

2.5.1. Proposition. *Pour $n \in \mathbf{N}_{>0}$, l'homomorphisme de groupes $\mu_n : \widehat{\mathrm{U}}(E) \to \widehat{\mathrm{U}}(E)$ défini par $\hat{\alpha} \mapsto n\hat{\alpha}$ est surjectif, et son noyau est d'ordre n. Par conséquent, chaque angle a n pré-images pour μ_n.*

Preuve. Comme $\widehat{\mathrm{U}}(E) \cong \mathbf{R}/2\pi\mathbf{Z}$, il suffit de montrer ces propriétés pour l'homomorphisme μ'_n : $\mathbf{R}/2\pi\mathbf{Z} \to \mathbf{R}/2\pi\mathbf{Z}$ donné par $\varphi + 2\pi\mathbf{Z} \mapsto n\varphi + 2\pi\mathbf{Z}$. Pour la surjectivité, considérons un élément C

de $\mathbf{R}/2\pi\mathbf{Z}$, une classe qui est de la forme $\psi + 2\pi\mathbf{Z}$; alors $\mu'_n(\frac{\psi}{n} + 2\pi\mathbf{Z}) = C$ (quel pré-image on trouve ici dépend de l'écriture choisi pour C, mais le seul point à montrer était l'existence de (au moins) un pré-image). Le noyau de μ'_n est l'ensemble des classes de la forme $2\pi\frac{k}{n} + 2\pi\mathbf{Z}$ avec $k \in \mathbf{Z}$, quel ensemble contient n classes distincts, car $k, k' \in \mathbf{Z}$ définissent la même classe si et seulement si $k \equiv k' \pmod{n}$. □

Pour $n = 2$, les deux pré-images d'un angle $\widehat{\vec{u}\vec{v}}$, c'est-à-dire les solutions pout l'équation $2\hat{\alpha} = \widehat{\vec{u}\vec{v}}$ en $\hat{\alpha} \in \widehat{\mathrm{U}}(E)$, sont faciles à construire géométriquement. La droite médiatrice $\mathcal{D}$ de $\vec{u}$ et $\vec{v}$ passe par $\vec{0}$ et est donc une droite vectoriel; soit $\vec{w}_1, \vec{w}_2$ ses deux vecteurs unitaires. Alors comme $s_\mathcal{D}(\vec{u}) = \vec{v}$ et $\vec{w}_1, \vec{w}_2$ sont fixes par $s_\mathcal{D}$, on a $\widehat{\vec{u}\vec{w}_i} = -\widehat{\vec{v}\vec{w}_i} = \widehat{\vec{w}_i\vec{v}}$ pour $i = 1, 2$ d'après l'équation (18b), d'où $2\widehat{\vec{u}\vec{w}_i} = \widehat{\vec{u}\vec{w}_i} + \widehat{\vec{w}_i\vec{v}} = \widehat{\vec{u}\vec{v}}$, et $\widehat{\vec{u}\vec{w}_1}$ et $\widehat{\vec{u}\vec{w}_2}$ sont les deux solutions cherchées. C'est d'ailleurs essentiellement le seul cas où une telle construction existe (mis à part $n = 2^k$ qu'on résout en itérant cette "bissection d'angle" k fois); le cas $n = 3$ (trisection de l'angle) est un problème classique pour lequel on a longtemps cherché, en vain, une méthode géométrique pour construire un pré-image, et dont il est finalement prouvé au 19ème siècle par des méthodes algébriques que (pour une interprétation précise du terme "méthode géométrique") aucune telle méthode existe (ni pour d'autres nombres n avec au moins un facteur premier impair).

Le noyau du morphisme μ_2 de doublement d'angle est $\{\hat{0}, \hat{\varpi}\}$, où $\hat{\varpi}$ est appellé l'angle plat; il a mesure π quelle que soit l'orientation de E, et s'écrit $\widehat{\vec{u}\vec{v}}$ avec $\vec{v} = -\vec{u}$. La bissection de l'angle plat donne deux angles droits opposés: la droite médiatrice des vecteurs $\vec{u}$ et $-\vec{u}$ est orthogonal à $\vec{u}$, donc les angles droits s'écrivent $\widehat{\vec{u}\vec{v}}$ avec $\vec{u} \perp \vec{v}$. Si E est un plan orienté, alors l'angle droit $\widehat{\vec{u}\vec{v}}$ a comme mesure $\frac{\pi}{2}$ si $(\vec{u}, \vec{v})$ est une base orthonormée directe, et il a comme mesure $-\frac{\pi}{2}$ si $(\vec{u}, \vec{v})$ est une base orthonormée indirecte.

Si on veut parler de l'angle orienté entre deux droites vectorielles $\mathcal{D}, \mathcal{D}'$ de E, il est important d'observer qu'il y a toujours *deux* rotations distinctes qui transforment $\mathcal{D}$ en $\mathcal{D}'$, car la rotation $-\mathrm{id}_E$ par l'angle plat $\hat{\varpi}$ envoie toute droite vectorielle sur elle même; par conséquent si une rotation par $\hat{\alpha}$ envoie $\mathcal{D}$ vers $\mathcal{D}'$, la rotation par $\hat{\alpha} + \hat{\varpi}$ le fait aussi. Cette ambiguïté peut être levée par le choix dans chaque droite d'un des deux vecteurs unitaires (ou demi-droites), qui permettra de prendre l'angle défini par ceux-ci; mais sans tels éléments pour guider un choix, rien ne permet de préférer l'un des deux angles possibles à l'autre. Donc pour avoir une notion d'angle entre droites vectorielles bien définie, on est obligé de considérer le quotient du groupe $\widehat{\mathrm{U}}(E)$ des angles orientés par le sous-groupe $\{\hat{0}, \hat{\varpi}\}$.

2.5.2. Définition. *Dans un plan vectoriel euclidien E, le groupe $\widehat{\mathcal{D}}(E)$ des angles orientés de droites est le groupe quotient $\widehat{\mathrm{U}}(E)/\{\hat{0}, \hat{\varpi}\}$. Pour deux droites vectorielles $\mathcal{D}, \mathcal{D}'$ de E, l'angle $\widehat{\mathcal{D}\mathcal{D}'} \in \widehat{\mathcal{D}}(E)$ est l'image par la projection canonique de l'angle $\widehat{\vec{x}\vec{y}} \in \widehat{\mathrm{U}}(E)$ de deux vecteurs non nuls $\vec{x} \in \mathcal{D}$ et $\vec{y} \in \mathcal{D}'$, quelle image est indépendante des choix de $\vec{x}, \vec{y}$.*

On constate que le groupe $\{\hat{0}, \hat{\varpi}\}$ par lequel on quotiente ici est aussi le noyau du morphisme μ_2 de doublement d'angle. Ceci permet de former un *isomorphisme* entre $\widehat{\mathrm{U}}(E)$ et $\widehat{\mathcal{D}}(E)$, non pas bien sûr par la projection canonique, mais en prenant pour l'angle $\hat{\beta} \in \widehat{\mathrm{U}}(E)$ correspondant à un angle de droites $\hat{\alpha} \in \widehat{\mathcal{D}}(E)$ son *double*, c'est-à-dire $\hat{\beta} = \mu_2(\hat{\alpha}_0)$ ou $\hat{\alpha}_0$ est un angle représentant $\hat{\alpha}$ (un pré-image par la projection canonique). deux choix possibles pour $\hat{\alpha}_0$ donnent le même résultat pour $\hat{\beta}$, d'où cette application est bien définie et bijective. Formulé de façon informelle: la bissection d'un angle orienté donne un angle orienté de droites bien défini, et le doublement d'un angle orienté de droites donne un angle orienté (de demi-droites) unique. Cette dernière correspondance a une description directe.

