

Université
de Liège



Faculté des Sciences
Département de Mathématique

Géométrie

Notes provisoires

Pierre Mathonet

Première année d'études de bachelier en sciences physiques
Année académique 2014-2015

Introduction

L'objet de ce cours est l'étude de la géométrie des espaces affines en général et des espaces affines euclidiens en particulier.

On a étudié dans l'enseignement secondaire la géométrie plane et la géométrie dans l'espace (les espaces de dimension 2 et 3). Les notions de point, droite, plan, d'angle de parallélisme, de perpendicularité sont donc familières dans ces contextes. Cependant la science moderne impose souvent un travail dans des espaces de dimension supérieure à 3 ; citons à titre d'exemple l'espace de Minkowski, de dimension 4, cadre naturel de la relativité restreinte, ou encore les espaces de Hilbert, qui sont des espaces vectoriels euclidiens de dimension finie ou infinie, et fournissent un cadre mathématique pour la mécanique quantique.

Un des objectifs principaux de ce cours sera donc l'extension des définitions et de certaines propriétés vues dans l'enseignement secondaire à des espaces de dimension quelconque, mais le plus souvent finie. Cette extension pourrait être menée à bien en conservant l'approche suivie dans l'enseignement secondaire, à savoir en définissant des objets de base que sont les points, les droites, les plans et leurs analogues de dimension supérieure et en adoptant des *axiomes* qui déterminent le comportement de ces différents objets (intersections, parallélisme, perpendicularité, ...). Cette approche atteint cependant vite certaines limites et nous adopterons l'approche plus moderne qui consiste à étudier en premier les propriétés des espaces vectoriels et à les transposer en propriétés géométriques. Il est cependant bien entendu que l'intuition fournie par les éléments de géométrie vus dans l'enseignement secondaire guidera notre démarche, et que les définitions que nous adopterons dans le cas général fourniront une géométrie équivalente à celle qui a été vue jusqu'à présent, dans les cas particuliers des espaces de dimension 2 et 3.

Après avoir introduit et étudié de manière abstraite (mais pas sans exemples) les espaces vectoriels, nous introduirons les espaces affines, qui modélisent le mieux l'espace géométrique dans lequel nous vivons (si on considère qu'il est plat). Ensuite nous introduirons les notions métriques dans les espaces vectoriels (angles de vecteurs, produits vectoriels, produits mixtes) à partir de la donnée d'un produit scalaire. Cette approche où le produit scalaire n'est pas unique est nécessaire pour aborder les espaces plus généraux, courbes, qui sont le cadre de la relativité générale. Nous tirerons les conséquences de ces définitions dans les espaces affines euclidiens.

Nous terminerons notre étude par une introduction à la théorie des courbes dans des espaces affines de dimension 2 et 3, et par un mot sur les surfaces, ces deux thèmes constituant une introduction à la géométrie différentielle.

Ce cours est une adaptation libre des cours dispensés par le professeur P. Lecomte aux étudiants en mathématique et aux étudiants ingénieurs civils. J'ai aussi pu profiter des notes de cours et des conseils du professeur M. Rigo qui m'a précédé aux commandes de ce cours en première année de bachelier en Sciences Physiques. Je les remercie tous les deux chaleureusement pour leur aide.

Ces notes de cours sont encore fort incomplètes et ne sont pas encore aussi claires que je le voudrais, mais il m'a semblé qu'elles pourraient vous être utiles en leur état actuel. C'est pourquoi je vous les livre et vous prie d'être indulgents durant votre lecture.

Je vous souhaite bonne lecture et bon travail.

P. Mathonet

Chapitre 1

Espaces vectoriels réels

1.1 Introduction

La notion de vecteur vue dans l'enseignement secondaire repose sur des a priori qui n'ont pas été définis rigoureusement. Nous adoptons une définition axiomatique qui ne dit pas ce que sont les vecteurs (des vecteurs devrait-on dire), mais qui indique les opérations que l'on doit pouvoir effectuer sur les vecteurs. On pourra cependant garder l'intuition qui a été donnée dans l'enseignement secondaire dans "le plan" ou dans "l'espace", mais en ayant bien conscience qu'il s'agit d'une "représentation" des définitions rigoureuses que nous posons.

L'intuition des "flèches" que vous avez vues dans l'enseignement secondaire peut être conservée, mais ce que nous en retiendrons est que l'on peut additionner des vecteurs et les multiplier par des nombres, et que tout se passe bien, c'est-à-dire que ces opérations ont des propriétés semblables à celles des opérations définies sur les nombres réels.

1.2 Premières définitions et exemples

Nous définissons donc la notion d'espace vectoriel plutôt que celle de vecteur. Il s'agit d'un ensemble muni des opérations d'addition et de multiplication par des nombres réels, encore appelée multiplication scalaire^a. Ces opérations doivent jouir de propriétés qui permettent de traiter les vecteurs un peu comme on traite les nombres, lorsqu'il s'agit de les additionner ou les multiplier par des nombres.

Définition 1.2.1. Un espace vectoriel réel est un ensemble E non vide (dont les éléments sont appelés vecteurs), muni de deux opérations, l'addition $+$ et la multiplication par des nombres réels \cdot satisfaisant les propriétés suivantes.

1. L'addition est interne : on a

$$+ : E \times E \rightarrow E : (u, v) \mapsto u + v,$$

et la multiplication scalaire associe un vecteur à un nombre et un vecteur : on a

$$\cdot : \mathbb{R} \times E \rightarrow E : (\lambda, u) \mapsto \lambda \cdot u;$$

2. L'addition est associative : on a $(u + v) + w = u + (v + w)$ pour tous u, v, w dans E ;
3. L'addition admet un élément neutre $0 \in E$, tel que $0 + u = u + 0 = u$ pour tout $u \in E$;
4. Tout élément $u \in E$ admet un opposé $-u$ pour l'addition, tel que $u + (-u) = (-u) + u = 0$;
5. L'addition est commutative : on a $u + v = v + u$ pour tous $u, v \in E$;
6. La multiplication distribue l'addition des vecteurs : on a $a \cdot (u + v) = a \cdot u + a \cdot v$ pour tous $u, v \in E$ et $a \in \mathbb{R}$;

a. A ne pas confondre avec la notion de produit scalaire, qui sera définie plus tard.

7. La multiplication distribue l'addition des réels : on a $(a + b) \cdot u = a \cdot u + b \cdot u$ pour tous $a, b \in \mathbb{R}$ et $u \in E$;
8. La multiplication définit une action : on a $a \cdot (b \cdot u) = (ab) \cdot u$ pour tous $u \in E$ et $a, b \in \mathbb{R}$;
9. La multiplication définit une action : on a $1 \cdot u = u$ pour tout $u \in E$.

Avant de passer aux exemples, voici quelques remarques sur cette définition.

Remarque 1.1. 1. La même définition permet d'introduire la notion d'espace vectoriel complexe, il suffit juste de remplacer \mathbb{R} par \mathbb{C} . Nous n'utiliserons cependant pas cette notion plus générale dans ce cours de géométrie. On peut aussi considérer d'autres ensembles de nombres (comme l'ensemble des nombres rationnels \mathbb{Q}), pour autant qu'on puisse les multiplier et additionner comme les nombres réels. De tels ensembles de nombres sont appelés *champs*.

2. Les propriétés qui sont énoncées sont naturelles, puisqu'elles sont valides pour les nombres réels. Elles nous permettent donc de manipuler l'addition des vecteurs et la multiplication par des nombres comme nous le ferons pour l'addition et la multiplication des nombres réels.
3. Les éléments de E sont appelés vecteurs.
4. On définira la soustraction $u - v$ par l'opération $u + (-v)$. Cependant pour que la soustraction ainsi définie ait un sens, il est souhaitable de prouver que l'opposé d'un élément $v \in E$ est unique. Nous le ferons d'ici peu. De même, la division d'un vecteur u par un réel a non nul sera définie comme le résultat de l'opération $\frac{1}{a} \cdot u$.
5. Vous pourriez être étonnés de voir des vecteurs sans flèches. Vous avez en effet noté jusqu'ici \vec{u} le "vecteur u ". Dans ce cours, comme nous allons le voir d'ici peu, une grande variété d'objets mathématiques sont des éléments d'espaces vectoriels, et donc peuvent porter le nom de vecteur. C'est le cas par exemple de la fonction cosinus \cos . Je suppose qu'il ne vous viendrait pas à l'idée de la noter comme ceci : $\vec{\cos}$. On n'indiquera donc en général pas dans ce chapitre la flèche, mais on *écrira* "soit le vecteur u ", ou "considérons le vecteur ξ , etc...

Tirons maintenant les premières conséquences de la définition des espaces vectoriels.

Proposition 1.2.2. *Soit E un espace vectoriel.*

1. L'élément neutre pour l'addition est unique ;
2. Tout vecteur $u \in E$ admet un unique opposé^b ;
3. Pour tout $u \in E$, on a $0 \cdot u = 0$ ^c ;
4. Pour tout $u \in E$, on a $-u = (-1) \cdot u$ ^d ;
5. Pour tout $\lambda \in \mathbb{R}$, on a $\lambda \cdot 0 = 0$;
6. Si $\lambda \in \mathbb{R}$ et $u \in E$ satisfont $\lambda \cdot u = 0$, alors on a $u = 0$ ou $\lambda = 0$.

Démonstration. Supposons qu'il existe deux éléments neutres pour l'addition de E , 0 et $0'$. Alors puisque 0 est neutre, on a $0 + 0' = 0'$, mais puisque $0'$ est neutre, on a aussi $0 + 0' = 0$, donc on a $0 = 0'$.

Si u' et u'' sont deux opposés de u , alors on a

$$u' = u' + (u + u'') = (u' + u) + u'' = u''$$

et donc on a $u' = u''$.

Pour tout $u \in E$, on a

$$u = 1 \cdot u = (1 + 0) \cdot u = 1 \cdot u + 0 \cdot u = u + 0 \cdot u.$$

b. Cela justifie l'introduction de la notation $-u$.

c. Les zéros de part et d'autre de l'égalité n'ont pas le même sens.

d. Les signes - de part et d'autre de l'égalité n'ont pas le même sens.

En ajoutant $-u$ aux deux membres de cette identité, on obtient $0 = 0 \cdot u$.

Pour tout $u \in E$, on a

$$u + (-1) \cdot u = 1 \cdot u + (-1) \cdot u = (1 + (-1)) \cdot u = 0 \cdot u = 0$$

On a de la même manière $(-1) \cdot u + u = 0$, ce qui montre que $(-1) \cdot u$ est l'opposé de u .

Pour tout $u \in E$, on a $\lambda \cdot u = \lambda \cdot (u + 0) = \lambda \cdot u + \lambda \cdot 0$. En ajoutant $-(\lambda \cdot u)$ aux deux membres de cette identité, on obtient $\lambda \cdot 0 = 0$.

Enfin, si $\lambda \cdot u = 0$ et si $\lambda \neq 0$, alors $\frac{1}{\lambda}$ est bien défini et on a

$$u = \frac{1}{\lambda} \cdot (\lambda \cdot u) = \frac{1}{\lambda} \cdot 0 = 0,$$

ce qui achève la preuve. □

Remarque 1.2. 1. Nous avons énoncé et démontré les propriétés principales de l'addition et de la multiplication scalaire. Toutes les autres propriétés dont nous aurions besoin peuvent (et doivent) être démontrées de la même façon à partir des conditions intervenant dans la définition, ou à partir de ce que nous avons déjà démontré.

2. Dans la suite, sauf si cela peut engendrer des confusions, nous ne noterons plus la multiplication du vecteur u par le réel λ par $\lambda \cdot u$, mais simplement par λu . On note par exemple $\frac{\pi}{2}u$ la multiplication du vecteur u par le nombre $\frac{\pi}{2}$.
3. Dans beaucoup d'ouvrages, on note les vecteurs par un type de lettres particulier, ou avec une flèche par dessus, ou encore en gras. Vu la diversité des espaces vectoriels que nous allons définir, il est difficile de maintenir l'une ou l'autre de ces notations.

Il est maintenant largement temps de donner quelques exemples d'espaces vectoriels. Pour rompre dans un premier temps avec la notion de vecteur introduite dans l'enseignement secondaire et pour bien illustrer la généralité de la définition que nous avons posée, nous commençons par des exemples particulièrement abstraits.

Exemple 1.2.3. 1. L'ensemble $E =]0, +\infty[^2$ muni des opérations d'addition \oplus et de multiplication \odot définies par

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \oplus \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 y_1 \\ x_2 y_2 \end{pmatrix}, \quad \forall x_1, x_2, y_1, y_2 \in]0, +\infty[, \quad (1.1)$$

et

$$\lambda \odot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1^\lambda \\ x_2^\lambda \end{pmatrix}, \quad \forall x_1, x_2 \in]0, +\infty[, \lambda \in \mathbb{R}, \quad (1.2)$$

est un espace vectoriel.

2. Soit $\mathcal{F}(\mathbb{R})$ l'ensemble des fonctions définies sur \mathbb{R} et à valeurs dans \mathbb{R} . On définit l'addition et la multiplication scalaire par

$$\begin{cases} (f + g)(x) &= f(x) + g(x); \\ (\lambda f)(x) &= \lambda f(x); \end{cases} \quad \forall x \in \mathbb{R}, \quad (1.3)$$

pour toutes fonctions f et g et tout $\lambda \in \mathbb{R}$.

3. Soit \mathcal{P} l'ensemble des fonctions polynomiales de \mathbb{R} dans \mathbb{R} : Une fonction polynomiale de \mathbb{R} dans \mathbb{R} , de degré inférieur ou égal à n , est une fonction P telle qu'il existe des coefficients $c_0, \dots, c_n \in \mathbb{R}$ satisfaisant

$$P(x) = c_n x^n + \dots + c_1 x + c_0, \quad \forall x \in \mathbb{R}.$$

On peut démontrer que pour une telle fonction, les coefficients c_0, \dots, c_n sont uniques. Une fonction est polynomiale, s'il existe $n \in \mathbb{N}$ tel qu'elle soit polynomiale de degré inférieur ou égal à n . Si la somme de telles fonctions et la multiplication scalaire sont définies par (1.3), alors \mathcal{P} est un espace vectoriel.

4. L'ensemble $C_0(\mathbb{R})$ des fonctions continues de \mathbb{R} dans \mathbb{R} , muni des opérations définies par (1.3), est un espace vectoriel.
5. L'ensemble \mathcal{P}_2 des fonctions polynomiales de degré inférieur ou égal à 2, muni encore des mêmes opérations, est un espace vectoriel.
6. L'ensemble $F = \{f \in \mathcal{F}(\mathbb{R}) : f(1) = 0\}$, muni des opérations (1.3), est un espace vectoriel ;
7. L'ensemble $F = \{f \in \mathcal{F}(\mathbb{R}) : f(1) = 3\}$, muni des opérations (1.3), n'est pas un espace vectoriel, car l'addition n'est pas interne (par exemple).

Passons maintenant à l'exemple fondamental pour la suite du cours.

Exemple 1.2.4 (L'espace vectoriel \mathbb{R}^n). L'ensemble \mathbb{R}^n est fait des n -uplets de nombres que l'on note verticalement

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

L'addition et la multiplication scalaires sont définies de la manière la plus simple qui soit. On définit

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} \quad \text{et} \quad \lambda \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}$$

pour tous $\lambda, x_1, \dots, x_n, y_1, \dots, y_n \in \mathbb{R}$. On démontre facilement que cet ensemble muni de ces opérations est un espace vectoriel.

Quand on considérera \mathbb{R}^n , sans mentionner les opérations, ce sera toujours celles qui viennent d'être définies qui seront prises en considération. De plus quand aucune confusion ne pourra en

résulter, afin de gagner de la place, on notera plutôt $(x_1, \dots, x_n) \sim$ l'élément $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ de \mathbb{R}^n .

L'opération \sim , appelée transposition, correspond donc simplement à transformer une ligne en une colonne (et vice-versa).

Dans la même veine, nous pouvons considérer l'espace vectoriel des matrices.

Exemple 1.2.5 (L'espace vectoriel \mathbb{R}_q^p). Cet exemple n'est pas fondamentalement différent du précédent, puisque les éléments de l'espace vectoriel considéré sont des tableaux de nombres. Une *matrice* de type (p, q) est un tableau rectangulaire de nombre réels, ayant p lignes de nombres et q colonnes.^e Alors \mathbb{R}_q^p est l'ensemble de toutes ces matrices. Voici quelques exemples plus concrets. On a

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in \mathbb{R}_3^2, \quad B = \begin{pmatrix} 1 & 2 & 3 \\ 0 & \pi & e \\ \sqrt{2} & \sqrt{3} & 2 \end{pmatrix} \in \mathbb{R}_3^3, \quad C = (1 \ 2 \ 3) \in \mathbb{R}_3^1.$$

Pour tout $A \in \mathbb{R}_q^p$ et tout $i \leq p, j \leq q$, l'élément i, j de la matrice A , noté $(A)_{i,j}$ (ou simplement $A_{i,j}$), est le nombre situé sur i -ème ligne et la j -ème colonne de A . Par exemple, on a pour les matrices définies ci-dessus

$$(A)_{1,3} = 3, \quad (B)_{3,2} = \sqrt{3}, \quad (B)_{2,3} = e, \dots$$

Une matrice quelconque A de \mathbb{R}_q^p a donc la forme

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,q} \\ \vdots & \ddots & \vdots \\ a_{p,1} & \cdots & a_{p,q} \end{pmatrix},$$

e. On dit aussi simplement une matrice à p lignes et q colonnes.

et on a alors $(A)_{i,j} = a_{i,j}$. Pour gagner de la place, on note, quand c'est possible A par $(a_{i,j})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}}$. Notons encore que les matrices n'ayant qu'une seule ligne sont appelées matrices lignes ou vecteurs lignes, et celles n'ayant qu'une colonne sont appelées matrices colonnes ou vecteurs colonnes.

Les opérations d'espace vectoriel ne sont pas différentes de celles introduites pour \mathbb{R}^n : on additionne et on multiplie par un nombre "élément à élément". Plus formellement, on a

$$(A + B)_{i,j} = (A)_{i,j} + (B)_{i,j}, \quad (\lambda A)_{i,j} = \lambda(A)_{i,j}$$

pour tous $A, B \in \mathbb{R}_q^p$ et $\lambda \in \mathbb{R}$. On démontre comme dans l'exemple précédent qu'il s'agit bien d'un espace vectoriel.

1.2.1 Une représentation intuitive et commode

Rappelons que les couples de nombres réels sont classiquement représentés dans l'enseignement secondaire au moyen d'un système d'axes. On trace deux droites perpendiculaires et orientées a_1 et a_2 , qui se coupent en O . On porte sur ces droites les mêmes unités. Le couple $(a, b)^\sim$ est alors représenté par le point A obtenu en portant la valeur a sur le premier axe et la valeur b sur le second, et en traçant des parallèles aux axes comme indiqué à la figure 1.1. Nous pouvons utiliser cette représentation graphique et convenir de représenter l'élément $(a, b)^\sim$ de l'espace vectoriel \mathbb{R}^2 par une flèche de O à A , comme indiqué sur cette même figure.

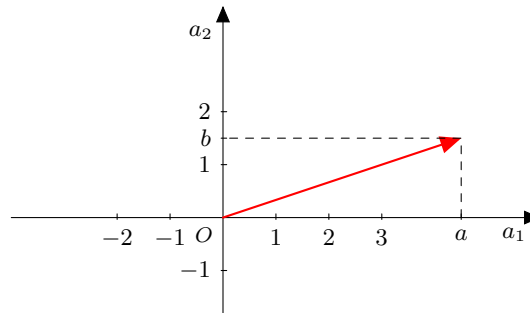


FIGURE 1.1 – Une représentation graphique du vecteur $(2, 3)^\sim$ de \mathbb{R}^2 .

Attention ! Aucune des notions qui ont été utilisées pour obtenir cette représentation graphique n'a encore été définie dans ce cours. Cette représentation n'est donc pas vraiment définie non plus et elle ne le sera que dans le prochain chapitre. Elle permet cependant de guider notre intuition, par exemple quand nous définirons les droites vectorielles (il est utile pour cela de représenter $2(a, b)^\sim$ sur ce même graphique).

Cette représentation graphique peut s'étendre tant bien que mal pour les éléments de \mathbb{R}^3 : on trace trois droites perpendiculaires deux à deux et orientées, en utilisant une certaine perspective. On y marque les mêmes unités, et cela permet de représenter graphiquement les éléments de \mathbb{R}^3 par des flèches.

1.2.2 Combinaisons linéaires

En utilisant les deux opérations qui définissent un espace vectoriel E , à partir de vecteurs u_1, \dots, u_p et de coefficients réels $\lambda_1, \dots, \lambda_p$, on peut définir le vecteur

$$u = \lambda_1 u_1 + \dots + \lambda_p u_p.$$

Définition 1.2.6. Le vecteur défini par

$$u = \lambda_1 u_1 + \dots + \lambda_p u_p$$

est appelé *combinaison linéaire* des vecteurs u_1, \dots, u_p , avec les coefficients $\lambda_1, \dots, \lambda_p$.

Par exemple, dans \mathbb{R}^2 , si on choisit

$$u_1 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad u_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \lambda_1 = 2, \quad \lambda_2 = 3, \quad \lambda_3 = -1,$$

on a directement

$$\lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 u_3 = 2u_1 + 3u_2 - u_3 = \begin{pmatrix} 3 \\ 9 \end{pmatrix}.$$

Notons qu'a priori, un même vecteur peut s'écrire de plusieurs façons comme combinaison linéaire de u_1, \dots, u_p . Dans l'exemple précédent, on a aussi $u = 9u_2 + 3u_3$. Je laisse au lecteur le soin de trouver une expression de u comme combinaison linéaire de u_1 et u_3 par exemple. Il est aussi utile de déterminer toutes les façons d'exprimer u comme combinaison linéaire de u_1, u_2 et u_3 , ou comme combinaison linéaire de u_1 et u_3 .

Voici encore quelques exemples.

Exemple 1.2.7. 1. Dans \mathbb{R}^3 , si $u_1 = (1, 2, 3)^\sim$, $u_2 = (-1, 0, 1)^\sim$ et $u_3 = (4, 2, 1)^\sim$, alors on a

$$3u_1 - u_2 + u_3 = \begin{pmatrix} 8 \\ 8 \\ 9 \end{pmatrix}.$$

2. Dans $(]0, +\infty[^2, \oplus, \odot)$, si $u_1 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ et $u_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, alors on a

$$3 \odot u_1 \oplus 2 \odot u_2 = \begin{pmatrix} 2^3 \\ 3^3 \end{pmatrix} \oplus \begin{pmatrix} 1 \\ 2^2 \end{pmatrix} = \begin{pmatrix} 8 \\ 27 \end{pmatrix} \oplus \begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 8 \\ 108 \end{pmatrix}.$$

1.3 Sous-espaces vectoriels

La définition d'espace vectoriel est parfois assez longue à mettre en oeuvre : il faut vérifier un certain nombre de propriétés.

La construction que nous introduisons ici permet d'obtenir de nombreux nouveaux exemples d'espaces vectoriels, assez facilement. De plus, les sous-espaces vectoriels, puisque c'est de cela qu'il s'agit, vont nous permettre dans le prochain chapitre, de modéliser la généralisation des droites et plans bien connus en géométrie.

Enfin, c'est un premier contact avec la construction de sous-structure, très présente en mathématique, et que le lecteur rencontrera sans doute plus tard (avec les sous-groupes, sous-algèbres, sous-variétés, etc...)

Définition 1.3.1. Soit E un espace vectoriel. Un sous-espace vectoriel de E est une partie non vide V de E qui contient les combinaisons linéaires de ses éléments.

En d'autres termes, une partie V non vide de E est un sous-espace vectoriel si, et seulement si,

$$\left\{ \begin{array}{l} p \in \mathbb{N}_0 \\ \lambda_1, \dots, \lambda_p \in \mathbb{R} \\ u_1, \dots, u_p \in V \end{array} \right\} \Rightarrow \lambda_1 u_1 + \dots + \lambda_p u_p \in V.$$

Avant de présenter des exemples, il est commode de donner une caractérisation plus simple des sous-espaces vectoriels.

Proposition 1.3.2. Une partie non vide V d'un espace vectoriel E est un sous-espace vectoriel si, et seulement si, pour tous $u, v \in V$ et tout $\lambda \in \mathbb{R}$, λu et $u + v$ appartiennent à V .

Cette proposition nous permet essentiellement de simplifier les notations et les conditions à vérifier pour démontrer qu'une partie V est un sous-espace vectoriel.

Démonstration. Si V est un sous-espace vectoriel, alors pour tout $u, v \in V$, $u + v$ est une combinaison linéaire particulière de u et v , et est donc un élément de V . Il en va de même pour λu pour tout $\lambda \in \mathbb{R}$.

Réciproquement, si V satisfait les conditions de l'énoncé, alors pour tous $p \in \mathbb{N}_0$, $\lambda_1, \dots, \lambda_p \in \mathbb{R}$ et tous $u_1, \dots, u_p \in V$, on a $\lambda_1 u_1, \dots, \lambda_p u_p \in V$. On a ensuite

$$\lambda_1 u_1 + \dots + \lambda_p u_p = (\dots (\lambda_1 u_1 + \lambda_2 u_2) + \lambda_3 u_3) + \dots) + \lambda_n u_n.$$

En appliquant $p - 1$ fois la condition sur la somme donnée dans l'énoncé, on conclut que cette combinaison linéaire est dans V . \square

Remarquons qu'étant donné une partie V de E , pour démontrer qu'il s'agit d'un sous-espace vectoriel, il faut d'abord démontrer qu'elle est non vide. C'est en fait très simple : si V est un sous-espace vectoriel, alors pour $u \in V$, $0u = 0$ doit aussi être dans V . Donc tout sous-espace vectoriel contient le vecteur nul. C'est aussi une façon de démontrer qu'un ensemble n'est pas un sous-espace vectoriel. Voici quelques exemples.

Exemple 1.3.3. Soit $E = \mathbb{R}^3$ muni de sa structure d'espace vectoriel standard. Alors

$$V_1 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in E : x_1 + 2x_2 - 4x_3 = 0 \right\}$$

est un sous-espace vectoriel de E . De même l'ensemble

$$V_2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in E : \begin{cases} x_1 + 2x_2 - 4x_3 = 0 \\ x_1 - 5x_2 = 0 \end{cases} \right\}$$

est un sous-espace vectoriel de \mathbb{R}^3 .

Dans ce type d'exemple, le sous-espace vectoriel est décrit au moyen d'*équations*. Remarquons qu'il est crucial que le second membre des équations soit nul (on dit que les équations sont *homogènes*), sinon le vecteur nul ne serait pas dans le sous-ensemble.

Exemple 1.3.4. Soit E un espace vectoriel et $u \in E$. Alors,

$$V_3 = \{\lambda u : \lambda \in \mathbb{R}\}$$

est un sous-espace vectoriel de E . De même, si u et v sont dans E , alors

$$V_4 = \{\lambda u + \mu v : \lambda, \mu \in \mathbb{R}\}$$

est un sous-espace vectoriel de E .

Les sous-espaces V_3 et V_4 sont définis à partir de certains vecteurs. Chaque valeur des *paramètres* λ et μ définit alors un élément du sous-espace. Nous verrons d'ici peu que, quand E est de dimension finie (voir la section 1.4), tout sous-espace vectoriel admet les deux descriptions ci-dessus. Nous verrons également comment passer d'une description à l'autre. Voici encore un exemple moins simple.

Exemple 1.3.5. Considérons l'espace vectoriel $\mathcal{F}(\mathbb{R})$ des fonctions de \mathbb{R} dans \mathbb{R} , muni de son addition et de sa multiplication scalaire habituelles (voir l'exemple 1.2.3). L'ensemble $C_0(\mathbb{R})$ des fonctions continues de \mathbb{R} dans \mathbb{R} n'est pas vide : il contient la fonction nulle. On sait que la somme et les multiples de fonctions continues sont des continues. Donc $C_0(\mathbb{R})$ est un sous-espace vectoriel de $\mathcal{F}(\mathbb{R})$. On montre de la même façon que l'ensemble \mathcal{P} des fonctions polynomiales de \mathbb{R} dans \mathbb{R} est un sous-espace vectoriel de $\mathcal{F}(\mathbb{R})$ et de $C_0(\mathbb{R})$.

Exemple 1.3.6. Le sous-ensemble $F = \{f : \mathbb{R} \rightarrow \mathbb{R} : f(2) = 3\}$ n'est pas un sous-espace vectoriel de $\mathcal{F}(\mathbb{R})$.

Proposition 1.3.7. Si $(E, +, 0, \cdot)$ est un espace vectoriel et si V est un sous-espace vectoriel de E , alors $(V, +, 0, \cdot)$ est un espace vectoriel.

Démonstration. Il suffit de vérifier toutes les conditions de la définition. Elles sont satisfaites dans V car elles le sont dans E . \square

Ce résultat permet de construire des exemples d'espaces vectoriels, comme annoncé dans l'introduction de cette section. On constate alors que si V est un sous-espace vectoriel de E , alors V est un espace vectoriel inclus à E . Mais la réciproque n'est pas vraie : ce n'est pas suffisant pour définir un sous-espace vectoriel. Il est important de remarquer que les opérations dans V sont les restrictions des opérations correspondantes dans E . Ainsi, si on considère $E = \mathbb{R}^2$, muni des opérations standards et $V =]0, +\infty[^2$ muni des opérations \oplus et \odot (voir l'exemple 1.2.3), on a $V \subset E$ et V est lui-même un espace vectoriel, mais V n'est pas un sous-espace vectoriel : les multiples (au sens des opérations de E) des éléments de V ne sont pas tous dans V . Cela est dû au fait que les opérations de V ne sont pas les restrictions de celles de E .

Passons maintenant à des constructions classiques de sous-espaces vectoriels, que sont l'intersection, l'enveloppe linéaire, et la somme.

Définition 1.3.8. Si V_1 et V_2 sont des sous-espaces vectoriels de E , l'intersection de V_1 et V_2 est l'ensemble

$$V_1 \cap V_2 = \{x \in E : x \in V_1 \text{ et } x \in V_2\}.$$

On a alors le résultat suivant.

Proposition 1.3.9. Soient V_1 et V_2 des sous-espaces vectoriels de E . Alors $V_1 \cap V_2$ est un sous-espace vectoriel de E .

Démonstration. Considérons $u, v \in V_1 \cap V_2$ et $\lambda \in \mathbb{R}$, alors λu appartient à V_1 et à V_2 car ce sont des sous-espaces vectoriels, donc λu appartient à $V_1 \cap V_2$. On montre de la même façon que $u + v$ appartient à $V_1 \cap V_2$. \square

On pourrait être tenté de faire de même pour l'union de sous-espaces vectoriels, mais en général, si V_1 et V_2 sont des sous-espaces vectoriels de E , $V_1 \cup V_2$ n'est pas un sous-espace vectoriel. Par exemple, dans \mathbb{R}^2 , on peut considérer $V_1 = \{(x, 0)^\sim : x \in \mathbb{R}\}$ et $V_2 = \{(0, y)^\sim : y \in \mathbb{R}\}$. Ce sont des sous-espaces vectoriels. Mais par exemple $(3, 0)^\sim$ et $(0, 4)^\sim$ appartiennent à $V_1 \cup V_2$, mais leur somme $(3, 4)^\sim$ n'est ni dans V_1 , ni dans V_2 . L'union de V_1 et V_2 est donc trop petite pour être un sous-espace vectoriel. Si on veut trouver un sous-espace vectoriel contenant V_1 et V_2 , on doit inventer une nouvelle construction.

Définition 1.3.10. Soit E un espace vectoriel et A une partie non vide de E . Le sous-espace vectoriel engendré par A , ou l'enveloppe linéaire de A est l'ensemble

$$\rangle A \langle = \{\lambda_1 a_1 + \dots + \lambda_p a_p : p \in \mathbb{N}, \lambda_1, \dots, \lambda_p \in \mathbb{R}, a_1, \dots, a_p \in A\}.$$

Cette définition est justifiée par le résultat suivant.

Proposition 1.3.11. Pour toute partie A non vide de E , $\rangle A \langle$ est un sous-espace vectoriel de E . Il contient A et est inclus dans tout sous-espace vectoriel V contenant A . C'est l'unique sous-espace vectoriel ayant cette propriété.

Démonstration. Si u est un élément de $\rangle A \langle$, alors il existe $\lambda_1, \dots, \lambda_p$ réels et $a_1, \dots, a_p \in A$ tels que $u = \lambda_1 a_1 + \dots + \lambda_p a_p$. Si λ est réel, alors

$$\lambda u = \lambda(\lambda_1 a_1 + \dots + \lambda_p a_p) = (\lambda \lambda_1) a_1 + \dots + (\lambda \lambda_p) a_p$$

appartient aussi à $\langle A \rangle$. Si v est aussi un élément de $\langle A \rangle$, il existe $b_1, \dots, b_q \in A$ et $\mu_1, \dots, \mu_q \in \mathbb{R}$ tels que $v = \mu_1 b_1 + \dots + \mu_q b_q$. On a donc

$$u + v = \lambda_1 a_1 + \dots + \lambda_p a_p + \mu_1 b_1 + \dots + \mu_q b_q$$

et on constate que c'est un élément de $\langle A \rangle$. On a également $A \subset \langle A \rangle$, puisque pour tout $u \in A$, on a $u = 1 \cdot u$. Puisque A n'est pas vide, $\langle A \rangle$ ne l'est pas non plus. Nous avons donc démontré que $\langle A \rangle$ est un sous-espace vectoriel qui contient A . Si V est un sous-espace vectoriel et contient A , alors il contient les combinaisons linéaires des éléments de A et donc il contient $\langle A \rangle$. Enfin, l'unicité est évidente. \square

On résume généralement cette propriété en disant que $\langle A \rangle$ est *le plus petit sous-espace vectoriel contenant A* .

Nous pouvons maintenant définir la somme de deux sous-espaces vectoriels.

Définition 1.3.12. Soient V_1 et V_2 deux sous-espaces vectoriels d'un espace vectoriel E . La somme de V_1 et V_2 , notée $V_1 + V_2$ est l'ensemble

$$V_1 + V_2 = \{u_1 + u_2 : u_1 \in V_1, u_2 \in V_2\}.$$

La proposition suivante fait alors le lien avec la notion d'enveloppe linéaire.

Proposition 1.3.13. Si V_1 et V_2 sont deux sous-espaces vectoriels d'un espace vectoriel E , alors $V_1 + V_2$ est un sous-espace vectoriel de E . On a de plus $V_1 + V_2 = \langle V_1 \cup V_2 \rangle$.

Démonstration. On montre comme d'habitude que $V_1 + V_2$ est un sous-espace vectoriel. Il contient V_1 car 0 appartient à V_2 et il contient aussi V_2 . De plus tout sous-espace vectoriel contenant V_1 et V_2 contient visiblement $V_1 + V_2$. Cela suffit à cause de l'unicité de $\langle V_1 \cup V_2 \rangle$. \square

Dans ce qui suit, nous utiliserons très souvent la notion d'enveloppe linéaire d'un ensemble fini. Si $A = \{u_1, \dots, u_n\}$, on note l'enveloppe linéaire $\langle u_1, \dots, u_n \rangle$ au lieu de $\langle \{u_1, \dots, u_n\} \rangle$. On peut alors facilement montrer l'égalité

$$\langle u_1, \dots, u_n \rangle = \left\{ \sum_{i=1}^n \lambda_i u_i : \lambda_1, \dots, \lambda_n \in \mathbb{R} \right\}.$$

Enfin, terminons sur un cas particulier de somme de sous-espaces vectoriels.

Définition 1.3.14. On dit que la somme des sous-espaces vectoriels V_1 et V_2 est directe si $V_1 \cap V_2 = \{0\}$. On note alors cette somme $V_1 \oplus V_2$.

Ce cas est particulier à cause de la proposition suivante.

Proposition 1.3.15. Soient V_1 et V_2 deux sous-espaces vectoriels de E . La somme de V_1 et V_2 est directe si, et seulement si, tout vecteur u de $V_1 + V_2$ se décompose de manière unique en $u = u_1 + u_2$ où $u_1 \in V_1$ et $u_2 \in V_2$.

Démonstration. Supposons que l'on a $V_1 \cap V_2 = \{0\}$ et soit $u \in V_1 + V_2$. Par définition de $V_1 + V_2$, il existe $u_1 \in V_1$ et $u_2 \in V_2$ tels que $u = u_1 + u_2$. Supposons que l'on ait aussi $u = u'_1 + u'_2$ avec $u'_1 \in V_1$ et $u'_2 \in V_2$. De l'égalité $u_1 + u_2 = u'_1 + u'_2$, on tire alors $u_1 - u'_1 = u'_2 - u_2$. Ce vecteur est dans $V_1 \cap V_2$, donc il est nul, et on a donc $u_1 = u'_1$ et $u_2 = u'_2$.

Passons à la réciproque. Supposons que tout vecteur de $V_1 + V_2$ se décompose de manière unique en somme d'un élément de V_1 et d'un élément de V_2 et supposons qu'il existe un vecteur v non nul dans $V_1 \cap V_2$. On a alors deux décompositions possibles de 0 : en effet

$$0 = 0 + 0, \quad \text{et} \quad 0 = v + (-v).$$

C'est contraire à l'hypothèse. \square

f. On eut aussi démontrer qu'alors tout élément de $V_1 + V_2$ se décompose d'au moins deux façons comme somme d'un élément de V_1 et d'un élément de V_2 .

1.4 Bases et dimension

Nous introduisons ici les notions fondamentales d'ensembles de vecteurs linéairement indépendants et générateurs d'un espace vectoriel. Nous définissons ces notions pour des ensembles quelconques, mais pour que l'exposé reste simple, nous nous limiterons assez rapidement à des sous ensembles contenant un nombre fini d'éléments.

Définition 1.4.1. Un ensemble de vecteurs $G \subset E$ est une *partie génératrice* de E si tout vecteur de E est une combinaison linéaire des éléments de G .

En d'autres termes, G est une partie génératrice de E si, et seulement si $\langle G \rangle = E$.

Exemple 1.4.2. Voici quelques exemples qu'on traitera comme exercice. On commence avec \mathbb{R}^2 :

1. Dans \mathbb{R}^2 , les vecteurs $u = (1, 0)^\sim$ et $v = (0, 1)^\sim$ forment une partie génératrice (i.e. $G = \{u, v\}$ est une partie génératrice).
2. Dans \mathbb{R}^2 , les vecteurs $u = (1, 2)^\sim$ et $v = (0, 1)^\sim$ forment une partie génératrice.
3. Dans \mathbb{R}^2 , les vecteurs $u = (1, 2)^\sim$, $v = (0, 1)^\sim$ et $w = (3, 7)^\sim$ forment une partie génératrice.
4. Dans \mathbb{R}^2 , les vecteurs $u = (1, 2)^\sim$ et $v = (4, 5)^\sim$ forment une partie génératrice.
5. Dans \mathbb{R}^2 , les vecteurs $u = (1, 2)^\sim$ et $v = (2, 4)^\sim$ ne forment pas une partie génératrice.

Voici quelques exemples dans \mathbb{R}^3 :

1. Dans \mathbb{R}^3 , les vecteurs $u = (1, 0, 0)^\sim$ et $v = (0, 1, 0)^\sim$ ne forment pas une partie génératrice.
2. Dans \mathbb{R}^3 , les vecteurs $u = (1, 0, 0)^\sim$, $v = (0, 1, 0)^\sim$ et $w = (0, 0, 1)^\sim$ forment une partie génératrice.
3. Dans \mathbb{R}^3 , les vecteurs $u = (1, 2, 0)^\sim$, $v = (0, 1, 3)^\sim$ et $w = (0, 0, 1)^\sim$ forment une partie génératrice.

L'intérêt d'une partie génératrice est qu'elle permet de remplacer des vecteurs abstraits, ou des vecteurs d'espaces vectoriels bizarres par des nombres, en particulier si on a une partie génératrice contenant un nombre fini d'éléments : Si $G = \{u_1, \dots, u_p\}$ est une partie génératrice de E , on fixe un ordre dans ces vecteurs, obtenant ainsi la *famille* (u_1, \dots, u_p) . On sait alors que tout vecteur u de E se décompose en

$$u = \lambda_1 u_1 + \dots + \lambda_p u_p,$$

et on associe au vecteur u le p -uple de nombres $(\lambda_1, \dots, \lambda_p)^\sim \in \mathbb{R}^p$.

Cette remarque vaut bien une définition.

Définition 1.4.3. Un espace vectoriel E est de dimension finie s'il admet une partie génératrice contenant un nombre fini d'éléments.

Nous avons déjà vu quelques exemples. Voici un contre-exemple.

Exemple 1.4.4. L'espace vectoriel des fonctions polynomiales n'est pas de dimension finie.

Voici maintenant un exemple simple qui montre que l'association que nous avons introduite ci-dessus est imparfaite.

Exemple 1.4.5. Dans l'espace vectoriel \mathbb{R}^2 , les vecteurs

$$u_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad \text{et} \quad u_3 = \begin{pmatrix} 5 \\ 4 \end{pmatrix}.$$

forment une partie génératrice. En effet, on peut décomposer tout élément $\begin{pmatrix} a \\ b \end{pmatrix}$ de \mathbb{R}^2 comme combinaison linéaire de ces vecteurs : la condition

$$\begin{pmatrix} a \\ b \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 2 \\ 2 \end{pmatrix} + z \begin{pmatrix} 5 \\ 4 \end{pmatrix}$$

est équivalente à

$$\begin{cases} a &= x + 2y + 5z \\ b &= 2y + 4z, \end{cases}$$

où les inconnues sont x, y, z . On peut résoudre ce système linéaire, et on obtient la forme la plus simple de solution :

$$\begin{cases} x &= a - b - z \\ y &= \frac{b-4z}{2}. \end{cases}$$

Pour toute valeur de z , on obtient une valeur de x et y qui satisfont les équations. Par exemple, pour $a = 2$ et $b = 6$, on peut prendre $z = 0$ et on obtient

$$\begin{pmatrix} 2 \\ 6 \end{pmatrix} = -4 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 2 \\ 2 \end{pmatrix} = -4u_1 + 3u_2 + 0u_3.$$

Pour z quelconque, on aura

$$\begin{pmatrix} 2 \\ 6 \end{pmatrix} = (-4 - z) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + (3 - 2z) \begin{pmatrix} 2 \\ 2 \end{pmatrix} + z \begin{pmatrix} 5 \\ 4 \end{pmatrix} = (-4 - z)u_1 + (3 - 2z)u_2 + zu_3.$$

On pourrait donc associer à $(2, 6)^\sim$ le triplet $(-4, 3, 0)^\sim$, mais aussi le triplet $(-5, 1, 1)^\sim$, ou une infinité d'autres triplets.