2.5.3. Proposition. *Il existe un isomorphisme naturel de groupes $\widetilde{2} : \widehat{\mathcal{D}}(E) \to \widehat{\mathrm{U}}(E)$ de doublement d'angle, qui envoie $\widehat{\mathcal{D}\mathcal{D}'} \in \widehat{\mathcal{D}}(E)$ vers l'angle de rotation de la composée de réflexions $s_{\mathcal{D}'} \circ s_\mathcal{D}$.*

Preuve. La composée indiquée donne un élément $\rho \in \mathbf{O}^+(E)$ bien défini, qui correspond donc à un angle orienté unique. La rotation ρ ne dépend pas du choix de $\mathcal{D}$ et $\mathcal{D}'$, car (comme pour les angles ordinaires) les seules autres façons de présenter l'angle des droites $\widehat{\mathcal{D}\mathcal{D}'}$ est d'appliquer une même rotation ρ' à $\mathcal{D}$ et à $\mathcal{D}'$; ce changement aura l'effet de conjuguer $s_\mathcal{D}$ et $s_{\mathcal{D}'}$, et donc leur composée ρ, par ρ', mais les rotations commutant entre elles le résultat de la conjugaison de ρ restera ρ. Si $\vec{u} \in \mathcal{D}$ est un vecteur unitaire, son image $\vec{v} = \rho(\vec{u})$ est égale à $s_{\mathcal{D}'}(\vec{u})$ (car $s_\mathcal{D}(\vec{u}) = \vec{u}$), et ρ est une rotation d'angle $\widehat{\vec{u}\vec{v}}$. Cet angle est le double de $\widehat{\mathcal{D}\mathcal{D}'}$, ou plus précisément de chacun des angles dans $\widehat{\mathrm{U}}(E)$ dont $\widehat{\mathcal{D}\mathcal{D}'}$ est l'image par la projection canonique, car si $\vec{w} \in \mathcal{D}'$ est un vecteur unitaire on a

$\widehat{\vec{u}\vec{v}} = \widehat{\vec{u}\vec{w}} + \widehat{\vec{w}\vec{v}} = 2\widehat{\vec{u}\vec{w}}$ (par (18b), car $s_{\mathcal{D}'}(\vec{w}) = \vec{w}$). Le fait qu'il s'agit d'un homomorphisme de groupes découle de la règle de Chasles dans $\widehat{\mathcal{D}}(E)$: $\widehat{\mathcal{D}\mathcal{D}'} + \widehat{\mathcal{D}'\mathcal{D}''} = \widehat{\mathcal{D}\mathcal{D}''}$, car la rotation correspondante sera $s_{\mathcal{D}''} \circ s_{\mathcal{D}'} \circ s_{\mathcal{D}'} \circ s_{\mathcal{D}} = s_{\mathcal{D}''} \circ \circ s_{\mathcal{D}}$. Pour montrer que c'est un isomorphisme on construit l'application réciproque: étant donné $\widehat{\vec{u}\vec{v}}$ avec $\vec{u}, \vec{v}$ unitaires, soit $\mathcal{D}$ l'unique droite vectorielle telle que $s_{\mathcal{D}}(\vec{u}) = \vec{v}$ (c'est la médiatrice de $\vec{v}$ et $\vec{u}$ si $\vec{u} \neq \vec{v}$, sinon c'est $\langle\vec{u}\rangle$); alors on voit que $\widehat{\vec{u}\vec{v}}$ est l'image de $\widehat{\langle\vec{u}\rangle\mathcal{D}}$. □

Il est à remarquer que la bissection d'angles $\widetilde{2}^{-1} : \widehat{\mathrm{U}}(E) \to \widehat{\mathcal{D}}(E)$ donnée par l'isomorphisme réciproque dans cette proposition est la seule forme de bissection avec des valeurs uniques: la bissection dans $\widehat{\mathrm{U}}(E)$ donne deux résultats qui diffèrent par l'angle plat, pendant que la bissection dans $\widehat{\mathcal{D}}(E)$ donne deux résultats qui diffèrent par l'angle de droites droit, qui est défini par deux droites orthogonales.

Ces notions d'angle ont des notions correspondantes dans un plan affine euclidien, définis par passage aux vecteurs. Par exemple, l'angle orienté d'un triangle ABC au sommet B, noté $\widehat{ABC}$, est par définition égal à $\widehat{\overrightarrow{BA}\overrightarrow{BC}}$. On définit aussi l'angle orienté de deux demi-droites du plan affine euclidien comme l'angle de demi-droites vectorielles correspondantes; concrètement la demi-droite affine $[AB)$ déterminée par deux points distincts A, B est définie comme $\{ A + \lambda\overrightarrow{AB} \mid \lambda \in \mathbf{R}_{\geq 0} \}$ où $\overrightarrow{[AB)} = \{ \lambda\overrightarrow{AB} \mid \lambda \in \mathbf{R}_{>0} \}$ est la demi-droite vectorielle correspondante (ici on a enlevé $\vec{0}$, pour que tout élément de la demi-droite vectorielle puisse la représenter dans la formation d'angles). On voit que $\widehat{ABC}$ est aussi égal à l'angle entre les demi-droites affines $[BA)$ et $[BC)$ (en considérant des angles entre demi-droites affines, on les prend le plus souvent issues du même point, bien que la définition ne l'exige pas). Or la demi-droite vectorielle $\overrightarrow{[BA)} = \mathbf{R}_{>0} \cdot \overrightarrow{BA}$ est égale à $-\overrightarrow{[AB)} = \mathbf{R}_{<0} \cdot \overrightarrow{AB}$; en voici une simple application.

2.5.4. Proposition. *Pour trois points distincts A, B, C dans le plan affine euclidien, on a*

$$\widehat{ABC} + \widehat{BCA} + \widehat{CAB} = \hat{\varpi}.$$

Preuve. On applique successivement des rotations par les trois angles du premier membre au demi-droite vectorielle $\overrightarrow{[BA)}$. La rotation par $\widehat{ABC}$ la transforme en $\overrightarrow{[BC)} = -\overrightarrow{[CB)}$, puis la rotation par $\widehat{BCA}$ transforme ce résultat en $-\overrightarrow{[CA)} = \overrightarrow{[AC)}$, et finalement la rotation par $\widehat{CAB}$ transforme ce résultat en $\overrightarrow{[AB)} = -\overrightarrow{[BA)}$; au total la composée des trois rotations a transformé la demi-droite en sa demi-droite opposée, donc c'est une rotation par l'angle plat $\hat{\varpi}$. □

Pour un angle de deux demi-droites affines $[BA)$ et $[BC)$ issues du même point B, l'angle de droites qui le bissecte se construit ainsi. On prend des points P, Q sur $[BA)$ et $[BC)$ à la même distance $r > 0$ de B; la médiatrice $\mathcal{D}$ de P et Q passe par B, et on vérifie facilement que $\widetilde{2}(\widehat{\mathcal{D}_{A,B}\mathcal{D}}) = \widehat{ABC}$. La droite $\mathcal{D}$ s'appelle la médiatrice de l'angle formé par $[BA)$ et $[BC)$; c'est l'unique droite telle que $s_{\mathcal{D}}$ envoie $[BA)$ vers $[BC)$. Pour deux droites sécantes $\mathcal{D}, \mathcal{D}'$ dans le plan, il existe deux bissectrices: si on fixe le choix d'une demi-droite de $\mathcal{D}$ issue du point d'intersection, alors la construction ci-dessus donne des réflections uniques qui l'envoient sur chacune des demi-droites de $\mathcal{D}'$ issues du même point. Or la composition des deux réflexions envoie chacune des demi-droites de $\mathcal{D}'$ vers son opposée, donc c'est une rotation par l'angle plat; par conséquence le deux bissectrices, qui sont les axes de ces deux réflexions, sont à un angle droit.

2.5.5. Définition/Proposition. *La distance d'un point A à un sous-espace $\mathcal{V}$ d'un espace affine euclidien est définie par $d(A, \mathcal{V}) = d(A, p_{\mathcal{V}}(A)$. Pour $B \in \mathcal{V}$ on a $d(A, B) \geq d(A, \mathcal{V})$ avec égalité ssi $B = p_{\mathcal{V}}(A)$.*

Preuve. Posant $B = p_{\mathcal{V}}(A) + \vec{v}$, on a $d(A, B) = \sqrt{d(A, \mathcal{V})^2 + \|\vec{v}\|^2} \geq d(A, \mathcal{V})$, avec égalité ssi $\vec{v} = \vec{0}$. □

Comme la distance entre deux points, la distance d'un point à un sous-espace est invariante par les isométries. Donc en particulier pour deux droites sécantes $\mathcal{D}, \mathcal{D}'$ de plan, tout point de l'une des deux bissectrices a la même distances aux deux droites, à cause de la réflexion par la bissectrice qui fixe le point et envoie $\mathcal{D}$ vers $\mathcal{D}'$. Cette propriété peut être utilisée pour caractériser les bissectrices:

2.5.6. Proposition. *Si $\mathcal{D}, \mathcal{D}'$ sont deux droites non parallèles d'un plan affine euclidien, alors*

$$\{ A \in \mathcal{A} \mid d(A, \mathcal{D}) = d(A, \mathcal{D}') \}$$

est égal à la réunion des deux bissectrices (othogonales entre elles) de $\mathcal{D}$ et $\mathcal{D}'$.

Preuve. On vient de voir que tout point d'un des deux bissectrices appartient à l'ensemble indiqué. Soit réciproquement A un point tel que $d(A, \mathcal{D}) = d(A, \mathcal{D}')$, et $P = p_{\mathcal{D}}(A)$, $P' = p_{\mathcal{D}'}(A)$. Dans le cas où $P = P'$, le vecteur $\overrightarrow{AP} = \overrightarrow{AP'}$ est à la fois orthogonal à $\mathcal{D}$ et $\mathcal{D}'$, et comme ces droites non sont pas parallèles, ceci n'est possible que si $A = P$, l'unique point qui appartient aux deux bissectrices à la fois. Dans le cas restant, où $P \neq P'$, on montrera que la médiatrice de P et P' est l'une des bissectrices de $\mathcal{D}$ et $\mathcal{D}'$, et comme par hypothèse A se trouve sur cette médiatrice, cela prouvera la proposition. La réflexion par la médiatrice envoie par définition P vers P'; pour montrer qu'elle envoie $\mathcal{D}$ vers $\mathcal{D}'$, et que la médiatrice est donc bissectrice, il suffit de montrer que la réflexion envoie un autre point de $\mathcal{D}$ vers $\mathcal{D}'$. On montrera en fait que le point d'intersection M de $\mathcal{D}$ et $\mathcal{D}'$ est fixé par la réflexion, c'est-à-dire se trouve sur la médiatrice. Pour cela il suffit d'appliquer le théorème de Pythagore: $d(M, P) = \sqrt{d(M, A)^2 - d(A, P)^2} = \sqrt{d(M, A)^2 - d(A, P')^2} = d(M, P')$. □

2.5.7. Définition. *Dans un triangle A, B, C du plan affine euclidien, la bissectrice intérieure issue du sommet B est la bissectrice des demi-droites $[BA)$ et $[BC)$, et les bissectrices intérieures des autres sommets sont définies de façon similaire. La bissectrice extérieure issue d'un sommet est l'autre bissectrice des deux côtés du triangle (des droites) passant par ce sommet.*

Voici un autre critère pour distinguer les bissectrices intérieure et extérieure issues du même sommet.

2.5.8. Proposition. *Soit P un point d'une bissectrice issue d'un sommet du triangle A, B, C et distinct de ce sommet. Alors P est sur la bissectrice intérieure si et seulement si les coordonnées barycentriques de P correspondantes aux autre deux sommets, dans le repère affine $\mathcal{S} = (A, B, C)$, ont le même signe.*

Preuve. Pour simplifier on suppose que le sommet concerné est A. On considère la droite $\mathcal{D}$ passant par P et parallèle à l'autre bissectrice; comme les bissectrices sont évidemment sécantes avec $\mathcal{D}_{A,B}$ et $\mathcal{D}_{A,C}$, c'est aussi le cas de $\mathcal{D}$. Les deux points d'intersection Q, R de $\mathcal{D}$ avec ces côtés sont échangés par la réflexion par rapport à la bissectrice qui passe par P, donc $P = \mathrm{bar}(Q, R)$. En écrivant $Q = (1 - \lambda, \lambda, 0)_{\mathcal{S}}$ et $R = (1 - \mu, 0, \mu)_{\mathcal{S}}$, cette bissectrice est intérieure si et seulement si λ et μ ont le même signe. Or on calcule facilement $P = (1 - \frac{\lambda}{2} - \frac{\mu}{2}, \frac{\lambda}{2}, \frac{\mu}{2})_S$, ce qui établit la caractérisation dans l'énoncé. □

2.5.9. Théorème [du cercle inscrit]. *Chaque paire de bissectrices d'un triangle issues de sommets différents est sécante, et les deux bissectrices sont concourantes avec l'une des bissectrices issues du troisième sommet. Pour un tel triplet de bissectrices concourantes le nombre de bissectrices intérieures est impair, le point de concours est au centre d'un cercle qui est tangent à chacun des côtés du triangle, appelé le cercle inscrit du triangle si les trois bissectrices sont intérieures, et un cercle exinscrit si seulement une des bissectrices est intérieure. Le cercle inscrit est contenu dans l'enveloppe convexe des sommets.*

Preuve. Pour la première partie on utilise la caractérisation de la proposition 2.5.6. Si deux bissectrices issues de sommets différents étaient parallèles, la composée de leurs réflections serait une translation (cela découle d'un simple calcul); or l'un des côtés du triangle est envoyé par cette composée sur un autre côté, ce qu'une translation ne saura faire (car ces côtés ne sont pas parallèles). Deux telles bissectrices se coupent donc, et leur point d'intersection se trouve à distance égale des côtés, à la fois pour deux paires de côtés différentes. Comme ces deux paires ont forcément un côté en commun, ce point est à distance égale des trois côtés, et donc sur l'une des bissectrices issues du sommet restant. Il est aussi clair qu'un cercle de centre ce point d'intersection et de rayon égal à la distance commune aux trois côtés sera tangent à ces trois cotés. Pour trois bissectrices concourantes, la proposition 2.5.8 permet de déterminer lesquelles sont intérieures en comparant les signes des coordonnées barycentriques du point d'intersection deux à deux; on trouvera égalité des signes soit pour une paire, soit pour toutes les trois paires. Dans le cas où les signes sont tous égaux (intersection P des trois bissectrices intérieures) ils sont positifs, car la somme des coordonnées barycentriques est 1. L'enveloppe convexe des sommets (c'est-à-dire l'intérieur du triangle) est l'ensemble des points à coordonnées barycentriques toutes positives; on montrera pour chaque coordonnées barycentrique que sa valeur et partout positive sur le cercle inscrit Γ. Comme elle s'annule sur un côté du triangle qui est tangent à Γ, disons en P, la coordonnée est de signe constant sur l'ensemble connexe $\Gamma \setminus \{P\}$, et ce signe est celui qu'elle prend au centre du cercle, c'est-à-dire positif. □

2.6. *Cocyclicité.*

La notion d'angle de droites a une application importante concernant la question de "cocyclicité", c'est à dire s'il existe un cercle passant par un nombre de points donnés. Pour trois points dans le plan cette condition est toujours vérifiée, sauf s'ils sont alignés (et distincts). En fait on a en dimension n:

2.6.1. Proposition. *Soit $(A_0, \ldots, A_n)$ un repère affine dans un espace affine euclidien $\mathcal{A}$, alors il existe une unique sphère qui contient tous les points A_i du repère (elle est de rayon $r > 0$ si $n = \dim \mathcal{A} > 0$).*

Preuve. Il suffit de montrer l'existence d'un point unique $\Omega \in \mathcal{A}$ telle que toutes les distances $d(\Omega, A_i)$ soient égales pour $i = 0, 1, \ldots, n$: si r est cette distance, la sphère cherchée sera $\mathcal{S}(\Omega, r)$. La condition sur les distances veut dire que Ω est dans chacun des hyperplans médiateurs de A_0 et A_i, pour $1 \leq i \leq n$. La condition pour être dans l'hyperplan médiateur de A_0 et A_i s'écrit $\langle \overrightarrow{A_0\Omega}, \overrightarrow{A_0A_i} \rangle = \frac{d(A_0,A_i)^2}{2}$. Comme les $\overrightarrow{A_0A_i}$ forment une base de E, ces équations forment un système de Cramer, avec une solution unique. □

2.6.2. Corollaire/Définition. *Pour un triangle du plan affine euclidien, il existe un cercle unique passant par les trois sommets, appelé le cercle circonscrit du triangle. Son centre est l'intersection des médiatrices (des paires de sommets) du triangle.*

La cocyclicité est donc une question intéressante pour 4 points dans le plan affine euclidien, non alignés trois à trois, et revient à demander si le quatrième point est sur le cercle déterminé par les trois premiers. L'énoncé qui caractérise la cocyclicité, qu'on donnera ci-dessous, ne mentionne que des angles entre des droites définis par les points concernés. Mais pour arriver à ce résultat, on commence avec l'analyse d'une situation un peu différente, que voici.

2.6.3. Lemme. *Soit $\Gamma = \mathcal{S}(\Omega, r)$ un cercle dans le plan affine euclidien de centre Γ et de rayon $r > 0$, $P \in \Gamma$, et $\mathcal{D}$ la tangente à Γ en P. Alors pour tout $\mathcal{A} \in \Gamma \setminus \{P\}$ on a $\widehat{P\Omega A} = \widetilde{2}(\widehat{\mathcal{D}\mathcal{D}_{P,A}})$.*

Preuve. La médiatrice de P et A passe par Ω, donc la réflexion par rapport à cette médiatrice montre que $\widehat{AP\Omega} = -\widehat{PA\Omega} = \widehat{\Omega AP}$ (égalité d'angles d'un triangle isocèle). Puis la proposition 2.5.4 donne $\widehat{P\Omega A} = \hat{\varpi} - \widehat{\Omega AP} - \widehat{AP\Omega} = \hat{\varpi} - 2\widehat{AP\Omega} = \widetilde{2}(\widehat{\mathcal{D}\mathcal{D}_{P,\Omega}}) - \widetilde{2}(\widehat{\mathcal{D}_{A,P}\mathcal{D}_{P,\Omega}}) = \widetilde{2}(\widehat{\mathcal{D}\mathcal{D}_{P,\Omega}} + \widehat{\mathcal{D}_{P,\Omega}\mathcal{D}_{P,A}}) = \widetilde{2}(\widehat{\mathcal{D}\mathcal{D}_{P,A}})$. □

De manière informelle, cette proposition dit que si A parcourt le cercle Γ à une certain vitesse angulaire constante, alors la droite $\mathcal{D}_{P,A}$ tourne aussi avec une vitesse angulaire constante, mais qui est la moitié de celle de A (vu du centre Ω); par conséquent quand le point A a fait un tour complet, la droite $\mathcal{D}_{A,P}$ aura fait un demi-tour. La tangente $\mathcal{D}$ prend la place de la droite $\mathcal{D}_{A,P}$ pour le cas où $A = P$. Ainsi le mouvement de la droite $\mathcal{D}_{A,P}$ devient continu au passage de P, ce qui est possible grâce au fait qu'on la considère comme une droite sans direction positive choisie ; par contre, la demi-droite engendrée par le vecteur $\overrightarrow{PA}$ change de façon discontinue au passage de P.