Exemple 1.4.6. Toujours dans \mathbb{R}^2 , les vecteurs

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{et} \quad v_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

forment une partie génératrice. En effet, tout vecteur $(a, b)^\sim$ de \mathbb{R}^2 se décompose comme combinaison linéaire de v_1 et v_2 , car la condition

$$\begin{pmatrix} a \\ b \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

est équivalente à

$$\begin{cases} a &= x + 2y \\ b &= y \end{cases}$$

On trouve facilement

$$\begin{cases} x &= a - 2b \\ y &= b \end{cases}$$

Contrairement à l'exemple précédent, la solution est ici unique. Pour reprendre le même cas particulier, on a

$$\begin{pmatrix} 2 \\ 6 \end{pmatrix} = -10 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 6 \begin{pmatrix} 2 \\ 1 \end{pmatrix} = -10v_1 + 6v_2.$$

La différence entre les deux exemples est que dans le premier, il existe une *dépendance linéaire* entre les vecteurs u_1, u_2 et u_3 . En effet, on a

$$u_1 + 2u_2 - u_3 = 0.$$

C'est ce qui permet d'ajouter cette combinaison linéaire à n'importe quel vecteur sans en changer la valeur. *On a donc une combinaison linéaire nulle de u_1, u_2, u_3 , dont les coefficients ne sont pas tous nuls.*

Cette *dépendance linéaire* s'exprime aussi par le fait qu'un des vecteurs dépend linéairement des deux autres, puisqu'on a

$$u_3 = u_1 + 2u_2.$$

Par contre, il n'y a pas de dépendance linéaire entre les vecteurs du deuxième exemple. En effet, on remarque tout de suite que v_1 n'est pas multiple de v_2 et v_2 n'est pas multiple de v_1 non plus. Si on essaie d'obtenir le vecteur nul comme combinaison linéaire de ces vecteurs, on obtient l'équation

$$\lambda_1 v_1 + \lambda_2 v_2 = 0,$$

où λ_1, λ_2 sont réels. On développe cette équation pour obtenir

$$\begin{cases} \lambda_1 + 2\lambda_2 & = & 0 \\ \lambda_2 & = & 0 \end{cases}$$

qui donne l'unique solution $\lambda_1 = \lambda_2 = 0$. La seule combinaison linéaire nulle que l'on peut former avec v_1 et v_2 est donc celle dont les coefficients sont nuls.

Définition 1.4.7. Une famille (non vide^g) de vecteurs $D \subset E$ est linéairement dépendante s'il existe une combinaison linéaire nulle de vecteurs de D dont les coefficients ne sont pas tous nuls. Dans ces cas, on dit aussi que les éléments de D sont *linéairement dépendants*.

Une famille L de E qui n'est pas linéairement dépendante est dite *libre* ou *linéairement indépendante*. Les éléments de L sont alors dits *linéairement indépendants*.

Il est utile de faire le lien entre cette définition et le fait qu'un des vecteurs de la famille soit combinaison linéaire des autres, comme nous l'avons fait plus haut.

Proposition 1.4.8. Une famille D de vecteurs de E est linéairement dépendante si, et seulement si, l'un des vecteurs de D est combinaison linéaire des autres.

Démonstration. Soit $D \subset E$ une famille linéairement dépendante. Il existe alors des vecteurs u_1, \dots, u_n de D et des coefficients **non tous nuls** $\lambda_1, \dots, \lambda_n$ tels que

$$\lambda_1 u_1 + \dots + \lambda_n u_n = 0.$$

Alors si $\lambda_i \neq 0$, on a

$$u_i = -\frac{1}{\lambda_i}(\lambda_1 u_1 + \dots + \widehat{i} + \dots + \lambda_n u_n).$$

Donc l'un des vecteurs est combinaison linéaire des autres.

Réciproquement, si $u \in D$ est combinaison linéaire des vecteurs u_1, \dots, u_n de D , il existe des nombres réels $\lambda_1, \dots, \lambda_n$ tels que

$$u = \lambda_1 u_1 + \dots + \dots + \lambda_n u_n.$$

En posant $\lambda_0 = -1$, on obtient

$$\lambda_0 u + \lambda_1 u_1 + \dots + \lambda_n u_n = 0,$$

avec des coefficients qui ne sont pas tous nuls. □

On obtient également, par **contraposition**, le résultat suivant concernant l'indépendance linéaire.

Proposition 1.4.9. Une famille L de vecteurs de E est linéairement indépendante si, et seulement si, pour $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, et tous $u_1, \dots, u_n \in L$, l'égalité

$$\lambda_1 u_1 + \dots + \lambda_n u_n = 0$$

a lieu **seulement** si on a $\lambda_1 = \dots = \lambda_n = 0$.

^g. Dans la suite, sauf mention explicite du contraire, nous ne travaillerons qu'avec des familles non vides

Nous pouvons maintenant passer à la définition du concept de *base* d'un espace vectoriel.

Définition 1.4.12. Une base d'un espace vectoriel E est une famille de vecteurs qui est à la fois libre et génératrice.

Regardons quelques exemples, en commençant par celui qui sera fondamental dans la suite, car c'est le plus simple.

Exemple 1.4.13 (Base canonique de \mathbb{R}^n). Dans l'espace vectoriel \mathbb{R}^n , les vecteurs

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

forment une base, appelée *base canonique de \mathbb{R}^n* .

Exemple 1.4.14. 1. Dans \mathbb{R}^2 , les vecteurs $u_1 = (1, 0)^\sim$ et $u_2 = (1, 1)^\sim$ forment une base $\mathcal{B} = (u_1, u_2)$.

2. Dans \mathbb{R}^3 , les vecteurs $u_1 = (1, 0, 1)^\sim$, $u_2 = (0, 1, 1)^\sim$ et $(0, 0, 1)^\sim$ forment une base $\mathcal{B} = (u_1, u_2, u_3)$.

3. Les mêmes vecteurs forment une autre base $\mathcal{B} = (u_2, u_1, u_3)$.

4. Dans l'espace vectoriel \mathcal{P}_2 des polynômes de degré inférieur ou égal à 2, les fonctions P_0, P_1, P_2 définies par

$$P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = x^2$$

pour tout $x \in \mathbb{R}$ forment une base $\mathcal{B} = (P_0, P_1, P_2)$ ou une base $\mathcal{B}' = (P_2, P_1, P_0)$.

5. Dans le même espace vectoriel, les fonctions Q_0, Q_1, Q_2 définies par

$$Q_0(x) = 1, \quad Q_1(x) = x - 1, \quad Q_2(x) = (x - 1)^2$$

pour tout $x \in \mathbb{R}$ forment une base $\mathcal{B} = (Q_0, Q_1, Q_2)$.

6. Dans l'espace \mathcal{P} des fonctions polynomiales de \mathbb{R} dans \mathbb{R} , on définit pour tout $n \in \mathbb{N}$ la fonction P_n par

$$P_n(x) = x^n, \quad \forall x \in \mathbb{R}.$$

L'ensemble ordonné (la suite) $(P_n : n \in \mathbb{N})$ est alors une base de \mathcal{P} . On peut démontrer que \mathcal{P} n'est pas de dimension finie.

7. Dans l'espace vectoriel $(]0, +\infty[^2, \oplus, \odot)$, les vecteurs

$$u_1 = \begin{pmatrix} e \\ 1 \end{pmatrix}, \quad \text{et} \quad u_2 = \begin{pmatrix} 1 \\ e \end{pmatrix}$$

forment une base $\mathcal{B} = (u_1, u_2)$.

La définition de base est fondamentale à cause du résultat suivant.

Proposition 1.4.15. Une famille \mathcal{B} est une base de E si, et seulement si, tout élément de E s'écrit de manière unique comme combinaison linéaire des éléments de \mathcal{B} .

Démonstration. Soit \mathcal{B} une base de E . Puisque \mathcal{B} est une partie génératrice, tout élément u de E s'écrit au moins d'une façon comme combinaison linéaire de vecteurs u_1, \dots, u_n appartenant à \mathcal{B} . On a donc

$$u = \lambda_1 u_1 + \dots + \lambda_n u_n.$$

Supposons maintenant qu'un élément u s'écrive de deux façons distinctes comme combinaison linéaire d'éléments de \mathcal{B} . On a alors

$$u = \lambda_1 u_1 + \cdots + \lambda_n u_n, \quad \text{et} \quad u = \mu_1 u'_1 + \cdots + \mu_m u'_m$$

avec au moins un vecteur différent dans la deuxième combinaison, ou un coefficient différent. On en déduit

$$0 = u - u = (\lambda_1 u_1 + \cdots + \lambda_n u_n) - (\mu_1 u'_1 + \cdots + \mu_m u'_m).$$

On a donc une combinaison nulle des vecteurs $u_1, \dots, u_n, u'_1, \dots, u'_m$ dont les coefficients ne sont pas tous nuls, ce qui contredit l'hypothèse.

Passons maintenant à la réciproque. Nous supposons qu'une partie \mathcal{B} satisfait les conditions de l'énoncé et nous montrons que c'est une base. C'est une partie génératrice par hypothèse. C'est également une partie libre. En effet, supposons que ce ne soit pas le cas. Alors par la proposition 1.4.8, il existe des éléments u_0, u_1, \dots, u_n de \mathcal{B} tels que

$$u_0 = \lambda_1 u_1 + \cdots + \lambda_n u_n.$$

Alors u_0 s'écrit de deux manières différentes comme combinaison linéaire des éléments de \mathcal{B} , ce qui contredit l'hypothèse. \square

Voici encore quelques résultats et définitions utiles.

Proposition 1.4.16. *Soit E un espace vectoriel de dimension finie^j. Alors E admet une base contenant un nombre fini d'éléments. Toutes les bases de E contiennent le même nombre d'éléments.*

Démonstration. Soit $G = \{u_1, \dots, u_n\}$ une partie génératrice finie. Si (u_1, \dots, u_n) est une famille libre, alors c'est une base, par définition. Si elle n'est pas libre, alors un des éléments, disons u_i , est combinaison linéaire des autres. On peut le retirer pour obtenir $G_{(1)} = (u_1, \dots, \hat{u}_i, \dots, u_n)$. C'est encore une partie génératrice par la proposition 1.4.10. Si ce n'est pas une base, on lui retire encore un élément. On procède de la même façon jusqu'à obtenir une base. Ce processus s'arrête puisqu'on a un nombre fini de vecteurs.

Soit une base $\mathcal{B} = (u_1, \dots, u_n)$ et \mathcal{B}' une autre base. Par le théorème de Steinitz, si l'ensemble \mathcal{B}' contient strictement plus de n éléments, il ne peut former une partie libre. Si \mathcal{B}' contient strictement moins de n éléments, alors les vecteurs u_1, \dots, u_n sont linéairement dépendants (on utilise le même raisonnement), ce qui est contraire à l'hypothèse. \square

Définition 1.4.17. Dans un espace vectoriel E de dimension finie, le nombre d'éléments d'une base quelconque est appelé *la dimension* de E , et noté $\dim(E)$, ou $\dim E$.

Je vous laisse le soin de trouver la dimension des espaces vectoriels \mathbb{R}^n , $(]0, +\infty[, \oplus, \odot)$, \mathcal{P}_2 . La notion de dimension est importante dans tout cours de géométrie, mais dans l'immédiat, elle va nous permettre de gagner du temps pour démontrer que des vecteurs forment une base, à travers le résultat suivant.

Proposition 1.4.18. *Soit E un espace vectoriel de dimension n .*

1. toute partie génératrice contient au moins n éléments,
2. toute partie libre contient au plus n éléments, et si elle contient exactement n éléments, c'est une base ;
3. toute partie génératrice contenant exactement n éléments est une base ;
4. toute partie libre contenant exactement n éléments est une base ;
5. Toute partie génératrice contient une base ;
6. Toute partie libre est incluse dans une base.

^j. Nous supposerons que $E \neq \{0\}$ pour éviter des discussions pénibles.

Démonstration. Soit $\mathcal{B} = (u_1, \dots, u_n)$ une base de E .

1. Soit $G = \{v_1, \dots, v_m\}$ une partie génératrice de E . On remarque que tous les éléments de \mathcal{B} sont combinaisons linéaires de v_1, \dots, v_m . Si $n > m$, par le théorème de Steinitz, les éléments de \mathcal{B} sont linéairement dépendants. C'est contraire à la définition de \mathcal{B} , donc on a $m \geq n$.
2. Soit $L = (v_1, \dots, v_m)$ une partie libre de E . On remarque que tous les éléments de L sont combinaisons linéaires des éléments de \mathcal{B} . Si $m > n$, par le théorème de Steinitz, les éléments de L sont linéairement dépendants. C'est contraire à l'hypothèse sur L , donc on a $m \leq n$.
3. Soit $G = \{v_1, \dots, v_n\}$ une partie génératrice de E . Si elle n'est pas libre, un des vecteurs est combinaison linéaire des autres, disons v_i . On peut retirer ce vecteur et obtenir une autre partie génératrice $G' = \{v_1, \dots, \widehat{v}_i, \dots, v_n\}$. Mais cette partie génératrice contient moins de n éléments, ce qui est impossible, vu le premier point.
4. Soit $L = (v_1, \dots, v_n)$ une famille libre de E . Pour tout $u \in E$, $G' = (v_1, \dots, v_n, u)$ est linéairement dépendante par le point 2. La proposition 1.4.10 montre alors que x est combinaison linéaire des éléments de L , donc L est une base.
5. Des raisonnements similaires permettent de prouver les derniers points. Je vous en fais grâce.

□

Nous pouvons encore donner un résultat sur la dimension de sous-espaces vectoriels.

Proposition 1.4.19. *Soit E un espace vectoriel réel de dimension n et soit F un sous-espace vectoriel (non réduit à $\{0\}$). Alors F est de dimension finie et $\dim(F) \leq n$. De plus si $\dim(F) = n$, alors $F = E$.*

Démonstration. On considère l'ensemble des parties libres $L = (v_1, \dots, v_r)$ formées de vecteurs de F , où r est maximal. On sait par le théorème de Steinitz que $r \leq n$, et puisque r est maximal, ces familles forment des bases. Elles sont donc génératrices, et la dimension de F est r . □

Dans la suite de ce cours, sauf mention explicite du contraire, nous ne considérerons que des espaces vectoriels de dimension finie. Un certain nombre de considérations s'étendent aux espaces de dimension infinie (que vous verrez en deuxième année), mais j'ai choisi de ne pas alourdir le cours avec les discussions sur la dimension.

1.5 Composantes

Nous avons introduit le concept de base pour obtenir une correspondance entre les vecteurs d'un espace vectoriel d'une part et des n -uples de nombres d'autre part. Formalisons cette correspondance dans la définition suivante.

Définition 1.5.1. Soit $\mathcal{B} = (b_1, \dots, b_n)$ une base de E . Pour tout $x \in E$, il existe un unique n -uplet $(x_1, \dots, x_n)^\sim \in \mathbb{R}^n$ tel que $x = \sum_{i=1}^n x_i b_i$. Ce n -uplet est appelé *vecteur de composantes* de x dans la base \mathcal{B} . On note $\Phi_{\mathcal{B}}$ l'application de E dans \mathbb{R}^n qui à tout $x \in E$ associe son vecteur de composantes $(x_1, \dots, x_n)^\sim$ dans la base \mathcal{B} .

Proposition 1.5.2. *Pour toute base \mathcal{B} de E l'application $\Phi_{\mathcal{B}}$ est une bijection. C'est de plus une application linéaire. En d'autres termes, elle satisfait les conditions suivantes :*

1. pour tous $u \in E$ et $\lambda \in \mathbb{R}$, on a $\Phi_{\mathcal{B}}(\lambda u) = \lambda \Phi_{\mathcal{B}}(u)$,
2. pour tous $u, v \in E$, on a $\Phi_{\mathcal{B}}(u + v) = \Phi_{\mathcal{B}}(u) + \Phi_{\mathcal{B}}(v)$.

En particulier, les composantes d'une combinaison linéaire de vecteurs sont obtenues en formant la combinaison linéaire correspondante des vecteurs de composantes, dans \mathbb{R}^n .

Démonstration. Pour montrer que l'application $\Phi_{\mathcal{B}}$ est une bijection, il suffit de montrer qu'elle peut être inversée, c'est-à-dire qu'il existe une application $\Psi_{\mathcal{B}}$ de \mathbb{R}^n dans E telle que $\Phi_{\mathcal{B}} \circ \Psi_{\mathcal{B}}$ est l'identité de \mathbb{R}^n et $\Psi_{\mathcal{B}} \circ \Phi_{\mathcal{B}}$ est l'identité de E . L'application $\Psi_{\mathcal{B}}$ doit donc consister à reconstruire un vecteur à partir de ses composantes dans la base \mathcal{B} . On définit donc naturellement

$$\Psi_{\mathcal{B}} : \mathbb{R}^n \rightarrow E : (x_1, \dots, x_n)^{\sim} \mapsto \sum_{i=1}^n x_i b_i, \quad (1.4)$$

et on constate que cette application a les propriétés recherchées.

La linéarité de $\Phi_{\mathcal{B}}$ se vérifie directement : soient $u \in E$ et $\lambda \in \mathbb{R}$. Notons $\Phi_{\mathcal{B}}(u) = (u_1, \dots, u_n)^{\sim}$. On a donc $u = \sum_{i=1}^n u_i b_i$, et par suite $\lambda u = \sum_{i=1}^n \lambda u_i b_i$ et finalement

$$\Phi_{\mathcal{B}}(\lambda u) = (\lambda u_1, \dots, \lambda u_n)^{\sim} = \lambda \Phi_{\mathcal{B}}(u),$$

ce qui démontre le premier point. On démontre le second de la même manière.

En appliquant ces deux points à plusieurs reprises, on obtient le cas général, à savoir

$$\Phi_{\mathcal{B}}\left(\sum_{i=1}^r \lambda_i a_i\right) = \sum_{i=1}^r \lambda_i \Phi_{\mathcal{B}}(a_i), \quad (1.5)$$

pour tout naturel r , tous réels $\lambda_1, \dots, \lambda_r$ et tous $a_1, \dots, a_r \in E$. \square

Voici un exemple.

Exemple 1.5.3. Les vecteurs $b_1 = (1, 1)^{\sim}$ et $b_2 = (1, -1)^{\sim}$ de \mathbb{R}^2 forment une base $\mathcal{B} = (b_1, b_2)$ de \mathbb{R}^2 . Dans cette base, le vecteur $x = (4, 6) \in \mathbb{R}^2$ admet pour composantes $(5, -1)^{\sim}$, puisque $x = 5b_1 - b_2$. De manière plus générale, le vecteur $x = (x_1, x_2)^{\sim} \in \mathbb{R}^2$ admet $(\frac{x_1+x_2}{2}, \frac{x_1-x_2}{2})^{\sim}$ pour composantes dans la base \mathcal{B} .

Nous étudions maintenant les liens qui unissent les composantes d'un même vecteur $x \in E$ dans deux bases différentes $\mathcal{B} = (b_1, \dots, b_n)$ et $\mathcal{B}' = (b'_1, \dots, b'_n)$.

Proposition 1.5.4. Les composantes $(x_1, \dots, x_n)^{\sim}$ et $(x'_1, \dots, x'_n)^{\sim}$ d'un même élément x de E dans des bases $\mathcal{B} = (b_1, \dots, b_n)$ et $\mathcal{B}' = (b'_1, \dots, b'_n)$ sont liées par les formules

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A_{\mathcal{B}', \mathcal{B}} \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = A_{\mathcal{B}, \mathcal{B}'} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad (1.6)$$

où $A_{\mathcal{B}, \mathcal{B}'}$ et $A_{\mathcal{B}', \mathcal{B}}$ sont des matrices à n lignes et n colonnes. Les colonnes de $A_{\mathcal{B}, \mathcal{B}'}$ sont les composantes des éléments de la base \mathcal{B} dans la base \mathcal{B}' . Enfin, les matrices $A_{\mathcal{B}, \mathcal{B}'}$ et $A_{\mathcal{B}', \mathcal{B}}$ sont inverses l'une de l'autre.

Démonstration. On a par exemple pour la première formule, $x = \sum_{i=1}^n x'_i b'_i$. On utilise la linéarité du passage aux composantes dans la base \mathcal{B} (voir la formule (1.5) ci-dessus) pour obtenir directement

$$(x_1, \dots, x_n)^{\sim} = \Phi_{\mathcal{B}}(x) = \sum_{i=1}^n x'_i \Phi_{\mathcal{B}}(b'_i).$$

On obtient donc la première formule avec

$$A_{\mathcal{B}', \mathcal{B}} = (\Phi_{\mathcal{B}}(b'_1), \dots, \Phi_{\mathcal{B}}(b'_n)).$$

On obtient la deuxième formule de la même façon. Enfin, en combinant les deux formules de (1.6), on obtient

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A_{\mathcal{B}', \mathcal{B}} A_{\mathcal{B}, \mathcal{B}'} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = A_{\mathcal{B}, \mathcal{B}'} A_{\mathcal{B}', \mathcal{B}} \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix},$$

pour tous $x_1, \dots, x_n, x'_1, \dots, x'_n$ réels. Cela montre que les produits $A_{\mathcal{B}', \mathcal{B}} A_{\mathcal{B}, \mathcal{B}'}$ et $A_{\mathcal{B}, \mathcal{B}'} A_{\mathcal{B}', \mathcal{B}}$ sont la matrice identité. \square

Nous pouvons reprendre l'exemple précédent.

Exercice 1.5.5. Ecrire les matrices de changement de base entre la base canonique de \mathbb{R}^2 et la base \mathcal{B} donnée dans l'exemple 1.5.3.

1.6 Equations de sous-espaces vectoriels

On se place dans un espace vectoriel E de dimension finie. Nous allons définir plusieurs descriptions possibles d'un sous-espace vectoriel F de E et on montrera comment passer de l'une à l'autre. Nous avons déjà rencontré ces deux descriptions dans la section sur les espaces vectoriels.

La description paramétrique permet de générer des éléments de F en attribuant des valeurs à certains paramètres. Si on fixe une base \mathcal{B} de E , on peut alors tout exprimer en composantes dans \mathcal{B} et on obtient une description *paramétrique cartésienne*. Enfin, en *éliminant* les paramètres, dans cette dernière description, on obtient un moyen de déterminer si un élément de E est dans F , à partir de conditions que l'on peut vérifier sur ses composantes dans \mathcal{B} . C'est la description de F au moyen d'*équations cartésiennes*. Enfin, en résolvant ces équations, on peut récupérer une description paramétrique de F .

1.6.1 Equations paramétriques

Puisque E est de dimension finie, il en est de même pour F . Etant lui-même un espace vectoriel, il admet une base (u_1, \dots, u_p) . Ceci nous permet de décrire F comme une enveloppe linéaire : on a $F = \langle u_1, \dots, u_p \rangle$. Vu la définition d'une enveloppe linéaire, on a directement le résultat suivant.

Proposition 1.6.1. *Tout sous-espace vectoriel F de E est une enveloppe linéaire $F = \langle u_1, \dots, u_p \rangle$. Si x appartient à E , on a alors $x \in F$ si, et seulement si,*

$$\exists \lambda_1, \dots, \lambda_p \in \mathbb{R} : x = \lambda_1 u_1 + \dots + \lambda_p u_p \quad (1.7)$$

La condition (1.7) est appelé **une** *équation paramétrique vectorielle* de F .

Remarque 1.3. 1. Puisqu'un sous-espace vectoriel peut être écrit comme une enveloppe linéaire de plusieurs façons^k (en fait une façon par partie génératrice ordonnée), il y a également plusieurs équations paramétriques vectorielles pour un sous-espace vectoriel.

2. Dans la proposition ci-dessus, il n'est pas nécessaire que les vecteurs u_1, \dots, u_p soient linéairement indépendants. En ajoutant un vecteur combinaison linéaire des autres, on ajoute un paramètre, obtenant ainsi une équation paramétrique vectorielle différente, plus compliquée, mais équivalente.

3. Il est important de ne pas oublier " $\exists \lambda_1, \dots, \lambda_p \in \mathbb{R}$ " dans la condition (1.7) ci-dessus. Elle disparaît cependant souvent dans les copies lors des examens. Cela revient à peu près à considérer que les phrases "Je mange une pomme" et "Je une pomme" sont identiques.

Vous conviendrez que cela nuit à la compréhension du texte.

Nous avons vu que le choix d'une base $\mathcal{B} = (b_1, \dots, b_n)$ dans un espace vectoriel E de dimension n permet une identification de E à l'espace vectoriel \mathbb{R}^n . On travaille alors en composantes dans \mathcal{B} , ce qui est bien plus concret. Nous pouvons faire de même en ce qui concerne la condition (1.7), et obtenir le résultat suivant.

Proposition 1.6.2. *Soit une base $\mathcal{B} = (b_1, \dots, b_n)$ de E et $F = \langle u_1, \dots, u_p \rangle$. Si les composantes de u_i dans \mathcal{B} sont $(u_{1,i}, \dots, u_{n,i})^\sim$, alors un vecteur x de composantes $(x_1, \dots, x_n)^\sim$ appartient à F si, et seulement si,*

$$\exists \lambda_1, \dots, \lambda_p \in \mathbb{R} : \begin{cases} x_1 &= \lambda_1 u_{1,1} + \dots + \lambda_p u_{1,p} \\ &\vdots \\ x_n &= \lambda_1 u_{n,1} + \dots + \lambda_p u_{n,p} \end{cases} \quad (1.8)$$

k. Sauf si $F = \{0\}$, mais ce cas n'est pas extrêmement intéressant ici.

La condition qui apparaît dans (1.10) est appelée une équation paramétrique cartésienne du sous-espace vectoriel F dans la base \mathcal{B} . Elle permet entre autres, en assignant des valeurs quelconques aux paramètres, de générer des composantes de vecteurs appartenant à F . Voici quelques exemples.

Exemple 1.6.3. Dans \mathbb{R}^2 , on considère la droite vectorielle $F = \langle u \rangle$ où $u = (1, 2)^\sim$. On a alors

$$x \in F \Leftrightarrow \exists \lambda \in \mathbb{R} : x = \lambda u.$$

Dans la base canonique de \mathbb{R}^2 , si $x = (x_1, x_2)^\sim$, alors x appartient à F si, et seulement si

$$\exists \lambda \in \mathbb{R} : \begin{cases} x_1 = \lambda \\ x_2 = 2\lambda. \end{cases}$$

Exemple 1.6.4. Dans \mathbb{R}^3 , on considère la droite vectorielle $F = \langle u \rangle$ où $u = (1, 2, 3)^\sim$. On a alors

$$x \in F \Leftrightarrow \exists \lambda \in \mathbb{R} : x = \lambda u.$$

Dans la base canonique, si $x = (x_1, x_2, x_3)^\sim$, alors x appartient à F si, et seulement si

$$\exists \lambda \in \mathbb{R} : \begin{cases} x_1 = \lambda \\ x_2 = 2\lambda \\ x_3 = 3\lambda. \end{cases}$$

Exemple 1.6.5. Dans \mathbb{R}^3 , on considère la droite vectorielle $F = \langle u \rangle$ où $u = (1, 2, 0)^\sim$. On a alors

$$x \in F \Leftrightarrow \exists \lambda \in \mathbb{R} : x = \lambda u.$$

Dans la base canonique, si $x = (x_1, x_2, x_3)^\sim$, alors x appartient à F si, et seulement si

$$\exists \lambda \in \mathbb{R} : \begin{cases} x_1 = \lambda \\ x_2 = 2\lambda \\ x_3 = 0. \end{cases}$$

Exemple 1.6.6. Dans \mathbb{R}^3 , on considère $F = \langle u, v \rangle$ où $u = (1, 2, 3)^\sim$ et $v = (2, 4, 6)^\sim$. On a alors

$$x \in F \Leftrightarrow \exists \lambda, \mu \in \mathbb{R} : x = \lambda u + \mu v. \quad (1.9)$$

Dans la base canonique, si $x = (x_1, x_2, x_3)^\sim$, alors x appartient à F si, et seulement si

$$\exists \lambda, \mu \in \mathbb{R} : \begin{cases} x_1 = \lambda + 2\mu \\ x_2 = 2\lambda + 4\mu \\ x_3 = 3\lambda + 6\mu. \end{cases}$$

On constate que tout s'exprime en fonction du paramètre $\nu = \lambda + 2\mu$. A quoi est-ce dû ? Quelle est la dimension de F ?

Exemple 1.6.7. Dans \mathbb{R}^3 , on considère $F = \langle u, v \rangle$ où $u = (1, 2, 0)^\sim$ et $v = (1, 0, 1)^\sim$. On a alors

$$x \in F \Leftrightarrow \exists \lambda, \mu \in \mathbb{R} : x = \lambda u + \mu v.$$

Dans la base canonique, si $x = (x_1, x_2, x_3)^\sim$, alors x appartient à F si, et seulement si

$$\exists \lambda, \mu \in \mathbb{R} : \begin{cases} x_1 = \lambda + \mu \\ x_2 = 2\lambda \\ x_3 = \mu. \end{cases}$$

Quelle est la dimension de F ?

Exemple 1.6.8. Dans $E = \mathcal{P}_3$, on considère $F = \langle u, v \rangle$ où $u(x) = x^2 + 1$ et $v(x) = 2x^3 + x$, pour tout $x \in \mathbb{R}$. On a alors

$$p \in F \Leftrightarrow \exists \lambda, \mu \in \mathbb{R} : p(x) = \lambda u(x) + \mu v(x), \quad \forall x \in \mathbb{R}.$$

Dans la base $\mathcal{B} = \{1, x, x^2, x^3\}$, un polynôme p de composantes (x_1, x_2, x_3, x_4) est dans F si, et seulement si,

$$\exists \lambda, \mu \in \mathbb{R} : \begin{cases} x_1 = \lambda \\ x_2 = \mu \\ x_3 = \lambda \\ x_4 = 2\mu. \end{cases}$$

Remarquons que nous n'avons pas utilisé x pour noter un vecteur pour éviter la confusion avec la lettre x souvent utilisée pour décrire les fonctions polynômes. On aurait aussi pu noter cette variable par une autre lettre.

Considérant un sous-espace vectoriel, disons celui de l'exemple 1.6.7, et étant donné le vecteur x de composantes $(7, 8, 3)^\sim$, comment déterminer si x est dans F . La réponse est claire, il faut et il suffit que l'on puisse trouver des valeurs des paramètres λ et μ tels que $x = \lambda u + \mu v$. Dans ce cas, on trouve $x = 4u + 3v$ et on conclut que x est dans F . Il serait cependant plus simple d'obtenir une fois pour toutes des conditions que l'on peut vérifier sur les composantes de x dans la base \mathcal{B} pour que x soit dans F , sans devoir chercher à chaque fois les valeurs des paramètres λ, μ qui conviennent. Cela amène la définition des équations cartésiennes.

Définition 1.6.9. Des équations cartésiennes du sous-espace vectoriel F dans la base \mathcal{B} sont des conditions nécessaires et suffisantes sur les composantes d'un vecteur x dans la base \mathcal{B} pour qu'il soit dans F .

Trouver de telles conditions à partir des équations paramétriques cartésiennes consiste à *éliminer les paramètres* dans les équations paramétriques. On peut le faire "à la main" sur des exemples simples. Reprenons l'exemple 1.6.7. Étant donné un vecteur x de composantes $(x_1, x_2, x_3)^\sim$, s'il existe $\lambda, \mu \in \mathbb{R}$ tels que

$$\begin{cases} x_1 = \lambda + \mu \\ x_2 = 2\lambda \\ x_3 = \mu \end{cases}, \quad (1.10)$$

alors les deux dernières équations donnent les valeurs de λ et μ en fonction de x_1, x_2 et x_3 : on a nécessairement

$$\lambda = \frac{x_2}{2}, \quad \text{et} \quad \mu = x_3.$$

Mais puisque λ et μ doivent satisfaire **les trois** équations dans (1.10), on obtient une *condition de compatibilité*

$$x_1 = \frac{x_2}{2} + x_3,$$

que l'on écrit plus proprement

$$2x_1 - x_2 - 2x_3 = 0.$$

Inversement, si $(x_1, x_2, x_3)^\sim$ satisfait cette dernière équation, alors on peut *choisir* $\lambda = \frac{x_2}{2}$ et $\mu = x_3$ et ce choix permet alors de satisfaire toutes les conditions de (1.10).

Nous avons alors obtenu une équation cartésienne de F au sens de la définition 1.6.9. Pour décider si $(7, 8, 3)^\sim$ est dans F , il suffit de vérifier s'il satisfait l'équation, et on a en effet

$$2 \cdot 7 - 8 - 2 \cdot 3 = 0.$$

Puisqu'on a obtenu des conditions **nécessaires et suffisantes**, on montre alors par exemple facilement que $(2, 5, 4)^\sim$ n'appartient pas à F car on a

$$2 \cdot 2 - 5 - 2 \cdot 4 = -9 \neq 0.$$

On peut alors reprendre les exemples et appliquer cette méthode.

- Pour l'exemple 1.6.3, on obtient l'équation $x_2 = 2x_1$;
- Pour l'exemple 1.6.4, on obtient le système d'équations

$$\begin{cases} x_2 = 2x_1 \\ x_3 = 3x_1. \end{cases}$$

- Pour l'exemple 1.6.5, on obtient le système d'équations

$$\begin{cases} x_2 = 2x_1 \\ x_3 = 0. \end{cases}$$

- Pour l'exemple 1.6.6, on constate qu'il n'est pas possible d'obtenir λ et μ en fonction de x_1 , x_2 et x_3 , mais si λ et μ satisfont cet ensemble de conditions, alors on trouve λ en fonction de x_1 et μ par $\lambda = x_1 - 2\mu$. On a alors nécessairement

$$\begin{cases} x_2 = 2x_1 \\ x_3 = 3x_1. \end{cases}$$

Si x_1 , x_2 et x_3 satisfont ces conditions, alors en donnant une valeur quelconque à μ , on choisit $\lambda = x_1 - 2\mu$ et on constate que λ et μ satisfont les trois conditions dans (1.9).

- Pour l'exemple 1.6.8, on obtient le système d'équations

$$\begin{cases} x_3 = x_1 \\ x_4 = 2x_2. \end{cases}$$

Il est toutefois intéressant d'avoir une méthode générale pour obtenir des équations d'un sous-espace vectoriel dans une base \mathcal{B} donnée. Heureusement, le prix pour obtenir cette procédure automatique pour obtenir des équations cartésiennes est payé en algèbre (voir appendice).

Proposition 1.6.10. *Soit un espace vectoriel $F = \langle u_1, \dots, u_p \rangle$. Si les composantes de u_i dans \mathcal{B} sont $U_i = (u_{1,i}, \dots, u_{n,i})^\sim$, alors un vecteur x de composantes $X = (x_1, \dots, x_n)^\sim$ est dans F si, et seulement si,*

$$\text{rg}(U_1, \dots, U_p) = \text{rg}(U_1, \dots, U_p | X). \quad (1.11)$$

Démonstration. On peut réécrire les équations paramétriques cartésiennes que nous avons obtenues plus haut en

$$x \in F \Leftrightarrow \exists \lambda_1, \dots, \lambda_p \in \mathbb{R} : \begin{pmatrix} u_{1,1} & \cdots & u_{1,p} \\ \vdots & \ddots & \vdots \\ u_{n,1} & \cdots & u_{n,p} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Il s'agit donc d'étudier la condition de compatibilité du système linéaire en les paramètres $\lambda_1, \dots, \lambda_p$. On obtient donc le résultat annoncé. \square

Remarque 1.4. Il n'est pas nécessaire que les vecteurs u_1, \dots, u_p soient linéairement indépendants pour appliquer la proposition précédente. Tous les cas sont pris en compte lors du calcul du rang. Dans le cas de l'exemple 1.6.6, on obtient des équations cartésiennes en explicitant

$$\text{rg} \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{pmatrix} = \text{rg} \begin{pmatrix} 1 & 2 & x_1 \\ 2 & 4 & x_2 \\ 3 & 6 & x_3 \end{pmatrix}$$

Le rang dans le membre de gauche est 1, puisque les deux colonnes sont multiples l'une de l'autre. La condition est donc

$$\text{rg} \begin{pmatrix} 1 & 2 & x_1 \\ 2 & 4 & x_2 \\ 3 & 6 & x_3 \end{pmatrix} = 1$$

En utilisant la méthode des déterminants bordés et en tenant compte du fait que les deux première colonnes sont linéairement dépendantes, on arrive à la condition

$$\det \begin{pmatrix} 1 & x_1 \\ 2 & x_2 \end{pmatrix} = \det \begin{pmatrix} 1 & x_1 \\ 3 & x_3 \end{pmatrix} = 0,$$

ce qui donne effectivement ce que nous avons trouvé à la main, mais cette fois-ci, de manière automatique.

On peut tirer les conséquences du résultat précédent en termes de nombre d'équations. Intuitivement, la relation est claire : dans un espace de dimension n , l'ensemble des vecteurs dont les composantes satisfont une équation (non triviale) est un sous-espace plus petit. L'équation représente une contrainte, et on a donc un sous-espace de dimension $n - 1$. L'ensemble des vecteurs dont les composantes satisfont deux équations (indépendantes, sinon une des deux ne sert à rien), est un espace plus petit. On a deux contraintes, et donc la dimension est diminuée de deux unités. Plus généralement, on a le résultat suivant.

Proposition 1.6.11. *Dans un espace vectoriel de dimension n muni d'une base $\mathcal{B} = (b_1, \dots, b_n)$, un sous-espace vectoriel F de dimension p admet des équations cartésiennes formées d'un système de $n - p$ équations linéaires indépendantes.*

Démonstration. On choisit une base f_1, \dots, f_p de F (formée de vecteurs de $F \subset E$). On a donc

$$F = \langle f_1, \dots, f_p \rangle.$$

Il suffit alors d'écrire la condition (1.11) en utilisant la règle déterminants bordés ; On constate alors facilement que cela donne $n - p$ équations indépendantes. \square

Ce résultat admet la réciproque suivante.

Proposition 1.6.12. *Soient E un espace vectoriel de dimension n et $\mathcal{B} = (b_1, \dots, b_n)$ une base de E . L'ensemble des vecteurs x de E dont les composantes $X = (x_1, \dots, x_n)^\sim$ dans \mathcal{B} satisfont un système d'équations linéaires homogènes de rang r est un sous-espace vectoriel de E de dimension $n - r$.*

Démonstration. Le passage au composantes est linéaire. On se ramène alors directement à un résultat les systèmes linéaires, que nous avons admis (voir appendice). \square

Voici quelques conséquences pour des ensembles bien connus. Mais généralisons d'abord quelques définitions.

Définition 1.6.13. Dans un espace vectoriel E de dimension n :

1. une droite (vectorielle) est un sous-espace vectoriel de dimension 1 ;
2. un plan (vectoriel) est un sous-espace vectoriel de dimension 2 (si $n \geq 2$) ;
3. un hyperplan (vectoriel) est un sous-espace vectoriel de dimension $n - 1$.

On peut alors donner le nombre d'équations nécessaires pour déterminer de tels sous-espaces.

- Dans un espace de dimension 2, une droite vectorielle correspond à un système de rang 1, i.e. une équation homogène, non triviale.
- Dans un espace de dimension 3, une droite vectorielle correspond à un système d'équations homogènes, de rang 2.
- Dans un espace de dimension 3, un plan est déterminé par une équation homogène.
- Un hyperplan dans un espace de dimension n , est déterminé par 1 équation.
- ...

Exemple 1.6.14. 1. Dans \mathbb{R}^4 , muni d'une base \mathcal{B} , l'équation

$$x_1 - x_2 + 3x_3 = 0$$

décrit un sous-espace vectoriel de dimension 3. Trouvez-en une base.

2. Dans \mathbb{R}^5 , le système d'équations

$$\begin{cases} x_1 - 2x_3 + x_5 = 0 \\ 2x_4 - x_5 = 0 \end{cases}$$

détermine un sous-espace vectoriel de dimension 3. Trouvez-en une base.

1.6.2 Cas particuliers : droites et hyperplans

Dans le cas des droites et des hyperplans, respectivement les sous-espaces vectoriels non triviaux de dimension minimale et maximale, la recherche des équations cartésiennes est particulièrement simple. Revenons un peu sur ces situations.

On considère un espace vectoriel E de dimension n , muni d'une base \mathcal{B} . Soit d une droite et $u : (u_1, \dots, u_n)^\sim \in d \setminus \{0\}$. On appelle un tel vecteur "vecteur directeur" de d . On a alors $d = \rangle u \langle$. Alors on a $x : (x_1, \dots, x_n)^\sim \in d$ si et seulement si

$$\exists \lambda \in \mathbb{R} : \begin{cases} x_1 = \lambda u_1 \\ \vdots \\ x_n = \lambda u_n. \end{cases} \quad (1.12)$$

– Si $u_1 \cdots u_n \neq 0$, alors on élimine facilement le paramètre :

$$d \equiv \frac{x_1}{u_1} = \cdots = \frac{x_n}{u_n}.$$

– Si $u_i = 0$ (pour un $i \leq n$), on revient à (1.12), et on trouve

$$d \equiv \begin{cases} \frac{x_1}{u_1} = \cdots \hat{i} \cdots = \frac{x_n}{u_n} \\ x_i = 0 \end{cases}$$

– Cela fonctionne quand plusieurs composantes de u sont nulles.

Exemple 1.6.15. 1. Dans \mathbb{R}^4 muni de la base canonique, écrire des équations cartésiennes de

$$d_1 = \rangle \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \langle, d_2 = \rangle \begin{pmatrix} 1 \\ 0 \\ 3 \\ 4 \end{pmatrix} \langle, d_3 = \rangle \begin{pmatrix} 0 \\ 0 \\ 3 \\ 4 \end{pmatrix} \langle, d_4 = \rangle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 3 \end{pmatrix} \langle.$$

2. Dans l'espace des polynômes de degré inférieur ou égal à deux, muni de sa base $\mathcal{B} = \{1, x, x^2\}$, écrire une équation cartésienne de la droite vectorielle d engendrée par P défini par $P(x) = 3x^2 - 2x + 1$.

Dans le cas des hyperplans, il n'y a qu'une équation simple à obtenir.

Proposition 1.6.16. Soit un hyperplan vectoriel $F = \rangle f_1, \dots, f_{n-1} \langle$. Si les composantes de f_i dans \mathcal{B} sont $F_i = (f_{1,i}, \dots, f_{n,i})^\sim$, alors un vecteur x de composantes $X = (x_1, \dots, x_n)^\sim$ est dans F si, et seulement si,

$$\det(F_1, \dots, F_{n-1}, X) = 0. \quad (1.13)$$

Exemple 1.6.17. 1. Dans \mathbb{R}^2 muni de sa base canonique, écrire une équation cartésienne de la droite

$$d_1 = \rangle \begin{pmatrix} 1 \\ 2 \end{pmatrix} \langle, d_2 = \rangle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \langle, d_3 = \rangle \begin{pmatrix} 5 \\ 6 \end{pmatrix} \langle.$$

2. Dans \mathbb{R}^3 muni de sa base canonique, écrire une équation cartésienne du plan

$$\pi_1 = \rangle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \langle, \pi_2 = \rangle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \langle, \pi_3 = \rangle \begin{pmatrix} 5 \\ 6 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \langle.$$

3. Dans \mathbb{R}^4 muni de sa base canonique, écrire une équation cartésienne de l'hyperplan

$$\pi_4 = \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\rangle, \pi_5 = \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle.$$

1.6.3 Faisceaux de plans en dimension 3

Terminons cette section sur les équations de sous-espaces vectoriels par un résultat sur les faisceaux de plans.