2.6.4. Théorème [de l'angle inscrit]. *Soit $\Gamma = \mathcal{S}(\Omega, r)$ un cercle dans le plan affine euclidien de centre Γ et de rayon $r > 0$, des points $A, B \in \Gamma$ sur le cercle avec $A \neq B$, et $P \in \mathcal{A} \setminus \{A, B\}$. Alors*

$$P \in \Gamma \iff \widehat{A\Omega B} = \widetilde{2}(\widehat{\mathcal{D}_{P,A}\mathcal{D}_{P,B}}).$$

Preuve. Si $P \in \Gamma$, soit $\mathcal{D}$ la tangente à Γ en P. Alors on a, d'après le lemme ci-dessus,

$$\widehat{A\Omega B} = \widehat{A\Omega P} + \widehat{P\Omega B} = \widetilde{2}(\widehat{\mathcal{D}\mathcal{D}_{P,B}}) - \widetilde{2}(\widehat{\mathcal{D}\mathcal{D}_{P,A}}) = \widetilde{2}(\widehat{\mathcal{D}_{P,A}\mathcal{D}_{P,B}}).$$

Supposons maintenant que réciproquement $\widehat{A\Omega B} = \widetilde{2}(\widehat{\mathcal{D}_{P,A}\mathcal{D}_{P,B}})$. On montre d'abord que $P \notin \mathcal{D}_{A,B}$: on a $A \neq B$ donc $\hat{0} \neq \widehat{A\Omega B} = \widetilde{2}(\widehat{\mathcal{D}_{P,A}\mathcal{D}_{P,B}})$, c'est-à-dire $\mathcal{D}_{P,A} \not\parallel \mathcal{D}_{P,B}$ (car $\widetilde{2}$ est injectif d'après la proposition 2.5.3) et $P \notin \mathcal{D}_{A,B}$. Ensuite on montre que $\mathcal{D}_{P,A}$ n'est pas tangent (en A) à Γ: en cas de tangence, le lemme donnerait $\widehat{A\Omega B} = \widetilde{2}(\widehat{\mathcal{D}_{P,A}\mathcal{D}_{A,B}})$ et donc $\mathcal{D}_{P,B} \parallel \mathcal{D}_{A,B}$ contredisant $P \notin \mathcal{D}_{A,B}$. La droite $\mathcal{D}_{A,P}$ coupe donc Γ en un deuxième point $P' \neq A$; on veut montrer que $P' = P$. On applique l'implication déjà montrée à $P' \in \Gamma$, ce qui donne $\widehat{A\Omega B} = \widetilde{2}(\widehat{\mathcal{D}_{P',A}\mathcal{D}_{P',B}})$. Or $\mathcal{D}_{P',A} = \mathcal{D}_{P,A}$, donc d'après l'hypothèse on a $\mathcal{D}_{P,B} \parallel \mathcal{D}_{P',B}$ et donc $\mathcal{D}_{P,B} = \mathcal{D}_{P',B}$ et $P' \in \mathcal{D}_{P,A} \cap \mathcal{D}_{P,B} = \{P\}$ car $P \notin \mathcal{D}_{A,B}$. □

2.6.5. Corollaire [cocyclicité]. *Soit $\mathcal{A}$ un plan euclidien, $A,B,C,D \in \mathcal{A}$ avec $\{A,B\} \cap \{C,D\} = \emptyset$, alors il existe un cercle ou une droite contentant A,B,C,D si et seulement si $\widehat{\mathcal{D}_{A,C}\mathcal{D}_{B,C}} = \widehat{\mathcal{D}_{A,D}\mathcal{D}_{B,D}}$.*

Preuve. Si $A = B$ la seconde condition est trivialement vraie, et la première est aussi toujours vraie d'après la proposition 2.6.1. Si A,B,C sont distincts et alignés, aucun cercle passe par les trois points (l'intersection d'un cercle et une droite ne peut pas voir plus de deux points, d'après la proposition 2.3.2). Dans ce cas les deux conditions affirment que $D \in \mathcal{D}_{A,B}$, et sont donc équivalentes. Il reste le cas le plus intéressant où A,B,C ne sont pas alignés; la première condition devient alors si $D \in \Gamma$ où Γ est le cercle circonscrit de A,B,C. Le théorème de l'angle inscrit donne $\widehat{A\Omega B} = \widetilde{2}(\widehat{\mathcal{D}_{A,C}\mathcal{D}_{B,C}})$, et appliqué pour D il dit que $D \in \Gamma$ si et seulement si $\widetilde{2}(\widehat{\mathcal{D}_{A,C}\mathcal{D}_{B,C}}) = \widetilde{2}(\widehat{\mathcal{D}_{A,D}\mathcal{D}_{B,D}})$. Par l'injectivité de l'application $\widetilde{2}$, cette dernière équation est équivalente à $\widehat{\mathcal{D}_{A,C}\mathcal{D}_{B,C}} = \widehat{\mathcal{D}_{A,D}\mathcal{D}_{B,D}}$, ce qui établit la preuve pour ce dernier cas. □

On donne maintenant une application typique de cocyclicité, concernant les hauteurs d'un triangle et le "triangle orthique" associé. Cela montrera comment on utilise la cocyclicité dans un sens pour une paire d'angles de droites égales (en fait des angles droits) pour obtenir la cocyclicité de points, et ensuite dans l'autre sens pour obtenir l'égalité d'une autre paire d'angles de droites. On verra aussi un avantage du fait qu'il s'agit d'une comparaison d'angles dans $\widehat{\mathcal{D}}(E)$ et non pas dans $\widehat{\mathrm{U}}(E)$: si on sait que trois points X,Y,Z sont alignés et distincts, on peut utiliser $\mathcal{D}_{X,Y} = \mathcal{D}_{Y,Z}$ sans se soucier de la question si $\overrightarrow{XY}$ et $\overrightarrow{YZ}$ engendrent la même demi-droite ou deux demi-droites opposées.

Dans un triangle A,B,C du plan, les *hauteurs* sont les trois droites h_A, h_B, h_C qui passent par un sommet et sont orthogonaux au côté passant par les deux autres sommets. L'intersection de la hauteur h_A avec le côté "opposé" $\mathcal{D}_{B,C}$ est formé par un point H_A appelé le pied de cette hauteur. Une hauteur peut être confondue avec l'un des côtés passant par le même sommet, auquel cas ce côté forme un angle droit avec le côté passant par les deux autres sommets; si cela arrive on appelle A,B,C un triangle rectangle (en le sommet dont l'angle est droit). Dans un triangle rectangle en C on a $h_A = \mathcal{D}_{A,C}$, $h_B = \mathcal{D}_{B,C}$, et $H_A = H_B = C$. Dans le cas contraire d'un triangle non rectangle, les pieds des trois hauteurs sont distincts des sommets et distincts entre eux, et forment le *triangle orthique* du triangle.

2.6.6. Théorème [triangle orthique]. *Soit A,B,C un triangle non rectangle du plan affine euclidien. Alors les bissectrices du triangle orthique H_A, H_B, H_C sont les côtés et les hauteurs du triangle A,B,C.*

Preuve. Comme le triangle n'est pas rectangle, les pieds sont distincts des sommets, et on peut décrire les côtés comme par exemple $\mathcal{D}_{B,C} = \mathcal{D}_{B,H_A} = \mathcal{D}_{H_A,C}$, quel type de changement on appliquera librement. Pour chaque paire de sommets ces sommets et les pieds correspondants sont cocycliques:, par exemple $\widehat{h_A\mathcal{D}_{B,C}} = \widehat{\mathcal{D}_{A,C}h_B}$ car les deux angles sont droits, donc A,B,H_A,H_B sont cocycliques (en fait H_A et H_B sont sur le cercle de diamètre $[A,B]$ d'après (14)). En appliquant le corollaire 2.6.5 dans le sens opposé pour les points B,H_B,A,H_A (dans cet ordre), on trouve $\widehat{\mathcal{D}_{A,B}\mathcal{D}_{A,C}} = \widehat{\mathcal{D}_{B,C}\mathcal{D}_{H_A,H_B}}$. Par le même argument pour les points cocycliques A,C,H_A,H_C on trouve $\widehat{\mathcal{D}_{A,B}\mathcal{D}_{A,C}} = \widehat{\mathcal{D}_{H_A,H_c}\mathcal{D}_{B,C}}$, ce qui montre que $\mathcal{D}_{B,C}$ est une bissectrice des côtés $\mathcal{D}_{H_A,H_B}$ et $\mathcal{D}_{H_A,H_c}$ du triangle orthique. L'autre bissectrice de ces droites est h_A, qui coupe $\mathcal{D}_{B,C}$ à angle droit en H_A. Les mêmes arguments s'appliquent pour H_B, H_C. □

On peut montrer que h_A est bissectrice de $\mathcal{D}_{H_A,H_B}$ et $\mathcal{D}_{H_A,H_C}$ directement par cocyclicité, mais l'argument est un petit peu plus long: en utilisant les mêmes ensembles cocycliques on obtient les égalités $\widehat{h_A\mathcal{D}_{H_A,H_B}} = \widehat{\mathcal{D}_{B,A}h_B}$ et $\widehat{\mathcal{D}_{H_A,H_C}h_A} = \widehat{h_C\mathcal{D}_{C,A}}$; pour voir que les deux seconds membres sont égaux, on peut soit appliquer une troisième cocyclicité (de $H_C, H_B.B, C$), soit ajouter des angles droites aux deux membres: $\widehat{\mathcal{D}_{B,A}h_B} + \widehat{h_B\mathcal{D}_{A,C}} = \widehat{\mathcal{D}_{A,B}\mathcal{D}_{A,C}} = \widehat{h_C\mathcal{D}_{C,A}} + \widehat{\mathcal{D}_{A,B}h_C}$ (ce qui illustre qu'on peut appliquer la relation de Chasles pour les angles de droites sans que ces droites soient concourantes; après tout, les angles sont définis par passage aux *directions* de ces droites). Si on admet que les trois hauteurs sont concourantes en H (voir ci-dessous), on peut également prouver directement $\widehat{\mathcal{D}_{B,A}h_B} = \widehat{\mathcal{D}_{H_A,H_C}h_A}$ par la cocyclicité de H_C, H, B, H_A, qui est une conséquence des angles droits à H_A et H_C.

D'après le théorème 2.5.9 du cercle inscrit, l'ensemble des six bissectrices mentionnées dans le théorème ci-dessus contient quatre triplets de droites concourantes; par définition deux côtés sont toujours concourants avec la hauteur issue de leur sommet commun, et cela épuise les combinaisons qui contiennent à la fois un côté et une hauteur, donc les trois hauteurs forment le quatrième triplet concourant. Le point d'intersection des trois hauteurs est appelé l'orthocentre H du triangle A, B, C. Sans information supplémentaire *on ne peut pas* dire lesquelles des bissectrices du triangle orthique sont intérieures et extérieures. Si le triangle est "acutangle", les trois pieds sont sur le segment entre deux sommets, leur enveloppe convexe (l'intérieur du triangle orthique) ne contient aucun des sommets A, B, C, et le centre du cercle inscrit du triangle orthique est forcément l'orthocentre H. Mais pour un triangle "obtusangle" deux parmi les pieds ne sont pas dans le segment du côté les contentant déterminé par deux sommets, et dans ce cas le centre du cercle inscrit du triangle orthique n'est pas H, mais le sommet où l'angle est obtus (qui se trouve donc à l'intérieur du triangle orthique).

2.7. *Quelques points remarquables du triangle.*

Résumons ce qu'on a vu sur les plus classiques parmi les triplets de droites concourantes associées à un triangle du plan affine euclidien (il y en a encore beaucoup plus, la littérature sur ce sujet est très riche).

On a vu dans la proposition 1.3.7 que les trois *médianes* du triangle sont concourantes en le *barycentre* du triangle, c'est-à-dire l'isobarycentre des trois sommets. C'est un résultat purement affine, et essentiellement le seul de ce type à l'être (on pourra dire que le théorème de Ceva (1.3.8) donne une infinité de résultats de ce type, mais comme les droites concernées dépendent d'un paramètre continue, il est difficile de les classer comme "remarquables").

Les *médiatrices* du triangle sont concourantes en le *centre du cercle circonscrit* du triangle, comme on l'a vu dans le corollaire 2.6.2.

Les *bissectrices intérieures* du triangle sont concourantes en le *centre du cercle inscrit* du triangle, et deux *bissectrices extérieures* sont concourantes avec la bissectrice intérieure issue du sommet restant en le *centre d'un cercle exinscrit*; c'est le théorème 2.5.9. On remarque que les preuves données pour les médiatrices et pour les bissectrices sont très similaires, utilisant l'égalité des distances vers les sommets respectivement les côtés du triangle.

Les *hauteurs* du triangle sont concourantes en l'*orthocentre* du triangle; c'était une conséquence du théorème sur le triangle orthique. En fait il existe une démonstration plus élémentaire de ce fait, qui le réduit au cas des médiatrices.

2.7.1. Théorème. *Soit A, B, C un triangle du plan affine euclidien. On considère le triangle A', B', C' où $A' = \mathrm{bar}((A, -1), (B, 1), (C, 1))$, $B' = \mathrm{bar}((A, 1), (B, -1), (C, 1))$, et $C' = \mathrm{bar}((A, 1), (B, 1), (C, -1))$. Alors les deux triangles ont le même barycentre P, et le second est l'image du premier par l'homothétie $h_{P,-2}$. Les hauteurs de A, B, C sont les médiatrices de A', B', C', et l'orthocentre H de A, B, C est le centre du cercle circonscrit de A', B', C'. Le centre du cercle circonscrit de A, B, C est $h_{P,-\frac{1}{2}}(H)$.*

Preuve. On a $A = \mathrm{bar}(B', C')$ par l'associativité des barycentres, ainsi que $\mathrm{bar}(A, A') = \mathrm{bar}(B, C)$, donc la médiane de A', B', C' issue de A' est $\mathcal{D}_{A',A}$ qui est aussi médiane de A, B, C issue de A; les triangles ont donc le même barycentre P. D'après proposition 1.3.7 on a $\overrightarrow{A'P} : \overrightarrow{PA} = 2 : 1$ d'où $h_{P,-2}(A) = A'$ et pareil pour B, C. Cette homothétie montre que $\mathcal{D}_{B,C} \parallel \mathcal{D}_{B',C'}$; la médiatrice de B' et C' est orthogonal à ces droites et passe par $\mathrm{bar}(B', C') = A$, donc c'est la hauteur h_A de A, B, C. L'orthocentre H est le point d'intersection des hauteurs de A, B, C qui sont les médiatrices de A', B', C', donc c'est le centre du cercle circonscrit de ce dernier. L'homothétie réciproque $h_{P,-\frac{1}{2}}$ de $h_{P,-2}$ envoie ce centre sur le centre du cercle circonscrit de A, B, C, ce qui donne le dernier énoncé. □

On voit donc que l'orthocentre H, le barycentre P, et le centre Ω du cercle circonscrit sont alignés, et que $\overrightarrow{PH} : \overrightarrow{P\Omega} = 2 : 1$. On peut montrer que les trois ne sont confondus que si A, B, C est un triangle équilatère; dans le cas contraire le trois centres définissent une droite appelé la *droite d'Euler* du triangle A, B, C. On remarque d'ailleurs que dans cette configuration le cercle de diamètre $[H, A']$, contient les points B, C, car $\overrightarrow{BH} \perp \overrightarrow{A'B}$ et $\overrightarrow{CH} \perp \overrightarrow{A'C}$, quel cercle est donc le cercle circonscrit au triangle A', C, B, triangle congruent au triangle A, B, C par la symétrie centrale en $\mathrm{bar}(B, C)$. Ce cercle circonscrit est donc du même rayon que le cercle circonscrit de A, B, C, comme le sont les cercles circonscrits de A, B', C et A, B, C' (ce qui confirme que les diamètres $[H, A']$, $[H, B']$, $[H, C']$ sont de la même longueur). On a ainsi quatre cercles de la même taille qui se coupent trois à trois, en quatre points A, B, C, H, et qui forment une configuration connue comme les "cercles de Johnson".

2.7.2. Corollaire. *L'image du cercle circonscrit d'un triangle A, B, C par la réflexion en un des trois côtés du triangle, où encore par la symétrie centrale en le milieu de ce côté, passe par son orthocentre.*

2.7.3. Corollaire. *Une hauteur d'un triangle coupe le cercle circonscrit, outre en le sommet dont elle est issue, en un point qui est l'image de l'orthocentre par la réflexion en le côté opposé au sommet.*

Preuve. L'orthocentre H se trouve sur l'image du cercle circonscrit Γ par cette réflexion, et l'image de H par cette réflexion se trouve donc sur Γ. La hauteur issue du sommet opposé est globalement stable par cette réflexion, et contient donc aussi cette image de H. Le point le plus délicat est de montrer que cette image n'est pas confondue avec le sommet dont est issu la hauteur, quel sommet est aussi à l'intersection de la hauteur et le cercle circonscrit. En fait, les deux peuvent être confondus, mais uniquement quand la hauteur est tangente au cercle circonscrit, auquel cas il n'y a donc qu'un seul point d'intersection: si la réflection par un côté envoie H vers un sommet de A, B, C, disons vers A, alors le triangle A, B, H est isocèle (tout comme A, C, H), et le centre de son cercle circonscrit se trouve sur la droite $\mathcal{D}_{B,C}$; la symétrie centrale par rapport à $\operatorname{bar}(A, B)$ envoie ce centre vers le centre Ω de Γ, qui vérifie donc $\mathcal{D}_{\Omega,A} \parallel \mathcal{D}_{B,C}$, et la hauteur h_A est effectivement la tangente à Γ en A. $\square$

Une vérification directe de la cocyclicité de $P = s_{\mathcal{D}_{B,C}}(H)$ avec A, B, C est évidemment possible: par la symétrie on a $\widehat{\mathcal{D}_{P,B}\mathcal{D}_{P,C}} = \widehat{h_C h_B}$, qui est égal à $\widehat{\mathcal{D}_{A,B}\mathcal{D}_{A,C}}$ par cocyclicité de H_C, H_B, H, A, ou plus simplement encore par $h_c \perp \mathcal{D}_{A,B}$ et $h_B \perp \mathcal{D}_{A,C}$. On remarque que, en excluant le cas d'un triangle rectangle où H est confondu avec l'un des sommets, les quatre points A, B, C, H sont tels que pour chaque triangle formé de 3 d'entre eux, le quatrième est l'orthocentre du triangle (c'est un conséquence immédiate de la définition).