Proposition 1.6.18. *Dans un espace vectoriel E de dimension 3, deux plans vectoriels distincts se coupent suivant une droite vectorielle.*

Démonstration. Fixons une base et considérons les équations des deux plans dans cette base. Appelons les plans π_1 et π_2 . On a donc, vu les résultats précédents :

$$\pi_1 \equiv a_1x_1 + a_2x_2 + a_3x_3 = 0,$$

et

$$\pi_2 \equiv b_1x_1 + b_2x_2 + b_3x_3 = 0.$$

De plus on a $(a_1, a_2, a_3) \neq (0, 0, 0)$ et $(b_1, b_2, b_3) \neq (0, 0, 0)$, pour que les équations soient de rang 1. Enfin, puisque les plans sont distincts, ces triplets ne sont pas multiples l'un de l'autre. L'intersection $\pi_1 \cap \pi_2$ admet pour équation

$$\begin{cases} a_1x_1 + a_2x_2 + a_3x_3 = 0 \\ b_1x_1 + b_2x_2 + b_3x_3 = 0 \end{cases}$$

Ce système est de rang 2, donc l'intersection $\pi_1 \cap \pi_2$ est de dimension $3-2=1$. \square

Ayant une droite déterminée par deux plans en dimension 3, on peut également trouver des équations cartésiennes de tous les plans contenant cette droite. L'ensemble de tous les plans contenant la droite vectorielle d est appelé faisceau de plans d'axe d .

Proposition 1.6.19. *Dans un espace vectoriel E de dimension 3, si une droite vectorielle d admet pour équation(s) dans une base \mathcal{B} le système*

$$\begin{cases} a_1x_1 + a_2x_2 + a_3x_3 = 0 \\ b_1x_1 + b_2x_2 + b_3x_3 = 0, \end{cases}$$

alors un plan π contient d si, et seulement si, il existe $(\lambda, \mu) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ t.q.

$$\pi \equiv \lambda(a_1x_1 + a_2x_2 + a_3x_3) + \mu(b_1x_1 + b_2x_2 + b_3x_3) = 0.$$

Démonstration. Remarquons tout d'abord que $d \subset \pi$ est équivalent à $d \cap \pi = d$, ou encore, puisque $d \cap \pi$ est inclus dans d , à $\dim(d \cap \pi) = 1$. Si π admet pour équation

$$c_1x_1 + c_2x_2 + c_3x_3 = 0$$

alors

$$d \cap \pi \equiv \begin{cases} a_1x_1 + a_2x_2 + a_3x_3 = 0 \\ b_1x_1 + b_2x_2 + b_3x_3 = 0 \\ c_1x_1 + c_2x_2 + c_3x_3 = 0 \end{cases}$$

Ce système définit un sous-espace de dimension 1 s'il est de rang 2. Puisque les deux premières lignes sont indépendantes, cela équivaut au fait que la troisième en soit combinaison linéaire, c'est-à-dire

$$(c_1, c_2, c_3) = \lambda(a_1, a_2, a_3) + \mu(b_1, b_2, b_3),$$

pour $(\lambda, \mu) \neq (0, 0)$, ce qui donne le résultat annoncé. \square

Chapitre 2

Espaces affines

La notion d'espace affine permet de modéliser rigoureusement l'espace dans lequel nous vivons, que nous considérons comme plat. Cette notion est basée sur celle d'espace vectoriel, que nous venons d'étudier. Intuitivement, les vecteurs pourraient représenter les forces, et les éléments d'un espace affine les objets, supposés ponctuels, auxquelles elles s'appliquent.

Vous pouvez vous représenter un espace affine (de dimension 2) comme une table de billard infinie, sur laquelle les boules sont les points, et qui permettent d'appliquer des forces : en jouant, on envoie une boule d'un point à un autre.

Les espaces affines de dimension 2 et 3 sont les ensembles dans lesquels vous avez étudié la géométrie, vous en avez donc une représentation intuitive que vous pouvez conserver. Cependant, il est important de bien distinguer la nature des objets avec lesquels on travaille.

Enfin, je précise que dans ce chapitre, il n'y a aucune notion euclidienne qui apparaît. Ces notions sont reléguées au chapitre 4, car elles nécessitent la généralisation des longueurs, angles, produits scalaires.

2.1 Définitions

Commençons par la définition.

Définition 2.1.1. Un espace affine \mathcal{A} modelé sur un espace vectoriel E est un ensemble dont les éléments sont appelés points, muni d'une opération de translation

$$t : \mathcal{A} \times E \rightarrow \mathcal{A} : (A, u) \mapsto t(A, u) = A + u,$$

satisfaisant les conditions suivantes :

1. On a $(A + u) + v = A + (u + v)$ pour tous $A \in \mathcal{A}$ et tous $u, v \in E$;
2. Pour tous $A, B \in \mathcal{A}$, il existe un **unique** $u \in E$ tel que $B = A + u$.

L'idée de la définition est donc bien que les éléments de E (les forces) agissent^a sur \mathcal{A} , par translation. Comme pour les espaces vectoriels, nous commençons par un exemple abstrait.

Exemple 2.1.2. Soit E l'espace vectoriel des fonctions polynomiales de degré au plus 2, qui prennent la valeur 0 en 1, i.e.

$$E = \{f : \mathbb{R} \rightarrow \mathbb{R} : f(x) = ax^2 + bx + c \text{ pour tout } x \in \mathbb{R}, \text{ et } a + b + c = 0\}.$$

Soit \mathcal{A} l'ensemble des fonctions polynomiales de degré au plus 2, qui prennent la valeur 3 en 1, i.e.

$$\mathcal{A} = \{F : \mathbb{R} \rightarrow \mathbb{R} : F(x) = ax^2 + bx + c \text{ pour tout } x \in \mathbb{R}, \text{ et } a + b + c = 3\}.$$

On définit la translation du point F de \mathcal{A} par rapport au vecteur f de E comme étant l'addition des fonctions f et F (dans l'espace vectoriel de toutes les fonctions). On peut alors vérifier que toutes les conditions sont satisfaites :

a. Vous verrez au cours de vos études la définition d'une action d'un groupe sur un ensemble. Celle-ci en est un cas particulier.

1. Pour tous $F \in \mathcal{A}$ et $f \in E$, on a $F + f \in \mathcal{A}$, puisque par définition on a $(F + f)(1) = F(1) + f(1) = 3$;
2. Pour tous $F \in \mathcal{A}$ et $f, g \in E$, on a $(F + f) + g = F + (f + g)$, puisque l'addition des fonctions est associative;
3. Pour tous $F, G \in \mathcal{A}$, alors nécessairement $f = G - F$ est un élément de E , et c'est l'unique élément $f \in E$ tel que $G = F + f$.

Passons maintenant à des exemples auxquels vous êtes peut-être plus habitués.

Exemple 2.1.3. 1. Si E est un espace vectoriel, alors $\mathcal{A} = E$ est un espace affine modelé sur E . La translation de $A \in \mathcal{A}$ par $u \in E$ est simplement l'addition de A et u dans E .

2. Si E est un espace vectoriel de dimension finie et si F est un sous-espace vectoriel, pour tout $u \in E$, l'ensemble $\{u + f : f \in F\}$ est un espace affine modelé sur F .

Remarque 2.1. 1. Pour des raisons qui deviendront claires d'ici peu, l'espace vectoriel définissant un espace affine \mathcal{A} est aussi noté $\overrightarrow{\mathcal{A}}$.

2. Quand cela ne prête pas à confusion, on parlera d'un espace affine \mathcal{A} , sans mentionner l'espace vectoriel sur lequel il est modelé.

3. On notera dans la mesure du possible les éléments de \mathcal{A} par des lettres capitales.

Voici quelques propriétés que l'on peut déduire de la définition.

Proposition 2.1.4. *Soit \mathcal{A} un espace affine.*

1. Si $u, v \in E$ satisfont $P + u = P + v$ pour un $P \in \mathcal{A}$, alors $u = v$;
2. On a $P + 0 = P$ pour tout $P \in \mathcal{A}$;
3. Si $P + u = Q + u$ pour $P, Q \in \mathcal{A}$ et $u \in E$, alors $P = Q$;
4. L'expression $P + u_1 + \dots + u_r$, $P \in \mathcal{A}$, $u_1, \dots, u_r \in E$ est bien définie : elle ne dépend pas de l'ordre dans lequel on effectue les opérations.

Démonstration. Prouvons point après point :

1. Si on appelle Q le point $P + u$. On a aussi $P + v = Q$. On utilise alors l'unicité (point 2 de la définition), et on a $u = v$.
2. Si Q est un point de \mathcal{A} , alors il existe u tel que $P = Q + u$. Alors on a

$$P + 0 = (Q + u) + 0 = Q + (u + 0) = Q + u = P.$$

3. Il existe v tel que $Q = P + v$. On a alors par hypothèse $P + u = (P + v) + u = P + (v + u)$. Par le point 1., on a $v + u = u$, donc $v = 0$. Dès lors, $Q = P + 0 = P$.

Le dernier point résulte directement de la définition et des propriétés correspondantes des espaces vectoriels. □

Puisqu'il n'y a qu'un seul vecteur dont la translation applique un point A sur un point B , il est naturel de lui donner un nom.

Définition 2.1.5. Soient A, B deux points de \mathcal{A} . On note \overrightarrow{AB} l'unique élément de E satisfaisant $A + \overrightarrow{AB} = B$.

Il est important de noter que si les points A et B sont des éléments de \mathcal{A} , on représente le vecteur \overrightarrow{AB} par une flèche allant de A à B , mais ce vecteur est bel et bien un élément de E . Ayant cette notion de vecteur, on se retrouve dans les situations classiques de géométrie, que vous connaissez bien, et dont je vous rappelle quelques propriétés. La proposition suivante contient alors une relation célèbre.

Proposition 2.1.6 (Relation de Chasles^b). *Pour tous points P, Q, R de \mathcal{A} , on a $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$. En particulier, on a $\overrightarrow{PP} = 0$ pour tout $P \in \mathcal{A}$ et $\overrightarrow{PQ} = -\overrightarrow{QP}$ pour tous $P, Q \in \mathcal{A}$.*

^b. Michel Chasles (1793-1880), Mathématicien français.

Démonstration. Pour la première partie, on a

$$P + (\overrightarrow{PQ} + \overrightarrow{QR}) = (P + \overrightarrow{PQ}) + \overrightarrow{QR} = Q + \overrightarrow{QR} = R$$

et

$$P + \overrightarrow{PR} = R,$$

on peut donc utiliser la proposition précédente. On a ensuite $\overrightarrow{PP} + \overrightarrow{PP} = \overrightarrow{PP}$, donc $\overrightarrow{PP} = 0$. Enfin, $\overrightarrow{PQ} + \overrightarrow{QP} = \overrightarrow{PP} = 0$. \square

La définition suivante est également classique.

Définition 2.1.7. On dit qu'un quadruple de points (A, B, C, D) forme un parallélogramme si $\overrightarrow{AB} = \overrightarrow{DC}$.

Attention cependant : on peut l'appliquer dans des situations plus générales que celles rencontrées dans l'enseignement secondaire, comme le montre l'exemple suivant.

Exemple 2.1.8. Dans l'espace vectoriel des fonctions polynomiales de degré inférieur ou égal à 4, vu comme un espace affine, les points

$$A : x^3 - 3x + 2, \quad B : 4x^3 - 2x + 6, \quad C : x + 1 \quad D : -3x^3 - 3$$

déterminent un parallélogramme.

2.2 Vecteurs liés et vecteurs libres

L'espace vectoriel E sur lequel un espace affine \mathcal{A} est modelé est généralement distinct de \mathcal{A} . Il existe deux façons naturelles d'associer aux vecteurs de E des objets construits à partir des points de \mathcal{A} . Ce sont les vecteurs liés en un point et les vecteurs libres. Ces deux notions sont celles qui ont permis de définir les vecteurs rencontrés dans l'enseignement secondaire, dans différents contextes.

Définition 2.2.1. Soit O un point de \mathcal{A} . On appelle vecteur lié en O tout couple (O, P) où P est un point de \mathcal{A} . On note \mathcal{A}_O l'ensemble des vecteurs liés en O .

On peut munir l'ensemble des vecteurs liés en O d'une structure d'espace vectoriel *isomorphe* à E . Cela veut dire que l'ensemble \mathcal{A}_O des vecteurs liés en O peut d'une part être mis en bijection (correspondance) avec E , mais qu'il sera possible d'additionner et de multiplier par des nombres les éléments de \mathcal{A}_O , de façon telle que la correspondance soit linéaire.

Proposition 2.2.2. L'application f qui à tout couple (O, P) associe le vecteur \overrightarrow{OP} est une bijection. Elle est linéaire pour les opérations sur \mathcal{A}_O définies par

$$(O, P) \oplus (O, Q) = (O, O + \overrightarrow{OP} + \overrightarrow{OQ}), \quad \text{et} \quad \lambda \odot (O, P) = (O, O + \lambda \overrightarrow{OP}),$$

pour tous $P, Q \in \mathcal{A}$ et tout $\lambda \in \mathbb{R}$.

Démonstration. Montrons que f est injective. Si $f(O, P) = f(O, P')$, alors $\overrightarrow{OP} = \overrightarrow{OP'}$, donc $O + \overrightarrow{OP} = O + \overrightarrow{OP'}$ et finalement $P = P'$. Montrons que f est surjective : il faut montrer que tout vecteur u s'écrit \overrightarrow{OP} pour au moins un P . Il suffit de choisir $P = O + u$. On définit alors les opérations sur \mathcal{A}_O pour que f soit linéaire. \square

On est habitué à la règle du parallélogramme pour additionner les vecteurs liés en un point. Rien ne change comme le montre la remarque suivante.

Remarque 2.2. L'addition dans \mathcal{A}_O suit la règle du parallélogramme. Quels que soient P et Q , le quadruple de points

$$O, \quad P, \quad O + \overrightarrow{OP} + \overrightarrow{OQ}, \quad Q$$

est un parallélogramme.

Passons maintenant à la définition des vecteurs libres.

Définition 2.2.3. Deux couples (A, B) et (C, D) sont dits équipollents si $\overrightarrow{AB} = \overrightarrow{CD}$. On appelle vecteur libre représenté par (A, B) l'ensemble^c de tous les couples équipollents à (A, B) . On le note^d $[(A, B)]$. Nous noterons encore $\mathcal{A}_{\mathcal{L}}$ l'ensemble des vecteurs libres de \mathcal{A} .

Le couple (A, B) est donc équipollent au couple (D, C) si le quadruple (A, B, D, C) forme un parallélogramme.

Les résultats sur les vecteurs libres sont similaires à ceux concernant les vecteurs liés en un point.

Proposition 2.2.4. L'application $f' : \mathcal{A}_{\mathcal{L}} \rightarrow E$ qui à $[(A, B)]$ associe \overrightarrow{AB} est une bijection. Elle est linéaire pour les opérations suivantes sur $\mathcal{A}_{\mathcal{L}}$ par

$$[(A, B)] \oplus [(C, D)] = [(A, A + \overrightarrow{AB} + \overrightarrow{CD})], \quad \text{et} \quad \lambda \odot [(A, B)] = [(A, A + \lambda \overrightarrow{AB})],$$

pour tous $A, B, C, D \in \mathcal{A}$ et tout $\lambda \in \mathbb{R}$.

Démonstration. La démonstration suit les mêmes lignes que celle qui concernait les vecteurs libres. Je vous en fais grâce. \square

2.3 Combinaisons affines

Dans les espaces vectoriels, on dispose naturellement d'un moyen de former des combinaisons linéaires de vecteurs, puisque la somme et la multiplication scalaire des vecteurs sont données a priori. Ce n'est pas le cas dans les espaces affines, puisqu'il n'y a pas d'opération définie sur un espace affine \mathcal{A} , autre que la translation par les éléments de l'espace vectoriel $\vec{\mathcal{A}}$ sur lequel il est modelé. Une idée intuitive pour définir la combinaison de deux points A et B avec les coefficients λ et μ pourrait cependant être la suivante. On fixe un point O dans \mathcal{A} on calcule la combinaison $\lambda \overrightarrow{OA} + \mu \overrightarrow{OB}$ et on translate O par ce vecteur pour obtenir la somme

$$\lambda A + \mu B = O + \lambda \overrightarrow{OA} + \mu \overrightarrow{OB}.$$

Le problème de cette construction est qu'elle dépend du choix de O . Si mon beau-frère Raoul^e choisit un point O' comme origine, il calculera le vecteur $\lambda \overrightarrow{O'A} + \mu \overrightarrow{O'B}$, puis traduira O' par ce vecteur et déclarera

$$\lambda A + \mu B = O' + \lambda \overrightarrow{O'A} + \mu \overrightarrow{O'B}.$$

La somme $\lambda A + \mu B$ n'est pas intrinsèquement définie, puisqu'elle dépend du choix d'une origine (en physique, on dirait sans doute que la construction dépend de l'observateur). Il y a donc autant de résultats en général qu'il y a de choix possibles d'origine O choisie pour faire le calcul. Cependant, nous avons déjà défini un certain nombre d'objets géométriques associés aux points A et B : le vecteur \overrightarrow{AB} , les multiples de ce vecteurs, toutes les translations de A (ou B) par les multiples de ce vecteur. Il y a donc possibilité de définir de nouveaux objets géométriques à partir de A et B . La façon simple de faire est de déterminer quand mon beau-frère Raoul et moi sommes d'accord.

Lemme 2.3.1. On a $O + \lambda \overrightarrow{OA} + \mu \overrightarrow{OB} = O' + \lambda \overrightarrow{O'A} + \mu \overrightarrow{O'B}$ si, et seulement si, $O = O'$ ou $\lambda + \mu = 1$.

c. Plus précisément, on devrait dire la classe d'équivalence

d. Cette notation n'est pas classique. Elle est souvent utilisée quand on a une relation d'équivalence, pour décrire un point de l'ensemble quotient associé, mais ceci est une autre histoire.

e. Comme chacun sait, mon beau-frère Raoul est un imbécile. Il ne fait rien comme tout le monde.

Démonstration. Il suffit de tout exprimer à partir de O par exemple. On a donc

$$O' + \lambda \overrightarrow{O'A} + \mu \overrightarrow{O'B} = O + \overrightarrow{OO'} + \lambda(\overrightarrow{O'O} + \overrightarrow{OA}) + \mu(\overrightarrow{O'O} + \overrightarrow{OB}).$$

Par la proposition 2.1.4, ce point vaut $O + \lambda \overrightarrow{OA} + \mu \overrightarrow{OB}$ si, et seulement si

$$\overrightarrow{OO'} + \lambda \overrightarrow{O'O} + \mu \overrightarrow{O'O} = 0,$$

ce qui donne le résultat annoncé. \square

D'après ce résultat, les points que l'on peut obtenir à l'aide de la construction ci-dessus sont les combinaisons

$$P = O + (1 - \mu) \overrightarrow{OA} + \mu \overrightarrow{OB}, \quad \mu \in \mathbb{R}$$

Ce n'est pas étonnant, puisqu'en développant, on a

$$P = O + \overrightarrow{OA} + \mu(\overrightarrow{OB} - \overrightarrow{OA}) = A + \mu \overrightarrow{AB}, \quad \mu \in \mathbb{R}.$$

On peut généraliser à des combinaisons quelconques.

Proposition 2.3.2. *Pour $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ tels que $\lambda_1 + \dots + \lambda_n = 1$, le point*

$$P = O + \lambda_1 \overrightarrow{OP_1} + \dots + \lambda_n \overrightarrow{OP_n} \tag{2.1}$$

est indépendant du point O .

Démonstration. On généralise la preuve du lemme 2.3.1. Si O' est un autre point de \mathcal{A} , on a

$$\begin{aligned} & O' + \lambda_1 \overrightarrow{O'P_1} + \dots + \lambda_n \overrightarrow{O'P_n} \\ &= O + \overrightarrow{OO'} + \lambda_1(\overrightarrow{O'O} + \overrightarrow{OP_1}) + \dots + \lambda_n(\overrightarrow{O'O} + \overrightarrow{OP_n}) \\ &= O + \lambda_1 \overrightarrow{OP_1} + \dots + \lambda_n \overrightarrow{OP_n} + (1 - \lambda_1 - \dots - \lambda_n) \overrightarrow{OO'}, \end{aligned}$$

ce qui suffit, vu l'hypothèse. \square

Définition 2.3.3. Le point P défini par la relation (2.1) est la *combinaison affine* des points P_1, \dots, P_n de coefficients $\lambda_1, \dots, \lambda_n$. Elle est aussi appelée *barycentre*. On la note

$$\lambda_1 P_1 + \dots + \lambda_n P_n.$$

Voici quelques exemples classiques, dont celui de segment $[A, B]$ et de milieu de ce segment, que nous redéfinissons.

Exemple 2.3.4. Le segment $[A, B]$ est l'ensemble

$$\{(1 - \lambda)A + \lambda B : \lambda \in [0, 1]\}.$$

Le milieu du segment $[A, B]$ est le point $M = \frac{1}{2}A + \frac{1}{2}B$.

On rappelle que $(1 - \lambda)A + \lambda B = A + \lambda \overrightarrow{AB}$, et que $\frac{1}{2}A + \frac{1}{2}B = A + \frac{1}{2} \overrightarrow{AB}$. On représentera le segment $[A, B]$ comme d'habitude.

Quand les coefficients λ et $1 - \lambda$ sont positifs, on peut interpréter la combinaison affine $(1 - \lambda)A + \lambda B$ comme le centre de masse des points A et B , avec les masses respectives $1 - \lambda$ et λ .

Comme nous l'avons pressenti le vecteur \overrightarrow{AB} et tous ses multiples sont des objets géométriques associés à A et B . Il est donc naturel d'essayer de les construire à partir de \overrightarrow{OA} et \overrightarrow{OB} . Nous connaissons déjà la réponse :

$$\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA}, \quad \forall O \in \mathcal{A}.$$

On pense donc naturellement à la généralisation de la proposition 2.3.2.

Proposition 2.3.5. Pour $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ tels que $\lambda_1 + \dots + \lambda_n = 0$, le vecteur

$$u = \lambda_1 \overrightarrow{OP_1} + \dots + \lambda_n \overrightarrow{OP_n} \in \overrightarrow{\mathcal{A}} \quad (2.2)$$

est indépendant du choix du point O .

Démonstration. La preuve est similaire à celle que nous avons déjà vue plus haut. Si O' est un autre point de \mathcal{A} , on a

$$\begin{aligned} & \lambda_1 \overrightarrow{O'P_1} + \dots + \lambda_n \overrightarrow{O'P_n} \\ &= \lambda_1 (\overrightarrow{O'O} + \overrightarrow{OP_1}) + \dots + \lambda_n (\overrightarrow{O'O} + \overrightarrow{OP_n}) \\ &= \lambda_1 \overrightarrow{OP_1} + \dots + \lambda_n \overrightarrow{OP_n} + (\lambda_1 + \dots + \lambda_n) \overrightarrow{O'O}, \end{aligned}$$

ce qui suffit, vu l'hypothèse. \square

Définition 2.3.6. Si $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ satisfont $\lambda_1 + \dots + \lambda_r = 0$, alors le vecteur $u = \lambda_1 \overrightarrow{OP_1} + \dots + \lambda_r \overrightarrow{OP_r}$ (pour tout $O \in \mathcal{A}$) est noté $\lambda_1 P_1 + \dots + \lambda_r P_r$.

Il est important de pouvoir passer d'une écriture à l'autre, en "simplifiant". C'est l'objet du résultat suivant, qui insiste sur des propriétés qui découlent en fait directement de la définition des combinaisons affines.

Proposition 2.3.7. Si $\lambda_1 + \dots + \lambda_n = 1$, alors les assertions suivantes sont équivalentes :

1. $A = \lambda_1 P_1 + \dots + \lambda_n P_n$;
2. Pour tout $O \in \mathcal{A}$ on a $\overrightarrow{OA} = \lambda_1 \overrightarrow{OP_1} + \dots + \lambda_n \overrightarrow{OP_n}$;
3. Il existe $O \in \mathcal{A}$ tel que $\overrightarrow{OA} = \lambda_1 \overrightarrow{OP_1} + \dots + \lambda_n \overrightarrow{OP_n}$.

Démonstration. Montrons que la première assertion implique la deuxième. Par définition, si $A = \lambda_1 P_1 + \dots + \lambda_n P_n$, on a pour tout O , $A = O + \lambda_1 \overrightarrow{OP_1} + \dots + \lambda_n \overrightarrow{OP_n}$. Mais puisqu'on a aussi $A = O + \overrightarrow{OA}$, on peut conclure par la proposition 2.1.4. La deuxième assertion implique visiblement la troisième. La troisième implique

$$A = O + \overrightarrow{OA} = O + \lambda_1 \overrightarrow{OP_1} + \dots + \lambda_n \overrightarrow{OP_n} = \lambda_1 P_1 + \dots + \lambda_n P_n,$$

ce qui termine la preuve. \square

Voici la propriété similaire quand la somme des coefficients est nulle.

Proposition 2.3.8. Si $\lambda_1 + \dots + \lambda_n = 0$, alors les assertions suivantes sont équivalentes :

1. On a $u = \lambda_1 P_1 + \dots + \lambda_n P_n$;
2. Pour tout $O \in \mathcal{A}$ on a $u = \lambda_1 \overrightarrow{OP_1} + \dots + \lambda_n \overrightarrow{OP_n}$;
3. Il existe $O \in \mathcal{A}$ tel que $u = \lambda_1 \overrightarrow{OP_1} + \dots + \lambda_n \overrightarrow{OP_n}$.

Démonstration. La preuve découle directement de la définition 2.3.6. \square

Ces deux propositions permettent d'énoncer des règles de calcul pratique pour les combinaisons du type $\lambda_1 P_1 + \dots + \lambda_n P_n$.

On peut manipuler de telles combinaisons pour autant que l'on s'assure que le résultat a toujours un sens, c'est-à-dire que la somme des coefficients soit toujours soit 0, soit 1.

Si on prend soin de n'avoir que des sommes de coefficients égales à 0 ou à 1, on peut toujours se ramener à une propriété des vecteurs en utilisant les propositions précédentes.

On peut remarquer également que si on voulait écrire un résultat avec une somme des coefficients inadmissible, c'est-à-dire différente de 0 ou 1, il suffit pour la rendre admissible de diviser chaque coefficient par leur somme, comme le montre l'exemple suivant.

Exemple 2.3.9. Soient A, B, C, D quatre points de l'espace affine \mathcal{A} . Si on a $D = A + B - C$, alors on a également $D - A = B - C$, c'est-à-dire $\overrightarrow{AD} = \overrightarrow{CB}$. On ne peut cependant pas écrire $D + C = A + B$, car la somme des coefficients des combinaisons dans chaque membre vaut 2. On divise alors chaque coefficient par 2 et on peut écrire $\frac{1}{2}D + \frac{1}{2}C = \frac{1}{2}A + \frac{1}{2}B$, qui a un sens géométrique, et qui est équivalente à la première égalité.

Si vous souhaitez justifier ces calculs, vous considérez un point O quelconques, écrivez des vecteurs liés en O , et effectuez les mêmes transformations d'équations.

Terminons cette section par deux résultats concernant l'interprétation et le calcul des barycentres. Le premier est bien connu et son interprétation physique est évidente.

Proposition 2.3.10. Si $\lambda_1 + \dots + \lambda_n = 1$, alors on a $Q = \sum_{i=1}^n \lambda_i P_i$ si, et seulement si, on a $\sum_{i=1}^n \lambda_i \overrightarrow{QP_i} = 0$.

Démonstration. On utilise simplement la proposition 2.3.7 avec $O = Q$. □

A titre d'exemple, on peut considérer les point A et B d'un espace affine et leur associer les poids $\frac{1}{3}$ et $\frac{2}{3}$. Le barycentre est alors $M = \frac{1}{3}A + \frac{2}{3}B$. Ceci est équivalent à $\frac{1}{3}\overrightarrow{MA} + \frac{2}{3}\overrightarrow{MB} = 0$.

On sait aussi en physique que l'on peut calculer le centre de masse d'un système par étapes. En mathématiques, ce fait se généralise, mais il faut être prudent sur les sommes de coefficients, qui peuvent s'annuler, car on accepte des coefficients négatifs.

Proposition 2.3.11. Si $\sum_{i=1}^n \lambda_i = 1$, si P_1, \dots, P_n sont des points de \mathcal{A} , et si $m < n$ est tel que $\alpha = \sum_{i=1}^m \lambda_i \notin \{0, 1\}$, alors on a

$$\sum_{i=1}^n \lambda_i P_i = \alpha B_1 + (1 - \alpha) B_2,$$

où $B_1 = \sum_{i=1}^m \frac{\lambda_i}{\alpha} P_i$ et $B_2 = \sum_{i=m+1}^n \frac{\lambda_i}{1-\alpha} P_i$.

Démonstration. Utilisons encore la proposition 2.3.7 : on a $A = \sum_{i=1}^n \lambda_i P_i$ si, et seulement si

$$\overrightarrow{OA} = \sum_{i=1}^n \lambda_i \overrightarrow{OP_i}$$

où O est un point de \mathcal{A} . Mais on a

$$\sum_{i=1}^n \lambda_i \overrightarrow{OP_i} = \alpha \left(\sum_{i=1}^m \frac{\lambda_i}{\alpha} \overrightarrow{OP_i} \right) + (1 - \alpha) \left(\sum_{i=m+1}^n \frac{\lambda_i}{1 - \alpha} \overrightarrow{OP_i} \right) = \alpha \overrightarrow{OB_1} + (1 - \alpha) \overrightarrow{OB_2},$$

toujours en utilisant la même proposition 2.3.7. On obtient ainsi le résultat annoncé. □

2.4 Variétés affines

Les variétés affines sont aux espaces affines ce que les sous-espaces vectoriels sont aux espaces vectoriels. Elles sont d'ailleurs parfois appelées sous-espaces affines. Elles sont les généralisations des droites et plans que vous avez déjà rencontrés. Nous adoptons naturellement une définition similaire à celle des sous-espaces vectoriels.

Définition 2.4.1. Soit \mathcal{A} un espace affine. Une variété affine de \mathcal{A} est un sous ensemble non vide $\mathcal{V} \subset \mathcal{A}$ qui contient les combinaisons affines de ses éléments.

Voici quelques exemples simples.

- Dans $\mathcal{A} = \mathbb{R}^2$, $\mathcal{V} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x + 2y = 1 \right\}$ est une variété affine ;
- Dans tout espace affine \mathcal{A} , si $A, B \in \mathcal{A}$, $\{(1 - \lambda)A + \lambda B : \lambda \in \mathbb{R}\}$ est une variété affine ;
- Dans tout espace affine \mathcal{A} , les singletons \mathcal{A} sont des variétés affines.

La proposition suivante donne un exemple fondamental, parce que d'une part il fournit toute une famille d'exemples, et d'autre part, il donne naissance à la notion de sous-espace vectoriel directeur, qui donne un lien étroit entre variétés affines et sous-espaces vectoriels, que nous utiliserons intensivement pour écrire des équations cartésiennes.

Proposition 2.4.2 (Exemple fondamental). *Si P est un point de \mathcal{A} et si F un sous-espace vectoriel de $\overrightarrow{\mathcal{A}}$, alors*

$$\mathcal{V} = P + F = \{P + u : u \in F\}$$

est une variété affine de \mathcal{A} .

Démonstration. L'ensemble \mathcal{V} n'est pas vide puisqu'il contient P . Soient maintenant $P_i = P + u_i$, $i \leq n$ et λ_i des coefficients tels que $\sum_{i=1}^n \lambda_i = 1$. Montrons que le point $Q = \sum_{i=1}^n \lambda_i P_i$ appartient à \mathcal{V} . On a

$$Q = P + \overrightarrow{PQ} = P + \sum_{i=1}^n \lambda_i \overrightarrow{PP_i} = P + \sum_{i=1}^n \lambda_i u_i.$$

On arrive à la conclusion puisque $\sum_{i=1}^n \lambda_i u_i$ appartient à F . □

Passons à une caractérisation des variétés affines, qui montre que toute variété affine est de la forme $P + F$, pour un certain P et un certain F .

Proposition 2.4.3. *Soit \mathcal{V} une variété affine de \mathcal{A} et P un point de \mathcal{V} . L'ensemble*

$$V_P = \{\overrightarrow{PQ} : Q \in \mathcal{V}\}$$

est un sous-espace vectoriel de E . On a $X \in \mathcal{V}$ si, et seulement si, $\overrightarrow{PX} \in V_P$ et aussi $\mathcal{V} = P + V_P$. De plus, V_P est l'unique sous-espace vectoriel ayant cette propriété. Enfin, V_P est indépendant du choix de P dans \mathcal{V} .

Démonstration. Pour démontrer que V_P est un sous-espace vectoriel, il y a trois conditions à vérifier.

- L'ensemble V_P contient le vecteur nul car $\overrightarrow{PP} = 0$.
- Soient u_1 et u_2 deux éléments de V_P . Il existe Q_1, Q_2 dans \mathcal{V} tels que $u_1 = \overrightarrow{PQ_1} = Q_1 - P$ et $u_2 = \overrightarrow{PQ_2} = Q_2 - P$. On a alors $u_1 + u_2 = Q_1 - P + Q_2 - P = \overrightarrow{PQ}$ si $Q = Q_1 + Q_2 - P$. Le point Q ainsi défini est dans \mathcal{V} , donc $u_1 + u_2$ appartient à V_P .
- Soit $u_1 \in V_P$ et $\lambda \in \mathbb{R}$. Il existe $Q_1 \in \mathcal{V}$ tel que $u = \overrightarrow{PQ_1} = Q_1 - P$. Alors $\lambda u = \lambda Q_1 - \lambda P = \overrightarrow{PQ}$ si $Q = \lambda Q_1 + (1 - \lambda)P$. Ce point est encore dans \mathcal{V} puisque c'est une combinaison affine de deux points de \mathcal{V} .

Ensuite, par définition, on a $\overrightarrow{PX} \in V_P$ si, et seulement si il existe $Q \in \mathcal{V}$ tel que $\overrightarrow{PX} = \overrightarrow{PQ}$. Cette dernière égalité est équivalente à $X = Q$. On a donc $\overrightarrow{PX} \in V_P$ si, et seulement si il existe $Q \in \mathcal{V}$ tel que $X = Q$.

On a $X \in P + V_P$ si, et seulement si il existe $u \in V_P$ tel que $X = P + u$. Dans ce cas u ne peut-être que \overrightarrow{PX} . On a donc $X \in P + V_P$ si, et seulement si $\overrightarrow{PX} \in V_P$, ce qui permet de conclure par le point précédent.

Si V' est un autre sous-espace vectoriel satisfaisant $\mathcal{V} = P + V'$, alors pour tout $u' \in V'$, on a $P + u' \in \mathcal{V}$, donc il existe $u \in V_P$ tel que $P + u' = P + u$. Donc $u' = u$ et u' appartient à V_P . On a donc $V' \subset V_P$ et par symétrie $V_P \subset V'$.

Soit P' un autre point de \mathcal{V} . Démontrons que $V_{P'} = V_P$. Soit $u \in V_{P'}$. Il existe $Q \in \mathcal{V}$ tel que $u = \overrightarrow{P'Q}$. Alors on a $u = \overrightarrow{PQ} - \overrightarrow{PP'}$ et u appartient à V_P . Donc on a $V_{P'} \subset V_P$ et par symétrie, $V_P \subset V_{P'}$. □

Définition 2.4.4. Si \mathcal{V} est une variété affine, le sous-espace vectoriel défini par la proposition précédente est le sous-espace vectoriel directeur de \mathcal{V} . On le note $\overrightarrow{\mathcal{V}}$.

Remarque 2.3. On aurait pu aussi naturellement poser $\overrightarrow{\mathcal{V}} = \{\overrightarrow{AB} : A, B \in \mathcal{V}\}$. Je vous laisse comme exercice de montrer que les deux définitions coïncident.

En calculant le sous-espace vectoriel directeur de l'espace affine \mathcal{A} , on obtient bien $\overrightarrow{\mathcal{A}} = E$.

Avant d'aller plus loin, voici quelques points de vocabulaire.

Définition 2.4.5. Un espace affine \mathcal{A} est de dimension n si l'espace vectoriel E sur lequel il est modelé est de dimension n . On note alors $\dim \mathcal{A} = n$; La dimension d'une variété affine \mathcal{V} de \mathcal{A} est la dimension de $\vec{\mathcal{V}}$. On la note $\dim \mathcal{V}$. Une droite est une variété affine de dimension 1; Un plan est une variété affine de dimension 2 (si $\dim \mathcal{A} \geq 2$); Un hyperplan est une variété affine de dimension $n - 1$, si $\dim \mathcal{A} = n$.

Rappelons maintenant, à titre d'exemples, le cas particulier de droites et plans, que vous connaissez bien. Attention cependant : nous sommes dans un espace affine de dimension n quelconque.

Proposition 2.4.6 (Droite déterminée par deux points). *Soit \mathcal{A} un espace affine de dimension $n \geq 1$, et soient $A, B \in \mathcal{A}$, distincts. Il existe une unique droite d (notée AB) qui contient A et B , on a $\vec{d} = \langle \vec{AB} \rangle$. La droite AB est l'ensemble de toutes les combinaisons affines de A et B . De plus, si P et Q sont deux points distincts dans AB , alors $AB = PQ$.*

Démonstration. Si d est une droite contenant A et B , alors on peut écrire $d = A + \vec{d}$ et on a $\vec{AB} \in \vec{d}$. On a donc $\langle \vec{AB} \rangle \subset \vec{d}$. Ces sous-espaces vectoriels ayant la même dimension, ils sont égaux. Donc le seul choix possible est $d = A + \langle \vec{AB} \rangle$, qui heureusement convient. On a donc prouvé l'existence et l'unicité de d et caractérisé \vec{d} . On a aussi

$$AB = A + \langle \vec{AB} \rangle = \{A + \lambda \vec{AB} : \lambda \in \mathbb{R}\} = \{(1 - \lambda)A + \lambda B : \lambda \in \mathbb{R}\}.$$

Enfin, si P et Q appartiennent à AB , alors AB est une droite contenant P et Q . Mais PQ est l'unique droite contenant PQ . On a donc $PQ = AB$, vu l'unicité. \square

Remarque 2.4. Tout vecteur non nul u de \vec{d} est appelé un vecteur directeur de d car si u est un tel vecteur, on a $\vec{d} = \langle u \rangle$.

La proposition précédente se généralise au cas des plans.

Proposition 2.4.7 (Plan déterminé par trois points). *Soit \mathcal{A} un espace affine tel que $\dim \mathcal{A} \geq 2$, et soient $A, B, C \in \mathcal{A}$, non alignés.^f Il existe un unique plan π (encore noté ABC) qui contient A, B et C . On a $\vec{\pi} = \langle \vec{AB}, \vec{AC} \rangle = \{\lambda \vec{AB} + \mu \vec{AC} : \lambda, \mu \in \mathbb{R}\}$. De plus π est l'ensemble des combinaisons affines de A, B et C .*

Démonstration. Remarquons tout d'abord que $A \neq B$, sinon A, B, C seraient alignés. De plus on a $C \in AB = A + \langle \vec{AB} \rangle$ si et seulement si \vec{AC} est multiple de \vec{AB} . Par hypothèse, ce n'est pas le cas et on a $\dim \langle \vec{AB}, \vec{AC} \rangle = 2$. On procède alors comme plus haut : si π contient A, B, C , alors on écrit $\pi = A + \vec{\pi}$ et on a $\vec{AB} \in \vec{\pi}$ et $\vec{AC} \in \vec{\pi}$, donc $\langle \vec{AB}, \vec{AC} \rangle \subset \vec{\pi}$. Ces sous-espaces ayant la même dimension, ils sont égaux, et on a nécessairement

$$\pi = A + \langle \vec{AB}, \vec{AC} \rangle.$$

Ce plan convient et $\vec{\pi}$ a la forme annoncée. Enfin, on a

$$\pi = A + \langle \vec{AB}, \vec{AC} \rangle = \{A + \lambda \vec{AB} + \mu \vec{AC} : \lambda, \mu \in \mathbb{R}\} = \{(1 - \lambda - \mu)A + \lambda B + \mu C : \lambda, \mu \in \mathbb{R}\},$$

ce qui achève la preuve. \square

L'idée qui vient d'être développée dans les propositions précédentes est de définir une variété affine par un ensemble de points qu'elle doit contenir. Elle est généralisée par la définition suivante, qui est également similaire à la définition que nous avons posée pour les enveloppes dans le cas des espaces vectoriels.

f. Des points sont alignés s'ils appartiennent à une même droite.

Définition 2.4.8. Soient $P_1, \dots, P_r \in \mathcal{A}$. L'enveloppe affine de P_1, \dots, P_r est l'ensemble de toutes les combinaisons affines de ces points. On la note $\rangle P_1, \dots, P_r \langle_{\mathcal{A}}$.

Comme toujours, nous prouvons que cet ensemble a les propriétés adéquates.

Proposition 2.4.9. Soient \mathcal{A} un espace affine et $P_1, \dots, P_r \in \mathcal{A}$. Alors, on a $\rangle P_1, \dots, P_r \langle_{\mathcal{A}} = P_1 + \rangle \overrightarrow{P_1 P_2}, \dots, \overrightarrow{P_1 P_r} \langle$. L'enveloppe affine $\rangle P_1, \dots, P_r \langle_{\mathcal{A}}$ est une variété affine qui est incluse dans toute variété affine contenant P_1, \dots, P_r . On a de plus

$$\dim \rangle P_1, \dots, P_r \langle_{\mathcal{A}} = \dim \rangle \overrightarrow{P_1 P_2}, \dots, \overrightarrow{P_1 P_r} \langle.$$

Démonstration. On a $\rangle P_1, \dots, P_r \langle_{\mathcal{A}} \subset P_1 + \rangle \overrightarrow{P_1 P_2}, \dots, \overrightarrow{P_1 P_r} \langle$. En effet, si $\lambda_1 + \dots + \lambda_r = 1$, alors

$$\lambda_1 P_1 + \dots + \lambda_r P_r = P_1 + (\lambda_1 \overrightarrow{P_1 P_1} + \dots + \lambda_r \overrightarrow{P_1 P_r}) = P_1 + (\lambda_2 \overrightarrow{P_1 P_2} + \dots + \lambda_r \overrightarrow{P_1 P_r}).$$

Pour l'autre inclusion, on note que pour tous $\mu_1, \dots, \mu_r \in \mathbb{R}$, on a

$$P_1 + (\mu_2 \overrightarrow{P_1 P_2} + \dots + \mu_r \overrightarrow{P_1 P_r}) = (1 - \mu_2 - \dots - \mu_r) P_1 + \mu_2 P_2 + \dots + \mu_r P_r.$$

L'enveloppe affine des points P_1, \dots, P_r est donc une variété affine (voir l'exemple fondamental) et on a également son sous-espace vectoriel directeur et sa dimension. Elle est incluse dans toute variété affine qui contient P_1, \dots, P_r , par définition. \square

La dimension de $\rangle \overrightarrow{P_1 P_2}, \dots, \overrightarrow{P_1 P_r} \langle$ étant égale au nombre de vecteurs linéairement indépendants parmi $\overrightarrow{P_1 P_2}, \dots, \overrightarrow{P_1 P_r}$, il est naturel de poser la définition suivante.