2.8. *Le groupe des isométries.*

Dans les sections précédentes on s'est beaucoup servi des isométries particulière d'un espace affine euclidien, notamment des réflexions et des rotations. Il est parfois utile d'avoir des informations concernant la totalité du groupe $\mathbf{Isom}(\mathcal{A})$ des isométries d'une espace affine, par exemple pour savoir s'il existe une isométrie qui transformé une figure donnée en une autre, et sous quelle forme on peut décrire une telle isométrie. On sait déjà qu'une application affine $f : \mathcal{A} \to \mathcal{A}$ est un isométrie si et seulement si $\overrightarrow{f}$ est une isométrie vectorielle, ce qui veut dire que sa matrice dans une base orthonormée est orthogonale. Cela est utile pour avoir une idée globale de combien d'isométries on dispose; par exemple si $n = \dim \mathcal{A}$ on aura besoin d'au moins $\frac{n(n-1)}{2}$ paramètres réels pour décrire la matrice de $\overrightarrow{f}$ (plus un paramètre “booléen” pour distinguer les composantes connexes $\mathbf{O}^+(E)$ et $\mathbf{O}^-(E)$) et encore n pour spécifier l'image $f(A)$ d'un point particulier. La “dimension” $\frac{n^2+n}{2}$ de $\mathbf{Isom}(\mathcal{A})$ ainsi déterminée (attention, il ne s'agit pas pour autant d'un espace vectoriel) est précisément la moitié de celle de $\mathbf{GA}(\mathcal{A})$, qui est $n^2 + n$.

Mais dans la pratique il est beaucoup plus utile d'avoir une classification géométrique des isométries. Une telle classification est relativement facile à établir, en termes de l'ensemble $\operatorname{Fix}(f)$ des points fixes de f. On aura besoin d'un résultat préalable sur le compositions d'une isométrie avec une réflexion.

2.8.1. Lemme. *Soit $f \in \mathbf{O}(E)$ une isométrie d'un espace vectoriel euclidien E, et $f' = s \circ f$ où $s \in \mathbf{O}(E)$ est une réflexion (par rapport à un hyperplan vectoriel $\mathcal{H}$). Alors $|\dim \operatorname{Fix}(f) - \dim \operatorname{Fix}(f')| \leq 1$.*

Preuve. On a $\mathcal{H} \cap \operatorname{Fix}(f) = \mathcal{H} \cap \operatorname{Fix}(f')$, car les deux inclusions se vérifient de façon directe. Or on a soit $\operatorname{Fix}(f) \subseteq \mathcal{H}$, soit $\mathcal{H} \cap \operatorname{Fix}(f)$ est un hyperplan dans $\operatorname{Fix}(f)$, d'après la proposition 1.2.10. La même chose s'applique pour $\operatorname{Fix}(f')$, et on a $\{\dim \operatorname{Fix}(f), \dim \operatorname{Fix}(f')\} \subseteq \{d, d+1\}$ où $d = \dim(\mathcal{H} \cap \operatorname{Fix}(f))$. $\square$

2.8.2. Lemme. *Soit $f \in \mathbf{Isom}(\mathcal{A})$ une isométrie d'un espace affine euclidien $\mathcal{A}$ de dimension n. Si $\operatorname{Fix}(f) \neq \emptyset$ on pose $d = \dim \operatorname{Fix}(f)$, sinon on pose $d = -1$. Alors f s'écrit comme une composée d'au plus $n - d$ réflexions. Si $\operatorname{Fix}(f) \neq \emptyset$ ce nombre est aussi minimal: aucune composée de moins de $n - d$ réflexions n'est égale à f.*

Preuve. Par récurrence sur $n - d$. Si $n - d = 0$, tous les points sont fixes, donc on a $f = \mathrm{id}_{\mathcal{A}}$, qui est par convention la composée d'une suite vide de réflexions. Supposons ensuite $n - d > 0$, ce qui garantit l'existence d'un point $A \in \mathcal{A} \setminus \mathrm{Fix}(f)$ non fixé par f. Pour un choix d'un tel point on considère l'hyperplan médiateur $\mathcal{H}$ de A et $f(A)$, et la composée $f' = s_H \circ f$. On a aussi $f = s_\mathcal{H} \circ f'$, donc pour trouver l'écriture cherchée de f, il suffira d'écrire f' comme composée d'au plus $n - d - 1$ réflexions. Cela découlera de l'hypothèse de récurrence pourvu qu'on ait $\dim \mathrm{Fix}(f') \geq d + 1$. Par construction on a $f'(A) = s_\mathcal{H}(f(A)) = A$ et donc $A \in \mathrm{Fix}(f')$; en particulier $\mathrm{Fix}(f') \neq \emptyset$ et $\dim \mathrm{Fix}(f')$ est bien défini. Or $\mathrm{Fix}(f) \subseteq \mathrm{Fix}(f')$ car pour $P \in \mathrm{Fix}(f)$ on a $d(P, A) = d(f(P), f(A)) = d(P, f(A)$ car f est isométrie, donc $P \in \mathcal{H}$ et $f'(P) = s_\mathcal{H}(f(P)) = P$. Le sous-espace affine $\mathrm{Fix}(f')$ qui contient $\mathrm{Fix}(f)$ et aussi $A \notin \mathrm{Fix}(f)$ est de dimension au moins $d+1$, comme voulu. Il reste à démontrer que si $\mathrm{Fix}(f) \neq \emptyset$, le nombre de réflexions trouvé est minimal. Pour cela on utilise l'application $\overrightarrow{f}$, dont on sait que $\dim \mathrm{Fix}(\overrightarrow{f}) = \dim \mathrm{Fix}(f)$ (car $\mathrm{Fix}(f) = P + \mathrm{Fix}(\overrightarrow{f})$ quand $P \in \mathrm{Fix}(f)$) et que tout écriture de f comme composée de m réflexions donne une écriture de $\overrightarrow{f}$ comme composée de m réflexions dans $\mathbf{O}(E)$. Il découle du lemme précédent que pour une telle écriture on aura $\dim \mathrm{Fix}(\overrightarrow{f}) \geq n - m$, ce qui donne une contradiction si $m < n - d$. □

Pour la dernière partie de la preuve, on observera l'importance du fait que les isométries vectorielles ont toujours un point fixe (à savoir $\vec{0}$), sans lequel la démonstration du lemme préalable n'aurait pas pu se faire. On pourra détailler les résultats des lemmes précédents ainsi.

2.8.3. Lemme. *Soit $f, f' \in \mathbf{Isom}(\mathcal{A})$ avec $f = s_\mathcal{H} \circ f'$ où s_H est la réflexion par rapport à un hyperplan affine $\mathcal{H}$. On a $|\dim \mathrm{Fix}(\overrightarrow{f}) - \dim \mathrm{Fix}(\overrightarrow{f'})| = 1$, et si $\mathrm{Fix}(f)$ et $\mathrm{Fix}(f')$ sont tous les deux non vides on a aussi $|\dim \mathrm{Fix}(f) - \dim \mathrm{Fix}(f')| = 1$. Si $\mathrm{Fix}(f) = \emptyset$ et $\mathrm{Fix}(f') \neq \emptyset$, alors $\dim \mathrm{Fix}(\overrightarrow{f}) = \dim \mathrm{Fix}(f') + 1$.*

Preuve. Le lemme 2.8.1 donne $|\dim \mathrm{Fix}(\overrightarrow{f}) - \dim \mathrm{Fix}(\overrightarrow{f'})| \leq 1$. On peut interpréter les applications linéaires comme des applications affines, qui auront toujours au moins un point fixe; alors le lemme 2.8.2 interprète $n - \dim \mathrm{Fix}(\overrightarrow{f})$ comme le nombre minimal de réflexions qui puissent produire $\overrightarrow{f}$, et la parité de ce nombre dépend de la partie $\mathbf{O}^+(E)$ ou $\mathbf{O}^-(E)$ qui contient $\overrightarrow{f}$. Cette parité est opposée pour $\overrightarrow{f'}$, donc $\dim \mathrm{Fix}(\overrightarrow{f}) \neq \dim \mathrm{Fix}(\overrightarrow{f'})$. Si $\mathrm{Fix}(f) \neq \emptyset$ on a $\dim \mathrm{Fix}(f) = \dim \mathrm{Fix}(\overrightarrow{f})$ comme dans la preuve du lemme 2.8.2, et pareil pour f'. Finalement pour que $\mathrm{Fix}(f) = \emptyset$ et $\mathrm{Fix}(f') \neq \emptyset$, il faut que $\mathcal{H}$ soit disjoint de $\mathrm{Fix}(f')$, et donc faiblement parallèle à $\mathrm{Fix}(f')$; alors $\overrightarrow{\mathcal{H}} \supseteq \mathrm{Fix}(\overrightarrow{f'})$ et donc $\mathrm{Fix}(\overrightarrow{f}) \supset \mathrm{Fix}(\overrightarrow{f'})$. □

2.8.4. Corollaire. *Soit $\dim \mathcal{A} = n$ et $g_0, g_1, \ldots, g_m \in \mathbf{Isom}(\mathcal{A})$, avec $g_0 = \mathrm{id}_\mathcal{A}$, et telle que, pour $0 < i \leq m$ on ait soit $g_i = s_{\mathcal{H}_i} \circ g_{i-1}$ soit $g_i = g_{i-1} \circ s_{\mathcal{H}_i}$. Alors si $\mathrm{Fix}(g_m) \neq \emptyset$ on a $\dim \mathrm{Fix}(g_m) \geq n - m$, et sinon $\dim \mathrm{Fix}(\overrightarrow{g_m}) \geq n - m + 2$. Si pour $0 < i \leq m$ l'hyperplan $\mathcal{H}_i$ qui ne contient pas $\mathrm{Fix}(g_{i-1})$ tout entier, les inégalités sont des égalités, et g_m ne s'écrit pas comme composée de moins de m réflexions.*

Preuve. Le lemme 2.8.3 s'applique à (g_i, g_{i-1}) si $g_i = s_{\mathcal{H}_i} \circ g_{i-1}$, et à (g_i^{-1}, g_{i-1}^{-1}) si $g_i = g_{i-1} \circ s_{\mathcal{H}_i}$. Alors les inégalités en découlent par récurrence, ainsi que les égalités, dont l'hypothèse entraîne $\mathrm{Fix}(g_i) \neq \emptyset$ pour $i < m$. La minimalité de m dans ce cas est une redite des inégalités, en fixant $g = g_m$ pour m variable. □

Le procédé récursif du lemme 2.8.2 produit toujours une écriture de f comme une composée d'un nombre minimal de réflexions, car par construction $A \in \mathrm{Fix}(f') \setminus \mathcal{H}$; seulement, si $\mathrm{Fix}(f) = \emptyset$ la borne $n - d = n + 1$ peut être supérieure à ce minimum. Par exemple, il est aisé de vérifier que la composée de deux réflexions par rapport à des hyperplans parallèles donne une translation: si $\vec{v}$ est un vecteur orthogonal au premier hyperplan $\mathcal{H}$ tel que le second hyperplan est $\mathcal{H} + \vec{v}$, c'est la translation $t_{2\vec{v}}$. Comme une translation par un vecteur non nul n'a pas de points fixes, la borne $n + 1$ donnée par ce lemme peut être améliorée à 2 dans ce cas. En fait, le corollaire montre que pour $\mathrm{Fix}(f) = \emptyset$ le nombre minimal de réflexions nécessaires est $n - \dim \mathrm{Fix}(\overrightarrow{f}) + 2$ (pour les translations on a $\dim \mathrm{Fix}(\overrightarrow{f}) = n$).

2.8.5. Proposition. *Pour tout $f \in \mathbf{Isom}(\mathcal{A})$ il existe $f_0 \in \mathbf{Isom}(\mathcal{A})$ et $\vec{v} \in E$ tels que $\mathrm{Fix}(f_0) \neq \emptyset$, $\vec{v} \in \mathrm{Fix}(\overrightarrow{f})$ et $f = t_{\vec{v}} \circ f_0$. De plus, ce couple $(f_0, \vec{v})$ est unique, et $t_{\vec{v}}$ et f_0 commentent entre eux.*

Preuve. On considère l'application $g : \mathcal{A} \to E = \overrightarrow{\mathcal{A}}$ définie par $g : A \mapsto \overrightarrow{Af(A)}$. C'est une application affine (en considérant E comme espace affine) dont l'application linéaire associée est donnée par $\overrightarrow{g}(\vec{x}) = g(A+\vec{x}) - g(A) = \overrightarrow{(A + \vec{x})f(A + \vec{x})} - \overrightarrow{Af(A)} = \overrightarrow{f}(\vec{x}) - \vec{x}$ (par l'équation (8)), indépendamment de $A \in \mathcal{A}$ comme il le faut. L'image de $\vec{g}$ est orthogonal à $V = \mathrm{Fix}(\overrightarrow{f})$, car pour tout $\vec{y} \in V$ le fait que $\overrightarrow{f} \in \mathbf{O}(E)$ donne $\langle \vec{x}, \vec{y} \rangle = \langle \overrightarrow{f}(\vec{x}), \overrightarrow{f}(\vec{y}) \rangle = \langle \overrightarrow{f}(\vec{x}), \vec{y} \rangle$, et donc $\overrightarrow{f}(\vec{x}) - \vec{x} \perp \vec{y}$. Par conséquent on a pour la composée $h = p_V \circ g$ que $\overrightarrow{h}(\vec{x}) = p_V(\overrightarrow{f}(\vec{x}) - \vec{x}) = \vec{0}$, et c'est donc une application constante. Si $(f_0, \vec{v})$ est comme dans l'énoncé, et $P \in \mathrm{Fix}(f_0)$, on a $g(P) = \overrightarrow{P(f_0(P) + \vec{v})} = \vec{v} = p_V(\vec{v}) = h(P)$ donc $\vec{v}$ doit être la valeur constante de h, et $f_0 = t_{-\vec{v}} \circ f$. Ceci établit l'unicité de ce couple, et il reste à vérifier que le couple indiqué possède bien les propriétés requises. La commutation de $t_{\vec{v}}$ et f_0 est aisé à prouver: $f_0(A + \vec{v}) = t_{-\vec{v}}(f(A) + \overrightarrow{f}(\vec{v})) = f(A) + \overrightarrow{f}(\vec{v}) - \vec{v} = f(A)$ car $\vec{v} \in \mathrm{Fix}(\overrightarrow{f})$, donc $f_0 \circ t_{\vec{v}} = f = t_{\vec{v}} \circ f_0$. Pour montrer que f_0 possède un point fixe, on observe que par construction $f_0(A) = f(A) - p_V(\overrightarrow{Af(A)}) = A + p_{V^\perp}(\overrightarrow{Af(A)}) \in A + V^\perp$ pour tout $A \in \mathcal{A}$, donc $A + V^\perp$ est globalement stable par f_0. Or la restriction $f_0|_{A+V^\perp}$ est telle que $\mathrm{Fix}(\overrightarrow{f_0|_{A+V^\perp}}) = \mathrm{Fix}(\overrightarrow{f_0}) \cap V^\perp = \{\vec{0}\}$ car $\overrightarrow{f_0} = \overrightarrow{f}$; alors cette restriction possède précisément un point fixe d'après la proposition 1.8.7, qui est aussi point fixe de f_0. □

2.8.6. Théorème. *Soit $f \in \mathbf{Isom}(\mathcal{A})$ une isométrie d'un espace affine euclidien $\mathcal{A}$ de dimension n. En posant $d = \dim \mathrm{Fix}(\overrightarrow{f})$, on peut écrire f comme la composée de $n - d$ réflexions si $\mathrm{Fix}(f) \neq \emptyset$, et de $n - d + 2$ réflexions sinon; dans les deux cas ce nombre est le minimum possible. Dans le dernier cas on pourra choisir l'écriture telle que la composée des $n - d$ premières réflexions donne f_0 avec $\overrightarrow{f_0} = \overrightarrow{f}$ et $\mathrm{Fix}(f_0) \neq \emptyset$, et que la composée des 2 dernières réflexions donne une translation $t_{\vec{v}} \neq \mathrm{id}_{\mathcal{A}}$ où $\vec{v} \in \mathrm{Fix}(\overrightarrow{f})$.*

Preuve. Presque tout a déjà été prouvé; pour le cas $\mathrm{Fix}(f) = \emptyset$ il suffit d'appliquer la proposition précédente pour trouver f_0 et $\vec{v}$, et d'écrire f_0 et $t_{\vec{v}}$ comme composée de $n-d$ respectivement 2 réflexions. □

Les arguments ci-dessus sont un peu long, mais on les a donnés dans leur intégralité pour indiquer que la classification des isométries par leur écriture minimale comme composée de réflexions peut se faire complètement, et de façon indépendante de la dimension n de l'espace. L'importance pratique ces résultats est surtout de permettre par la suite une classification explicite en petite dimension; on donne maintenant cette classification pour $n = 2$ et $n = 3$. On donnera dans ces classifications aussi quelques compléments pour information, mais sans preuve détaillée; avec les faits déjà établis ces compléments sont relativement facile a obtenir, en considérant des écritures du type mentionné et les possibilités de les les varier. Il s'agit d'une description des donnés qui caractérisent un élément de chaque classe, du nombre de paramètres libres de chaque classe (c'est-à-dire leur dimension), et de quels paramètres caractérisent les classes de conjugaison dans $\mathbf{Isom}(\mathcal{A})$.