Définition 2.4.10. Des points P_1, \dots, P_r d'un espace affine sont *affinement indépendants* si $(\overrightarrow{P_1 P_2}, \dots, \overrightarrow{P_1 P_r})$ est une famille *linéairement* indépendante.

On peut montrer que les points P_1, \dots, P_r sont affinement indépendants si aucun de ces points n'est combinaison affine des autres. Ainsi, un point P est affinement indépendant. Deux points distincts sont affinement indépendants, ainsi que trois points non alignés ou quatre points non coplanaires.

2.5 Intersection et parallélisme de variétés affines

Vous avez déjà étudié les positions relatives des droites et des plans. Ces positions sont, dans le contexte affine, définies par l'intersection et le parallélisme. Nous allons comme d'habitude poser les définitions générales et les particulariser en dimension 2 et 3, pour vérifier qu'elles coïncident avec l'intuition que nous en avons. En ce qui concerne l'intersection, nous n'avons pas besoin de définition autre que celle que nous connaissons déjà, puisque les variétés affines sont des ensembles.

Proposition 2.5.1. L'intersection de deux variétés affines \mathcal{V}_1 et \mathcal{V}_2 est soit vide soit une variété affine. Dans ce cas on a $\mathcal{V}_1 \cap \mathcal{V}_2 = \overrightarrow{\mathcal{V}_1} \cap \overrightarrow{\mathcal{V}_2}$.

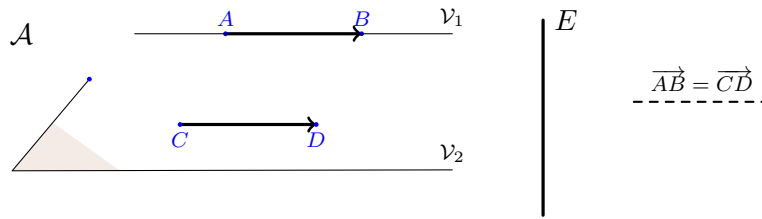
Démonstration. On doit prouver une alternative. On suppose donc qu'un des cas n'est pas vrai et on prouve que l'autre l'est. Si $\mathcal{V}_1 \cap \mathcal{V}_2$ n'est pas vide, cet ensemble contient visiblement les combinaisons affines de ses éléments, donc c'est une variété affine.

Soit $P \in \mathcal{V}_1 \cap \mathcal{V}_2$. On a $u \in \overrightarrow{\mathcal{V}_1 \cap \mathcal{V}_2}$ si et seulement si il existe $Q \in \mathcal{V}_1 \cap \mathcal{V}_2$ tel que $u = \overrightarrow{PQ}$. Dans ce cas, on a $Q \in \mathcal{V}_1$ et donc $u \in \overrightarrow{\mathcal{V}_1}$, et pour la même raison $u \in \overrightarrow{\mathcal{V}_2}$. Réciproquement, si $u \in \overrightarrow{\mathcal{V}_1} \cap \overrightarrow{\mathcal{V}_2}$, alors $Q = P + u$ est dans \mathcal{V}_1 car u appartient à $\overrightarrow{\mathcal{V}_1}$ et pour la même raison on a $Q \in \mathcal{V}_2$. Donc $u = \overrightarrow{PQ}$ pour un $Q \in \mathcal{V}_1 \cap \mathcal{V}_2$, et finalement $u \in \overrightarrow{\mathcal{V}_1 \cap \mathcal{V}_2}$. \square

En ce qui concerne le parallélisme, puisque la "direction" d'une variété affine est intuitivement donnée par son sous-espace vectoriel directeur, l'intuition commande de poser que des variétés affines sont parallèles si elles ont le même sous-espace vectoriel directeur. Cependant, cette définition est trop restrictive puisqu'elle ne permet pas de traduire le parallélisme de variétés affines de dimensions différentes.

Définition 2.5.2. Deux variétés affines \mathcal{V}_1 et \mathcal{V}_2 sont parallèles si $\vec{\mathcal{V}}_1 \subset \vec{\mathcal{V}}_2$ ou $\vec{\mathcal{V}}_2 \subset \vec{\mathcal{V}}_1$. On note alors $\mathcal{V}_1 // \mathcal{V}_2$.

Voici la représentation graphique à laquelle vous êtes habitués, du parallélisme d'une droite et d'un plan (ou d'une variété affine de dimension supérieure).



Les propriétés qui lient en général l'intersection et le parallélisme sont assez naturelles. Nous les regroupons dans le résultat suivant.

Proposition 2.5.3.

1. Si $\mathcal{V}_1 \cap \mathcal{V}_2 \neq \emptyset$ et $\mathcal{V}_1 // \mathcal{V}_2$ alors $\mathcal{V}_1 \subset \mathcal{V}_2$ ou $\mathcal{V}_2 \subset \mathcal{V}_1$;
2. Si $\dim \mathcal{V}_1 = \dim \mathcal{V}_2$, alors $\mathcal{V}_1 // \mathcal{V}_2$ si, et seulement si $\vec{\mathcal{V}}_1 = \vec{\mathcal{V}}_2$;
3. Dans ce dernier cas, si de plus $\mathcal{V}_1 \cap \mathcal{V}_2 \neq \emptyset$, alors $\mathcal{V}_1 = \mathcal{V}_2$;
4. Soient \mathcal{V} une variété affine et $A \in \mathcal{A}$. Il existe une unique variété affine \mathcal{V}_A satisfaisant les conditions $A \in \mathcal{V}_A$, $\mathcal{V}_A // \mathcal{V}$ et $\dim \mathcal{V}_A = \dim \mathcal{V}$.

Remarque 2.5. Le dernier point est une généralisation du plus célèbre axiome d'Euclide qui, dans notre approche, n'en est évidemment plus un.

Démonstration. Soit P un point de $\mathcal{V}_1 \cap \mathcal{V}_2$ et supposons, sans perte de généralité, qu'on a l'inclusion $\vec{\mathcal{V}}_1 \subset \vec{\mathcal{V}}_2$. Si Q appartient à \mathcal{V}_1 , alors $\vec{PQ} \in \vec{\mathcal{V}}_1$, donc $\vec{PQ} \in \vec{\mathcal{V}}_2$ et $Q \in \mathcal{V}_2$.

Si les variétés sont de même dimension, il en est de même pour leurs sous-espaces vectoriels directeurs, et l'inclusion des ces sous-espaces qui traduit le parallélisme est alors équivalente à leur égalité.

On démontre le troisième point comme le premier, mais on a deux inclusions pour les sous-espaces vectoriels, donc deux inclusions pour les variétés affines.

Enfin, si \mathcal{V}_A est la variété affine cherchée, on peut écrire $\mathcal{V}_A = A + \vec{\mathcal{V}}_A$, mais on a $\vec{\mathcal{V}}_A = \vec{\mathcal{V}}$. Donc on a nécessairement $\mathcal{V}_A = A + \vec{\mathcal{V}}$. On constate que cette variété affine répond bien à la question. \square

2.5.1 Positions relatives de droites en dimensions 2 et 3

Nous dirons que deux droites sont dites sécantes si leur intersection est un singleton^g. On a alors en dimension 2 le résultat bien connu^h.

Proposition 2.5.4. Soit \mathcal{A} un espace affine de dimension 2. Deux droites de \mathcal{A} sont soit parallèles soit sécantes.

Démonstration. Supposons que les droites $d_1 = A + \langle u \rangle$ et $d_2 = B + \langle v \rangle$ ne soient pas parallèles. Alors les vecteurs u et v sont linéairement indépendants et on a $\vec{\mathcal{A}} = \langle u, v \rangle$. On peut donc décomposer

$$\vec{AB} = \lambda u + \mu v,$$

où λ et μ sont des nombres réels. On a donc

$$B - \mu v = A + \lambda u.$$

Le membre de droite définit un point P de d_1 tandis que le membre de gauche définit un point de d_2 , donc l'intersection des droites d_1 et d_2 n'est pas vide. Enfin, si $d_1 \cap d_2$ contenait un autre point Q , alors on aurait $d_1 = d_2 = PQ$, et d_1 serait parallèle à d_2 . \square

g. Rappelons qu'un singleton est un ensemble qui contient exactement un élément.

h. Attention, dans ce cours, c'est une proposition, pas une définition du parallélisme.

Passons maintenant aux positions relatives de droites en dimension 3. La situation est en fait plus simple.

Définition 2.5.5. Dans un espace affine \mathcal{A} de dimension trois, deux droites sont gauches si elles ne sont ni parallèles, ni sécantes.

Puisque en vertu de la proposition 2.5.3, des droites ne peuvent pas être à la fois sécantes et parallèles, on a par définition trois cas exclusifs pour la position de deux droites en dimension 3 : elles sont soit parallèles, soit sécantes soit gauches. Il est cependant utile de pouvoir déterminer dans quel cas on se trouve quand les droites sont données par un point et un vecteur directeur. C'est l'objet du résultat suivant.

Proposition 2.5.6. Soit \mathcal{A} un espace affine de dimension trois et deux droites $d_1 = A+\rangle u\langle$, $d_2 = A+\rangle v\langle$.

1. Si la famille (u, v) est linéairement dépendante, alors $d_1 // d_2$;
2. Si (u, v) est linéairement indépendante et $(u, v, \overrightarrow{AB})$ est linéairement dépendante, alors d_1 et d_2 sont sécantes ;
3. Si la famille $(u, v, \overrightarrow{AB})$ est linéairement indépendante, alors d_1 et d_2 sont gauches.

Démonstration. Si (u, v) est linéairement dépendante, alors les vecteurs (non nuls) u et v sont multiples l'un de l'autre, et on a $\rangle u\langle = \rangle v\langle$, donc les droites sont parallèles.

Dans le deuxième cas, on déduit que \overrightarrow{AB} est combinaison linéaire de u et v et on montre, comme dans la proposition 2.5.4 que d_1 et d_2 ont un point en commun. Elles ne peuvent avoir plus d'un point en commun, sinon elles seraient égales et donc parallèles, et (u, v) serait une famille linéairement dépendante. \square

Passons maintenant aux positions relatives de deux plans et d'une droite et d'un plan. Les démonstrations sont similaires à celles qui viennent d'être faites et je les laisse donc comme exercice.

Proposition 2.5.7. Dans un espace affine \mathcal{A} de dimension 3, si deux plans ne sont pas parallèles, leur intersection est une droite.

Proposition 2.5.8. Soit d une droite et π un plan dans un espace affine de dimension 3. Soit d et π sont parallèles, soit leur intersection est un singleton.

2.6 Repères et coordonnées

Nous savons que dans un espace vectoriel de dimension finie, le choix d'une base permet d'associer à tout vecteur un n -uplet de nombres, un élément de \mathbb{R}^n . Cette correspondance linéaire identifie l'espace vectoriel abstrait à \mathbb{R}^n . On se propose d'étendre cette construction aux espaces affines de dimension finie. L'idée est extrêmement simple : si \mathcal{A} un espace affine de dimension n , une fois que l'on s'est fixé un point $O \in \mathcal{A}$ comme origine, chaque point P définit un vecteur lié en O (à savoir (O, P)) qui correspond lui même au vecteur \overrightarrow{OP} de $\overrightarrow{\mathcal{A}}$. Puisque $\overrightarrow{\mathcal{A}}$ est un espace vectoriel, il n'y a plus qu'à y fixer une base pour obtenir un n -uplet de nombres à partir du point P .

Définition 2.6.1. Un repère \mathcal{R} de \mathcal{A} est un couple (O, \mathcal{B}) où O est un point de \mathcal{A} appelé origine du repère et où \mathcal{B} est une base de $\overrightarrow{\mathcal{A}}$ appelée base du repère.

Les coordonnées d'un point X quelconque de \mathcal{A} dans le repère \mathcal{R} sont les composantes de \overrightarrow{OX} dans \mathcal{B} . On les note $\Phi_{\mathcal{R}}(X)$.

Remarque 2.6. Afin de ne pas alourdir l'exposé, et pour permettre des notations faciles dans les calculs, nous utiliserons parfois la notation $X : (x_1, \dots, x_n)^\sim$ pour indiquer que X admet les coordonnées $(x_1, \dots, x_n)^\sim$ dans le repère donné.

Nous avons construit la notion de repère pour que le passage aux coordonnées soit une bijection. Il n'est pas inutile de le montrer explicitement.

Proposition 2.6.2. *Soit un repère $\mathcal{R} = (O, (b_1, \dots, b_n))$ d'un espace affine. Alors P admet pour coordonnées $(p_1, \dots, p_n)^\sim$ dans \mathcal{R} si, et seulement si*

$$P = O + p_1 b_1 + \dots + p_n b_n. \quad (2.3)$$

En particulier, le passage aux coordonnées dans \mathcal{R} définit une bijection $\Phi_{\mathcal{R}}$ de \mathcal{A} dans \mathbb{R}^n .

Démonstration. Par définition, P admet pour coordonnées $(p_1, \dots, p_n)^\sim$ dans \mathcal{R} si, et seulement si \overrightarrow{OP} admet pour composantes $(p_1, \dots, p_n)^\sim$ dans \mathcal{B} . Par définition des composantes dans une base, ceci est équivalent à la relation

$$\overrightarrow{OP} = p_1 b_1 + \dots + p_n b_n,$$

qui est, comme nous le savons, équivalente à la relation de l'énoncé. \square

Dans les espaces vectoriels, nous avons montré que le passage aux composantes dans une base donnée est une application linéaire. Dans le cadre des espaces affines, la notion de combinaison linéaire n'est plus définie, mais nous avons défini deux façons de combiner des points pour former de nouveaux objets géométriques. Les propositions suivantes, qui sont fondamentales pour pouvoir mener des calculs dans les espaces affines, montrent que le passage aux coordonnées se comporte bien vis à vis de ces constructions.

Proposition 2.6.3. *Soit $\mathcal{R} = (O, \mathcal{B})$ un repère d'un espace affine de dimension n . L'application $\Phi_{\mathcal{R}}$ est affine : on a*

$$\Phi_{\mathcal{R}}\left(\sum_{i=1}^r \lambda_i P_i\right) = \sum_{i=1}^r \lambda_i \Phi_{\mathcal{R}}(P_i)$$

pour tous points P_1, \dots, P_r de \mathcal{A} et tous $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ tels que $\sum_{i=1}^r \lambda_i = 1$.

Démonstration. Par définition du passage aux coordonnées, on a

$$\Phi_{\mathcal{R}}(X) = \Phi_{\mathcal{B}}(\overrightarrow{OX}),$$

pour tout $X \in \mathcal{A}$. Soit maintenant $P = \sum_{i=1}^r \lambda_i P_i$. Par définition des combinaisons affines, cela signifie que

$$P = O + \sum_{i=1}^r \lambda_i \overrightarrow{OP_i},$$

ou encore

$$\overrightarrow{OP} = \sum_{i=1}^r \lambda_i \overrightarrow{OP_i}.$$

On a donc finalement

$$\Phi_{\mathcal{R}}(P) = \Phi_{\mathcal{B}}(\overrightarrow{OP}) = \Phi_{\mathcal{B}}\left(\sum_{i=1}^r \lambda_i \overrightarrow{OP_i}\right) = \sum_{i=1}^r \lambda_i \Phi_{\mathcal{B}}(\overrightarrow{OP_i}) = \sum_{i=1}^r \lambda_i \Phi_{\mathcal{R}}(P_i),$$

en utilisant la linéarité de $\Phi_{\mathcal{B}}$ et on obtient ce qui était demandé. \square

Proposition 2.6.4. *Soit $\mathcal{R} = (O, \mathcal{B})$ un repère d'un espace affine de dimension n . On a*

$$\Phi_{\mathcal{B}}\left(\sum_{i=1}^r \lambda_i P_i\right) = \sum_{i=1}^r \lambda_i \Phi_{\mathcal{R}}(P_i)$$

pour tous points P_1, \dots, P_r de \mathcal{A} et tous $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ tels que $\sum_{i=1}^r \lambda_i = 0$.

i. On peut retenir que les coordonnées d'une combinaison affine sont la combinaison affine des coordonnées.

Démonstration. La preuve est analogue à celle de la proposition précédente. Écrivons-la pour être complets. Par définition de la combinaison $\sum_{i=1}^r \lambda_i P_i$ quand $\sum_{i=1}^r \lambda_i = 0$, on a

$$\sum_{i=1}^r \lambda_i P_i = \sum_{i=1}^r \lambda_i \overrightarrow{OP_i}.$$

Dès lors on a

$$\Phi_{\mathcal{B}}\left(\sum_{i=1}^r \lambda_i P_i\right) = \Phi_{\mathcal{B}}\left(\sum_{i=1}^r \lambda_i \overrightarrow{OP_i}\right) = \sum_{i=1}^r \lambda_i \Phi_{\mathcal{B}}(\overrightarrow{OP_i}) = \sum_{i=1}^r \lambda_i \Phi_{\mathcal{R}}(P_i),$$

ce qui termine la preuve. \square

Voici quelques exemples.

Exemple 2.6.5. Soit un espace affine \mathcal{A} de dimension 2, muni d'un repère $\mathcal{R} = (O, (b_1, b_2))$. Soient les points A et B définis par leurs coordonnées dans $\mathcal{R} : A : (1, 2)^\sim$ et $B : (5, 8)^\sim$. alors le milieu de $[A, B]$ a pour coordonnées $(3, 5)^\sim$ dans le repère \mathcal{R} . En effet, le milieu de $[A, B]$ est par définition $\frac{1}{2}A + \frac{1}{2}B$, et a donc pour coordonnées $\frac{1}{2}(1, 2)^\sim + \frac{1}{2}(5, 8)^\sim$. De la même manière, le vecteur \overrightarrow{AB} a pour composantes $(4, 6)^\sim$ dans (b_1, b_2) , puisque $\overrightarrow{AB} = B - A$.

Exemple 2.6.6. Soit un espace affine \mathcal{A} de dimension 3, muni d'un repère $\mathcal{R} = (O, (b_1, b_2, b_3))$. Si dans ce repère, on a

$$A : \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad B : \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix}, \quad C : \begin{pmatrix} 3 \\ -4 \\ 2 \end{pmatrix}$$

alors le centre de gravité du triangle ABC a pour coordonnées

$$\frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} 3 \\ -4 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}.$$

Inutile donc, pour calculer le centre de gravité, de calculer l'équation des médianes et de résoudre un système d'équations pour obtenir les coordonnées (comme nous allons le voir dans la section suivante).

2.7 Equations de variétés affines

Dans le cadre des espaces vectoriels, nous avons vu comment caractériser un sous-espace vectoriel quelconque à partir d'équations paramétriques ou cartésiennes. Le but de cette section est de transposer ces résultats dans le cadre des espaces affines, muni d'un repère, pour caractériser les variétés affines. On considère donc un espace affine \mathcal{A} de dimension n , muni d'un repère $\mathcal{R} = (O; \mathcal{B})$, où $\mathcal{B} = (b_1, \dots, b_n)$ est une base de $\vec{\mathcal{A}}$, une variété affine $\mathcal{V} = P_0 + \vec{\mathcal{V}}$. On se donne les coordonnées de P_0 dans $\mathcal{R} : (p_1, \dots, p_n)^\sim$.

On souhaite obtenir des équations qui caractérisent \mathcal{V} en termes de coordonnées de ses points. Il y a donc deux questions.

1. Obtenir une description constructive des coordonnées des points de \mathcal{V} , à l'aide de paramètres ([des équations paramétriques](#));
2. Obtenir des conditions (nécessaires et suffisantes) exprimant qu'un point P est dans \mathcal{V} ([des équations cartésiennes](#)), sous forme sur ses coordonnées.

Le résultat fondamental pour se ramener aux sous-espaces vectoriels a déjà été démontré. Vu son caractère fondamental, je le rappelle ici.

Proposition 2.7.1. *On a $P \in \mathcal{V} = P_0 + \vec{\mathcal{V}}$ si, et seulement si $\overrightarrow{P_0P} \in \vec{\mathcal{V}}$.*

On obtient comme conséquence directe la description paramétrique des variétés affines.

Proposition 2.7.2 (Equations paramétriques cartésiennes). *Soit \mathcal{A} un espace affine muni d'un repère (O, \mathcal{B}) , \mathcal{V} une variété affine et P_0 un point de \mathcal{V} . Si $\vec{\mathcal{V}} = \langle v_1, \dots, v_r \rangle$ et si les composantes de v_i ($i \leq r$) dans \mathcal{B} sont $(v_{1,i}, \dots, v_{n,i})^\sim$ alors $P : (x_1, \dots, x_n)^\sim$ est dans \mathcal{V} si, et seulement si,*

$$\exists \lambda_1, \dots, \lambda_r \in \mathbb{R} : \begin{cases} x_1 - p_1 &= \lambda_1 v_{1,1} + \dots + \lambda_r v_{1,r} \\ &\vdots \\ x_n - p_n &= \lambda_1 v_{n,1} + \dots + \lambda_r v_{n,r}, \end{cases} \quad (2.4)$$

où $P_0 : (p_1, \dots, p_n)^\sim$ dans (O, \mathcal{B}) .

Exemple 2.7.3. dans un espace affine de dimension 3 muni d'un repère, si $A : (1, 2, 3)^\sim$ et $B : (6, 3, 7)^\sim$ alors

$$X : (x, y, z)^\sim \in AB \Leftrightarrow \exists \lambda \in \mathbb{R} : \begin{cases} x - 1 &= 5\lambda \\ y - 2 &= \lambda \\ z - 3 &= 4\lambda. \end{cases}$$

On peut donc écrire des équations paramétriques de toute variété affine dans un repère, à condition d'en connaître un point et suffisamment de vecteurs directeurs. Passons maintenant au cas des équations cartésiennes. Il n'est pas inutile d'en rappeler une définition.

Définition 2.7.4. Des équations cartésiennes de la variété affine \mathcal{V} dans le repère \mathcal{R} sont des conditions nécessaires et suffisantes sur les coordonnées d'un point P dans le repère \mathcal{R} pour qu'il soit dans \mathcal{V} .

Il suffit, pour obtenir des équations cartésiennes dans un repère donné, d'éliminer les paramètres dans des équations paramétriques cartésiennes. La méthode est rigoureusement identique à celle employée pour les sous-espaces vectoriels. On obtient donc directement le résultat suivant.

Proposition 2.7.5. *Soit une variété affine \mathcal{V} telle que $\vec{\mathcal{V}} = \langle v_1, \dots, v_r \rangle$ et P_0 un point de \mathcal{V} de coordonnées $X_0 = (p_1, \dots, p_n)^\sim$. Si les composantes de v_i dans \mathcal{B} sont $V_i = (v_{1,i}, \dots, v_{n,i})^\sim$, alors un point P de coordonnées $X = (x_1, \dots, x_n)^\sim$ est dans \mathcal{V} si, et seulement si,*

$$\text{rg}(V_1, \dots, V_r) = \text{rg}(V_1, \dots, V_r | X - X_0). \quad (2.5)$$

Cette méthode générale s'applique que v_1, \dots, v_r soient indépendants ou non.

Comme dans le cas des sous-espaces vectoriels également, on peut compter les équations indépendantes fournies par (2.5) et on obtient un résultat similaire.

Proposition 2.7.6. *Soit \mathcal{A} un espace affine de dimension n et $\mathcal{R} = (O; \mathcal{B})$ un repère. Toute variété affine de dimension p admet dans \mathcal{R} des équations cartésiennes formant un système linéaire de rang $n - p$.*

De plus, en comparant avec les équations de $\vec{\mathcal{V}}$ obtenues dans le cadre des espaces vectoriels, on peut passer des équations de $\vec{\mathcal{V}}$ dans \mathcal{B} à celles de \mathcal{V} dans \mathcal{R} , et vice-versa :

Si dans \mathcal{B} , $\vec{\mathcal{V}}$ admet des équations cartésiennes

$$\begin{cases} a_{1,1}y_1 + \dots + a_{1,n}y_n &= 0 \\ &\vdots \\ a_{p,1}y_1 + \dots + a_{p,n}y_n &= 0 \end{cases},$$

(ou $AY = 0$ dans l'écriture matricielle naturelle), alors \mathcal{V} admet dans \mathcal{R} les équations

$$\begin{cases} a_{1,1}(x_1 - x_{1,0}) + \dots + a_{1,n}(x_n - x_{n,0}) &= 0 \\ &\vdots \\ a_{p,1}(x_1 - x_{1,0}) + \dots + a_{p,n}(x_n - x_{n,0}) &= 0 \end{cases},$$

ou encore $AX = AX_0$, où $X_0 = (x_{1,0}, \dots, x_{n,0})^\sim$ est le vecteur de coordonnées d'un point P de \mathcal{V} .

On peut évidemment parcourir le chemin en sens inverse : si dans $\mathcal{R} = (O, \mathcal{B})$, une variété affine \mathcal{V} admet les équations

$$\begin{cases} a_{1,1}x_1 + \dots + a_{1,n}x_n & = & b_1 \\ & \vdots & \\ a_{p,1}x_1 + \dots + a_{p,n}x_n & = & b_p \end{cases},$$

alors son sous-espace vectoriel directeur admet, dans \mathcal{B} , l'équation

$$\begin{cases} a_{1,1}y_1 + \dots + a_{1,n}y_n & = & 0 \\ & \vdots & \\ a_{p,1}y_1 + \dots + a_{p,n}y_n & = & 0 \end{cases}.$$

Remarque 2.7. En règle générale, on ne change pas le nom des inconnues. Je l'ai fait ici pour bien marquer que dans un cas, les nombres x_1, \dots, x_n représentent les coordonnées d'un point dans un repère, tandis que les nombres y_1, \dots, y_n représentent les composantes d'un vecteur dans une base (la base associée au repère).

Plus généralement, si on a des équations d'un certain type, on peut en déduire que l'ensemble des points dont les coordonnées sont solutions est une variété affine et en calculer la dimension. La démonstration est technique (et relativement simple). Je vous en fais grâce.

Proposition 2.7.7. *Dans un espace affine \mathcal{A} de dimension n , l'ensemble \mathcal{V} des points P de \mathcal{A} dont les coordonnées $X = (x_1, \dots, x_n)^\sim$ dans \mathcal{R} satisfont un système d'équations linéaires compatible $AX = B$ est une variété affine de dimension $n - \text{rg}(A)$, dont le sous-espace vectoriel directeur $\vec{\mathcal{V}}$ admet pour équations $AX = 0$ (dans \mathcal{B}).*

Voici maintenant deux exemples :

- Dans un espace de dimension 6 muni d'un repère, le système d'équations

$$\begin{cases} x_1 + 2x_2 - x_3 = 3 \\ 2x_1 - x_4 + x_5 = 1 \end{cases}$$

détermine une variété affine de dimension 4. Le sous espace vectoriel directeur $\vec{\mathcal{V}}$ admet pour équations

$$\begin{cases} x_1 + 2x_2 - x_3 = 0 \\ 2x_1 - x_4 + x_5 = 0 \end{cases}$$

- Dans un espace de dimension 3 muni d'un repère, le système d'équations

$$\begin{cases} x_1 + 2x_2 - x_3 = 3 \\ 2x_1 - x_2 + x_3 = 1 \end{cases}$$

détermine une droite.

Comme dans le cas des espaces vectoriels, nous nous attardons plus longuement sur le cas des droites et des hyperplans, puis nous passerons en dimension 2 et 3.

Proposition 2.7.8. *Dans un espace affine \mathcal{A} de dimension n muni d'un repère \mathcal{R} , une droite d déterminée par un point $P_0 : (p_1, \dots, p_n)^\sim$ et un vecteur directeur $u : (u_1, \dots, u_n)^\sim \in \vec{d} \setminus \{0\}$ admet les équations cartésiennes*

$$\frac{x_1 - p_1}{u_1} = \dots = \frac{x_n - p_n}{u_n},$$

si $u_1 \cdots u_n \neq 0$. Si $u_i = 0$ pour un $i \leq n$, alors

$$d \equiv \begin{cases} \frac{x_1 - p_1}{u_1} = \dots = \widehat{i} = \dots = \frac{x_n - p_n}{u_n} \\ x_i = 0 \end{cases}$$

Cela fonctionne quand plusieurs composantes de u sont nulles.

Si la droite est donnée par deux points, on se ramène au cas précédent, en notant que $AB = A + \overrightarrow{AB}$. Passons maintenant au cas des hyperplans.

Proposition 2.7.9. *Dans un espace affine \mathcal{A} de dimension n , muni d'un repère \mathcal{R} , un hyperplan π est déterminé par un point P_0 de coordonnées $X_0 = (p_1, \dots, p_n)^\sim$ et par $n - 1$ vecteurs v_1, \dots, v_{n-1} (indépendants). Si les composantes de v_i dans \mathcal{B} sont $V_i = (v_{1,i}, \dots, v_{n,i})^\sim$, alors on a*

$$\pi \equiv \det(V_1, \dots, V_{n-1} | X - X_0) = 0. \quad (2.6)$$

Tous les autres façons de déterminer un hyperplan se ramènent à la proposition précédente. Voici quelques exercices.

Exercice 2.7.10. Dans un espace affine de dimension 3 muni d'un repère, déterminer une équation cartésienne du plan

- π_1 contenant $A : (1, 2, 3)^\sim$ et de vecteur directeur $v_1 : (-1, 0, 2)^\sim$ et $v_2 : (1, 1, 1)^\sim$;
- π_2 contenant les points $A : (1, 2, 3)^\sim$, $B : (-1, 0, 2)^\sim$ et $C : (1, 1, 1)^\sim$;
- π_3 contenant $A : (0, 3, 2)$, $B : (-1, 1, 2)$ et de v. dir. $v : (1, 1, 1)$.

2.7.1 Le cas des droites en dimension 2

Dans un espace affine \mathcal{A} de dimension deux, une droite est également un hyperplan. On a donc deux façons équivalentes d'obtenir des équations cartésiennes. On se donne un repère $\mathcal{R} = (O, (b_1, b_2))$.

- La droite $d = A + \langle u \rangle$, où $A : (a_1, a_2)^\sim$ et $u : (u_1, u_2)^\sim$ admet pour équation cartésienne

$$d \equiv \det \begin{pmatrix} x_1 - a_1 & u_1 \\ x_2 - a_2 & u_2 \end{pmatrix} = 0 \quad \text{ou} \quad d \equiv \frac{x_1 - a_1}{u_1} = \frac{x_2 - a_2}{u_2}$$

avec la convention habituelle, si $u_1 u_2 = 0$.

- La droite AB , où $A : (a_1, a_2)^\sim$ et $B : (b_1, b_2)^\sim$: $\overrightarrow{AB} : (b_1 - a_1, b_2 - a_2)$ admet pour équation cartésienne

$$AB \equiv \det \begin{pmatrix} x_1 - a_1 & b_1 - a_1 \\ x_2 - a_2 & b_2 - a_2 \end{pmatrix} = 0 \quad \text{ou} \quad AB \equiv \frac{x_1 - a_1}{b_1 - a_1} = \frac{x_2 - a_2}{b_2 - a_2}$$

avec la convention habituelle.

- Réciproquement, l'équation générale d'une droite est $a_1 x_1 + a_2 x_2 + b = 0$ ($(a_1, a_2) \neq (0, 0)$). On peut alors déterminer les données géométriques associées à la droite :

1. On trouve les coordonnées de points et composantes de vecteurs directeurs en résolvant l'équation (la solution va forcément dépendre d'un paramètre).
2. Le sous-vectoriel directeur a pour équation cartésienne $a_1 y_1 + a_2 y_2 = 0$.
3. Un vecteur directeur est donné par $(-a_2, a_1)$.

Cela permet d'exprimer le parallélisme de droites dans le plan.

Proposition 2.7.11. *Les droites d et d' définies par leurs équations $d \equiv a_1 x_1 + a_2 x_2 = b$ et $d' \equiv a'_1 x_1 + a'_2 x_2 = b'$ sont*

- parallèles si

$$\text{rg} \begin{pmatrix} a_1 & a_2 \\ a'_1 & a'_2 \end{pmatrix} = 1 \quad \text{ou} \quad \det \begin{pmatrix} a_1 & a_2 \\ a'_1 & a'_2 \end{pmatrix} = 0$$

- sécantes si

$$\text{rg} \begin{pmatrix} a_1 & a_2 \\ a'_1 & a'_2 \end{pmatrix} = 2 \quad \text{ou} \quad \det \begin{pmatrix} a_1 & a_2 \\ a'_1 & a'_2 \end{pmatrix} \neq 0.$$

On peut également utiliser des résultats d'algèbre linéaire pour obtenir l'équation du faisceau de droites déterminé par deux droites sécantes d_1 et d_2 .

Proposition 2.7.12. Soient les droites $d \equiv a_1x_1 + a_2x_2 = b$ et $d' \equiv a'_1x_1 + a'_2x_2 = b'$ sécantes en P . Une droite d'' contient P si et seulement si elle admet pour équation

$$d'' \equiv \lambda(a_1x_1 + a_2x_2 - b) + \mu(a'_1x_1 + a'_2x_2 - b') = 0, \quad (\lambda, \mu) \in \mathbb{R}^2 \setminus \{(0, 0)\}.$$

Démonstration. Il suffit de remarquer que la condition que d'' contienne l'intersection de d et d' est équivalente au fait que la matrice

$$\begin{pmatrix} a_1 & a_2 & b \\ a'_1 & a'_2 & b' \\ a''_1 & a''_2 & b'' \end{pmatrix}$$

soit de rang deux. Cela équivaut encore, puisque les deux premières lignes sont indépendantes, au fait que la troisième ligne soit une combinaison linéaire des deux premières. Cela conduit directement au résultat. \square

2.7.2 Le cas des droites en dimension 3

On se place dans un espace affine de dimension trois muni d'un repère $\mathcal{R} = (O; (b_1, b_2, b_3))$.

- La droite $d = A + \langle u \rangle$, où $A : (a_1, a_2, a_3)^\sim$ et $u : (u_1, u_2, u_3)^\sim$ admet pour équation cartésienne

$$d \equiv \text{rg} \begin{pmatrix} x_1 - a_1 & u_1 \\ x_2 - a_2 & u_2 \\ x_3 - a_3 & u_3 \end{pmatrix} = 2 \quad \text{ou} \quad d \equiv \frac{x_1 - a_1}{u_1} = \frac{x_2 - a_2}{u_2} = \frac{x_3 - a_3}{u_3}$$

avec la convention habituelle, si $u_1u_2u_3 = 0$.

- La droite AB , où $A : (a_1, a_2, a_3)^\sim$ et $B : (b_1, b_2, b_3)^\sim$: on a $\overrightarrow{AB} : (b_1 - a_1, b_2 - a_2, b_3 - a_3)$ admet pour équation cartésienne

$$AB \equiv \frac{x_1 - a_1}{b_1 - a_1} = \frac{x_2 - a_2}{b_2 - a_2} = \frac{x_3 - a_3}{b_3 - a_3}$$

avec la convention habituelle, si un des dénominateurs au moins est nul.

- Réciproquement, l'équation générale d'une droite est

$$d \equiv \begin{cases} a_1x_1 + a_2x_2 + a_3x_3 = b \\ a'_1x_1 + a'_2x_2 + a'_3x_3 = b' \end{cases} \quad \text{où} \quad \text{rg} \begin{pmatrix} a_1 & a_2 & a_3 \\ a'_1 & a'_2 & a'_3 \end{pmatrix} = 2$$

Le sous-vectoriel directeur associé a pour équation (dans \mathcal{B}) :

$$\vec{d} \equiv \begin{cases} a_1y_1 + a_2y_2 + a_3y_3 = 0 \\ a'_1y_1 + a'_2y_2 + a'_3y_3 = 0 \end{cases}$$

On trouve un point de d en résolvant complètement l'équation de d , ou en coupant par un plan (par exemple $x_3 = 0$). On trouve un vecteur directeur en résolvant l'équation de \vec{d} . Une solution non triviale est toujours donnée par

$$u : \left(\det \begin{pmatrix} a_2 & a_3 \\ a'_2 & a'_3 \end{pmatrix}, -\det \begin{pmatrix} a_1 & a_3 \\ a'_1 & a'_3 \end{pmatrix}, \det \begin{pmatrix} a_1 & a_2 \\ a'_1 & a'_2 \end{pmatrix} \right).$$

Les positions relatives de deux droites sont réglées par la proposition suivante.

Proposition 2.7.13. Soient les droites

$$d \equiv \begin{cases} a_1x_1 + a_2x_2 + a_3x_3 = b \\ a'_1x_1 + a'_2x_2 + a'_3x_3 = b' \end{cases} \quad \text{et} \quad d' \equiv \begin{cases} c_1x_1 + c_2x_2 + c_3x_3 = e \\ c'_1x_1 + c'_2x_2 + c'_3x_3 = e' \end{cases}$$

Posons

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ a'_1 & a'_2 & a'_3 \\ c_1 & c_2 & c_3 \\ c'_1 & c'_2 & c'_3 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} b \\ b' \\ e \\ e' \end{pmatrix}$$

Les droites d et d' sont

- Parallèles confondues si, et seulement si, $\text{rg}(A) = \text{rg}(A|B) = 2$;
- Parallèles distinctes si, et seulement si, $\text{rg}(A) = 2$ et $\text{rg}(A|B) = 3$;
- Sécantes si, et seulement si, $\text{rg}(A) = \text{rg}(A|B) = 3$;
- Gauches si, et seulement si, $\text{rg}(A|B) = 4$, ou $\det(A|B) \neq 0$.

2.7.3 Le cas des plans en dimension 3

On place encore dans un espace affine de dimension trois muni d'un repère $(O; (b_1, b_2, b_3))$.

- Le plan $\pi = A + \langle u, v \rangle$, où $A : \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, $u : \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ et $v : \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ admet pour équation cartésienne

$$\pi \equiv \det \begin{pmatrix} x_1 - a_1 & u_1 & v_1 \\ x_2 - a_2 & u_2 & v_2 \\ x_3 - a_3 & u_3 & v_3 \end{pmatrix} = 0.$$

- Le plan $\pi = ABC$ admet \overrightarrow{AB} et \overrightarrow{AC} comme vecteurs directeurs et a donc pour équation cartésienne

$$ABC \equiv \det \begin{pmatrix} x_1 - a_1 & b_1 - a_1 & c_1 - a_1 \\ x_2 - a_2 & b_2 - a_2 & c_2 - a_2 \\ x_3 - a_3 & b_3 - a_3 & c_3 - a_3 \end{pmatrix} = 0$$

- Pour le plan π contenant A, B et dont un vecteur directeur u est donné (indépendant avec \overrightarrow{AB} , considérer A, \overrightarrow{AB} et u , et faire comme plus haut.
- Réciproquement, un plan π en dimension trois admet une équation générale

$$a_1x_1 + a_2x_2 + a_3x_3 = b, \quad (a_1, a_2, a_3) \neq (0, 0, 0). \quad (2.7)$$

On trouve facilement des points de π en donnant des valeurs à deux des trois coordonnées.

On peut aussi résoudre : si par exemple $a_1 \neq 0$, alors (2.7) s'écrit

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{b}{a_1} \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -\frac{a_2}{a_1} \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -\frac{a_3}{a_1} \\ 0 \\ 1 \end{pmatrix}$$

On trouve ainsi un point et deux vecteurs directeurs.

On a un résultat analogue à celui des droites en dimension deux.

Proposition 2.7.14. *Les plans $\pi \equiv a_1x_1 + a_2x_2 + a_3x_3 = b$ et $\pi' \equiv a'_1x_1 + a'_2x_2 + a'_3x_3 = b'$ sont*

- parallèles si et seulement si

$$\text{rg} \begin{pmatrix} a_1 & a_2 & a_3 \\ a'_1 & a'_2 & a'_3 \end{pmatrix} = 1.$$

- confondus si et seulement si

$$\text{rg} \begin{pmatrix} a_1 & a_2 & a_3 & b \\ a'_1 & a'_2 & a'_3 & b' \end{pmatrix} = 1.$$

- sécants dans le dernier cas possible, selon la droite

$$d \equiv \begin{cases} a_1x_1 + a_2x_2 + a_3x_3 = b \\ a'_1x_1 + a'_2x_2 + a'_3x_3 = b' \end{cases} \quad (2.8)$$

On a encore un résultat d'algèbre concernant les faisceaux de plans.

Proposition 2.7.15 (Faisceaux de plans). *Un plan π'' contient d admettant l'équation (2.8) si et seulement si il admet une équation du type*

$$\lambda(a_1x_1 + a_2x_2 + a_3x_3 - b) + \mu(a'_1x_1 + a'_2x_2 + a'_3x_3 - b') = 0, \quad (\lambda, \mu) \in \mathbb{R}^2 \setminus \{(0, 0)\}.$$

2.7.4 Deux problèmes classiques

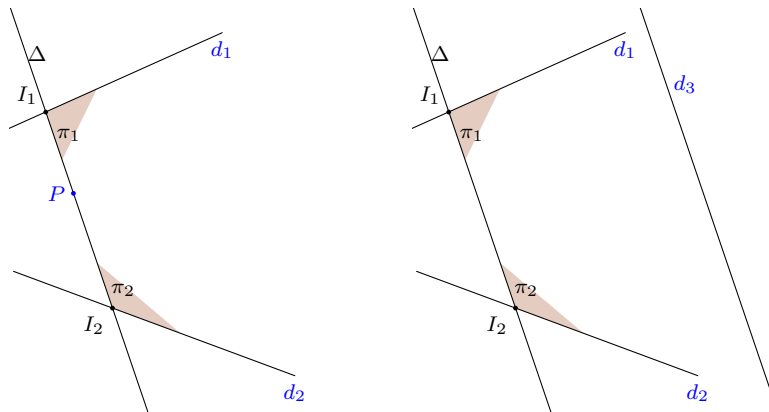
On se donne deux droites gauches d_1 et d_2 , un point P (extérieur à d_1 et d_2) et une droite d_3 . Les faisceaux de plans et leurs équations permettent de résoudre d'un point de vue analytique les deux constructions géométriques suivantes.

Proposition 2.7.16. *Il existe une unique droite Δ qui coupe d_1 et d_2 et qui contient P si et seulement si aucun des plans déterminés par P et l'une des droites n'est parallèle à l'autre droite. Cette droite est l'intersection des plans π_1 et π_2 déterminés par P et d_1 et P et d_2 , respectivement.*

De la même façon, on a le résultat suivant.

Proposition 2.7.17. *Il existe une unique droite Δ qui coupe d_1 et d_2 et qui est parallèle à d_3 si et seulement si d_1 , d_2 et d_3 ne sont pas parallèles à un même plan. Cette droite est l'intersection des plans π_1 contenant d_1 et parallèle à d_3 et π_2 contenant d_2 et parallèle à d_3 .*

Ces situations se représentent comme suit.



Chapitre 3

Espaces vectoriels euclidiens

Jusqu'à présent, vous avez sans doute rencontré *le produit scalaire*, présenté comme unique, dans le plan, l'espace, dans \mathbb{R}^2 ou \mathbb{R}^3 . Il est cependant nécessaire de généraliser cette notion, et de pouvoir définir plusieurs produits scalaires. Cette notion n'est pas utile qu'en mathématique. Citons par exemple le cas des *variétés riemanniennes* dont une généralisation est utilisée en relativité générale, et qui sont des espaces où un produit scalaire est associé à chaque point, ou les espaces de Hilbert, qui généralisent également les notions élémentaires que nous allons étudier et sont utiles pour formuler la mécanique quantique.

Dans l'enseignement secondaire, on [mesure](#) des angles entre demi-droites avec un rapporteur, dans le plan. On généralise cette mesure à l'espace, tant bien que mal, en remarquant que deux demi-droites sont toujours dans un même plan. On considère également que les mesures de distances sont des données, et cela permet de définir la norme de vecteurs. On en déduit les propriétés de perpendicularité, pour finalement arriver au produit scalaire défini par la célèbre formule

$$u \bullet v = |u| |v| \cos(\theta), \tag{3.1}$$

où θ est l'angle entre les vecteurs u et v et où $|u|$ et $|v|$ sont les normes de u et v .

Mais quel serait alors l'angle entre les deux matrices x et y définies par

$$x = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix} \quad \text{et} \quad y = \begin{pmatrix} -1 & -1 \\ 0 & \sqrt{2} \end{pmatrix}?$$

De même, quelle est la longueur de ces matrices ?

On constate que le produit scalaire défini par (3.1) encode toutes les informations concernant la norme des vecteurs (en considérant $\theta = 0$) et les angles, une fois qu'on a les normes.

La solution consiste donc à partir du produit scalaire, que l'on fixe comme on veut, pour autant qu'il ait de bonnes propriétés, puis prendre (3.1) non pas comme la définition du produit scalaire à partir des normes et des angles, mais comme une définition des normes et des angles [à partir du produit scalaire](#).

3.1 Les produits scalaires

Comme annoncé ci-dessus, on définit [un](#) produit scalaire par les propriétés qu'il doit avoir.

Définition 3.1.1. Un produit scalaire sur un espace vectoriel réel E est une application

$$\langle \cdot, \cdot \rangle : E \times E \rightarrow \mathbb{R} : (x, y) \mapsto \langle x, y \rangle$$

qui est

1. Symétrique : on a $\langle x, y \rangle = \langle y, x \rangle$ pour tous $x, y \in E$;
2. Bilinéaire : on a

$$\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle, \quad \langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle,$$

et

$$\langle \lambda x, y \rangle = \lambda \langle x, y \rangle = \langle x, \lambda y \rangle,$$

pour tous $x, y, z \in E$ et tout $\lambda \in \mathbb{R}$;

3. Définie positive : On a $\langle x, x \rangle \geq 0$ pour tout $x \in E$ et $\langle x, x \rangle = 0$ (si et **seulement** si $x = 0$). Un espace vectoriel E muni d'un produit scalaire est appelé espace vectoriel euclidien.

Comme premier exemple, nous généralisons à \mathbb{R}^n le produit scalaire de \mathbb{R}^2 et \mathbb{R}^3 introduit dans l'enseignement secondaire.

Exemple 3.1.2. Le produit scalaire préféré (encore appelé standard) de \mathbb{R}^n est défini par

$$\langle x, y \rangle = x_1 y_1 + \cdots + x_n y_n = \sum_{i=1}^n x_i y_i.$$

Vérifions que cette opération satisfait les conditions pour être un produit scalaire.

1. Elle est symétrique puisque $\langle y, x \rangle = \sum_{i=1}^n y_i x_i$;
2. Elle est bilinéaire puisqu'elle peut être exprimée sous la forme d'un produit matriciel :

$$\langle x, y \rangle = (x_1, \dots, x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix};$$

3. Elle est définie positive puisque $\langle x, x \rangle = \sum_{i=1}^n x_i^2$ ne peut être nul que si x_i est nul pour tout $i \leq n$.

Il y a cependant beaucoup d'autres exemples. En voici un assez simple.

Exemple 3.1.3. L'application suivante définit un produit scalaire sur \mathbb{R}^2 :

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = 7x_1 y_1 + 2x_2 y_2 + 2(x_1 y_2 + x_2 y_1).$$

En effet, elle est visiblement symétrique, elle est bilinéaire puisqu'elle peut être mise sous forme matricielle

$$\langle x, y \rangle = (x_1, x_2) \begin{pmatrix} 7 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

Enfin, elle est définie positive puisque

$$\langle x, x \rangle = 7x_1^2 + 2x_2^2 + 4x_1 x_2 = 2(x_1 + x_2)^2 + 5x_1^2$$

ne peut être nul que si $x_1 = x_2 = 0$.

Voici un exemple concernant les matrices.

Exemple 3.1.4. L'application définie sur \mathbb{R}_p^p par

$$\langle A, B \rangle = \text{tr}(B \sim A) = \sum_{i,j=1}^p A_{i,j} B_{i,j}, \quad \forall A, B \in \mathbb{R}_p^p$$

définit un produit scalaire sur \mathbb{R}_p^p .

Je vous laisse le soin de le vérifier. Ce produit scalaire est le produit scalaire standard des matrices (réelles). Vous pouvez également calculer le produit scalaire des matrices données dans l'introduction.

Remarque 3.1. 1. La notion de produit scalaire peut être étendue aux espaces vectoriels complexes, mais elle doit être légèrement modifiée pour conserver de bonnes propriétés.

2. L'opération définissant l'espace de Minkowski^a est donnée par

$$\left(\begin{array}{c} x_1 \\ x_2 \\ x_3 \\ t \end{array} \right), \left(\begin{array}{c} y_1 \\ y_2 \\ y_3 \\ t' \end{array} \right) \mapsto x_1 y_1 + x_2 y_2 + x_3 y_3 - t t'.$$

Elle est bilinéaire et symétrique. Ce n'est cependant pas un produit scalaire puisqu'elle n'est pas définie positive. On a en effet des vecteurs x tels que $\langle x, x \rangle = 0$ et même des vecteurs x tels que $\langle x, x \rangle < 0$. Cela donne lieu à trois types de vecteurs, mais ceci est une autre histoire.

3.2 Normes de vecteurs, inégalités célèbres et angles non orientés

Dans cette section, nous considérons un espace vectoriel euclidien E , et nous nous donnons un produit scalaire $\langle \cdot, \cdot \rangle$. Comme annoncé plus haut, ce produit scalaire donné nous permet de définir la longueur des vecteurs.

Définition 3.2.1. La norme du vecteur $x \in E$, encore appelée longueur ou module, est

$$|x| = \sqrt{\langle x, x \rangle}.$$

Elle est encore notée $\|x\|$.

Cette définition est licite puisque $\langle x, x \rangle \geq 0$. Elle satisfait aussi la première propriété demandée à une longueur, donnée dans la proposition suivante.

Proposition 3.2.2. *Pour tout $x \in E$ et tout $\lambda \in \mathbb{R}$, on a $|\lambda x| = |\lambda| |x|$. On a $|x| = 0$ si, et seulement si, $x = 0$.*

Démonstration. On applique la définition : on a

$$|\lambda x| = \sqrt{\langle \lambda x, \lambda x \rangle} = \sqrt{\lambda^2 \langle x, x \rangle} = |\lambda| \sqrt{\langle x, x \rangle} = |\lambda| |x|.$$

Par définition du produit scalaire, on a $|x| = 0$ si et seulement si $\langle x, x \rangle = 0$. Cette condition équivaut encore à $x = 0$. \square

Les deux inégalités qui suivent sont fondamentales. La première, celle de Cauchy-Schwarz, va nous permettre de définir les angles (non orientés), tandis que la deuxième, celle de Minkowski, nous permettra de définir les distances dans les espaces affines.

Proposition 3.2.3 (Cauchy-Schwarz). *Pour tous $x, y \in E$, on a*

$$|\langle x, y \rangle| \leq |x| |y|.$$

L'égalité entre ces deux quantités a lieu si, et seulement si, x et y sont linéairement dépendants.

Démonstration. Fixons $x, y \in E$ et traitons deux cas.

Si $y = 0$, par linéarité du produit scalaire sur le second facteur, on a $\langle x, y \rangle = 0$ et l'égalité est satisfaite. Dans ce cas x et y sont toujours linéairement dépendants et le résultat est prouvé.

Supposons $y \neq 0$ et considérons la fonction f qui à tout $\lambda \in \mathbb{R}$ associe $f(\lambda) = \langle x + \lambda y, x + \lambda y \rangle$. Il s'agit en fait d'une fonction du second degré :

$$f(\lambda) = \lambda^2 |y|^2 + 2\lambda \langle x, y \rangle + |x|^2.$$

a. H. Minkowski (1864–1909), mathématicien germano-lithuanien. L'espace-temps qui porte son nom est le cadre naturel de la relativité restreinte.

Cette fonction du second degré ne change pas de signe vu sa définition, donc on a

$$4\langle x, y \rangle^2 - 4|x|^2|y|^2 \leq 0,$$

qui donne

$$\langle x, y \rangle^2 \leq |x|^2|y|^2,$$

et par suite, l'inégalité annoncée.

Passons maintenant au cas d'égalité, toujours dans le cas $y \neq 0$. Si x est multiple de y , on vérifie directement qu'on a l'égalité. Si on a l'égalité, le réalisant du trinôme du second degré f est nul. Il existe donc $\lambda_0 \in \mathbb{R}$ tel que $f(\lambda_0) = 0$. On a alors $x + \lambda_0 y = 0$, donc x et y sont linéairement dépendants. \square

Nous pouvons maintenant prouver l'inégalité de Minkowski, encore appelée inégalité triangulaire.

Proposition 3.2.4 (Minkowski). *Pour tous $x, y \in E$, on a*

$$|x + y| \leq |x| + |y|.$$

L'égalité entre ces deux quantités a lieu si, et seulement si, x et y sont multiples l'un de l'autre par un facteur non négatif.

Démonstration. Fixons $x, y \in E$. On se débarrasse de la racine (qui apparaît dans la définition de la norme) en considérant d'abord $|x + y|^2$. On a alors

$$|x + y|^2 = \langle x + y, x + y \rangle \tag{3.2}$$

$$= |x|^2 + 2\langle x, y \rangle + |y|^2 \tag{3.3}$$

$$\leq |x|^2 + 2|\langle x, y \rangle| + |y|^2 \tag{3.4}$$

$$\leq |x|^2 + 2|x||y| + |y|^2 \tag{3.5}$$

$$= (|x| + |y|)^2, \tag{3.6}$$

où on a utilisé la relation $\langle x, y \rangle \leq |\langle x, y \rangle|$ pour passer de (3.3) à (3.4) et l'inégalité de Cauchy-Schwarz pour passer de (3.4) à (3.5). Puisque $|x + y|$ et $|x| + |y|$ sont des nombres positifs, on déduit l'inégalité annoncée.

On a l'égalité si, et seulement si, on a

$$|\langle x, y \rangle| = |x||y| \quad \text{et} \quad \langle x, y \rangle = |\langle x, y \rangle|.$$

La première condition est le cas d'égalité de l'inégalité de Cauchy-Schwarz, ce qui implique que x et y sont multiples l'un de l'autre. Si x est nul, alors il est toujours égal à $0y$. S'il n'est pas nul, alors $y = \lambda x$ pour un λ réel et la deuxième condition donne $|\lambda| = \lambda$. \square

Terminons cette section par l'introduction des angles. L'inégalité de Cauchy-Schwarz implique que pour tous vecteurs x, y non nuls, on a

$$-1 \leq \frac{\langle x, y \rangle}{|x||y|} \leq 1.$$

Il existe donc un seul angle $\alpha \in [0, \pi]$ tel que

$$\cos(\alpha) = \frac{\langle x, y \rangle}{|x||y|}$$

Définition 3.2.5. Pour tous vecteurs x, y non nuls, l'angle non orienté des vecteurs x et y est l'unique nombre $\alpha \in [0, \pi]$ satisfaisant

$$\cos(\alpha) = \frac{\langle x, y \rangle}{|x||y|}.$$

Comme annoncé dans l'introduction, on retrouve la formule (3.1). On peut donc récupérer ce qui a été développé dans l'enseignement secondaire à partir de cette formule, mais on s'est maintenant libéré des mesures à la latte et au rapporteur^b. La définition 3.2.5 est cependant valide dans des espaces de dimension quelconque.

Exemple 3.2.6. 1. Dans \mathbb{R}^4 muni de son produit scalaire préféré, les vecteurs $x = (1, 0, 0, 1)^\sim$ et $y = (1, 0, 1, 0)^\sim$ satisfont

$$|x| = \sqrt{2} = |y|, \quad \text{et} \quad \langle x, y \rangle = 1.$$

L'angle des vecteurs x et y est donc l'unique nombre dans $[0, \pi]$ dont le cosinus vaut $\frac{1}{2}$, ce qui donne un angle égal à $\pi/3$.

2. Dans \mathbb{R}^2 muni de son produit scalaire préféré, les vecteurs $x = (2, 2)$ et $y = (0, 1)$ satisfont les relations

$$|x| = \sqrt{8} = 2\sqrt{2}, \quad |y| = 1, \quad \text{et} \quad \langle x, y \rangle = 2.$$

L'angle des vecteurs x et y est donc l'unique $\alpha \in [0, \pi]$ tels que $\cos(\alpha) = \frac{\sqrt{2}}{2}$. On a donc $\alpha = \frac{\pi}{4}$, conformément à l'intuition.

3. Dans l'espace vectoriel \mathbb{R}^2 muni du produit scalaire défini à l'exemple 3.1.3, les vecteurs $x = (1, 0)^\sim$ et $y = (0, 1)$ satisfont

$$|x| = \sqrt{7}, \quad |y| = \sqrt{2}, \quad \text{et} \quad \langle x, y \rangle = 2.$$

Leur angle non orienté $\alpha \in [0, \pi]$ satisfait donc

$$\cos(\alpha) = \frac{2}{\sqrt{14}}.$$

Pour ce même produit scalaire, les vecteurs $x = (0, 1)^\sim$ et $y = (1, -1)^\sim$ satisfont $\langle x, y \rangle = 0$, et ont donc un angle non orienté égal à $\frac{\pi}{2}$.

La notion d'angle conduit naturellement à celle d'orthogonalité. J'introduis aussi une notion liée à la norme.

Définition 3.2.7. Des vecteurs u et v sont orthogonaux (on dit aussi perpendiculaires) si $\langle u, v \rangle = 0$. Un vecteur u est normé (ou unitaire) si sa norme est égale à 1.

Remarquez que la définition d'orthogonalité prend aussi en compte le vecteur nul : il est orthogonal à tous les vecteurs. Voici quelques exemples.

Exemple 3.2.8. 1. Dans l'espace \mathbb{R}^n muni de son produit scalaire préféré, les vecteurs de la base canonique sont deux à deux orthogonaux ;
 2. Dans l'espace vectoriel \mathbb{R}^3 , muni du produit scalaire préféré, les vecteurs $u = (1, 2, 3)^\sim$ et $v = (3, 0, -1)^\sim$ sont orthogonaux ;
 3. Nous venons de voir que dans l'espace euclidien de l'exemple 3.1.3, les vecteurs $u = (0, 1)^\sim$ et $v = (1, -1)^\sim$ sont orthogonaux.

3.3 Bases orthonormées

Les vecteurs orthogonaux ont des propriétés très intéressantes, comme le démontre la proposition suivante.

Proposition 3.3.1. Soit $\{u_1, \dots, u_p\}$ des vecteurs non nuls et orthogonaux deux à deux, i.e. tels que $\langle u_i, u_j \rangle = \delta_{i,j} |u_i|^2$ pour $i, j \leq p$. Alors les propriétés suivantes sont valides.

b. A condition de définir le cosinus indépendamment des ces mesures, ce qui peut être fait au moyen des séries, voir la fin du cours de mathématiques générales.

1. Si $u \in E$ est tel que $u = \sum_{i=1}^p \lambda_i u_i$, alors on a $\lambda_i = \frac{\langle u, u_i \rangle}{|u_i|^2}$ pour $i \leq p$;
2. Les vecteurs u_1, \dots, u_p sont linéairement indépendants.

Démonstration. Pour la première assertion, on multiplie scalairement l'égalité $u = \sum_{i=1}^p \lambda_i u_i$ par le vecteur u_j pour $j \leq p$. On obtient ainsi

$$\langle u, u_j \rangle = \left\langle \sum_{i=1}^p \lambda_i u_i, u_j \right\rangle = \sum_{i=1}^p \lambda_i \langle u_i, u_j \rangle = \sum_{i=1}^p \lambda_i \delta_{i,j} |u_j|^2 = \lambda_j |u_j|^2,$$

ce qui donne le résultat annoncé, puisqu'on peut diviser par $|u_j|^2$.

Pour la seconde assertion, si $\sum_{i=1}^p \lambda_i u_i = 0$, par le premier point, on a $\lambda_i = \frac{\langle 0, u_i \rangle}{|u_i|^2} = 0$, ce qu'il fallait démontrer. \square

Définition 3.3.2. Soit E un espace vectoriel euclidien de dimension n . Une base orthonormée de E est une base $\mathcal{B} = (u_1, \dots, u_n)$ dont les vecteurs sont orthogonaux deux à deux et normés ($\langle u_i, u_j \rangle = \delta_{i,j}$ pour $i, j \leq n$).

Exemple 3.3.3. Dans \mathbb{R}^n muni de son produit scalaire préféré, la base canonique (e_1, \dots, e_n) est orthonormée. Dans \mathbb{R}^2 , muni du même produit scalaire les vecteurs $u = \frac{\sqrt{2}}{2}(1, 1)^\sim$ et $v = \frac{\sqrt{2}}{2}(1, -1)^\sim$ définissent également une base orthonormée.

Remarque 3.2. Dans un espace de dimension n , pour démontrer que des vecteurs u_1, \dots, u_n forment une base orthonormée, il suffit de démontrer qu'ils sont orthogonaux et normés. La proposition 3.3.1 permet alors de démontrer qu'ils sont linéairement indépendants. Etant en nombre égal à la dimension, ils forment alors une base.

La proposition 3.3.1 admet le corollaire suivant.

Corollaire 3.3.4. Si (u_1, \dots, u_n) forment une base orthonormée, alors la décomposition d'un vecteur u dans cette base est donnée par

$$u = \sum_{i=1}^n \langle u, u_i \rangle u_i.$$

Une propriété importante des bases orthonormées est la suivante. Elle justifie la nécessité dans nos développements ultérieurs, d'utiliser une base orthonormée pour obtenir des formules simples.

Proposition 3.3.5. Une base $\mathcal{B} = (u_1, \dots, u_n)$ est une base orthonormée si, et seulement si, le produit scalaire des vecteurs $x = \sum_{i=1}^n x_i u_i$ et $y = \sum_{i=1}^n y_i u_i$ est donné par

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i. \tag{3.7}$$

Démonstration. Si \mathcal{B} est orthonormée, alors

$$\langle x, y \rangle = \left\langle \sum_{i=1}^n x_i u_i, \sum_{j=1}^n y_j u_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \langle u_i, u_j \rangle = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \delta_{i,j} = \sum_{i=1}^n x_i y_i.$$

Inversement, puisque le vecteur de composantes de u_j dans \mathcal{B} est e_j pour tout $j \leq n$, si le produit scalaire est donné par (3.7), on constate que les vecteurs u_1, \dots, u_n sont deux à deux orthogonaux et normés. \square

On peut formuler des formes faibles de la proposition précédente comme suit :

1. Le produit scalaire de deux vecteurs est la somme des produits des composantes des vecteurs, [prises dans une base orthonormée](#).

2. Le passage aux composantes dans une base orthonormée identifie tout espace vectoriel euclidien à l'espace vectoriel \mathbb{R}^n , muni de son produit scalaire préféré.

Il est donc particulièrement important d'avoir des bases orthonormées. Le résultat suivant montre comment transformer une base quelconque en une base orthonormée, et démontre au passage l'existence de telles bases, en général.

Proposition 3.3.6 (Processus de Gram-Schmidt). *Tout espace vectoriel euclidien E de dimension finie admet une base orthonormée. De plus si u_1, \dots, u_n est une base de E , il existe une base orthonormée (v_1, \dots, v_n) telle que $\langle v_1, \dots, v_k \rangle = \langle u_1, \dots, u_k \rangle$ pour tout $k = 1, \dots, n$.*

Démonstration. On procède par récurrence sur n . Pour $n = 1$, on a une base (u_1) et $v_1 = \frac{u_1}{|u_1|}$ définit à lui seul une base orthonormée. Supposons le résultat vrai pour les dimensions inférieures ou égales à $n - 1$ et prouvons le pour un espace de dimension n . Soit u_1, \dots, u_n une base de E . Alors $\langle u_1, \dots, u_{n-1} \rangle$ est un espace vectoriel euclidien pour le même produit scalaire.

Par hypothèse d'induction, il existe des vecteurs deux à deux orthogonaux et normés dans cet espace, soit v_1, \dots, v_{n-1} satisfaisant $\langle v_1, \dots, v_k \rangle = \langle u_1, \dots, u_k \rangle$ pour tout $k = 1, \dots, n - 1$. Il nous reste à construire v_n . Pour cela on considère

$$v'_n = u_n - \sum_{i=1}^{n-1} \langle u_n, v_i \rangle v_i.$$

On montre facilement que v'_n est orthogonal à v_1, \dots, v_{n-1} :

$$\langle v_n, v_j \rangle = \langle u_n - \sum_{i=1}^{n-1} \langle u_n, v_i \rangle v_i, v_j \rangle = \langle u_n, v_j \rangle - \sum_{i=1}^{n-1} \langle u_n, v_i \rangle \underbrace{\langle v_i, v_j \rangle}_{=\delta_{i,j}} = 0, \quad j \leq n - 1,$$

Le vecteur v'_n n'est pas nul : si tel était le cas, u_n serait dans $\langle v_1, \dots, v_{n-1} \rangle = \langle u_1, \dots, u_{n-1} \rangle$, et c'est contraire à l'hypothèse. On peut donc poser $v_n = \frac{v'_n}{|v'_n|}$ et les vecteurs v_1, \dots, v_n définissent une base orthonormée de E , et on a donc aussi $\langle v_1, \dots, v_n \rangle = \langle u_1, \dots, u_n \rangle$. \square

Il est important de remarquer que nous avons établi un théorème d'existence d'une base orthonormée, mais aussi un moyen construire une telle base, à partir d'une base donnée. En effet, la récurrence utilisée plus haut fournit la construction : étant donnée une base u_1, \dots, u_n , on construit successivement

$$\begin{array}{ll} v'_1 = u_1 & v_1 = \frac{v'_1}{|v'_1|} \\ v'_2 = u_2 - \langle u_2, v_1 \rangle v_1 & v_2 = \frac{v'_2}{|v'_2|} \\ \vdots & \vdots \\ v'_k = u_k - \sum_{i=1}^{k-1} \langle u_k, v_i \rangle v_i & v_k = \frac{v'_k}{|v'_k|} \end{array}$$

et pour $k = n$, le processus s'arrête.

Exemple 3.3.7. On considère l'espace vectoriel \mathbb{R}^2 muni de son produit scalaire préféré. Soit $u_1 = (1, 1)^\sim$ et $u_2 = (0, 1)^\sim$. On applique le procédé ci-dessus :

$$\begin{array}{ll} v'_1 = u_1 = (1, 1)^\sim & v_1 = \frac{\sqrt{2}}{2}(1, 1)^\sim \\ v'_2 = u_2 - \langle u_2, v_1 \rangle v_1 = (0, 1)^\sim - \frac{\sqrt{2}}{2}v_1 = \frac{1}{2}(-1, 1)^\sim & v_2 = \frac{\sqrt{2}}{2}(-1, 1)^\sim \end{array}$$

Il sera utile pour la suite de connaître les particularités du changement de base dans le cas des bases orthonormées. Je rappelle d'abord une définition de calcul matriciel.

Définition 3.3.8. Une matrice carrée $A \in \mathbb{R}_n^n$ est orthogonale si $A^\sim = A^{-1}$.

Donc une matrice A est orthogonale si $A^{\sim} A = A A^{\sim} = \text{Id}$. Cela s'exprime aussi par la relation

$$\sum_{k=1}^n A_{k,i} A_{k,j} = \delta_{i,j}, \quad \forall i, j \leq n,$$

où δ est le symbole de Kronecker. Une condition équivalente est que les colonnes, ou les lignes, de A soient orthonormées dans \mathbb{R}^n , muni de son produit scalaire standard.

Proposition 3.3.9. *Soit $\mathcal{B} = (b_1, \dots, b_n)$ une base orthonormée d'un espace vectoriel euclidien E . Soit $\mathcal{B}' = (b'_1, \dots, b'_n)$ une autre base de E . Alors \mathcal{B}' est orthonormée si, et seulement si, la matrice de changement de base entre \mathcal{B} et \mathcal{B}' est orthogonale.*

Démonstration. Les colonnes de la matrice de changement de base sont formées des composantes de la nouvelle base dans l'ancienne. Si la matrice de changement de base est A , on a donc

$$b'_i = \sum_{k=1}^n A_{k,i} b_k.$$

On en déduit le produit scalaire suivant, puisque \mathcal{B} est orthonormée :

$$\langle b'_i, b'_j \rangle = \sum_{k=1}^n A_{k,i} A_{k,j},$$

et le résultat est alors clair. □

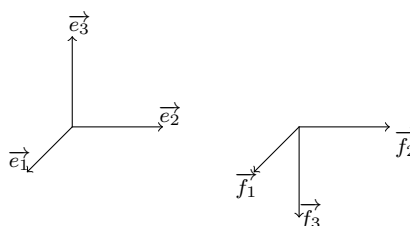
3.4 Produit mixte et produit vectoriel, angle orienté

Revenons quelques temps en dimension trois pour redéfinir deux objets géométriques bien connus. Le premier correspondra (au signe près) au volume dans les espaces affines euclidiens (voir le chapitre suivant), tandis que le deuxième est le produit vectoriel. Ici encore, je donne une définition générale. Le lecteur qui souhaite aller plus loin pourra se demander comment définir le produit mixte de quatre vecteurs en dimension quatre, ou le produit vectoriel de 3 vecteurs en dimension quatre. Avant cela, il est nécessaire d'introduire un concept supplémentaire, celui d'orientation, qui vaut pour tout espace vectoriel de dimension finie, euclidien ou non.

3.4.1 Orientation d'un espace vectoriel, changements de bases orthonormées

Si E est un espace vectoriel réel de dimension n , on peut y définir une relation entre les bases. Cette relation matérialise d'un point de vue mathématique la différence entre la main droite et la main gauche. Cette différence est importante dans l'expression des lois physiques que vous connaissez bien (notamment en électromagnétisme).

L'idée intuitive qui permet de définir l'orientation est qu'il n'est pas possible de "déformer continûment" une base représentée par la main droite en une base représentée par la main gauche. Bien sûr, si on est un peu contorsionniste, on peut le faire en changeant le "sens de son pouce". Mais ce retournement implique qu'on ne reste pas à tout instant dans l'ensemble des bases. Le pouce passe nécessairement à travers le plan déterminé par les autres doigts. A cet instant, on n'a plus une base. Voici une représentation de ce phénomène, avec une base gauche et une base droite.



J'ai dessiné des bases où les vecteurs semblent orthogonaux deux à deux. Mais vous pouvez décider qu'ils le sont, ou qu'ils ne le sont pas en changeant de produit scalaire : la notion d'orientation est indépendante de la notion de produit scalaire.

Définition 3.4.1. Deux bases $\mathcal{B} = (b_1, \dots, b_n)$ et $\mathcal{B}' = (b'_1, \dots, b'_n)$ ont la même orientation si la matrice de changement de base de \mathcal{B} à \mathcal{B}' a un déterminant positif.

Remarquons que la relation “avoir la même orientation” est symétrique : les matrices de changement de base de \mathcal{B} à \mathcal{B}' et de \mathcal{B}' à \mathcal{B} sont inverses l'une de l'autre. Leur déterminants aussi, et ils ont donc le même signe. Elle est réflexive : \mathcal{B} a la même orientation que \mathcal{B} , puisque l'identité a un déterminant positif. Enfin, elle est transitive : si \mathcal{B} a la même orientation que \mathcal{B}' et si \mathcal{B}' a la même orientation que \mathcal{B}'' , alors \mathcal{B} a la même orientation que \mathcal{B}'' . Les bases peuvent donc être réparties en ensembles de bases ayant la même orientation. La proposition suivante montre qu'il n'y a que deux tels “ensembles”.

Proposition 3.4.2. Soit E un espace vectoriel muni d'une base $\mathcal{B}_0 = (b_1, \dots, b_n)$. Alors $\mathcal{B}_1 = (b_1, \dots, b_{n-1}, -b_n)$ n'a pas la même orientation que \mathcal{B}_0 . Toute base \mathcal{B} a soit la même orientation que \mathcal{B}_0 soit la même orientation que \mathcal{B}_1 .

Démonstration. La matrice de changement de base de \mathcal{B}_0 et \mathcal{B}_1 est donnée par

$$A_0 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & 1 & 0 \\ 0 & \dots & 0 & -1 \end{pmatrix}.$$

Elle a visiblement un déterminant négatif. Si le changement de base de \mathcal{B} à \mathcal{B}_0 correspond à la matrice A , alors le changement de base de \mathcal{B} à \mathcal{B}_1 est obtenu en considérant le produit de A et A_0 , dont le déterminant est $-\det(A)$. \square

D'après ce que nous venons de voir, il n'y a que deux types de bases, en ce qui concerne l'orientation, c'est ce qui donne lieu au concept de base positive. Comme en ce qui concerne le produit scalaire, en mathématiques on a le choix pour privilégier un de ces deux types.

Définition 3.4.3. Une orientation d'un espace vectoriel est le choix d'une des deux classes de bases, qui sont déclarées positives. Les autres bases sont déclarées négatives.

3.4.2 Produit mixte

On se place dans un espace vectoriel euclidien orienté de dimension 3.

Définition 3.4.4. Soit $\mathcal{B} = (b_1, b_2, b_3)$ une base orthonormée positive de E . Le produit mixte des vecteurs u, v, w est défini par

$$[u, v, w] = \det \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix},$$

où $(u_1, u_2, u_3)^\sim, (v_1, v_2, v_3)^\sim$ et $(w_1, w_2, w_3)^\sim$ sont respectivement les composantes de u, v et w dans la base \mathcal{B} .

Comme d'habitude, quand nous définissons un objet géométrique à partir de composantes dans une base, nous devons montrer qu'il est intrinsèque, c'est-à-dire qu'il ne dépend pas de la base choisie. C'est l'objet du résultat suivant.

Proposition 3.4.5. Le produit mixte défini plus haut est indépendant du choix de la base orthonormée positive \mathcal{B} .

Démonstration. Supposons que nous avons à notre disposition deux bases orthonormées positives \mathcal{B} et \mathcal{B}' . Notons encore $(u_1, u_2, u_3)^\sim$, $(v_1, v_2, v_3)^\sim$ et $(w_1, w_2, w_3)^\sim$ les composantes de u, v et w dans la base \mathcal{B} , et $(u'_1, u'_2, u'_3)^\sim$, $(v'_1, v'_2, v'_3)^\sim$ et $(w'_1, w'_2, w'_3)^\sim$ les composantes de ces mêmes vecteurs dans la base \mathcal{B}' . Il existe alors une matrice orthogonale A telle que

$$\begin{pmatrix} u'_1 \\ u'_2 \\ u'_3 \end{pmatrix} = A \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \quad \begin{pmatrix} v'_1 \\ v'_2 \\ v'_3 \end{pmatrix} = A \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \quad \text{et} \quad \begin{pmatrix} w'_1 \\ w'_2 \\ w'_3 \end{pmatrix} = A \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$$

On peut alors calculer le produit mixte dans la base \mathcal{B}' :

$$\det \begin{pmatrix} u'_1 & v'_1 & w'_1 \\ u'_2 & v'_2 & w'_2 \\ u'_3 & v'_3 & w'_3 \end{pmatrix} = \det \left(A \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, A \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, A \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \right) = \det \left(A \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix} \right)$$

En utilisant les propriétés du déterminant, on obtient que cette quantité vaut

$$\det(A) \det \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix} = \det \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix},$$

puisque A est orthogonale et de déterminant positif. \square

Les propriétés du produit mixte se déduisent de celles du déterminant. Elles sont rassemblées dans la proposition suivante.

Proposition 3.4.6. *Le produit mixte a les propriétés suivantes.*

1. Il est trilinéaire, c'est-à-dire linéaire en chaque argument, les autres étant fixés ;
2. Il est alterné : il change de signe si on permute deux de ses arguments ;
3. On a $[u, v, w] = 0$ si, et seulement si, u, v et w sont linéairement dépendants ;
4. On a $[u, v, w] > 0$ (resp. < 0) si, et seulement si, (u, v, w) est une base positive.

Démonstration. Seul le dernier point n'est pas une propriété directe du déterminant et du passage aux composantes. Pour le démontrer, on considère une base \mathcal{B} orthonormée positive. Par le troisième point, si $[u, v, w] \neq 0$, alors (u, v, w) forme une base \mathcal{U} de E . Alors la matrice

$$\begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}$$

dont les colonnes sont les composantes de u, v et w dans la base \mathcal{B} n'est rien d'autre que la matrice de changement de base de \mathcal{U} à \mathcal{B} . Cette matrice a un déterminant positif (resp. négatif) si, et seulement si, \mathcal{U} et \mathcal{B} ont la même orientation (resp. une orientation opposée). \square

3.4.3 Le produit vectoriel

Nous pouvons maintenant définir le produit vectoriel. Pour ce faire, nous allons d'abord développer le produit mixte des vecteurs u, v et w en fonction de leurs composantes dans une base orthonormée positive, en utilisant la règle des cofacteurs sur la dernière colonne. On a

$$[u, v, w] = w_1 \det \begin{pmatrix} u_2 & v_2 \\ u_3 & v_3 \end{pmatrix} - w_2 \det \begin{pmatrix} u_1 & v_1 \\ u_3 & v_3 \end{pmatrix} + w_3 \det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} \quad (3.8)$$

Puisque nous avons considéré les composantes dans une base orthonormée, nous pouvons interpréter cette expression comme un produit scalaire du vecteur w avec un vecteur construit à l'aide de u et v . Plus précisément, on a le résultat suivant.

Proposition 3.4.7. *Pour tous $u, v \in E$, il existe un unique vecteur $a(u, v)$ satisfaisant la relation*

$$[u, v, w] = \langle a(u, v), w \rangle \quad \forall w \in E. \quad (3.9)$$

Dans toute base \mathcal{B} orthonormée positive, dans laquelle u et v ont pour composantes $(u_1, u_2, u_3)^\sim$ et $v : (v_1, v_2, v_3)^\sim$, le vecteur $a(u, v)$ admet pour composantes

$$\left(\det \begin{pmatrix} u_2 & v_2 \\ u_3 & v_3 \end{pmatrix}, -\det \begin{pmatrix} u_1 & v_1 \\ u_3 & v_3 \end{pmatrix}, \det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} \right)^\sim. \quad (3.10)$$

Démonstration. Pour l'existence, on fixe une base orthonormée positive \mathcal{B}_0 et on utilise (3.8) pour montrer que le vecteur $a(u, v)$ dont les composantes dans \mathcal{B}_0 sont données par (3.10) a la propriété demandée. Pour l'unicité, on note que si $a(u, v)$ et $a'(u, v)$ répondent à la question, alors on a

$$\langle a(u, v), w \rangle = [u, v, w] = \langle a'(u, v), w \rangle$$

pour tout vecteur w , qui donne directement

$$\langle a(u, v) - a'(u, v), w \rangle = 0, \quad \forall w \in E$$

et par suite $a(u, v) - a'(u, v) = 0$.

Dans toute base \mathcal{B} orthonormée positive le vecteur dont les composantes sont données dans \mathcal{B} en fonction de celles de u et v par (3.10) satisfait la condition (3.9). Vu l'unicité, ce vecteur est $a(u, v)$. \square

Définition 3.4.8. Le produit vectoriel des vecteurs u et v est le vecteur défini par la proposition précédente. Il est noté $u \wedge v$.

La proposition suivante rassemble les propriétés de l'opération produit vectoriel \wedge .

Proposition 3.4.9. *L'opération produit vectoriel a les propriétés suivantes :*

1. *Elle est antisymétrique ou alternée : on a $u \wedge v = v \wedge u$ pour tout $u, v \in E$.*
2. *Elle est bilinéaire, c'est-à-dire linéaire sur chacun de ses arguments, l'autre étant fixé ;*

Démonstration. On peut vérifier directement à partir de la formule (3.10) que ces propriétés sont satisfaites, puisqu'on travaille avec des déterminants. On peut aussi le montrer à l'aide de la caractérisation géométrique (Proposition 3.4.7). Montrons que $u \wedge v = -(v \wedge u)$ pour tous $u, v \in E$: pour tout $w \in E$, on a

$$\langle -(v \wedge u), w \rangle = -\langle v \wedge u, w \rangle = -[v, u, w] = [u, v, w].$$

Puisque $u \wedge v$ est le seul vecteur qui satisfait cette identité, on doit avoir $-(v \wedge u) = u \wedge v$. Montrons la linéarité sur le premier argument, on l'obtient sur le second par antisymétrie. Premièrement, on a $(u + v) \wedge w = u \wedge w + v \wedge w$. En effet, pour tout $z \in E$, on a

$$\langle u \wedge w + v \wedge w, z \rangle = \langle u \wedge w, z \rangle + \langle v \wedge w, z \rangle = [u, w, z] + [v, w, z] = [u + v, w, z]$$

Puisque $(u + v) \wedge w$ est le seul vecteur qui satisfait cette identité (pour tout z), on a le résultat annoncé. On fait de même pour démontrer l'identité $(\lambda u) \wedge v = \lambda(u \wedge v)$ pour tout $\lambda \in \mathbb{R}$. \square

Voici maintenant une caractérisation géométrique plus concrète du vecteur $u \wedge v$, qui fait le lien avec la définition admise habituellement.

Proposition 3.4.10. *Soient u, v des vecteurs de E . Alors*

1. *Le produit vectoriel $u \wedge v$ est orthogonal à u et à v ;*
2. *Le produit vectoriel $u \wedge v$ est nul si, et seulement si, u et v sont linéairement dépendants.*
3. *Si u et v sont non nuls, on a $|u \wedge v| = |u||v| \sin \alpha$ où α est l'angle non orienté entre u et v ;*

4. Si u et v sont linéairement indépendants, alors $(u, v, u \wedge v)$ est une base positive.

Démonstration. Montrons que $u \wedge v$ est orthogonal à u . On a

$$\langle u \wedge v, u \rangle = [u, v, u] = 0,$$

puisque u, v et u sont linéairement dépendants. On fait de même pour v .

Pour le deuxième point, d'abord si u et v sont linéairement dépendants, alors on a

$$\langle u \wedge v, w \rangle = [u, v, w] = 0$$

pour tout $w \in E$. On a alors $u \wedge v = 0$.

Réciproquement, supposons que $u \wedge v = 0$, que u et v soient linéairement indépendants et montrons une contradiction. Il existe alors $w \in E$ tel que (u, v, w) soit une base, donc tels que $[u, v, w] \neq 0$. Mais on a $[u, v, w] = \langle u \wedge v, w \rangle = 0$, ce qui donne une contradiction.

La troisième assertion est équivalente à $|u \wedge v|^2 = |u|^2|v|^2 \sin^2 \alpha$ ou encore à $|u \wedge v|^2 = |u|^2|v|^2(1 - \cos^2(\alpha))$. Cela s'écrit encore

$$|u \wedge v|^2 = |u|^2|v|^2 - \langle u, v \rangle^2.$$

Dans une base orthonormée arbitraire, le membre de gauche vaut

$$(u_2v_3 - u_3v_2)^2 + (u_1v_3 - u_3v_1)^2 + (u_1v_2 - u_2v_1)^2$$

tandis que le membre de droite vaut

$$(u_1^2 + u_2^2 + u_3^2)(v_1^2 + v_2^2 + v_3^2) - (u_1v_1 + u_2v_2 + u_3v_3)^2.$$

En développant l'une ou l'autre de ces expressions, on constate qu'elles coïncident.

Enfin, pour la dernière assertion, on calcule $[u, v, u \wedge v] = |u \wedge v|^2 > 0$. □

Puisque le produit mixte et le produit vectoriel sont multilinéaires (trilinéaire ou bilinéaire), il est utile de les exprimer en termes des composantes de vecteurs, dans une base orthonormée positive. Cela se fait naturellement à l'aide du symbole de Levi-Civita^c.

Définition 3.4.11. Le symbole^d de Levi-Civita est un symbole à trois indices défini par

$$\varepsilon_{i,j,k} = \begin{cases} 1 & \text{si } (i, j, k) = (1, 2, 3), (2, 3, 1), (3, 1, 2) \\ -1 & \text{si } (i, j, k) = (2, 1, 3), (1, 3, 2), (3, 2, 1) \\ 0 & \text{sinon.} \end{cases}$$

On peut résumer cette définition en disant que $\varepsilon_{i,j,k}$ vaut 1 si (i, j, k) est une permutation circulaire de $(1, 2, 3)$, -1 si c'est une permutation non circulaire et 0 sinon. On a alors le résultat suivant.

Proposition 3.4.12. Si dans une base orthonormée positive (b_1, b_2, b_3) , on a $u : (u_1, u_2, u_3)^\sim$, $v : (v_1, v_2, v_3)^\sim$ et $w : (w_1, w_2, w_3)^\sim$, alors on a

$$[u, v, w] = \sum_{i,j,k=1}^3 \varepsilon_{i,j,k} u_i v_j w_k$$

et

$$u \wedge v = \sum_{i,j,k=1}^3 \varepsilon_{i,j,k} u_i v_j b_k.$$

En particulier, on a $b_i \wedge b_j = \sum_{k=1}^3 \varepsilon_{i,j,k} b_k$, pour tous $i, j \leq 3$.

c. Tullio Levi-Civita (1873-1941), inventeur du calcul tensoriel (avec Ricci) vers 1900. Il eut une correspondance avec Einstein à propos de la relativité générale, l'aidant à corriger les erreurs initiales. En tant que mathématicien, j'apprécie particulièrement le passage suivant (d'une lettre d'Einstein) "I admire the elegance of your method of computation; it must be nice to ride through these fields upon the horse of true mathematics while the like of us have to make our way laboriously on foot".

d. En fait c'est un tenseur, mais ceci sort du cadre du cours.