Considérons le cas $n = 2$ du plan affine euclidien.

- L'identité est l'unique isométrie qui est composée de 0 réflexions. Elle est dans $\mathbf{Isom}^+(\mathcal{A})$.
- Les réflexions sont évidemment composées d'une seul réflexion; elles sont dans $\mathbf{Isom}^-(\mathcal{A})$. Une réflexion est déterminée par son axe (une droite qui est l'ensemble des points fixes), ce qui représente 2 paramètres libres. Toutes les réflexions sont conjuguées.
- Une composée de deux réflexions qui n'a pas de points fixe est une translation $t_{\vec{v}}$ avec $\vec{v} \neq \vec{0}$, d'après le théorème (on a $f_0 = \mathrm{id}_{\mathcal{A}}$); elle est dans $\mathbf{Isom}^+(\mathcal{A})$. Une translation s'écrit comme composée $t_{\vec{v}} = s_{\mathcal{D}+\frac{1}{2}\vec{v}} \circ s_{\mathcal{D}}$ de deux réflexions, où $\mathcal{D}$ est une droite orthogonale à $\vec{v}$. Le vecteur $\vec{v}$ représente 2 paramètres libres, et deux translations sont conjuguées si leurs vecteurs on la même norme.
- Une composée de deux réflexions distinctes qui a un point fixe en aura précisément un (on a $d = 0$ dans le théorème). Si $\mathcal{D}$ et $\mathcal{D}'$ sont des droites qui se coupent en P, la composée $s_{\mathcal{D}'} \circ s_D$ est une rotation de centre P et d'angle $\widetilde{2}(\widehat{\mathcal{D}\mathcal{D}'})$; elle est dans $\mathbf{Isom}^+(\mathcal{A})$. Une rotation détermine 3 paramètres libres: 2 pour le choix du centre, et 1 pour l'angle de rotation. Deux rotations sont conjuguées dans $\mathbf{Isom}(\mathcal{A})$ si les valeurs absolues de leurs angles (les angles géométriques) sont égales (elles sont conjuguées dans $\mathbf{Isom}^+(\mathcal{A})$ seulement si leurs angles orientés sont égaux).
- Une isométrie qui est une composée de 3 réflexions, et pas moins, n'a pas de points fixes (d'après le lemme 2.8.2). C'est une composée d'une réflexion (f_0 dans le théorème) et d'une translation $t_{\vec{v}}$ avec $\vec{v} \neq \vec{0}$ dans la direction de l'axe de la réflexion. Cette composée est appelée une *réflexion glissée*, et est dans $\mathbf{Isom}^-(\mathcal{A})$. Une réflexion glissée détermine 3 paramètres libres: 2 pour l'axe de la réflexion, et 1 pour le vecteur $\vec{v}$. Deux réflexions glissées sont conjuguées si leurs vecteurs ont la même norme.

On peut remarquer que la composé de trois réflexions dans le plan n'aura un point fixe que si le résultat est encore une réflexion, ce qui d'après le corollaire 2.8.4 arrive si et seulement si les axes de réflexion sont concourantes ou parallèles. Ceci permet une autre façon de prouver quelles combinaisons de bissectrices sont concourantes dans le théorème 2.5.9: la réflexion successive dans trois bissectrices stabilise globalement l'un des côtés du triangle, et on voit facilement que la direction de ce côté se retrouve inversée si et seulement si le nombre de bissectrices intérieures concernées est impair, auquel cas il y a un point fixe sur ce côté, et des bissectrices sont concourantes. Dans le cas d'un nombre pair de bissectrices intérieures, la composée est au contraire une réflexion glissée (le côté étant l'axe), et les bissectrices ne sont pas concourantes.

Dans le cas d'un espace affine euclidien de dimension 3, on retrouve tous ces types d'isométries, avec quelques petites différences.

- L'identité est l'unique isométrie qui est composée de 0 réflexions. Elle est dans $\mathbf{Isom}^+(\mathcal{A})$.
- Les réflexions sont évidemment composées d'une seul réflexion; elles sont dans $\mathbf{Isom}^-(\mathcal{A})$. On parle du plan d'une réflexion pour l'ensemble des points fixes (plutôt que des son axe), qui représente 3 paramètres libres. Toutes les réflexions sont conjuguées.
- Une composée de deux réflexions qui n'a pas de points fixe est une translation $t_{\vec{v}}$ avec $\vec{v} \neq \vec{0}$; elle est dans $\mathbf{Isom}^+(\mathcal{A})$. Le vecteur $\vec{v}$ représente 3 paramètres libres, et deux translations sont conjuguées si leurs vecteurs on la même norme.
- Une composée de deux réflexions distinctes qui a un point fixe est une rotation, qui a en fait toute une droite de points fixes, son axe; elle est dans $\mathbf{Isom}^+(\mathcal{A})$. Une rotation détermine 5 paramètres libres: 4 pour le choix de son axe, et 1 pour l'angle de rotation. La valeur absolue de cet angle est le double l'angle géométrique entre des vecteurs orthogonaux aux plans de réflexion. Pour associer une mesure d'angle à cette rotation, il faudra choisir l'une des deux orientations du plan orthogonal à l'axe. C'est la raison pour laquelle une rotation est conjuguée, même dans $\mathbf{Isom}^+(\mathcal{A})$, avec une rotation par l'angle opposé: deux rotations sont conjuguées dans $\mathbf{Isom}(\mathcal{A})+$ (et donc aussi dans $\mathbf{Isom}(\mathcal{A})$) si les valeurs absolues de leurs angles.
- Une isométrie qui est une composée de 3 réflexions, et pas moins, et qui n'a pas de points fixes est une réflexion glissée, qui appartient à $\mathbf{Isom}^-(\mathcal{A})$. C'est la composée d'une réflexion et d'une translation par un vecteur dans la direction du plan de la réflexion. Elle détermine 5 paramètres libres : 3 pour le plan et 2 pour la translation. Deux réflexions glissées sont conjuguées si leurs vecteurs de translation ont la même norme.

Il reste deux nouveaux types d'isométrie, qui n'ont pas d'équivalent en dimension 2.

- Une isométrie qui est une composée de 3 réflexions, et pas moins, peut aussi avoir précisément un point fixe, si les trois plans de réflexion ont un seul point en commun (qui sera donc le point fixe); elle appartient à $\mathbf{Isom}^-(\mathcal{A})$. On pourra toujours l'écrire sous la forme de la composée d'une rotation et d'une réflexion par rapport à un plan orthogonal à l'axe de la rotation (d'où les deux commutent), l'axe et le plan passent par le point fixe. Cet écriture est unique *sauf* si la rotation est d'un demi-tour (symétrie par rapport à son axe), auquel cas la composée est la symétrie par rapport au point fixe. Pour les autres cas les termes "*anti-rotation*" ou "roto-réflexion" peuvent décrire ce type d'isométrie. L'axe de rotation et le plan de réflexion d'une anti-rotation sont l'unique droite et l'unique plan globalement stable par l'anti-rotation. Comme pour les rotations, l'axe représente 4 paramètres libres, et l'angle de rotation et le choix du plan orthogonale donnent un total de 6 paramètres. Deux anti-rotations sont conjuguées si leur angles de rotation sont égaux au signe (non déterminé) près.
- Une isométrie qui est une composée de 4 réflexions, et pas moins, n'a pas de points fixes; c'est la composée d'une rotation et d'une translation dans la direction de l'axe de rotation, ce qui est appelé un *vissage* (ou "rotation glissée"); il appartient à $\mathbf{Isom}^+(\mathcal{A})$. La direction de translation permet d'orienter tout plan orthogonal à l'axe de rotation, et on peut associer au vissage une mesure d'angle unique modulo 2π. L'axe représente 4 paramètres libres; avec l'angle de rotation et le vecteur de translation, on a au total 6 paramètres. Deux vissages sont conjugués dans $\mathbf{Isom}^+(\mathcal{A})$ si les normes de leurs vecteurs de translation et leurs angles de rotation sont les mêmes; si les angles sont opposés, les vissages sont seulement conjugués dans $\mathbf{Isom}(\mathcal{A})$ (une réflexion change le "sens" du vissage).

Table de matières.