Ayant trois vecteurs à sa disposition, on peut calculer deux fois le produit vectoriel, à condition de placer des parenthèses. On obtient alors un *double produit vectoriel* de la forme $u \wedge (v \wedge w)$. Il est intéressant de noter que le résultat de cette opération dépend du produit scalaire que l'on a choisi, mais pas de l'orientation. On ne sera donc pas étonné que le résultat s'exprime en termes de produits scalaires, comme indiqué dans le résultat suivant.

Proposition 3.4.13 (Double produit vectoriel). *On a les formules*

$$u \wedge (v \wedge w) = \langle u, w \rangle v - \langle u, v \rangle w \quad \text{et} \quad (u \wedge v) \wedge w = \langle u, w \rangle v - \langle u, w \rangle u$$

pour tous u, v, w dans E .

Ces formules peuvent être retenues comme suit : “le double produit vectoriel vaut le vecteur du milieu multiplié par le produit scalaire des deux autres, moins le deuxième vecteur de la parenthèse, multiplié par le produit scalaire des deux autres.

Démonstration. Il suffit de montrer la première identité, la deuxième s'en déduit directement. Vu la linéarité des deux membres sur chacun des facteurs, il suffit de vérifier l'identité pour des vecteurs d'une base orthonormée. C'est alors une longue vérification. \square

Le produit vectoriel permet de définir un type d'équations linéaires particulier en dimension 3, appelé équation vectorielle. Nous terminons cette section par ce type d'équation.

Proposition 3.4.14. *Soient E un espace vectoriel euclidien orienté de dimension 3 et $a, b \in E$ tels que $a \neq 0$. L'équation*

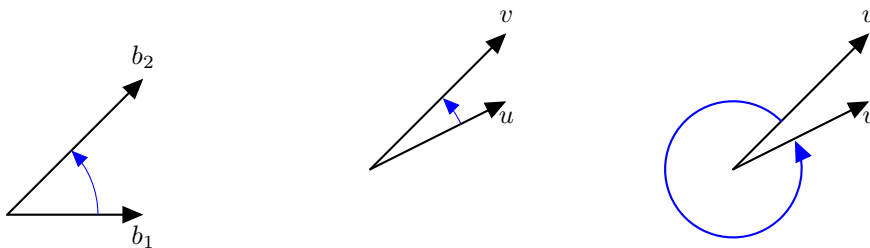
$$a \wedge x = b. \tag{3.11}$$

est compatible si, et seulement si, $\langle a, b \rangle = 0$. Dans ce cas l'ensemble des solutions est

$$S = \left\{ -\frac{1}{|a|^2} a \wedge b + \lambda a : \lambda \in \mathbb{R} \right\}.$$

3.4.4 Angle orienté dans le plan

On se place dans un espace vectoriel E euclidien orienté de dimension 2. Dans cette dimension particulière, la donnée de l'orientation définit un sens “de rotation”. Si $\mathcal{B} = (b_1, b_2)$ est une base positive, le sens positif consiste à “aller de b_1 vers b_2 ”. Cela permet de définir la notion d'angle orienté de deux vecteurs. D'un point de vue intuitif, on a la représentation suivante.



Pour calculer l'angle (u, v) , on va donc de u vers v comme on va de b_1 à b_2 . Cela amène naturellement la définition suivante.

Définition 3.4.15. Soit (u, v) un couple de vecteurs non nuls. L'angle orienté entre u et v est l'unique angle $\theta \in [0, 2\pi[$ tel que

- $\cos(\theta) = \cos(\alpha) = \frac{\langle u, v \rangle}{|u||v|}$ (α est l'angle non orienté entre u et v)
- $\sin(\theta)$ est positif (strictement) si, et seulement si (u, v) est une base positive.

Le signe du sinus est naturellement suffisant pour déterminer un angle entre 0 et 2π dont on connaît le cosinus. Il est facile d'exprimer le signe du sinus quand u et v sont donnés dans une base orthonormée positive.

Lemme 3.4.16. Si $\mathcal{B} = (b_1, b_2)$ est une base orthonormée positive, et si $u : (u_1, u_2)^\sim$ et $v : (v_1, v_2)^\sim$, alors $\sin((u, v))$ a le signe de $\det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}$.

Démonstration. Il suffit de remarquer que la matrice

$$\begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}$$

n'est rien d'autre que la matrice de changement de base de \mathcal{B} et (u, v) , du moins quand son déterminant est non nul. Par définition, elle a un déterminant positif si, et seulement si, (u, v) est une base positive (puisque \mathcal{B} est positive). \square

On peut même aller un peu plus loin et calculer exactement le sinus (le cosinus est déjà connu).

Proposition 3.4.17. Si $\mathcal{B} = (b_1, b_2)$ est une base orthonormée positive, et si $u : (u_1, u_2)^\sim$ et $v : (v_1, v_2)^\sim$, alors on a

$$\sin((u, v)) = \frac{\det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}}{|u||v|}.$$

Démonstration. On calcule le carré du sinus en utilisant la relation fondamentale de la trigonométrie :

$$\sin^2((u, v)) = 1 - \frac{(u_1v_1 + u_2v_2)^2}{(u_1^2 + u_2^2)(v_1^2 + v_2^2)} = \frac{(u_1^2 + u_2^2)(v_1^2 + v_2^2) - (u_1v_1 + u_2v_2)^2}{(u_1^2 + u_2^2)(v_1^2 + v_2^2)}.$$

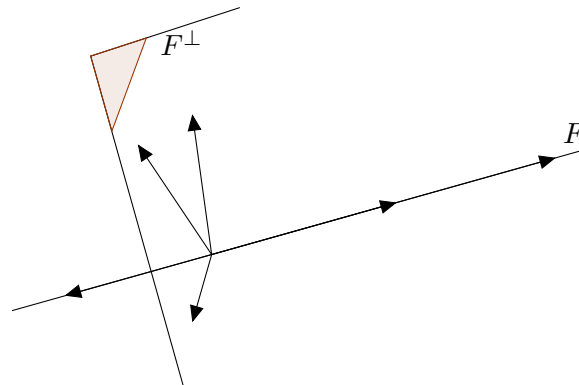
En simplifiant, on trouve directement

$$\sin^2((u, v)) = \frac{u_1^2v_2^2 + u_2^2v_1^2 - 2u_1v_1u_2v_2}{(u_1^2 + u_2^2)(v_1^2 + v_2^2)} = \frac{(u_1v_2 - u_2v_1)^2}{(u_1^2 + u_2^2)(v_1^2 + v_2^2)}.$$

On arrive au résultat puisqu'on connaît le signe du sinus. \square

3.5 Complément orthogonal et projection orthogonale

L'idée du complément orthogonal d'un sous-espace vectoriel est simple à concevoir si on se base sur les espaces de dimension 2 et 3, que l'on connaît bien. Voici la représentation intuitive de la situation en dimension 3 :



Le complément orthogonal de la droite vectorielle F est l'ensemble des vecteurs orthogonaux à tous les vecteurs de la droite. C'est un plan vectoriel F^\perp . Le complément orthogonal de ce plan est l'ensemble de tous les vecteurs orthogonaux à tous les vecteurs du plan. C'est $(F^\perp)^\perp$, qui n'est rien d'autre que F . Je vous invite à dessiner la situation en dimension 2.

Passons maintenant aux définitions générales. On se place ici encore dans un espace vectoriel euclidien de dimension n .

Définition 3.5.1. Pour tout sous-espace vectoriel F de E , le *complément orthogonal* de F est l'ensemble

$$F^\perp = \{u \in E : \langle u, f \rangle = 0, \quad \forall f \in F\}.$$

C'est donc l'ensemble des vecteurs orthogonaux à tous les vecteurs de F .

Cette définition se généralise quelque peu pour définir des sous-espaces orthogonaux.

Définition 3.5.2. Deux sous-espaces vectoriels F et G de E sont orthogonaux si on a $\langle f, g \rangle = 0$ pour tous $f \in F$ et $g \in G$. On note alors $F \perp G$.

De manière équivalente, on a $F \perp G$ si $F \subset G^\perp$ ou encore si $G \subset F^\perp$. On peut prouver en général ce que l'on a constaté en dimension 3. C'est l'objet du résultat suivant.

Proposition 3.5.3. Pour tout sous-espace vectoriel F de E ,

1. L'ensemble F^\perp est un sous-espace vectoriel de E ;
2. On a $E = F \oplus F^\perp$;
3. On a $\dim(F^\perp) = n - \dim(F)$;
4. On a $(F^\perp)^\perp = F$.

Démonstration. Pour le premier point, vu la définition de F^\perp , le vecteur nul appartient à F^\perp . Ensuite, on considère $u, v \in F^\perp$ et $\lambda \in \mathbb{R}$. Alors on a $u + v \in F^\perp$ car, pour tout $f \in F$, on a

$$\langle u + v, f \rangle = \langle u, f \rangle + \langle v, f \rangle = 0.$$

On montre de même que λu appartient à F^\perp .

Pour le deuxième point, on doit montrer que tout vecteur se décompose en la somme d'un élément de F et d'un élément de F^\perp et que $F \cap F^\perp = \{0\}$. Si $u \in F \cap F^\perp$, alors on a $\langle u, u \rangle = 0$, donc $u = 0$. Construisons une base (v_1, \dots, v_n) de E dont les p premiers éléments forment une base de F . Appliquons-lui le procédé de Gram-Schmidt pour obtenir une base (b_1, \dots, b_n) , orthonormée. Alors (b_1, \dots, b_p) est une base orthonormée de F , et pour tout $v \in E$, on a

$$v = \sum_{i=1}^p \langle v, b_i \rangle b_i + \sum_{i=p+1}^n \langle v, b_i \rangle b_i. \quad (3.12)$$

Le premier terme de cette somme appartient à F , tandis que le second terme appartient à F^\perp .

Pour le troisième point, en appliquant (3.12) aux éléments de F^\perp , on constate que (b_{p+1}, \dots, b_n) est une base de F^\perp , qui est donc de dimension $n - p$.

Pour le quatrième point, on sait que $(F^\perp)^\perp$ est de dimension p . Mais visiblement, on a $F \subset (F^\perp)^\perp$, donc ces deux sous-espaces vectoriels sont égaux. \square

Comme d'habitude, quand on a un objet unique, on lui donne un nom.

Définition 3.5.4. Pour tout $u \in E$, on a la décomposition unique $u = u_F + u_{F^\perp}$. La projection orthogonale de u sur F est u_F .

La relation (3.12), ainsi que la définition, nous donnent deux moyens de calculer la projection orthogonale d'un vecteur sur un sous-espace, une fois que l'on a une base orthonormée.

Proposition 3.5.5. *Si (f_1, \dots, f_p) est une base orthonormée de F , alors la projection orthogonale d'un vecteur u sur F est donnée par*

$$u_F = \langle u, f_1 \rangle f_1 + \dots + \langle u, f_p \rangle f_p.$$

De plus, la projection sur F^\perp est égale à $u - u_F$.

Remarque 3.3. 1. Cette proposition donne une interprétation du processus de Gram-Schmidt : on a simplement projeté chaque vecteur sur le complément orthogonal du sous-espace engendré par les précédents. Pour cela, il fallait une base orthonormée à chaque étape. Nous l'avons construite successivement en débutant par celle d'un espace de dimension 1.
2. Pour calculer la projection, il est parfois utile de choisir le sous-espace ayant la plus petite dimension (pour construire facilement une base orthonormée). C'est particulièrement évident dans le cas des hyperplans.

Terminons ce chapitre en particulierisant la proposition précédente dans le cas d'une droite et d'un hyperplan.

Proposition 3.5.6. *Si F est une droite vectorielle engendrée par f , alors la projection orthogonale de tout vecteur u sur F est*

$$u_d = \frac{\langle u, f \rangle}{|f|^2} f.$$

Démonstration. Il suffit de noter que $\frac{f}{|f|}$ est une base orthonormée de F . □

Pour les hyperplans, on se ramène aux droites via la définition suivante.

Définition 3.5.7. Si π est un hyperplan vectoriel, on appelle *normale* à π tout vecteur directeur de π^\perp .

Proposition 3.5.8. *Si π est un hyperplan vectoriel de normale N , alors la projection orthogonale de u sur π est*

$$u_\pi = u - \frac{\langle u, N \rangle}{|N|^2} N.$$

Remarque 3.4. Il arrive souvent que l'on considère a priori que la normale N est unitaire. La formule est alors plus simple.

Si un hyperplan π est donné par son équation dans une base orthonormée, il est facile de déterminer une normale, comme le montre le résultat suivant.

Proposition 3.5.9. *Dans un espace vectoriel euclidien E de dimension n muni d'une base orthonormée $\mathcal{B} = (b_1, \dots, b_n)$, le vecteur $N = (a_1, \dots, a_n)$ est normal à π si, et seulement si, l'hyperplan π admet l'équation cartésienne*

$$a_1 x_1 + \dots + a_n x_n = 0.$$

Démonstration. Par définition, le vecteur N est normal à π si, et seulement si, $\pi^\perp = \langle N \rangle$. Cette égalité est équivalente à $\pi = \langle N \rangle^\perp$, ou encore à

$$\pi = \{u \in E : \langle N, u \rangle = 0\}.$$

Cette description de π est visiblement équivalente à l'équation cartésienne de l'énoncé. □

L'équation cartésienne d'un hyperplan obtenue à partir d'une normale est appelée *équation normale* de cet hyperplan.

Chapitre 4

Espaces affines euclidiens

Dans ce chapitre, nous transposons dans les espaces affines les définitions et résultats que nous avons engrangés dans les espaces vectoriels euclidiens.

4.1 Espaces affines euclidiens et distances

Définition 4.1.1. Un espace affine euclidien est un espace affine modelé sur un espace euclidien. Un espace affine orienté est un espace affine modelé sur un espace orienté.

Dans un espace affine euclidien (resp. orienté), on peut introduire la notion de repère orthonormé (resp. positif).

Définition 4.1.2. Un repère orthonormé $(O, (b_1, \dots, b_n))$ (resp. positif) est un repère tel que la base (b_1, \dots, b_n) soit orthonormée (resp. positive).

Dans un repère orthonormé, il est simple d'obtenir les coordonnées d'un point P quelconque. Le résultat suivant est analogue à celui que nous avons rencontré dans les espaces vectoriels euclidiens.

Proposition 4.1.3. Les coordonnées d'un point P dans un repère orthonormé $(O, (b_1, \dots, b_n))$ sont données par

$$P : (\langle \overrightarrow{OP}, b_1 \rangle, \dots, \langle \overrightarrow{OP}, b_n \rangle)^\sim.$$

Démonstration. On se rappelle que les coordonnées de P sont les composantes de \overrightarrow{OP} dans (b_1, \dots, b_n) . \square

Définition 4.1.4. La distance euclidienne du point P au point Q de \mathcal{A} , notée $d(P, Q)$ est définie par

$$d(P, Q) = |\overrightarrow{PQ}|.$$

Les propriétés de la distance euclidienne sont rassemblées dans le résultat suivant.

Proposition 4.1.5. On a

1. $d(P, Q) = d(Q, P)$ pour tous $P, Q \in \mathcal{A}$;
2. $d(P, Q) \geq 0$ pour tous $P, Q \in \mathcal{A}$ et $d(P, P) = 0$;
3. $d(P, Q) + d(Q, R) \geq d(P, R)$ pour tous $P, Q, R \in \mathcal{A}$;
4. $d(P, Q) = 0$ si et seulement si $P = Q$.

Remarque 4.1. Les trois premières propriétés sont celles qui définissent en général une pseudo-distance. La quatrième est celle qui définit les distances parmi les pseudo-distances. La troisième propriété est appelée inégalité triangulaire. Elle exprime que dans un triangle, la somme des longueurs de deux côtés est supérieure à la longueur du troisième.

Démonstration. Ce sont des propriétés de la norme des vecteurs dans l'espace vectoriel euclidien définissant l'espace affine \mathcal{A} . \square

Passons maintenant au théorème de Pythagore et à sa généralisation.

Proposition 4.1.6. *Soit un triangle A, B, C . On a alors*

$$|\overrightarrow{AB}|^2 = |\overrightarrow{AC}|^2 + |\overrightarrow{CB}|^2 - 2|\overrightarrow{AC}||\overrightarrow{CB}|\cos(\alpha)$$

où α est l'angle non orienté entre \overrightarrow{CA} et \overrightarrow{CB} . En particulier, le triangle est rectangle en C si et seulement si on a

$$|\overrightarrow{AB}|^2 = |\overrightarrow{AC}|^2 + |\overrightarrow{CB}|^2.$$

Démonstration. On a

$$|\overrightarrow{AB}|^2 = \langle \overrightarrow{AB}, \overrightarrow{AB} \rangle = \langle \overrightarrow{AC} + \overrightarrow{CB}, \overrightarrow{AC} + \overrightarrow{CB} \rangle.$$

Après développement par bilinéarité et symétrie du membre de droite, on obtient

$$|\overrightarrow{AB}|^2 = |\overrightarrow{AC}|^2 + |\overrightarrow{CB}|^2 + 2\langle \overrightarrow{AC}, \overrightarrow{CB} \rangle = |\overrightarrow{AC}|^2 + |\overrightarrow{CB}|^2 - 2\langle \overrightarrow{CA}, \overrightarrow{CB} \rangle,$$

qui donne le résultat annoncé, vu la relation $\langle \overrightarrow{CA}, \overrightarrow{CB} \rangle = |\overrightarrow{CA}||\overrightarrow{CB}|\cos(\alpha)$. Le cas particulier est alors évident. \square

Enfin, il est important de pouvoir calculer la distance de deux points en fonction de leurs coordonnées.

Proposition 4.1.7. *Si les coordonnées de P et Q dans un repère orthonormé sont $(p_1, \dots, p_n)^\sim$ et $(q_1, \dots, q_n)^\sim$, alors la distance $d(P, Q)$ est donnée par*

$$d(P, Q) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}.$$

Démonstration. Nous devons calculer $\sqrt{\langle \overrightarrow{PQ}, \overrightarrow{PQ} \rangle}$. Les composantes de \overrightarrow{PQ} dans la base associée au repère sont par hypothèse $(q_1 - p_1, \dots, q_n - p_n)$. Puisqu'on est dans une base orthonormée, le produit scalaire est la somme des produits des composantes, et on obtient le résultat annoncé. \square

4.2 Variétés affines orthogonales

La notion d'orthogonalité pour les variétés affines se déduit aussi de la notion mère pour les sous-espaces vectoriels. Dans le cas des plans en dimension 3, elle ne doit cependant pas être confondue avec la notion de perpendicularité, que nous définirons en fin de chapitre.

Définition 4.2.1. Les variétés affines $A + F$ et $A' + F'$ sont dites orthogonales si leurs sous-espaces vectoriels directeurs F et F' sont orthogonaux.

Voici quelques cas particuliers intéressants. Il est utile de remarquer que l'orthogonalité de sous-espaces peut être vérifiée sur des parties génératrices de ces sous-espaces.

- Pour le cas des droites, si $d = A + \langle u \rangle$ et $d' = A' + \langle v \rangle$, alors $d \perp d'$ si, et seulement si $\langle u, v \rangle = 0$. On obtient donc le résultat suivant.
Si u et v admettent respectivement pour composantes $(u_1, \dots, u_n)^\sim$ et $(v_1, \dots, v_n)^\sim$ dans *une base orthonormée*, alors on a $d \perp d'$ si, et seulement si $u_1 v_1 + \dots + u_n v_n = 0$.

- En dimension 3, le seul cas intéressant en dehors de deux droites est celui d'une droite et d'un plan. On considère une droite $d = A + \langle u \rangle$ et un plan $\pi = B + \langle v, w \rangle$. On a alors $d \perp \pi$ si, et seulement si $\langle u, v \rangle = \langle u, w \rangle = 0$. On a aussi $d \perp \pi$ si, et seulement si $\vec{d} = \vec{\pi}^\perp$, ce qui s'exprime encore par le fait que u soit une normale à π .
Si dans un repère orthonormé, on a

$$\pi \equiv a_1x_1 + a_2x_2 + a_3x_3 = b$$

et si dans la base associée au repère, on a $u : (u_1, u_2, u_3)^\sim$, alors on a $d \perp \pi$ si et seulement si (u_1, u_2, u_3) est multiple de (a_1, a_2, a_3) .

- En dimension 3, deux plans ne sont jamais orthogonaux. En effet, si π et π' étaient deux plans orthogonaux, alors on aurait par exemple $\vec{\pi} \subset \vec{\pi}'^\perp$, mais le premier est de dimension 2, alors que le deuxième est de dimension 1.

4.3 Projections orthogonales

Ici encore, on se sert abondamment de la notion de projection orthogonale dans les espaces vectoriels. Je présente d'abord la proposition qui justifie la définition qui suit.

Proposition 4.3.1. *Soit $\mathcal{V} = A + F$ une variété affine et $P \in \mathcal{A}$. Il existe un unique point $P' \in \mathcal{V}$ tel que $\overrightarrow{PP'} \perp F$. De plus, on a $P' = A + p_{\vec{\mathcal{V}}}(\overrightarrow{AP}) = P - p_{\vec{\mathcal{V}}^\perp}(\overrightarrow{AP})$.*

Démonstration. Il suffit de décomposer le vecteur \overrightarrow{AP} , de manière unique, en la somme d'un vecteur de F et d'un vecteur de F^\perp . Il existe $u \in F$, $u' \in F^\perp$ uniques tels que $\overrightarrow{AP} = u + u'$. On a alors directement $A + u = P - u'$. Notons ce point P' . Il est dans \mathcal{V} car il s'écrit $A + u$. D'autre part, on a

$$\overrightarrow{PP'} = -u' \in F^\perp,$$

ce qui montre que P' satisfait les conditions de l'énoncé. Si P'' satisfait également les conditions de l'énoncé, alors

$$\overrightarrow{AP} = \overrightarrow{AP''} + \overrightarrow{P''P}, \quad \text{et} \quad \overrightarrow{AP} = \overrightarrow{AP'} + \overrightarrow{P'P}$$

sont deux décompositions de \overrightarrow{AP} en la somme d'un vecteur de F et d'un vecteur de F^\perp . On a donc $\overrightarrow{AP'} = \overrightarrow{AP''}$, qui donne $P' = P''$. \square

Définition 4.3.2. L'unique point P' défini par la proposition précédente est appelé projection orthogonale de P sur \mathcal{V} . Il est noté $p_{\vec{\mathcal{V}}}^\perp(P)$.

Voici maintenant une caractérisation géométrique.

Proposition 4.3.3. *La projection orthogonale de P sur $\mathcal{V} = A + F$ est l'unique point P' d'intersection des variétés affines orthogonales $A + F$ et $P + F^\perp$.*

Démonstration. On a montré dans la proposition précédente que ces variétés affines ont une intersection non vide. C'est donc une variété affine, dont le sous-espace vectoriel directeur est $F \cap F^\perp = \{0\}$. C'est donc un singleton. \square

Proposition 4.3.4. *Soit π l'hyperplan déterminé par le point A et le vecteur normal N , c'est-à-dire*

$$\pi = \{X \in \mathcal{A} : \langle \overrightarrow{AX}, N \rangle = 0\}$$

et $P \in \mathcal{A}$. Alors on a

$$p_\pi^\perp(P) = P' = P - \frac{\langle \overrightarrow{AP}, N \rangle}{|N|^2} N.$$

Démonstration. On applique la proposition 4.3.1 et la formule donnant la projection orthogonale sur $\vec{\pi}^\perp = \langle N \rangle$. \square

Nous pouvons également déterminer la projection orthogonale d'un point sur une droite.

Proposition 4.3.5. *Soit $d = A + \langle u \rangle$ et soit $P \in \mathcal{A}$. La projection orthogonale de P sur d est le point*

$$p_d^\perp(P) = P' = A + \frac{\overrightarrow{\langle AP, u \rangle}}{|u|^2}u.$$

Démonstration. On applique la proposition 4.3.1 et la formule donnant la projection orthogonale sur la droite vectorielle engendrée par u . \square

4.4 Distances de variétés affines, et cas des dimensions deux et trois

On a étudié la distance de deux droites, d'un point à une droite, à un plan dans l'enseignement secondaire, sans toujours avoir une idée claire de la définition. En fait la distance entre deux ensembles peut être définie de manière simple, quels que soit les ensembles considérés.

Définition 4.4.1. La distance entre deux sous-ensembles \mathcal{E} et \mathcal{F} non vides de \mathcal{A} est définie par

$$d(\mathcal{E}, \mathcal{F}) = \inf(\mathcal{D}), \quad \text{où } \mathcal{D} = \{d(A, B) : A \in \mathcal{E}, B \in \mathcal{F}\}.$$

Il est important de remarquer que la distance entre deux ensembles est toujours bien définie, car le sous-ensemble \mathcal{D} de \mathbb{R} est fait de nombres supérieurs ou égaux à 0. Il a donc une borne inférieure, qui est de plus supérieure ou égale à 0.

On dit que la distance entre \mathcal{E} et \mathcal{F} est réalisée si il existe un couple de points $A \in \mathcal{E}$ et $B \in \mathcal{F}$ tels que $d(A, B) = d(\mathcal{E}, \mathcal{F})$. Ce n'est pas toujours le cas. A titre d'exemple, considérons l'espace affine euclidien \mathbb{R}^2 muni de son produit scalaire préféré, et soient $\mathcal{E} = \{(x, y)^\sim \in \mathbb{R}^2 : y = 0\}$, et $\mathcal{F} = \{(x, \frac{1}{x})^\sim : x \neq 0\}$. On montre assez facilement que $d(\mathcal{E}, \mathcal{F}) = 0$, mais il n'existe aucun couple de points A et B satisfaisant $d(A, B) = 0$.

Le cas des variétés affines est différent. La distance est toujours réalisée, et on peut même la calculer facilement dans de nombreux cas.

Il est utile de se remémorer ce que l'on connaît de la distance de deux droites dans le plan, d'un point à une droite dans le plan, de deux plans en dimension 3, d'un point et d'un plan...

Une constante se dégage et je vous la livre dans le théorème général suivant, que nous appliquerons pour récupérer tous les cas que nous connaissons.

Théorème 4.4.2. *Soit $\mathcal{V} = A + F$ et $\mathcal{V}' = B + F'$ deux variétés affines. Il existe un unique vecteur $\overrightarrow{PP'}$ satisfaisant les conditions suivantes :*

1. $P \in \mathcal{V}$ et $P' \in \mathcal{V}'$;
2. $\overrightarrow{\langle PP' \rangle}$ est orthogonal à F et à F' .

On a $\overrightarrow{PP'} = p_{(F+F')^\perp}(\overrightarrow{AB})$. La distance de \mathcal{V} et \mathcal{V}' vaut $|\overrightarrow{PP'}|$. Si $F \cap F' = \{0\}$, alors la distance est réalisée par un unique couple (P, P') .

Démonstration. Démontrons l'unicité du vecteur $\overrightarrow{PP'}$. S'il satisfait les conditions, on a alors

$$\overrightarrow{AB} = (\overrightarrow{AP} + \overrightarrow{P'B}) + \overrightarrow{PP'}.$$

Le premier terme est alors dans $F + F'$ et le second dans $(F + F')^\perp$. C'est donc nécessairement $p_{(F+F')^\perp}(\overrightarrow{AB})$, vu l'unicité de la décomposition de \overrightarrow{AB} .

Démontrons l'existence d'un couple (P, P') satisfaisant la première propriété. On a

$$\overrightarrow{\mathcal{A}} = (F + F') \oplus (F + F')^\perp.$$

On peut donc décomposer \overrightarrow{AB} de manière unique en

$$\overrightarrow{AB} = u_{F+F'} + u_{(F+F')^\perp},$$

et on décompose $u_{F+F'}$ en $u_F + u_{F'}$, où $u_F \in F$ et $u_{F'} \in F'$. On a alors

$$B - A - u_F - u_{F'} = u_{(F+F')^\perp}$$

et il suffit de poser $P = A + u_F \in \mathcal{V}$ et $P' : B - u_{F'} \in \mathcal{V}'$. Alors on a

$$\overrightarrow{PP'} = u_{(F+F')^\perp} = p_{(F+F')^\perp}(\overrightarrow{AB}),$$

donc en particulier $\overrightarrow{PP'}$ est orthogonal à F et à F' .

Pour tous $X \in \mathcal{V}$ et $X' \in \mathcal{V}'$, on a alors

$$|\overrightarrow{XX'}|^2 = |\overrightarrow{XP} + \overrightarrow{PP'} + \overrightarrow{P'X'}|^2 = |\overrightarrow{PP'}|^2 + |\overrightarrow{XP} + \overrightarrow{P'X'}|^2 \geq |\overrightarrow{PP'}|^2. \quad (4.1)$$

On a donc démontré que $|\overrightarrow{XX'}| \geq |\overrightarrow{PP'}|$, et donc que $|\overrightarrow{PP'}|$ réalise la distance de \mathcal{V} à \mathcal{V}' .

Enfin, supposons qu'un coupe (X, X') réalise la distance. L'équation (4.1) donne alors $|\overrightarrow{XP} + \overrightarrow{P'X'}|^2 = 0$, donc $\overrightarrow{XP} = -\overrightarrow{P'X'}$. Ce vecteur est dans $F \cap F'$, donc il est nul. \square

Remarque 4.2. Ce théorème implique par exemple l'existence d'une perpendiculaire commune (unique) à deux droites gauches.

Particularisons au cas où une des variétés affines est un singleton.

Corollaire 4.4.3. *La distance d'un singleton $\{B\}$ à une variété affine \mathcal{V} est réalisée : on a $d(B, \mathcal{V}) = d(B, B')$ où B' est la projection orthogonale de B sur \mathcal{V} . La distance est réalisée de manière unique.*

Démonstration. On applique le théorème précédent, avec $F = \{0\}$. On a $d(B, \mathcal{V}) = d(P, P')$ où $P = B$, $P' \in \mathcal{V}$ et $\overrightarrow{PP'}$ est orthogonal à \mathcal{V} . Le point P' est donc la projection orthogonale de B sur \mathcal{V} . L'unicité résulte aussi du théorème général. \square

Dans le cas où la variété est un hyperplan π , on peut calculer facilement la distance.

Corollaire 4.4.4. *Soient $\pi = A + \vec{\pi}$ un hyperplan, N une normale à cet hyperplan et $B \in \mathcal{A}$. On a alors*

$$d(B, \pi) = \frac{|\langle \overrightarrow{AB}, N \rangle|}{|N|}$$

En particulier, si dans un repère orthonormé π admet pour équation

$$\pi \equiv c_1x_1 + \cdots + c_nx_n = b$$

et si B a pour coordonnées $(p_1, \dots, p_n)^\sim$, alors on a

$$d(P, \pi) = \frac{|c_1p_1 + \cdots + c_np_n - b|}{\sqrt{c_1^2 + \cdots + c_n^2}}.$$

Démonstration. On applique le théorème général avec $F = \vec{\pi}$ et $F' = \{0\}$. La distance est donnée par $|p_{\pi^\perp}(\overrightarrow{AB})|$. Puisque $\pi^\perp = \langle N \rangle$, on a

$$p_{\pi^\perp}(\overrightarrow{AB}) = \frac{\langle \overrightarrow{AB}, N \rangle}{|N|^2} N,$$

et on arrive à la conclusion du premier point en prenant la norme.

Si A a pour coordonnées (a_1, \dots, a_n) et si P et π sont comme dans l'énoncé, alors le vecteur N de composantes $(c_1, \dots, c_n)^\sim$ est une normale à π et on a $\overrightarrow{AP} : (p_1 - a_1, \dots, p_n - a_n)^\sim$. On a donc

$$\frac{|\langle \overrightarrow{AP}, N \rangle|}{|N|} = \frac{|c_1(p_1 - a_1) + \cdots + c_n(p_n - a_n)|}{\sqrt{c_1^2 + \cdots + c_n^2}}.$$

Puisque A est dans π , on obtient le résultat demandé. \square

Un autre corollaire d'importance est le suivant, qui donne la distance d'un point à une droite.

Corollaire 4.4.5. *Soit \mathcal{D} la droite déterminée par le point A et le vecteur directeur u et soit $B \in \mathcal{A}$. Alors on a*

$$d(B, \mathcal{D}) = |\overrightarrow{AB}| \sin(\alpha),$$

où α est l'angle non orienté entre \overrightarrow{AB} et u .

En particulier, pour un espace affine euclidien de dimension 3, on a

$$d(P, \mathcal{D}) = \frac{|\overrightarrow{AB} \wedge u|}{|u|},$$

où le produit vectoriel est calculé dans une orientation quelconque.

Démonstration. Il s'agit de calculer $d(P, \mathcal{D})$, où $\overrightarrow{PP'} = p_{\gamma_{u^\perp}}(\overrightarrow{AB})$. Il est plus facile de calculer la projection sur γ_{u^\perp} : elle vaut

$$\frac{\langle \overrightarrow{AB}, u \rangle}{|u|^2} u.$$

On a donc

$$\overrightarrow{AB} = \overrightarrow{PP'} + \frac{\langle \overrightarrow{AB}, u \rangle}{|u|^2} u,$$

donc

$$|\overrightarrow{PP'}|^2 = |\overrightarrow{AB}|^2 - \left(\frac{\langle \overrightarrow{AB}, u \rangle}{|u|}\right)^2 = |\overrightarrow{AB}|^2 - (|\overrightarrow{AB}| \cos(\alpha))^2.$$

On peut prendre la racine carrée des deux membres et arriver à la première formule puisque $\sin(\alpha)$ est positif.

Pour le cas particulier, on se rappelle la formule donnant la norme du produit vectoriel : $|\overrightarrow{AB} \wedge u| = |\overrightarrow{AB}| |u| \sin(\alpha)$. \square

Terminons cette section par l'analyse de la distance de deux variétés affines dans les espaces affines euclidiens de dimension deux et trois.

4.4.1 Le cas de la dimension 2

Outre les cas que nous avons considérés, il ne reste en dimension deux que le cas de deux droites. Le résultat est alors sans surprise.

Proposition 4.4.6. *Si \mathcal{D} et \mathcal{D}' sont sécantes, alors $d(\mathcal{D}, \mathcal{D}') = 0$. Si \mathcal{D} et \mathcal{D}' sont parallèles, alors on a $d(\mathcal{D}, \mathcal{D}') = d(P, P')$ pour tout $P \in \mathcal{D}$, si P' est la projection de P sur \mathcal{D}' .*

4.4.2 Le cas de la dimension 3

On a en dimension trois le cas de deux plans ou le cas de deux droites. Pour le cas de deux plans, le résultat et la preuve sont analogues à ceux de la proposition 4.4.6. Elle est laissée comme exercice.

Proposition 4.4.7. *Si π et π' sont sécants, alors $d(\pi, \pi') = 0$. Si π et π' sont parallèles, alors on a $d(\pi, \pi') = d(P, P')$ pour tout $P \in \pi$, si P' est la projection de P sur π' .*

On terminera en traitant le cas de deux droites. Dans le cas où elles sont parallèles ou sécantes, le résultat et la preuve sont encore analogues au précédent. Traitons le cas de deux droites gauches.

Proposition 4.4.8. *Si les droites $\mathcal{D} = A + \gamma_u$ et $\mathcal{D}' = B + \gamma_v$ sont gauches, alors la distance de \mathcal{D} à \mathcal{D}' est réalisée de manière unique, en des points $P \in \mathcal{D}$, et $P' \in \mathcal{D}'$ tels que PP' soit une droite orthogonale à \mathcal{D} et \mathcal{D}' . Elle vaut*

$$\frac{|[\overrightarrow{AB}, u, v]|}{|u \wedge v|}.$$

Démonstration. La distance est réalisée de manière unique car $\rangle u \langle \cap \rangle v \langle = \{0\}$. On la calcule en calculant la norme de la projection orthogonale de \overrightarrow{AB} sur le complément orthogonal de $\rangle u, v \langle$. Une base de ce complément orthogonal est donnée par $u \wedge v$. Donc la projection est donnée par

$$\frac{\langle \overrightarrow{AB}, u \wedge v \rangle}{|u \wedge v|^2} u \wedge v = \frac{[\overrightarrow{AB}, u v]}{|u \wedge v|^2} u \wedge v.$$

La norme de ce vecteur donne le résultat annoncé. \square

4.5 Angles de droites et plans

Nous commençons par la définition de l'angle de deux droites. Intuitivement, deux droites sécantes déterminent un seul plan et dans ce plan déterminent deux angles supplémentaires, au sens où cela a été défini dans l'enseignement secondaire.

Plus précisément, si u et u' sont des vecteurs directeurs des droites \mathcal{D} et \mathcal{D}' , alors ces angles (non orientés) sont les angles α_1 et α_2 déterminés d'une part par u et u' et d'autre part par $-u$ et u' . On a donc

$$\cos(\alpha_1) = \frac{\langle u, u' \rangle}{|u||u'|} \quad \text{et} \quad \cos(\alpha_2) = \frac{\langle -u, u' \rangle}{|-u||u'|} = -\cos(\alpha_1).$$

Par convention, on décide que l'angle des deux droites est le plus petit des angles α_1, α_2 , c'est-à-dire celui dont le cosinus est positif ou nul.

Cette définition ne dépend pas du fait que les droites soient sécantes ou non. On arrive donc à la définition suivante.

Définition 4.5.1. L'angle des droites $\mathcal{D} = A+\rangle u \langle$ et $\mathcal{D}' = A'+\rangle u' \langle$ est l'angle $\alpha \in [0, \frac{\pi}{2}]$ déterminé par

$$\cos(\alpha) = \frac{|\langle u, u' \rangle|}{|u||u'|}.$$

On remarque que l'angle de deux droites ne dépend que de leur direction et non de leur position précise. En effet, si \mathcal{D}' est parallèle à \mathcal{D}'' , alors \mathcal{D}' et \mathcal{D}'' partagent les mêmes vecteurs directeurs et l'angle de \mathcal{D} et \mathcal{D}' est égal à l'angle de \mathcal{D} et \mathcal{D}'' .

Pour introduire l'angle d'une droite et d'un hyperplan ou l'angle de deux hyperplans, on se ramène au cas précédent.

Définition 4.5.2. L'angle d'une droite \mathcal{D} et d'un hyperplan π vaut $\beta = \frac{\pi}{2} - \alpha$ où α est l'angle entre \mathcal{D} et une normale quelconque \mathcal{N} à π . On a donc

$$\sin(\beta) = \frac{|\langle u, N \rangle|}{|u||N|},$$

où u est un vecteur directeur de \mathcal{D} et N une normale à π .

Définition 4.5.3. L'angle de deux hyperplans π et π' est l'angle α entre deux de leurs droites normales quelconques. On a donc

$$\cos(\alpha) = \frac{|\langle N, N' \rangle|}{|N||N'|}$$

où N et N' sont des normales à π et π' respectivement.

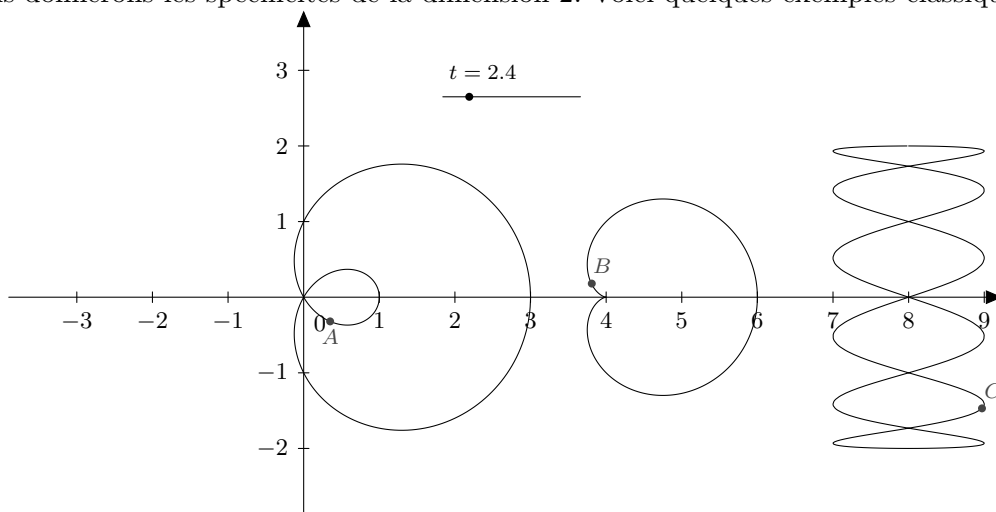
Chapitre 5

Théorie des courbes

Dans ce chapitre, nous présentons quelques points simples de la théorie des courbes. L'objet de cette théorie est de formaliser le mouvement d'un point dans un espace affine. La position du point dépend d'un paramètre qui peut être le temps, mais pas uniquement.

On va s'intéresser à la trajectoire parcourue par le point, c'est-à-dire à l'*arc de courbe* géométrique, mais aussi à la manière de parcourir cet arc de courbe, c'est-à-dire le *paramétrage* de cet arc.

Nous considérons un espace affine euclidien orienté \mathcal{A} (bien que le produit scalaire ou l'orientation ne soient pas utiles pour certaines notions), et nous fixerons éventuellement un repère orthonormé positif dans cet espace. Enfin nous étudierons en détail le cas de la dimension 3, puis donnerons les spécificités de la dimension 2. Voici quelques exemples classiques :



Les coordonnées des points A , B et C dépendent d'un paramètre t (qui peut par exemple être le temps). Ce paramètre varie dans un intervalle de \mathbb{R} , que nous noterons Ω . Nous devrons calculer la dérivée des coordonnées, et il est donc commode de supposer que cet intervalle est ouvert. Dans le repère représenté sur la figure ci-dessus, les points A , B et C ont pour coordonnées

$$\begin{aligned} A(t) &: ((1 + 2 \cos(t)) \cos(t), (1 + 2 \cos(t)) \sin(t))^\sim, & t \in]0, 2\pi[\\ B(t) &: (4 + (1 + \cos(t)) \cos(t), (1 + \cos(t)) \sin(t))^\sim, & t \in]0, 2\pi[\\ C(t) &: (8 + \sin(t), 2 \cos(t))^\sim, & t \in]0, 2\pi[. \end{aligned}$$

5.1 Quelques points techniques

Nous donnons ici les définitions nécessaires pour le calcul différentiel pour les fonctions à valeurs vectorielles ou à valeurs dans un espace affine. L'idée simple est de choisir une base ou un repère selon le cas, et d'étendre les définitions du cours d'analyse, composante à composante.

Ensuite, on montre que cette définition est indépendante du choix de la base ou du repère choisi. On considère des applications définies sur un ouvert $\Omega \subset \mathbb{R}$.

Définition 5.1.1. Une application P définie sur Ω et à valeurs dans un espace affine \mathcal{A} est continue (resp. dérivable, de classe C_p) si il existe un repère $\mathcal{R} = (O, (b_1, \dots, b_n))$ tel que les coordonnées de P soient des fonctions continues (resp. dérivables, de classe C_p).

La définition est identique pour une fonction à valeurs dans un espace vectoriel E . On demande l'existence d'une base dans laquelle les composantes de P ont la propriété demandée.

La définition est indépendante du choix d'un repère, en ce sens que si elle est vraie dans un repère, elle est vraie dans tous les repères, comme le montre la propriété suivante.

Proposition 5.1.2. Soit $P : \Omega \rightarrow \mathcal{A}$ une application. Si les coordonnées de P dans un repère $\mathcal{R} = (O, (b_1, \dots, b_n))$, sont continues (resp. dérivables, de classe C_p), alors elles le sont dans tout repère.

Démonstration. Si dans le repère \mathcal{R} l'application $P : \Omega \rightarrow \mathcal{A} : u \mapsto P(u)$ s'exprime en coordonnées par

$$P(u) : \begin{pmatrix} x_1(u) \\ \vdots \\ x_n(u) \end{pmatrix} \quad (5.1)$$

avec des fonctions x_1, \dots, x_n continues (resp. dérivables, de classe C_p) de Ω dans \mathbb{R} , alors dans un autre repère \mathcal{R}' , on a

$$P(u) : \begin{pmatrix} y_1(u) \\ \vdots \\ y_n(u) \end{pmatrix}.$$

Les fonctions y_1, \dots, y_n sont obtenues en fonction de x_1, \dots, x_n par la formule de changement de repère :

$$\begin{cases} y_1(u) &= a_{11}x_1(u) + \dots + a_{1n}x_n(u) + b_1 \\ \vdots &= \vdots \\ y_n(u) &= a_{n1}x_1(u) + \dots + a_{nn}x_n(u) + b_n, \end{cases} \quad \forall u \in \Omega, \quad (5.2)$$

où les coefficients $a_{11}, \dots, a_{nn}, b_1, \dots, b_n$ sont des constantes. Les théorèmes classiques du cours d'analyse permettent alors de montrer que les fonctions y_1, \dots, y_n sont continues (resp. dérivables, de classe C_p). \square

Remarque 5.1. La même construction permet de définir une fonction dérivable, différentiable ou de classe C_p définie sur un ouvert Ω de \mathbb{R}^k et de traiter ainsi des fonctions de plusieurs paramètres.

Si l'application P est dérivable en un point $u_0 \in \Omega$, d'après notre définition, c'est que ses coordonnées admettent des dérivées en ce point, dans tout repère. On peut donc définir la dérivée.

Définition 5.1.3. Soit $P : \Omega \rightarrow \mathcal{A}$ une application dérivable en un point $u_0 \in \Omega$, qui s'exprime dans un repère par (5.1), alors la dérivée de P en u_0 est l'élément de $\overrightarrow{\mathcal{A}}$ donné dans le repère $\mathcal{R} = (O, \mathcal{B})$ par

$$\overrightarrow{D_u P}(u_0) : \begin{pmatrix} D_u x_1(u_0) \\ \vdots \\ D_u x_n(u_0) \end{pmatrix} \quad (5.3)$$

Encore faut-il vérifier que cette définition est indépendante du choix du repère. Si P a des coordonnées x_1, \dots, x_n et y_1, \dots, y_n dans les repères \mathcal{R} et \mathcal{R}' , alors ces coordonnées sont liées

entre elles par la formule (5.2). La dérivée dans le repère \mathcal{R}' s'exprime alors en fonction de la dérivée dans le repère \mathcal{R} par la relation

$$\begin{pmatrix} D_u y_1(u_0) \\ \vdots \\ D_u y_n(u_0) \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} D_u x_1(u_0) \\ \vdots \\ D_u x_n(u_0) \end{pmatrix}.$$

On constate que c'est la formule de changement de base qui est utilisée ici, donc $\overrightarrow{D_u P}(u_0)$ est indépendant du choix du repère, si on précise que c'est en fait l'élément de $\overrightarrow{\mathcal{A}}$ dont les composantes dans la base \mathcal{B} du repère \mathcal{R} sont données par (5.3). C'est pour insister sur ce fait que l'on note la dérivée $\overrightarrow{D_u P}(u_0)$.

Remarque 5.2. Nous avons utilisé ici une approche en coordonnées assez classique : on a constaté que la dérivée d'une fonction à valeurs dans un espace affine devait être considérée comme un vecteur pour que la construction soit indépendante du choix du repère. C'est un procédé très courant, en particulier en physique, où on définit des objets au moyen de coordonnées, et où leur "loi de transformation" lors d'un changement de coordonnées définit le type d'objet qui a été défini (point, vecteur, forme différentielle, tenseur...). Ici nous avons constaté que la dérivée suit la loi de transformation des vecteurs.

On aurait en fait pu définir la limite en un point d'une fonction à valeurs dans un espace affine, et généraliser la construction faite au cours d'analyse pour définir la dérivée par

$$\overrightarrow{D_u P}(u_0) = \lim_{h \rightarrow 0} \frac{P(u_0 + h) - P(u_0)}{h}.$$

Le résultat aurait été le même, et on aurait vu directement qu'il s'agit d'un élément de $\overrightarrow{\mathcal{A}}$, il aurait cependant fallu alors expliquer comment calculer cette limite, et vous auriez été privés d'un raisonnement classique dans certaines branches de la physique.

Enfin, on remarquera que les mêmes raisonnements s'appliquent pour une fonction à valeurs dans un espace vectoriel E (ici encore, de dimension finie). La dérivée en un point u_0 est définie dans une base, composante à composante et l'objet ainsi défini suit la loi de transformation des vecteurs ; c'est donc un élément de E . Bien sûr, en laissant varier u_0 , on obtient une fonction à valeurs dans E .

5.1.1 Quelques résultats sur la dérivation des fonctions à valeurs vectorielles

Quand on a à sa disposition des fonctions sur un ouvert $\Omega \subset \mathbb{R}$ à valeurs dans un espace vectoriel muni d'opérations telles que le produit scalaire, le produit mixte..., on peut en construire de nouvelles par composition. Par exemple, si f et g sont des fonctions à valeurs dans un espace euclidien, on peut construire la fonction $\langle f, g \rangle$ qui à tout $u \in \Omega$ associe $\langle f(u), g(u) \rangle$. C'est une fonction de Ω dans \mathbb{R} , qu'il est utile de pouvoir dériver. C'est l'objet du résultat technique suivant.

Proposition 5.1.4. *Soient f, g, h des fonctions dérivables de $\Omega \subset \mathbb{R}$ dans un espace vectoriel E , alors*

- Si E est euclidien, on a

$$D_u(\langle f(u), g(u) \rangle) = \langle D_u f(u), g(u) \rangle + \langle f(u), D_u g(u) \rangle, \quad \forall u \in \Omega.$$

- En conséquence, si f et g ont un produit scalaire constant sur Ω , on a

$$\langle D_u f(u), g(u) \rangle = -\langle f(u), D_u g(u) \rangle, \quad \forall u \in \Omega.$$

En particulier, si la norme de f est constante sur Ω , alors $\langle D_u f(u), f(u) \rangle = 0$, pour tout $u \in \Omega$.

- Si E est euclidien orienté de dimension 3, alors

$$D_u(f(u) \wedge g(u)) = D_u f(u) \wedge g(u) + f(u) \wedge D_u g(u),$$

et

$$D_u[f(u), g(u), h(u)] = [D_u f(u), g(u), h(u)] + [f(u), D_u g(u), h(u)] \\ + [f(u), g(u), D_u h(u)].$$

Remarque 5.3. Ces résultats se retiennent facilement : les produits que nous avons défini (scalaire, mixte, vectoriel) vérifient la règle de Leibniz.

Démonstration. Il suffit d'exprimer le membre de gauche de chaque de l'égalité en fonction des composantes des fonctions f, g, h dans une base orthonormée de E , de dériver, en appliquant la règle de Leibniz classique et de constater que cela donne le membre de droite. En détail : Si $f(u) : (f_1(u), \dots, f_n(u))^\sim$ et si $g(u) : (g_1(u), \dots, g_n(u))^\sim$ dans une base orthonormée, alors on a

$$\langle f(u), g(u) \rangle = \sum_{k=1}^n f_k(u)g_k(u)$$

et

$$D_u \langle f(u), g(u) \rangle = D_u \left(\sum_{k=1}^n f_k(u)g_k(u) \right) = \sum_{k=1}^n (D_u f_k(u)g_k(u) + f_k(u)(D_u g_k(u))) \\ = \langle D_u f(u), g(u) \rangle + \langle f(u), D_u g(u) \rangle,$$

quel que soit $u \in \Omega$.

La conséquence est évidente, si f et g ont un produit scalaire constant, on a $D_u \langle f(u), g(u) \rangle = 0$. De même si f est de norme constante, alors $\langle f(u), f(u) \rangle$ est constant sur Ω , puisque c'est le carré de la norme de $f(u)$.

Pour les deux autres relations, on utilise la proposition 3.4.12 dans une base orthonormée positive (b_1, b_2, b_3) . Si dans cette base $f(u) : (f_1(u), f_2(u), f_3(u))^\sim$, $g(u) : (g_1(u), g_2(u), g_3(u))^\sim$ et $h(u) : (h_1(u), h_2(u), h_3(u))^\sim$, on a

$$[f(u), g(u), h(u)] = \sum_{i,j,k=1}^3 \varepsilon_{i,j,k} f_i(u)g_j(u)h_k(u)$$

et

$$f(u) \wedge g(u) = \sum_{i,j,k=1}^3 \varepsilon_{i,j,k} f_i(u)g_j(u)b_k.$$

On dérive ces deux expressions et on arrive au résultat annoncé, comme plus haut. \square

Il est important d'avoir une règle pour la dérivation des fonctions composées. En effet, si je décris la troisième trajectoire de l'exemple introductif par ses coordonnées :

$$C(t) : (8 + \sin(t), 2 \cos(t))^\sim, \quad t \in]0, 2\pi[$$

alors je peux décider d'exprimer t en fonction d'un autre "paramètre". Par exemple en "posant" $t = 2u$, j'obtiens une autre description

$$D(u) = C(2u) : (8 + \sin(2u), 2 \cos(2u))^\sim, \quad u \in \Omega' =]0, \pi[.$$

Je peux aussi poser $t = \ln(u)$, puisque la fonction \ln définit une bijection entre $]1, e^{2\pi}[$ et $]0, 2\pi[$. Je décrirai la même trajectoire en définissant

$$E(u) = C(\ln(u)) : (8 + \sin(\ln(u)), 2 \cos(\ln(u)))^\sim, \quad u \in \Omega' =]1, e^{2\pi}[.$$

Proposition 5.1.5. *Si P est une fonction de $\Omega \subset \mathbb{R}$ dans \mathcal{A} ou E , dérivable, et si $\varphi : \Omega' \rightarrow \Omega$ est dérivable, alors $P \circ \varphi$ est dérivable et on a, pour tout $u'_0 \in \Omega'$:*

$$\overrightarrow{D_{u'} P \circ \varphi}(u'_0) = \overrightarrow{D_u P}(u)|_{u=\varphi(u'_0)} D_{u'} \varphi(u'_0).$$

Démonstration. La preuve n'est pas surprenante : on écrit P en coordonnées dans un repère et on calcule les deux expressions : si $P(u) : (x_1(u), \dots, x_n(u))^\sim$, alors $P \circ \varphi(u') : (x_1 \circ \varphi(u'), \dots, x_n \circ \varphi(u'))^\sim$. On dérive cette expression, en u'_0 , composante à composante, en utilisant à chaque fois le théorème de dérivation des fonctions composées :

$$\begin{aligned} \overrightarrow{D_{u'} P \circ \varphi}(u'_0) : \begin{pmatrix} (D_{u'}(x_1 \circ \varphi))(u'_0) \\ \vdots \\ (D_{u'}(x_n \circ \varphi))(u'_0) \end{pmatrix} &= \begin{pmatrix} D_u x_1(u)|_{u=\varphi(u'_0)} D_{u'} \varphi(u'_0) \\ \vdots \\ D_u x_n(u)|_{u=\varphi(u'_0)} D_{u'} \varphi(u'_0) \end{pmatrix} \\ &= D_{u'} \varphi(u'_0) \begin{pmatrix} D_u x_1(u)|_{u=\varphi(u'_0)} \\ \vdots \\ D_u x_n(u)|_{u=\varphi(u'_0)} \end{pmatrix}, \end{aligned}$$

qui donne bien le résultat annoncé. □

Remarque 5.4. Bien que l'on indique généralement le multiple scalaire à gauche du vecteur, j'ai laissé l'énoncé dans cet ordre pour que l'on voie directement qu'il généralise le théorème de dérivation des fonctions de fonctions (ou fonctions composées) classique.

5.2 Arc réguliers de courbes : notions élémentaires

Nous pouvons maintenant passer à la théorie des courbes proprement dites. Nous allons définir ce qu'est un paramétrage d'un arc régulier de courbe, la notion de paramétrages équivalents, la tangente en un point d'un tel arc, l'orientation et la longueur d'arc, enfin, nous introduirons un type de paramétrages qui permet de parcourir tout arc régulier de courbe "à vitesse constante", à savoir les paramétrages naturels.

5.2.1 Paramétrages : définitions et exemples

Définition 5.2.1. Un paramétrage de classe C_p d'un arc régulier de courbe est un couple (Ω, P) où Ω est un intervalle ouvert de \mathbb{R} et $P : \Omega \rightarrow \mathcal{A}$ est une application de classe C_p telle que $\overrightarrow{D_u P}(u) \neq 0$ pour tout $u \in \Omega$. Un *arc régulier de courbe* est l'image d'un paramétrage.

La condition de non annulation de la dérivée dans la définition est importante. Elle va entre autres nous permettre de définir la tangente.

Exemple 5.2.2. Voici trois paramétrages de classe C_∞ d'un demi-cercle :

1. $\gamma_1 :]0, \pi[\rightarrow \mathcal{A} : \alpha \mapsto (R \cos(\alpha), R \sin(\alpha))$;
2. $\gamma_2 :]-R, R[\rightarrow \mathcal{A} : x \mapsto (x, \sqrt{R^2 - x^2})$;
3. $\gamma_3 :]-\frac{\pi}{2}, \frac{\pi}{2}[\rightarrow \mathcal{A} : \theta \mapsto (R \sin(\theta), R \cos(\theta))$.

Je vous laisse vérifier que la dérivée de ces fonctions ne s'annule pas. Voici encore un exemple fondamental, qui montre que la théorie des courbes plane généralise l'étude des fonctions dérivables. De manière générale, si f est une fonction de classe C_p sur un intervalle Ω de \mathbb{R} , alors on peut définir, dans un repère quelconque d'un plan affine \mathcal{A}

$$\gamma : \Omega \rightarrow \mathcal{A} : x \mapsto \gamma(x) : (x, f(x)).$$

La dérivée vaut alors $\overrightarrow{D_x \gamma}(x) : (1, D_x f'(x))$, et ne s'annule visiblement pas.

Par contre

$$B(t) : (4 + (1 + \cos(t)) \cos(t), (1 + \cos(t)) \sin(t))^\sim, t \in]0, 2\pi[$$

ne définit pas un paramétrage car

$$\overrightarrow{D_t B}(t) : \begin{pmatrix} -\sin(t)(1 + 2\cos(t)) \\ \cos(t) + \cos(2t) \end{pmatrix}$$

s'annule en $t = \pi$. On a quand même deux paramétrages de deux arcs réguliers définis par la même fonction, mais sur les intervalles $]0, \pi[$ et $]\pi, 2\pi[$. Dans le cas présent, en utilisant les propriétés des fonctions sinus et cosinus, on peut également recoller les morceaux et considérer la fonction B sur l'intervalle $]-\pi, \pi[$. On obtient ainsi un paramétrage de classe C_∞ d'un arc régulier de courbe.

5.2.2 Paramétrages équivalents

Nous avons déjà rencontré cette notion : je suis en train de considérer avec intérêt le paramétrage

$$C(t) : (8 + \sin(t), 2\cos(t))^\sim, t \in \Omega :]0, 2\pi[$$

quand mon voisin Raoul vient m'embêter avec le paramétrage qu'il trouve fascinant :

$$E(u) : (8 + \sin(\ln(u)), 2\cos(\ln(u)))^\sim, u \in \Omega' =]1, e^{2\pi}[$$

Après quelques secondes de réflexion et ayant constaté que nous avons dessiné le même arc régulier, j'essaie de le convaincre qu'il m'a plagié : il a juste posé $t = \ln(u)$ et remplacé t en fonction de u dans mon paramétrage. C'est alors logique qu'il obtienne le même arc que moi. Mais il n'en démord pas : il dit que je l'ai copié et que j'ai posé $u = e^t$ et remplacé u par cette valeur dans son paramétrage.

Nous avons en fait défini deux paramétrages équivalents du même arc régulier de courbe, puisque le logarithme et l'exponentielle définissent un *changement de variables* de classe C_∞ entre Ω et Ω' . Voici la définition exacte.

Définition 5.2.3. Deux paramétrages de classe C_p (Ω, P) et (Ω', Q) d'un arc régulier de courbe sont équivalents s'il existe un changement de variables (de classe C_p) $\varphi : \Omega' \rightarrow \Omega$ tel que

$$Q(t) = P \circ \varphi(t) = P(\varphi(t)) \quad \forall t \in \Omega'.$$

Remarque 5.5. Si aucune confusion n'est possible, il se peut qu'on ne change pas la dénomination du paramétrage : on écrit $P(u)$ et $P(t)$. Il est alors entendu qu'on exprime u en fonction de t en "posant" $u = \varphi(t)$. On voit aussi la notation

$$P = P(u), \quad u = u(u') \quad \text{donc} \quad P = P(u') = P(u(u')).$$

Exemple 5.2.4. Tous les paramétrages du demi-cercle que nous avons rencontrés plus haut sont équivalents : on peut voir facilement la correspondance entre deux valeurs du paramètre qui définissent le même point de l'arc :

1. On a $\gamma_1(\alpha) = \gamma_2(R \cos(\alpha))$ et donc $\gamma_2(x) = \gamma_1(\arccos(\frac{x}{R}))$. La correspondance est donc donnée par

$$x = \varphi(\alpha) = R \cos(\alpha) \quad \text{et} \quad \alpha = \arccos\left(\frac{x}{R}\right) = \varphi^{-1}(x).$$

La fonction φ définit bien un changement de variables entre les intervalles considérés.

2. De même, on a $\gamma_1(\alpha) = \gamma_3(\frac{\pi}{2} - \alpha)$ et $\gamma_3(\theta) = \gamma_1(\frac{\pi}{2} - \theta)$;
3. Et par transitivité^a $\gamma_2(x) = \gamma_3(\arcsin(\frac{x}{R}))$; et $\gamma_3(\theta) = \gamma_2(R \sin(\theta))$.

On peut vérifier que toutes les fonctions permettant de passer d'un paramètre à l'autre sont des changements de variables entre les intervalles considérés.

Pour associer un objet géométrique à l'arc régulier de courbe défini par une classe de paramétrages équivalents, il faut vérifier que cet objet ne dépend pas du paramétrage choisi pour le définir, ou du moins étudier sa dépendance au paramétrage.

a. La relation "être équivalent" est une relation symétrique et transitive dans l'ensemble des paramétrages.

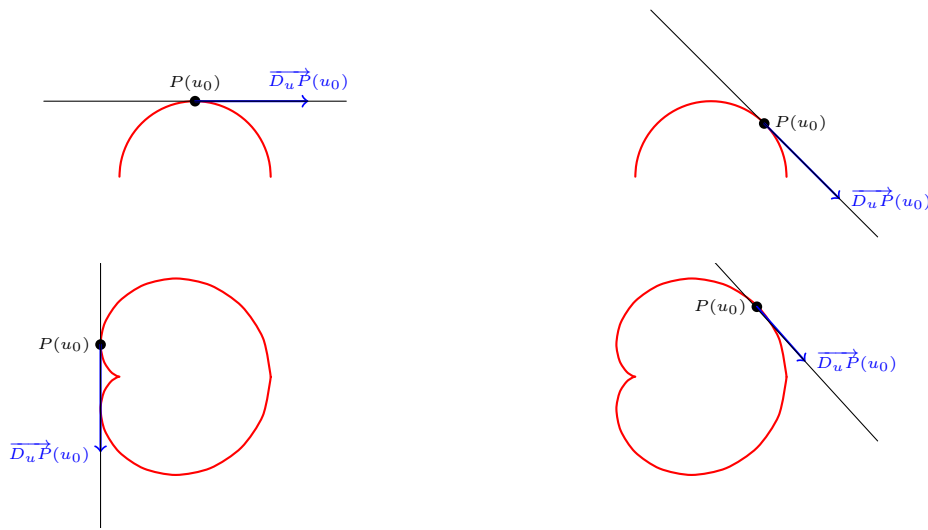
5.2.3 Tangente en un point d'un arc régulier

Nous connaissons tous la définition de la tangente en un point $(x_0, f(x_0))$ du graphe d'une fonction dérivable. C'est la droite passant par ce point et qui a pour pente $f'(x_0)$. En d'autres termes, cette droite admet comme vecteur directeur $(1, f'(x_0))$. C'est la valeur en x_0 de la dérivée du paramétrage. La définition géométrique de la dérivée du paramétrage suggère de généraliser la définition de la façon suivante.

Définition 5.2.5. La tangente en un point $P(u_0)$ d'un arc régulier de courbe Γ ayant pour paramétrage (Ω, P) est la droite passant par $P(u_0)$ et de vecteur directeur $\overrightarrow{D_u P}(u_0)$.

Remarque 5.6. D'un point de vue intuitif, $\overrightarrow{D_u P}(u_0)$ représente le vecteur vitesse instantané en u_0 . Si un point de l'arc correspond à plusieurs valeurs du paramètre (on dit alors que c'est un point multiple), il y a plusieurs tangentes en ce point, une pour chaque valeur du paramètre.

Voici quelques exemples.



Puisqu'on a les coordonnées d'un point et les composantes d'un vecteur directeur, on peut bien sûr calculer une équation cartésienne dans le repère qu'on s'est donné.

Exemple 5.2.6. Soit un paramétrage du cercle donné par $P(\theta) : (R \cos(\theta), R \sin(\theta))^\sim, \theta \in]0, 2\pi[$. Calculons une équation de la tangente en $P(\frac{\pi}{4})$. On a

$$P\left(\frac{\pi}{4}\right) : \left(\frac{R\sqrt{2}}{2}, \frac{R\sqrt{2}}{2}\right)^\sim, \quad \text{et} \quad \overrightarrow{D_u P}\left(\frac{\pi}{4}\right) : \left(-\frac{R\sqrt{2}}{2}, \frac{R\sqrt{2}}{2}\right)^\sim$$

On peut donc prendre $(-1, 1)$ comme vecteur directeur et la tangente a donc pour équation

$$\frac{x - \frac{R\sqrt{2}}{2}}{-1} = \frac{y - \frac{R\sqrt{2}}{2}}{1}$$

ou encore

$$x + y = R\sqrt{2}.$$

Avant de calculer une équation, il est important d'étudier la dépendance de cette définition au paramétrage.

Proposition 5.2.7. La définition de la tangente en un point d'un arc régulier de courbe est indépendante du paramétrage.

Démonstration. Soient (Ω, P) et (Ω', Q) deux paramétrages équivalents d'un même arc régulier de courbe. On sait qu'il existe un changement de variables φ de Ω' dans Ω tel que $P = Q \circ \varphi$. On se place en un point $P(u_0)$ de l'arc régulier. Ce point s'écrit aussi $Q(u'_0) = P(\varphi(u'_0))$, si

$u_0 = \varphi(u'_0)$. Dans le paramétrage P , la tangente contient $P(u_0)$ et a pour vecteur directeur $\overrightarrow{D_u P}(u_0)$. Dans le paramétrage Q , la tangente au même point contient $Q(u'_0) = P(u_0)$ et a pour vecteur directeur $\overrightarrow{D_{u'} Q}(u'_0)$. Mais on a, par la proposition 5.1.5 :

$$\overrightarrow{D_{u'} Q}(u'_0) = \overrightarrow{D_{u'} P \circ \varphi}(u'_0) = \overrightarrow{D_u P}(u)|_{u=\varphi(u'_0)} D_{u'} \varphi(u'_0) = \overrightarrow{D_u P}(u_0) D_{u'} \varphi(u'_0).$$

Les vecteurs directeurs sont multiples l'un de l'autre par un multiple non nul^b. Les droites sont donc égales. \square

5.2.4 Orientation

Les exemples de paramétrage du cercle montrent que les paramétrages définissent naturellement un sens de parcours de l'arc régulier. Puisque l'intervalle sur lequel la fonction φ est définie porte un ordre naturel (celui des réels), on le transporte sur l'arc de courbe.

Définition 5.2.8. Si (Ω, P) est un paramétrage, alors $P(u_1)$ est **avant** $P(u_2)$ si $u_1 < u_2$.

Si mon voisin considère le même arc de courbe mais avec un paramétrage (Ω', Q) équivalent au mien, il aura le même sens de parcours ou le sens opposé. Si (Ω', Q) définit le même sens de parcours et si $P(u_1) = Q(u'_1)$ et $P(u_2) = Q(u'_2)$, si la condition $u_1 < u_2$ est équivalente à la condition $u'_1 < u'_2$. Si φ est le changement de variables liant P à Q , on a $u_1 = \varphi(u'_1)$ et $u_2 = \varphi(u'_2)$. La condition précédente est satisfaite pour tout couple de points u'_1, u'_2 si φ est strictement croissante, ou si sa dérivée est positive. Cela motive la définition suivante^c.

Définition 5.2.9. Un paramétrage (Ω, P) et un paramétrage équivalent $(\Omega', Q = P \circ \varphi)$ définissent la même orientation si $D_v \varphi(v) > 0$ pour tout $v \in \Omega'$. Dans le cas contraire, ils définissent une orientation différente.

Exemple 5.2.10. Les paramétrages γ_1 et γ_2 du demi-cercle sont équivalents, comme nous l'avons déjà vu : on a $\gamma_1(\alpha) = \gamma_2(R \cos(\alpha))$. Le changement de variables est défini par

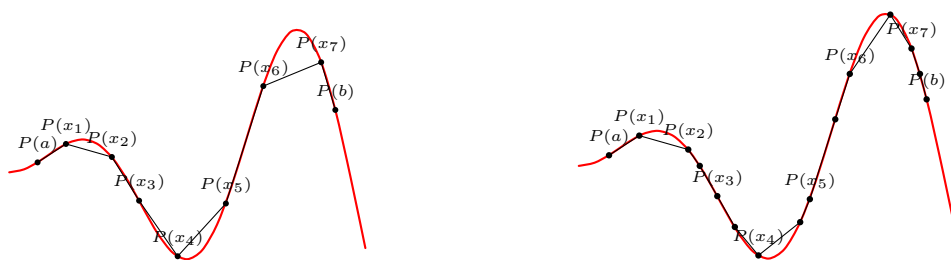
$$\varphi(\alpha) = R \cos(\alpha)$$

est tel que $D_\alpha \varphi(\alpha) = -R \sin(\alpha)$ est négatif sur $]0, \pi[$. Ces deux paramétrages ne définissent donc pas la même orientation sur le demi-cercle.

Vu le théorème de dérivation des fonctions de fonctions, on peut démontrer qu'il n'y a que deux orientations possibles : si mon voisin et moi avons des paramétrages ayant une orientation différente, et si le frère de mon voisin arrive avec un troisième paramétrage équivalent, il sera de la même orientation avec le mien, ou de la même orientation avec celui de mon voisin.

5.2.5 Longueur d'arc et paramétrage naturel

Quand on a un arc régulier de courbe dans un espace affine euclidien, il est assez naturel de vouloir calculer sa longueur. On peut faire la même chose pour la partie de l'arc comprise entre deux points $P(a)$ et $P(b)$. L'idée est d'approcher la longueur par la somme de longueurs de segments consécutifs dont les extrémités sont sur l'arc régulier. En voici une représentation graphique :



b. puisque φ est un changement de variables, sa dérivée ne s'annule pas.

c. Cette définition pourrait être une proposition, mais je ne l'ai pas présentée comme telle au cours.

Bien sûr, si on ajoute des points au découpage du segment $[a, b]$ (ou $[b, a]$ si $b < a$), on augmente la somme des longueurs des segments ainsi construits. La longueur d'arc correspond alors naturellement à la borne supérieure de ces sommes de longueurs, quand elle existe.

Définition 5.2.11. La longueur de l'arc de courbe déterminé par (Ω, P) entre $P(a)$ et $P(b)$ (si $a < b$) est le nombre

$$\sup_{D=[a=x_0, x_1 < \dots < x_n < b=x_{n+1}]} \sum_{i=1}^{n+1} |P(x_i) - P(x_{i-1})|$$

quand cette borne supérieure existe (auquel cas l'arc de courbe est dit rectifiable).

Le cas particulier des graphes de fonctions dans le plan est traité dans le cours d'analyse, comme application de l'intégrale de Riemann. Nous admettrons le résultat général suivant.

Proposition 5.2.12. Si (Ω, P) est un paramétrage de classe C_1 et si $a < b$ sont dans Ω , alors la longueur d'arc entre $P(a)$ et $P(b)$ vaut

$$\int_a^b |\overrightarrow{D_v P}(v)| dv.$$

Remarque 5.7. Si $b < a$ et dans les mêmes conditions, on calcule la longueur d'arc entre les deux points $P(a)$ et $P(b)$ par la formule

$$\int_b^a |\overrightarrow{D_v P}(v)| dv.$$

C'est facile de se le rappeler : la longueur doit être positive.

La notion de longueur d'arc permet de repérer un point sur un arc régulier de courbe : si on fixe $P(u_0)$ comme point de repère, on repère le point $P(u)$ par la longueur d'arc entre $P(u_0)$ et $P(u)$ si $u > u_0$ (c'est à dire si $P(u)$ est après $P(u_0)$ dans l'orientation définie par P), et l'opposé de cette longueur sinon.

Définition 5.2.13. Si (Ω, P) est un paramétrage de classe C_1 et u_0 in Ω . L'abscisse curviligne $s(u)$ de $P(u)$ est la longueur d'arc entre $P(u_0)$ et $P(u)$ si $u_0 < u$ et l'opposé de cette longueur d'arc si $u < u_0$.

Puisqu'on a une expression explicite de la longueur d'arc, on peut facilement calculer l'abscisse curviligne d'un point quelconque.

Proposition 5.2.14. Si (Ω, P) est un paramétrage de classe C_1 , alors l'abscisse curviligne de $P(u)$ vaut

$$s(u) = \int_{u_0}^u |\overrightarrow{D_v P}(v)| dv.$$

Démonstration. Si $u_0 < u$, $s(u)$ est la longueur d'arc entre $P(u_0)$ et $P(u)$ et on a son expression par la proposition 5.2.12. Si $u < u_0$, alors $s(u)$ est l'opposé de la longueur d'arc, c'est à dire

$$s(u) = - \int_u^{u_0} |\overrightarrow{D_v P}(v)| dv = \int_{u_0}^u |\overrightarrow{D_v P}(v)| dv,$$

ce qui termine la preuve. □

Etant donné un paramétrage P , on a défini l'abscisse curviligne associée à tout point $P(u)$. Il est utile de voir que cette fonction, qui à tout u associe $s(u)$, est un changement de variables.

Proposition 5.2.15. L'abscisse curviligne est une fonction à dérivée strictement positive :

$$D_u s(u) = |\overrightarrow{D_u P}(u)| > 0 \quad \forall u \in \Omega.$$

Elle définit donc un changement de variables $s : \Omega \rightarrow s(\Omega)$.

Le changement de variables inverse, noté $u = u(s)$ est tel que

$$D_s u(s) = \frac{1}{|\overrightarrow{D_u P}(u)|} \Big|_{u=u(s)}.$$

Démonstration. Le premier point est une conséquence du théorème fondamental du calcul intégral : $\int_{u_0}^u |\overrightarrow{D_v P}(v)| dv$ est l'unique primitive de $|\overrightarrow{D_u P}(u)|$ qui s'annule en u_0 . La dérivée de $s(u)$ est donc strictement positive et s définit un changement de variables (strictement croissant), par le théorème de la fonction inverse. Le deuxième point découle également du même théorème. \square

5.2.6 Paramétrages naturels

L'abscisse curviligne permet de définir un type de paramétrages particuliers, dit paramétrages naturels. L'idée intuitive est de parcourir l'arc régulier de courbe à vitesse constante, unitaire. La vitesse étant naturellement représentée par la norme de la dérivée du paramétrage, on arrive donc à la définition suivante.

Définition 5.2.16. Un paramétrage (Ω, P) est naturel si on a $|\overrightarrow{D_u P}(u)| = 1$, pour tout $u \in \Omega$.

Encore faut-il, pour que la notion soit utile, qu'il existe suffisamment de paramétrages naturels. Etant donné un arc régulier de courbe, donné par un paramétrage, on peut toujours lui associer un paramétrage naturel, comme le montre le résultat suivant.

Proposition 5.2.17. *Tout paramétrage (Ω, P) est équivalent à un paramétrage naturel, donné par l'abscisse curviligne.*

Démonstration. Soit (Ω, P) un paramétrage. On sait que l'abscisse curviligne définit un changement de variables, que nous notons $s(u)$. Si on note l'inverse $u = u(s)$, on obtient un paramétrage défini par $Q(s) = P(u(s))$. On a alors, par la proposition 5.1.5

$$\overrightarrow{D_s Q}(s) = \overrightarrow{D_u P}(u)|_{u=u(s)} D_s u(s).$$

On sait également par la proposition 5.2.15 que

$$D_s u(s) = \frac{1}{|\overrightarrow{D_u P}(u)|} |_{u=u(s)}.$$

On a donc

$$\overrightarrow{D_s Q}(s) = \frac{\overrightarrow{D_u P}(u)|_{u=u(s)}}{|\overrightarrow{D_u P}(u)|_{u=u(s)}}.$$

C'est visiblement un vecteur de norme unitaire. \square

On peut remarquer que le paramétrage naturel associé à (Ω, P) n'est pas unique : si on change de point de repère pour calculer l'abscisse curviligne, on lui ajoute une constante. Si on change d'orientation, l'abscisse curviligne change de signe. On peut démontrer que ce sont les seuls paramétrages naturels que l'on peut associer à (Ω, P) .

Exemple 5.2.18.

$$P(u) = (R \cos(u), R \sin(u)), \quad s \in]-\pi, \pi[$$

n'est pas naturel mais le paramétrage

$$\gamma(s) = \left(R \cos\left(\frac{s}{R}\right), R \sin\left(\frac{s}{R}\right) \right), \quad s \in]-\pi R, \pi R[,$$

est naturel.

5.3 Le trièdre de Frenet

Dans cette section, nous allons définir trois vecteurs normés associés à chaque point d'un arc régulier de courbe. Pour que cela puisse se faire et pour que cela ait une utilité, nous nous plaçons dans un espace affine euclidien orienté de dimension 3 (bref, la généralisation de l'espace usuel dans lequel nous vivons).

Le premier vecteur est simple à définir, c'est le vecteur tangent unitaire.

Définition 5.3.1. Soit (Ω, P) un paramétrage d'un arc régulier de courbe et $u_0 \in \Omega$. Le vecteur tangent unitaire en $P(u_0)$ est

$$\mathbf{t} = \frac{\overrightarrow{D_u P}(u_0)}{|\overrightarrow{D_u P}(u_0)|}.$$

Remarque 5.8. On devrait écrire $\mathbf{t}(u_0)$, pour signifier que ce vecteur dépend du point en lequel on se place. Cela alourdirait les notations.

Il faut bien sûr étudier la dépendance au paramétrage de ce vecteur. On voit sur les représentations qu'il y a deux vecteurs unitaires possibles. Chacun est associé à un sens de parcours de l'arc régulier de courbe, c'est à dire à une orientation.

Proposition 5.3.2. Si (Ω, P) et (Ω', Q) définissent des paramétrages équivalents, alors les vecteurs tangents en $P(u_0) = Q(u'_0)$ définis par ces deux paramétrages sont égaux ou opposés. Ils sont égaux si et seulement si les paramétrages définissent la même orientation.

Démonstration. Si Q est équivalent à P , il y a un changement de variables $\varphi : \Omega' \rightarrow \Omega$ tel que $Q(u') = P(\varphi(u'))$. On a $u_0 = \varphi(u'_0)$, et on utilise la formule habituelle :

$$\overrightarrow{D_{u'} Q}(u'_0) = \overrightarrow{D_{u'}(P \circ \varphi)}(u'_0) = \overrightarrow{D_u P}(u_0) D_{u'} \varphi(u'_0).$$

On divise alors ce vecteur par sa norme pour obtenir

$$\frac{\overrightarrow{D_{u'} Q}(u'_0)}{|\overrightarrow{D_{u'} Q}(u'_0)|} = \frac{\overrightarrow{D_u P}(u_0)}{|\overrightarrow{D_u P}(u_0)|} \frac{D_{u'} \varphi(u'_0)}{|D_{u'} \varphi(u'_0)|}$$

On constate que les deux vecteurs tangents unitaires sont multiples l'un de l'autre par un facteur qui vaut 1 quand $D_{u'} \varphi(u'_0) > 0$ et -1 dans le cas contraire. \square

Le vecteur tangent unitaire est facile à calculer quand on est dans un paramétrage naturel.

Proposition 5.3.3. Si on a un paramétrage naturel $(\Omega, s \mapsto P(s))$, alors on a

$$\mathbf{t}(s) = \dot{P}(s) = \overrightarrow{D_s P}(s).$$

Démonstration. Dans le cas d'un paramétrage naturel, le vecteur dérivé $\overrightarrow{D_s P}(s)$ est déjà normé. Il n'y a donc pas besoin de le diviser par sa norme. \square

Remarque 5.9. La notation utilisée dans le résultat précédent est générale : étant donnée une fonction f à valeurs dans un espace vectoriel, on note $f'(u)$ le vecteur dérivé $D_u f(u)$. Quand on est dans un paramétrage naturel, on insiste sur ce fait en notant la dérivée par un point. On note donc $\dot{f}(s)$ le vecteur $D_s f(s)$. On garde aussi souvent la lettre s pour le paramètre naturel.

Si on veut maintenant analyser la variation de \mathbf{t} en fonction de s , on est amené à calculer sa dérivée, comme d'habitude. Il est important de remarquer que la norme de \mathbf{t} étant constante, la dérivée de \mathbf{t} mesure sa variation "en direction".

Proposition 5.3.4. Le vecteur $\dot{\mathbf{t}}$ est orthogonal à \mathbf{t} .

Démonstration. C'est une application de la proposition 5.1.4. \square

Cette proposition nous permet d'introduire un autre vecteur important, le vecteur normal principal.

Définition 5.3.5. Soit (Ω, P) un paramétrage naturel d'un arc régulier de courbe. La courbure de l'arc régulier de courbe au point $P(s)$ est le nombre

$$\kappa(s) = |\dot{\mathbf{t}}(s)|.$$

Si $\kappa(s) \neq 0$, alors le vecteur **normal principal** en $P(s)$ est

$$\mathbf{n}(s) = \frac{\dot{\mathbf{t}}(s)}{|\dot{\mathbf{t}}(s)|}.$$

Remarquons que si le vecteur tangent unitaire change de signe si on change de paramétrage naturel en changeant d'orientation, ce n'est pas le cas du vecteur normal principal. Quand $\kappa(s) \neq 0$, on définit le rayon de courbure au point $P(s)$ par $R(s) = \frac{1}{\kappa(s)}$.

Ayant deux vecteurs orthogonaux, il est naturel d'en définir un troisième.

Définition 5.3.6. Soit (Ω, P) un paramétrage naturel d'un arc régulier de courbe. Le vecteur binormal, encore appelé la binormale au point $P(s)$ est le vecteur

$$\mathbf{b} = \mathbf{t} \wedge \mathbf{n}.$$

Remarque 5.10. Dans certains ouvrages on définit $\mathbf{b} = \mathbf{n} \wedge \mathbf{t}$. Il en résulte des changements de signe dans certaines formules.

Le vecteur \mathbf{b} est aussi un vecteur normé. Ces trois vecteurs et le point en lequel ils sont calculés permettent de définir trois plans. Ces trois plans forment le trièdre de Frenet.

Définition 5.3.7. 1. Le *plan osculateur* au point $P(s)$ est déterminé par le point $P(s)$ et les vecteurs $\mathbf{t}(s)$ et $\mathbf{n}(s)$;

2. Le *plan normal* au point $P(s)$ est déterminé par le point $P(s)$ et les vecteurs $\mathbf{n}(s)$ et $\mathbf{b}(s)$;

3. Le *plan rectifiant* au point $P(s)$ est déterminé par le point $P(s)$ et les vecteurs $\mathbf{t}(s)$ et $\mathbf{b}(s)$.

Nous allons maintenant établir les formules de Frenet. Ces formules expriment la dérivée des vecteurs \mathbf{t} , \mathbf{n} et \mathbf{b} dans la base orthonormée positive $(\mathbf{t}, \mathbf{n}, \mathbf{b})$. C'est important car la variation de ces vecteurs permet d'étudier la variation des trois plans définis plus haut.

Pour ce faire, nous devons encore introduire un paramètre, à savoir la torsion.

Définition 5.3.8. Soit (Ω, P) un paramétrage naturel d'un arc régulier de courbe. La torsion en un point $P(s)$ tel que $\kappa(s) \neq 0$ est définie par

$$\tau(s) = \langle \dot{\mathbf{n}}(s), \mathbf{b}(s) \rangle.$$

Théorème 5.3.9 (Formules de Frenet). Soit (Ω, P) un paramétrage naturel d'un arc régulier de courbe. Les dérivées des vecteurs \mathbf{t} , \mathbf{n} , \mathbf{b} en un point $P(s)$ en lequel la courbure n'est pas nulle sont données par les formules suivantes.

$$\begin{cases} \dot{\mathbf{t}} &= \kappa \mathbf{n}; \\ \dot{\mathbf{n}} &= -\kappa \mathbf{t} + \tau \mathbf{b}; \\ \dot{\mathbf{b}} &= -\tau \mathbf{n}. \end{cases}$$

Démonstration. La première formule est la conséquence de la définition de \mathbf{n} . Pour la deuxième et la troisième, on décompose $\dot{\mathbf{n}}$ et $\dot{\mathbf{b}}$ dans la base orthonormée $(\mathbf{t}, \mathbf{n}, \mathbf{b})$. On a ainsi

$$\dot{\mathbf{n}} = \langle \dot{\mathbf{n}}, \mathbf{t} \rangle \mathbf{t} + \langle \dot{\mathbf{n}}, \mathbf{n} \rangle \mathbf{n} + \langle \dot{\mathbf{n}}, \mathbf{b} \rangle \mathbf{b}.$$

On applique alors la proposition 5.1.4 : puisque \mathbf{n} est normé, on a $\langle \dot{\mathbf{n}}, \mathbf{n} \rangle = 0$. D'autre part, par définition on a $\langle \dot{\mathbf{n}}(s), \mathbf{b}(s) \rangle = \tau$. Enfin, puisque $\langle \mathbf{n}(s), \mathbf{t}(s) \rangle = 0$ pour tout s , on a

$$\langle \dot{\mathbf{n}}, \mathbf{t} \rangle = -\langle \mathbf{n}, \dot{\mathbf{t}} \rangle = -\langle \mathbf{n}, \kappa \mathbf{n} \rangle = -\kappa.$$

De la même manière, on obtient

$$\dot{\mathbf{b}} = \langle \dot{\mathbf{b}}, \mathbf{t} \rangle \mathbf{t} + \langle \dot{\mathbf{b}}, \mathbf{n} \rangle \mathbf{n} + \langle \dot{\mathbf{b}}, \mathbf{b} \rangle \mathbf{b},$$

avec

$$\begin{cases} \langle \dot{\mathbf{b}}, \mathbf{t} \rangle = -\langle \mathbf{b}, \dot{\mathbf{t}} \rangle = -\langle \mathbf{b}, \kappa \mathbf{n} \rangle = 0; \\ \langle \dot{\mathbf{b}}, \mathbf{n} \rangle = -\langle \mathbf{b}, \dot{\mathbf{n}} \rangle = -\langle \mathbf{b}, -\kappa \mathbf{t} + \tau \mathbf{b} \rangle = -\tau; \\ \langle \dot{\mathbf{b}}, \mathbf{b} \rangle = 0, \end{cases}$$

et le résultat est prouvé. □

Enfin, il est utile de pouvoir calculer les vecteurs du trièdre de Frenet dans un paramétrage qui n'est pas naturel. En effet, même si on sait qu'un tel paramétrage existe toujours, il n'est parfois (même souvent) pas facile de calculer l'abscisse curviligne, puis d'inverser le changement de variables qu'elle définit.

Théorème 5.3.10. *Soit Γ un arc régulier de courbe et un paramétrage quelconque (Ω, P) de Γ . On a alors*

- a) $\mathbf{t} = \frac{P'}{|P'|}$;
 b) $\kappa = \frac{|P' \wedge P''|}{|P'|^3}$;
 et, si $\kappa \neq 0$
 c) $\mathbf{b} = \frac{P' \wedge P''}{|P' \wedge P''|}$;
 d) $\mathbf{n} = \frac{P' \wedge P''}{|P' \wedge P''|} \wedge \frac{P'}{|P'|}$;
 e) $\tau = \frac{[P', P'', P''']}{|P' \wedge P''|^3}$.

Avant de nous lancer dans la preuve, il nous faut un résultat technique concernant les dérivées des fonctions définies sur le domaine d'un paramétrage. Si (Ω, P) est un paramétrage quelconque d'un arc régulier de courbe, on a défini l'abscisse curviligne $s : \Omega \mapsto s(\Omega) = \Omega'$. On note $u(s)$ le changement de variables réciproque (inverse). On note par la même lettre le paramétrage établi à partir de Ω' , mais on marque la différence en notant $P(s)$. On peut faire de même avec toute fonction définie sur Ω : si $f : \Omega \rightarrow E$ est une fonction à valeurs dans un ensemble E , espace vectoriel ou affine, alors on pose^d

$$f : \Omega' \rightarrow E : s \mapsto f(s) = f(u(s)).$$

Proposition 5.3.11. *On a la relation fondamentale*

$$\dot{f}(s) = \frac{f'(u)}{|P'(u)|} \quad (5.4)$$

Démonstration. C'est encore une fois la formule donnant la dérivée des fonctions composées (la proposition 5.1.5). On a

$$\dot{f}(s) = D_s f(s) = D_s f(u(s)) = D_u f(u)|_{u=u(s)} D_s(u(s)) = \frac{f'(u)}{|P'(u)|} \Big|_{u=u(s)}.$$

C'est le résultat annoncé, si on admet l'abus de langage consistant à laisser tomber la précision $u = u(s)$ dans la formule. \square

Démonstration du théorème 5.3.10. Pour \mathbf{t} , il s'agit de la définition. Calculons maintenant $\dot{\mathbf{t}}$. En utilisant (5.4) et la règle de dérivée des quotients (composante à composante), on obtient

$$\dot{\mathbf{t}} = \frac{\mathbf{t}'}{|P'|} = \frac{1}{|P'|} \frac{P''|P'| - |P'|'P'}{|P'|^2} = \frac{P''|P'| - |P'|'P'}{|P'|^3}.$$

Deux cas peuvent alors se présenter : soit on est en un point où $\dot{\mathbf{t}} = 0$ (i.e. $\kappa = 0$), soit on est en un point où $\dot{\mathbf{t}} \neq 0$ (i.e. $\kappa \neq 0$). Dans le premier cas, on constate par la formule ci-dessus que P' et P'' sont des vecteurs linéairement dépendants, et donc $|P' \wedge P''| = 0$. La formule annoncée pour κ est donc valide et on ne peut pas aller plus loin, puisque \mathbf{b} , \mathbf{n} et τ ne sont pas définis en ce point.

Dans le deuxième cas, on calcule

$$\mathbf{t} \wedge \dot{\mathbf{t}} = \frac{P'}{|P'|} \wedge \left(\frac{P''|P'| - |P'|'P'}{|P'|^3} \right) = \frac{P' \wedge P''}{|P'|^3}.$$

d. avec un léger abus de notation, puisqu'on utilise le même symbole pour cette nouvelle fonction.

En utilisant la première formule de Frenet, on trouve $\mathbf{t} \wedge \dot{\mathbf{t}} = \kappa \mathbf{b}$, donc on a

$$\kappa \mathbf{b} = \frac{P' \wedge P''}{|P'|^3},$$

qui donne d'une part

$$\kappa = |\kappa \mathbf{b}| = \frac{|P' \wedge P''|}{|P'|^3},$$

et d'autre part

$$\mathbf{b} = \frac{\kappa \mathbf{b}}{\kappa} = \frac{P' \wedge P''}{|P'|^3} \frac{|P'|^3}{|P' \wedge P''|} = \frac{P' \wedge P''}{|P' \wedge P''|}.$$

On a ensuite $\mathbf{n} = \mathbf{b} \wedge \mathbf{t}$ puisque $(\mathbf{t}, \mathbf{n}, \mathbf{b})$ est une base orthonormée positive, ce qui donne la valeur annoncée pour \mathbf{n} .

Il reste à établir la valeur de τ . Par la troisième formule de Frenet, on a $\tau = -\langle \dot{\mathbf{b}}, \mathbf{n} \rangle$. Il est alors naturel de calculer $\dot{\mathbf{b}}$, toujours de la même façon.

$$\dot{\mathbf{b}} = \frac{\mathbf{b}'}{|P'|} = \frac{1}{|P'|} \left(\frac{P' \wedge P''}{|P' \wedge P''|} \right)' = \frac{(P' \wedge P''')|P' \wedge P''| - |P' \wedge P''|'(P' \wedge P'')}{|P'| |P' \wedge P''|^2}.$$

Il est alors utile de remarquer que le dernier terme du numérateur est orthogonal à \mathbf{n} , vu l'expression de \mathbf{n} que nous avons démontrée (ou celle de \mathbf{b}), donc

$$\tau = -\langle \dot{\mathbf{b}}, \mathbf{n} \rangle = -\frac{\langle P' \wedge P''', (P' \wedge P'') \wedge P' \rangle}{|P'|^2 |P' \wedge P''|^2},$$

on conclut alors en utilisant la formule du double produit vectoriel. □

5.4 Normale et courbure signées des courbes planes

Dans cette section, on se place dans un espace affine \mathcal{A} orienté et de dimension 2, muni d'un repère orthonormé positif $(O, (b_1, b_2))$.

On se donne un arc régulier de courbe Γ et un paramétrage quelconque (Ω, P) . Ici encore, on peut au besoin utiliser l'abscisse curviligne pour obtenir un paramétrage naturel.

La définition du vecteur tangent unitaire \mathbf{t} en un point $P(u)$ de l'arc régulier de courbe est la même que celle donnée en général. Cependant, puisque le vecteur normal unitaire doit être normé et orthogonal à \mathbf{t} , il est déterminé au signe près. On adopte en dimension deux une autre convention pour définir le vecteur normal principal. Le vecteur binormal n'a plus de raison d'être en dimension deux et n'est donc pas défini.

Définition 5.4.1. La normale signée \mathbf{n}_2 en un point d'un arc régulier de courbe est l'unique vecteur tel que $(\mathbf{t}, \mathbf{n}_2)$ soit une base orthonormée positive. La courbure signée κ_2 en un point d'un arc régulier de courbe est le nombre tel que $\dot{\mathbf{t}} = \kappa_2 \mathbf{n}_2$.

Puisqu'on est en dimension 2, le complément orthogonal du sous-espace vectoriel engendré par \mathbf{t} est de dimension un, donc \mathbf{n}_2 et $\mathbf{n} = \frac{\dot{\mathbf{t}}}{|\dot{\mathbf{t}}|}$ sont multiples l'un de l'autre (par un facteur de module 1). Puisqu'on a les formules

$$\kappa \mathbf{n} = \dot{\mathbf{t}} = \kappa_2 \mathbf{n}_2,$$

on constate que deux cas (exclusifs) peuvent survenir du moins quand κ n'est pas nul :

1. $\mathbf{n}_2 = \mathbf{n}$, $\kappa_2 = \kappa$, $\kappa_2 > 0$, et \mathbf{n}_2 est orienté dans le sens de la concavité de la courbe (comme \mathbf{n}).
2. $\mathbf{n}_2 = \mathbf{n}$, $\kappa_2 = -\kappa$, $\kappa_2 < 0$, et \mathbf{n}_2 est orienté dans le sens contraire de la concavité de la courbe (comme $-\mathbf{n}$).

Il est donc utile de connaître le signe de κ_2 pour connaître le sens de concavité de la courbe. Le résultat suivante donne les éléments \mathbf{t} , \mathbf{n}_2 et κ_2 dans un paramétrage quelconque.

Proposition 5.4.2. *Pour tout paramétrage (Ω, P) d'un arc régulier de courbe, on a*

$$\mathbf{t} = \frac{P'}{|P'|} = \frac{\begin{pmatrix} x'(u) \\ y'(u) \end{pmatrix}}{|\sqrt{x'(u)^2 + y'(u)^2}|}, \quad \mathbf{n}_2 = \frac{\begin{pmatrix} -y'(u) \\ x'(u) \end{pmatrix}}{|P'(u)|}, \quad \kappa_2 = \frac{\det \begin{pmatrix} x'(u) & x''(u) \\ y'(u) & y''(u) \end{pmatrix}}{|P'(u)|^3}$$

Démonstration. Nous avons déjà donné la preuve pour \mathbf{t} (voir le trièdre de Frenet). Je vous la rappelle : on a

$$\mathbf{t} = \dot{P} = \frac{P'}{|P'|} = \frac{\begin{pmatrix} x'(u) \\ y'(u) \end{pmatrix}}{\sqrt{x'(u)^2 + y'(u)^2}}.$$

Pour \mathbf{n}_2 , on constate que l'expression proposée définit bien un vecteur normé et orthogonal à \mathbf{t} . Il est aussi tel que $(\mathbf{t}, \mathbf{n}_2)$ soit une base positive, puisque

$$\frac{\det \begin{pmatrix} x'(u) & -y'(u) \\ y'(u) & x'(u) \end{pmatrix}}{|\sqrt{x'(u)^2 + y'(u)^2}|} = 1.$$

Pour calculer κ_2 , on calcule encore $\dot{\mathbf{t}}$. Ce calcul a déjà été fait. Je le rappelle :

$$\dot{\mathbf{t}} = \frac{\mathbf{t}'}{|P'|} = \frac{1}{|P'|} \frac{P''|P'| - |P'|'P'}{|P'|^2} = \frac{|P'| \begin{pmatrix} x''(u) \\ y''(u) \end{pmatrix} - |P'|' \begin{pmatrix} x'(u) \\ y'(u) \end{pmatrix}}{|P'|^3}$$

On a toujours $\dot{\mathbf{t}} = \kappa_2 \mathbf{n}_2$ qui donne $\kappa_2 = \langle \dot{\mathbf{t}}, \mathbf{n}_2 \rangle$. Donc en prenant en compte le fait que le dernier terme du numérateur de la fraction ci-dessus est orthogonal à \mathbf{n}_2 , on obtient

$$\kappa_2 = \langle \dot{\mathbf{t}}, \mathbf{n}_2 \rangle = \frac{\left\langle \begin{pmatrix} x''(u) \\ y''(u) \end{pmatrix}, \begin{pmatrix} -y'(u) \\ x'(u) \end{pmatrix} \right\rangle}{|P'|^3}.$$

Cela donne le résultat annoncé, après calcul du produit scalaire. \square

Cette proposition fournit un moyen simple de calculer la courbure κ_2 , et donc de déterminer le sens de la concavité des courbes planes.

Le résultat suivant donne une interprétation de κ_2 comme une mesure de variation de la direction du vecteur \mathbf{t} . Cette variation de la direction est mesurée rigoureusement par la variation de l'angle orienté entre le vecteur de base b_1 et \mathbf{t} .

Proposition 5.4.3. *Si dans un paramétrage naturel (Ω, P) d'un arc de courbe plane, on écrit*

$$\mathbf{t}(s) = \cos(\psi(s))b_1 + \sin(\psi(s))b_2,$$

alors on a $\kappa_2(s) = D_s \psi(s) = \dot{\psi}$.

Démonstration. Il suffit d'utiliser les formules de la proposition précédente pour un paramétrage naturel. On obtient directement

$$\mathbf{t} = \begin{pmatrix} \dot{x}(s) \\ \dot{y}(s) \end{pmatrix}, \quad \mathbf{n}_2 = \begin{pmatrix} -\dot{y}(s) \\ \dot{x}(s) \end{pmatrix}$$

On a donc, selon l'hypothèse

$$\begin{cases} \cos(\psi(s)) = \dot{x}(s) \\ \sin(\psi(s)) = \dot{y}(s). \end{cases}$$

L'identité $\dot{\mathbf{t}} = \kappa_2 \mathbf{n}_2$ s'écrit alors

$$\begin{pmatrix} D_s(\cos(\psi(s))) \\ D_s(\sin(\psi(s))) \end{pmatrix} = \kappa_2 \begin{pmatrix} -\sin(\psi(s)) \\ \cos(\psi(s)) \end{pmatrix},$$

ce qui donne le résultat en calculant les dérivées dans le membre de gauche, vu que le facteur de κ_2 dans le membre de droite ne s'annule pas. \square

Chapitre 6

Appendice : Calcul matriciel

6.1 Éléments de calcul matriciel

Nous avons défini dans le premier chapitre les espaces vectoriels \mathbb{R}_q^p formés des matrices à p lignes et q colonnes. Nous savons donc comment additionner de telles matrices (de même type), et comment multiplier une matrice par un nombre. Nous allons voir ici comment multiplier des matrices entre elles. Le produit matriciel ainsi défini aura l'air quelque peu artificiel, mais on constate que ce produit ainsi défini sert directement dans les changements de base, mais sert également pour représenter des applications linéaires entre espaces vectoriels dans les cours d'algèbre et de physique. La définition s'expliquera alors naturellement en termes de compositions d'opérateurs linéaires.

6.1.1 Sous-matrices

Soit $A \in \mathbb{R}_n^m$. Une sous-matrice de A est la matrice obtenue en sélectionnant certaines rangées de A et en supprimant les autres. Par exemple, en sélectionnant les deux premières colonnes et la première et la troisième ligne d'une matrice à 3 lignes et 3 colonnes, ou en supprimant la troisième colonne et la deuxième ligne, comme ceci

$$A = \begin{pmatrix} 1 & 2 & a \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

on obtient la sous-matrice

$$A_{(1,3;1,2)} = \begin{pmatrix} 1 & 2 \\ 7 & 8 \end{pmatrix}.$$

On note entre parenthèses les lignes et les colonnes qui sont sélectionnées dans A . La sous-matrice n'est pas obligatoirement carrée, puisqu'on a

$$A_{(1;1,2,3)} = \begin{pmatrix} 1 & 2 & a \end{pmatrix}.$$

Un cas particulier important est celui où l'on ne sélectionne qu'une ligne et une colonne.

Définition 6.1.1. Soit $A \in \mathbb{R}_n^m$ une matrice. On note $A_{\hat{i},\hat{j}}$ la sous-matrice obtenue en supprimant la ligne i et la colonne j dans A .

Dans l'exemple précédent, on avait donc $A_{(1,3;1,2)} = A_{\hat{2},\hat{3}}$. De même dans la matrice

$$B = \begin{pmatrix} 0 & 1 & 2 & 0 \\ 3 & 4 & 6 & 1 \\ 0 & 4 & 2 & 1 \end{pmatrix},$$

on a

$$B_{\hat{1},\hat{3}} = \begin{pmatrix} 3 & 4 & 1 \\ 0 & 4 & 1 \end{pmatrix}.$$

6.1.2 Déterminants des matrices carrées : définitions

Nous allons maintenant définir l'application déterminant, qui à toute matrice carrée associe un nombre réel. Nous le ferons par *induction* (ou *récurrence*) sur la taille de la matrice.

Définition 6.1.2. Si $A \in \mathbb{R}_1^1 = \mathbb{R}$, alors $\det(A) = A$. Si $A \in \mathbb{R}_{n+1}^{n+1}$, alors

$$\det(A) = \sum_{j=1}^{n+1} a_{1,j} (-1)^{j+1} \det(A_{\hat{1},\hat{j}}).$$

On pourrait penser que cette définition est circulaire puisqu'elle définit l'application \det à partir d'elle-même. Mais c'est bien une définition par induction, puisque dans le membre de droite de l'égalité ci-dessus, les matrices dont on considère le déterminant sont dans \mathbb{R}_n^n . Avant de considérer des exemples, voici un peu de terminologie.

Définition 6.1.3. 1. Soit $A \in \mathbb{R}_n^m$ une matrice. Un *mineur d'ordre p* de A est un déterminant d'une sous-matrice de A à p lignes et p colonnes ;
 2. Si $A \in \mathbb{R}_m^m$, le mineur de l'élément $a_{i,j}$ est $\det(A_{\hat{i},\hat{j}})$;
 3. Si $A \in \mathbb{R}_m^m$, le cofacteur de l'élément $a_{i,j}$ est $\mathcal{A}_{i,j} = (-1)^{i+j} \det(A_{\hat{i},\hat{j}})$;

Puisque chaque élément d'une matrice carrée peut être associé à son cofacteur, il est naturel de ranger ces cofacteurs dans une matrice de même taille, la matrice des cofacteurs, notée \mathcal{A} ou $\text{cof}(A)$, et définie par

$$(\text{cof}(A))_{i,j} = (\mathcal{A})_{i,j} = (-1)^{i+j} \det(A_{\hat{i},\hat{j}}).$$

La formule que nous avons utilisée pour définir le déterminant s'appelle alors *règle des cofacteurs* (sur la première ligne) et s'écrit alors

$$\det(A) = \sum_{j=1}^{n+1} a_{1,j} \mathcal{A}_{1,j}.$$

Décrivons maintenant explicitement le cas des matrices de taille 2, 3 et 4.

1. Pour une matrice A à deux lignes et deux colonnes, on a

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11} \mathcal{A}_{11} + a_{12} \mathcal{A}_{12} = a_{11} a_{22} - a_{12} a_{21}.$$

Le calcul à entreprendre peut être représenté de manière graphique :

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

2. Pour une matrice à trois lignes et trois colonnes, on a

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= a_{11} \mathcal{A}_{11} + a_{12} \mathcal{A}_{12} + a_{13} \mathcal{A}_{13} \\ &= a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \\ &= a_{11}(a_{22} a_{33} - a_{32} a_{23}) - a_{12}(a_{21} a_{33} - a_{31} a_{23}) + a_{13}(a_{21} a_{32} - a_{31} a_{22}). \end{aligned}$$

Ici aussi le calcul peut être représenté de manière graphique. C'est la règle de Sarrus^a.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

a. Pierre Sarrus (1798-1861).

3. Enfin, pour une matrice à quatre lignes et quatre colonnes, on n'a plus la même interprétation graphique en termes de diagonales de la matrice dont on veut calculer le déterminant, mais on utilise simplement la définition. On a donc

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} &= a_{11}\mathcal{A}_{11} + a_{12}\mathcal{A}_{12} + a_{13}\mathcal{A}_{13} + a_{14}\mathcal{A}_{14} \\ &= a_{11} \det \begin{pmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{pmatrix} - a_{12} \det \begin{pmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{pmatrix} \\ &\quad + a_{13} \det \begin{pmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{pmatrix} - a_{14} \det \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{pmatrix}. \end{aligned}$$

On est ainsi ramené au calcul de quatre déterminants de matrices à trois lignes et trois colonnes.

6.1.3 Déterminants des matrices carrées : propriétés

Les propriétés caractéristiques du déterminant sont données dans la proposition suivante.

Proposition 6.1.4. *Le déterminant est une application multilinéaire sur les colonnes des matrices : on a pour tous $C_1, \dots, C_n, C'_i \in \mathbb{R}^n$, tout réel λ et tout $i \leq n$*

1. $\det(C_1 \dots C_i + C'_i \dots C_n) = \det(C_1 \dots C_i \dots C_n) + \det(C_1 \dots C'_i \dots C_n)$,
2. $\det(C_1 \dots \lambda C_i \dots C_n) = \lambda \det(C_1 \dots C_i \dots C_n)$.

Le déterminant est antisymétrique sur les colonnes : une matrice A et la matrice obtenue en permutant deux colonnes de A ont des déterminants opposés. Enfin, le déterminant de la matrice identité vaut 1.

Pour tout $n \in \mathbb{N}_0$, l'application déterminant est l'unique application définie sur \mathbb{R}_n^n à valeurs réelles et satisfaisant ces trois propriétés.

Démonstration. Nous avons défini le déterminant par induction. Il est donc naturel de démontrer cette proposition par induction. Pour $n = 1$, on a $\det(c_1 + c'_1) = c_1 + c'_1 = \det(c_1) + \det(c'_1)$ et $\det(\lambda c_1) = \lambda c_1 = \lambda \det(c_1)$. L'antisymétrie n'est pas définie puisqu'il n'y a qu'une colonne. On peut cependant montrer directement l'antisymétrie pour $n = 2$.

Supposons que le déterminant vérifie les propriétés pour des matrices de taille n et démontrons que c'est encore vrai pour des matrices de taille $n + 1$. Soient $C_1, \dots, C_{n+1}, C'_i$ des éléments de \mathbb{R}^{n+1} .

Pour tout $x \in \mathbb{R}^{n+1}$, notons \bar{x} l'élément de \mathbb{R}^n obtenu en supprimant la première composante de x . On a alors par définition, pour la matrice $A = (C_1 \dots C_i + C'_i \dots C_{n+1})$:

$$\begin{aligned} \det(A) &= \sum_{j=1, j \neq i}^{n+1} (C_j)_1 (-1)^{j+1} \det(\overline{C_1 \dots \hat{j} \dots C_i + C'_i \dots C_{n+1}}) \\ &\quad + (C_i + C'_i)_1 (-1)^{i+1} \det(\overline{C_1 \dots \hat{i} \dots C_{n+1}}) \end{aligned}$$

On applique alors l'hypothèse d'induction pour tous les termes de la première somme, et le fait que $(C_i + C'_i)_1 = (C_i)_1 + (C'_i)_1$ pour obtenir le résultat demandé.

On procède de la même façon pour démontrer l'identité

$$\det(C_1 \dots \lambda C_i \dots C_{n+1}) = \lambda \det(C_1 \dots C_i \dots C_{n+1}).$$

En ce qui concerne l'antisymétrie, cette propriété est équivalente au fait que le déterminant s'annule pour les matrices ayant deux colonnes identiques. Cette propriété est vraie pour $n = 2$ car on a

$$\det \begin{pmatrix} u & u \\ v & v \end{pmatrix} = uv - uv = 0.$$

Supposons que le déterminant soit antisymétrique pour les matrices de taille n et considérons maintenant une matrice A carrée de taille $n + 1$, ayant deux colonnes identiques, disons la i -ème et la j -ème, avec $i < j$. On a donc $A = (C_1 \dots C_{i-1} \underbrace{X}_i \dots \underbrace{X}_j \dots C_{n+1})$. Quand on applique la règle des cofacteurs, on calcule des déterminants de matrices de taille n ayant deux colonnes égales, sauf quand on considère les cofacteurs des éléments sur les colonnes i et j . On a donc

$$\begin{aligned} \det(A) &= X_1(-1)^{i+1} \det(\overline{C_1} \dots \overline{C_{i-1}} \overline{C_{i+1}} \dots \overline{C_{j-1}} \overline{X} \dots \overline{C_{n+1}}) \\ &+ X_1(-1)^{j+1} \det(\overline{C_1} \dots \overline{C_{i-1}} \overline{X} \overline{C_{i+1}} \dots \overline{C_{j-1}} \dots \overline{C_{n+1}}) \end{aligned}$$

On en déduit que $\det(A) = 0$, en utilisant l'antisymétrie pour les matrices de taille n .

Enfin, on montre directement par récurrence sur n que $\det(I_n) = 1$.

Supposons avoir une application f de \mathbb{R}_1^1 dans \mathbb{R} satisfaisant les trois propriétés de l'énoncé. On a alors $f(a) = af(1) = a = \det(a)$. On peut alors supposer que toute application satisfaisant les conditions de l'énoncé sur \mathbb{R}_n^n est égale au déterminant, et tenter de le montrer pour \mathbb{R}_{n+1}^{n+1} . Soit $A = (C_1, \dots, C_{n+1})$. Décomposons chaque colonne C_j en $a_{1,j}e_1 + C'_j$. On a alors

$$f(A) = f(a_{1,1}e_1 + C'_1, \dots, a_{1,n+1}e_1 + C'_{n+1}).$$

On peut alors développer cette expression en utilisant la multilinéarité de f . Cela donne en principe 2^{n+1} termes. On peut cependant remarquer que les termes contenant au moins deux colonnes multiples de e_1 sont nuls, par antisymétrie. On a donc

$$f(A) = f(C'_1, \dots, C'_{n+1}) + \sum_{j=1}^{n+1} a_{1,j} f(C'_1, \dots, C'_{j-1}, e_1, C'_{j+1}, \dots, C'_{n+1}).$$

Par antisymétrie, on a aussi

$$f(A) = f(C'_1, \dots, C'_{n+1}) + \sum_{j=1}^{n+1} a_{1,j} (-1)^{j+1} f(e_1, C'_1, \dots, C'_{j-1}, C'_{j+1}, \dots, C'_{n+1}).$$

Vu les propriétés de f , l'application qui a $(\overline{C_1}, \dots, \overline{C_{j-1}}, \overline{C_{j+1}}, \dots, \overline{C_{n+1}})$ associe $f(e_1, C'_1, \dots, C'_{j-1}, C'_{j+1}, \dots, C'_{n+1})$ satisfait les trois conditions de l'énoncé, sur des matrices de taille n . On a donc

$$f(e_1, C'_1, \dots, C'_{j-1}, C'_{j+1}, \dots, C'_{n+1}) = \det(\overline{C_1}, \dots, \overline{C_{j-1}}, \overline{C_{j+1}}, \dots, \overline{C_{n+1}}).$$

Les vecteurs C'_1, \dots, C'_{n+1} sont $n + 1$ combinaisons linéaires de e_2, \dots, e_{n+1} . Ils sont donc linéairement dépendants. L'un est combinaison linéaire des autres. Supposons que ce soit C'_{n+1} . On a donc $C'_{n+1} = \sum_{i=1}^n \lambda_i C'_i$. Dans ce cas, on a

$$f(C'_1, \dots, C'_{n+1}) = \sum_{i=1}^n \lambda_i f(C'_1, \dots, C'_n, C'_i) = 0.$$

par antisymétrie, et on peut conclure que $f(A) = \det(A)$ pour toute matrice $A \in \mathbb{R}_{n+1}^{n+1}$. \square

En corollaire de cette proposition, on obtient des propriétés supplémentaires du déterminant.

Corollaire 6.1.5. *Pour tout $A \in \mathbb{R}_n^n$, on a pour tout $i \leq n$*

$$\det(A) = \sum_{j=1}^n a_{i,j} \mathcal{A}_{i,j}$$

et pour tout $j \leq n$

$$\det(A) = \sum_{i=1}^n a_{i,j} A_{i,j}.$$

Ce sont les règles des cofacteurs pour les lignes et les colonnes. On a aussi $\det(A) = \det(A^\sim)$ et $\det(AB) = \det(A)\det(B)$ pour toute matrice $B \in \mathbb{R}_n^n$.

Démonstration. Pour les deux premières identités, on vérifie comme plus haut que ces expressions définissent une application multilinéaire, antisymétrique et qui donne la valeur 1 à la matrice identité. \square

Le déterminant trouve son utilité à travers la proposition suivante.

Proposition 6.1.6. *Les éléments C_1, \dots, C_n de \mathbb{R}^n sont linéairement dépendants si, et seulement si on a $\det(C_1, \dots, C_n) = 0$.*

Démonstration. Pour $n = 1$, le résultat est évident (et sans réelle utilité). Nous supposons dès lors $n \geq 2$. Si C_1, \dots, C_n sont linéairement dépendants, alors un de ces vecteurs est combinaison linéaire des autres. Supposons, pour ne pas alourdir les notations, que ce soit C_n . On a donc $C_n = \sum_{k=1}^{n-1} \lambda_k C_k$, où les coefficients $\lambda_1, \dots, \lambda_{n-1}$ sont réels.

Par linéarité du déterminant sur la dernière colonne, on a alors

$$\det(C_1, \dots, C_n) = \sum_{i=1}^{n-1} \lambda_i \det(C_1, \dots, C_{n-1}, C_i),$$

et donc $\det(C_1, \dots, C_n) = 0$, vu l'antisymétrie du déterminant.

Nous prouvons la réciproque par récurrence sur n . Pour $n = 1$, on a $\det(C_1) = C_1$ car C_1 est un nombre. Donc si $\det(C_1) = 0$, alors C_1 est nul et donc linéairement dépendant. Supposons le résultat vrai pour $n \in \{1, \dots, k-1\}$ ($k \geq 2$) et prouvons le pour $n = k$. Soient C_1, \dots, C_k des vecteurs de \mathbb{R}^k satisfaisant $\det(C_1, \dots, C_k) = 0$.

Si C_1, \dots, C_k sont combinaisons linéaires des vecteurs e_2, \dots, e_k de la base canonique de \mathbb{R}^k , alors ils sont linéairement dépendants, par le théorème de Steinitz. Dans le cas contraire, un de ces vecteurs a une première composante non nulle et, quitte à les renuméroter, on peut supposer que $C_{1,1} \neq 0$. On définit lors les vecteurs

$$C'_2 = C_2 - \frac{C_{2,1}}{C_{1,1}} C_1, \dots, C'_k = C_k - \frac{C_{k,1}}{C_{1,1}} C_1$$

et on a

$$\det(C_1, \dots, C_k) = \det(C_1, C'_2, \dots, C'_k) = 0$$

Par définition, ce dernier déterminant vaut aussi

$$C_{1,1} \det(\overline{C'_2}, \dots, \overline{C'_k})$$

où chaque $\overline{C'_j} \in \mathbb{R}^{k-1}$ est obtenu à partir de C'_j en supprimant la première composante (nulle) de ce vecteur.

Vu l'hypothèse de récurrence et puisque $C_{1,1} \neq 0$, les vecteurs $\overline{C'_2}, \dots, \overline{C'_k}$ sont linéairement dépendants dans \mathbb{R}^{k-1} : il existe des coefficients $\mu_2, \dots, \mu_{k-1} \in \mathbb{R}$, non tous nuls, tels que

$$\mu_2 \overline{C'_2} + \dots + \mu_{k-1} \overline{C'_k} = 0.$$

Cette relation reste vraie pour les vecteurs C'_2, \dots, C'_k , puisque leur première composante est nulle. On a alors

$$0 = \mu_2 C'_2 + \dots + \mu_{k-1} C'_{k-1} = \mu_2 C_2 + \dots + \mu_{k-1} C_{k-1} - \frac{\sum_{i=2}^{k-1} \mu_i C_{i,1}}{C_{1,1}} C_1$$

et on conclut que C_1, \dots, C_k sont linéairement dépendants. \square

6.1.4 Rang d'une matrice

La notion de rang d'une matrice nous sera particulièrement utile d'ici peu. Le rang peut être facilement calculé à l'aide de déterminants.

Définition 6.1.7. Soit $A \in \mathbb{R}_p^n$. Le rang de A est le nombre maximal de colonnes de A linéairement indépendantes (dans \mathbb{R}^n). On le note $\text{rg}(A)$ ou $\rho(A)$.

Par définition le rang de A est donc la dimension dans \mathbb{R}^n de l'enveloppe linéaire des colonnes de A . Le théorème suivant, dont la démonstration est quelque peu technique, permet un calcul du rang à l'aide des déterminants, et sera un premier pas fondamental pour obtenir des équations cartésiennes de sous-espaces vectoriels.

Proposition 6.1.8. Le rang d'une matrice $A \in \mathbb{R}_p^n$ est l'ordre du plus grand mineur non nul de A . C'est aussi le nombre maximal de lignes linéairement indépendantes de A .

On déduit de la définition du rang et de cette proposition que le rang d'une matrice est inférieur ou égal au nombre de colonnes et au nombre de lignes de cette matrice. On peut déjà faire des exemples simples.

Exemple 6.1.9. On a

$$\text{rg} \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{pmatrix} = 1$$

puisque le nombre maximal de colonnes linéairement indépendantes est 1.

On a

$$\text{rg} \begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \\ 3 & 4 & 7 \end{pmatrix} = 2$$

car le mineur $\det \begin{pmatrix} 2 & 2 \\ 3 & 4 \end{pmatrix}$ est non nul, tandis que le seul mineur d'ordre 3 dans cette matrice est nul. De la même manière, on a

$$\text{rg} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 2 & 0 & 3 \end{pmatrix} = 2.$$

6.2 Systèmes linéaires

Un *système d'équations linéaires à p équations et n inconnues* (que nous notons x_1, \dots, x_n) est un ensemble d'équations de la forme

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{p1}x_1 + \cdots + a_{pn}x_n = b_p \end{cases},$$

où les coefficients $a_{11}, \dots, a_{pn}, b_1, \dots, b_p$ sont réels. Nous noterons généralement (S) un tel système d'équations.

Une *solution* d'un système (S) d'équations est un n -uplet $(x_1, \dots, x_n)^\sim$ de \mathbb{R}^n qui satisfait **simultanément** toutes les équations de (S) .

Le système (S) est dit *compatible* s'il admet au moins une solution. Il est dit *incompatible* dans le cas contraire. Dans le cadre de ce cours et en particulier dans les pages qui suivent, nous nous intéresserons aux conditions de compatibilité des systèmes linéaires.

La matrice A de \mathbb{R}_n^p définie par

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pn} \end{pmatrix}$$

est appelée *matrice des coefficients* du système (S) tandis que le n -uplet $b = (b_1, \dots, b_p)^\sim$ de \mathbb{R}^p est le *terme indépendant* du système (S) . En notant x le n -uplet d'inconnues $(x_1, \dots, x_n)^\sim$, nous observons directement que le système (S) peut être mis sous forme *matricielle* :

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix},$$

ou encore $Ax = b$. On peut également lui donner une forme intermédiaire, dite *forme vectorielle* :

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{p1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{pn} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \quad (6.1)$$

Enfin, un système linéaire est dit *homogène* si son terme indépendant est nul. Etant donné un système (S) , le système homogène *associé* à (S) est le système obtenu en remplaçant le terme indépendant de (S) par 0.

Exemple 6.2.1. L'ensemble d'équations

$$(S) : \begin{cases} 2x_1 - x_2 + x_3 & = & 1 \\ 3x_1 - 5x_2 + 4x_3 & = & 0 \end{cases}$$

est un système linéaire de 2 équations à 3 inconnues. Il peut être mis sous forme matricielle en

$$\begin{pmatrix} 2 & -1 & 1 \\ 3 & -5 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Il s'agit d'un système compatible puisque le triplet $(5/7, 3/7, 0)$ est une solution de (S) . Ce système n'est pas homogène puisque son terme indépendant est $(1, 0)^\sim$, mais le système homogène associé à (S) est

$$(S_0) : \begin{cases} 2x_1 - x_2 + x_3 & = & 0 \\ 3x_1 - 5x_2 + 4x_3 & = & 0 \end{cases}$$

Voici un second exemple, avec un système incompatible.

Exemple 6.2.2. Le système d'équations

$$(S) : \begin{cases} x_1 + x_2 & = & 3 \\ 2x_1 + 2x_2 & = & 8 \end{cases} \quad (6.2)$$

est un système linéaire de 2 équations à 2 inconnues. Il est visiblement incompatible. Par contre le système homogène associé à (S) est compatible et admet même une infinité de solutions.

Le système d'équations

$$(S) : \begin{cases} x_1 + x_2 & = & 3 \\ 2x_1 + 2x_2 & = & 6 \end{cases} \quad (6.3)$$

est un système linéaire de 2 équations à 2 inconnues. Il est compatible. Il admet une infinité de solutions.

L'exemple précédent nous amène à poser une dernière définition.

Définition 6.2.3. Un système linéaire qui admet une solution unique est dit *déterminé*. S'il admet plusieurs solutions, il est dit *indéterminé*. Deux systèmes linéaires à n inconnues (S) et (S') sont dits *équivalents* s'ils ont les mêmes ensembles de solutions. On note alors $(S) \Leftrightarrow (S')$.

6.3 Compatibilité et structure des solutions

On pourrait penser de manière naïve que, pour analyser les solutions d'un système, il suffit de comparer le nombre d'inconnues (parfois appelés "degrés de liberté") du système avec le nombre d'équations. En effet, chaque équation amène une contrainte sur les inconnues et permet en principe d'exprimer une inconnue en fonction des autres. La conclusion serait alors que s'il n'y a pas assez d'équations, le système serait indéterminé, s'il y a trop d'équations, le système serait incompatible, et s'il y a le bon nombre, il serait déterminé.

Les systèmes considérés à l'exemple 6.2.2 montrent que cette vision des choses est trop simpliste. D'une part, il y a des contraintes contradictoires sur les équations, et d'autre part le comptage des équations doit se faire de manière précise.

En effet, de manière générale, on peut toujours "dupliquer" (recopier) une équation d'un système linéaire. On obtient alors un système équivalent à celui de départ :

$$\begin{pmatrix} 1 & 2 & 3 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{est équivalent à} \quad \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

De manière un peu plus intelligente, on peut toujours ajouter à un système linéaire une équation qui est combinaison linéaire des équations du système, sans changer l'ensemble des solutions : on a

$$(S_1) : \begin{pmatrix} 1 & 2 & 3 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad (S_2) : \begin{pmatrix} 1 & 6 & 7 \\ 1 & 2 & 3 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \quad (6.4)$$

En lisant cette équivalence dans l'autre sens, on voit qu'on peut enlever dans un système une équation qui est combinaison (linéaire) des autres et obtenir un système équivalent, qui comporte une équation de moins.

En conclusion, ce n'est pas le nombre d'équations qui est important, mais bien le nombre d'équations "qui ne sont pas combinaisons des autres", ou en d'autres termes, le nombre maximal d'équations linéairement indépendantes. On peut caractériser ce nombre à partir de la matrice des coefficients A et du terme indépendant b . Pour ce faire, on introduit une notation supplémentaire.

Définition 6.3.1. Si $A \in \mathbb{R}_n^p$ et $B \in \mathbb{R}_m^p$, on note $(A|B)$ la matrice dont les colonnes sont celles de A suivies de celles de B .

Le nombre maximal d'équations indépendantes dans un système linéaires est donc $\text{rg}(A|b)$. Il est donc naturel que ce nombre joue un rôle important dans les résultats qui suivent, puisque c'est le nombre de contraintes indépendantes (utiles) que l'on impose aux inconnues.

En ce qui concerne les contraintes contradictoires, considérons encore l'exemple 6.2.2. Le premier système n'admet pas de solution, car le membre de gauche de la deuxième équation est le double de celui de la première, mais cette relation n'est pas vraie pour les membres de droite. On constate plus facilement ce fait quand le système est mis sous forme matricielle : dans ce cas, on a

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \quad \text{et} \quad (A|b) = \begin{pmatrix} 1 & 1 & 3 \\ 2 & 2 & 8 \end{pmatrix}$$

Les lignes de A sont linéairement dépendantes, mais puisque 8 ne vaut pas $2 \cdot 3$, celles de $(A|b)$ ne le sont pas. En d'autres termes, l'ajout de la colonne b a fait augmenter le rang.

On peut également reprendre l'exemple des systèmes (S_1) et (S_2) de (6.4). Ils sont équivalents et compatibles puisque par exemple $x = (1/2, 1/4, 0)$ est solution du premier. Dans le système (S_2) , la première équation vaut 2 fois la deuxième plus la troisième. D'un point de vue matriciel, on a

$$\text{rg}(A) = \text{rg} \begin{pmatrix} 1 & 6 & 7 \\ 1 & 2 & 3 \\ -1 & 2 & 1 \end{pmatrix} = 2 = \text{rg} \begin{pmatrix} 1 & 6 & 7 & 2 \\ 1 & 2 & 3 & 1 \\ -1 & 2 & 1 & 0 \end{pmatrix} = \text{rg}(A|b).$$

Si nous modifions un peu cet exemple pour considérer le système

$$(S'_2) : \begin{pmatrix} 1 & 6 & 7 \\ 1 & 2 & 3 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \\ 0 \end{pmatrix},$$

nous concluons évidemment que ce système n'est pas compatible. En effet, les membres de gauche satisfont toujours la relation évoquée plus haut, mais ce n'est plus le cas des membres de droite correspondants. Dans ce cas, on a

$$\operatorname{rg}(A) = \operatorname{rg} \begin{pmatrix} 1 & 6 & 7 \\ 1 & 2 & 3 \\ -1 & 2 & 1 \end{pmatrix} = 2, \quad \text{et} \quad \operatorname{rg}(A|b) = \operatorname{rg} \begin{pmatrix} 1 & 6 & 7 & 4 \\ 1 & 2 & 3 & 1 \\ -1 & 2 & 1 & 0 \end{pmatrix} = 3.$$

Nous rassemblons maintenant ces différentes constatations dans un résultat dont nous admettrons la preuve.

Proposition 6.3.2. *Un système linéaire à p équations et n inconnues qui s'écrit $Ax = b$ est compatible si, et seulement si, on a $\operatorname{rg}(A) = \operatorname{rg}(A|b)$.*

Un système homogène à p équations et n inconnues $Ax = 0$ est toujours compatible. L'ensemble de ses solutions est un sous-espace vectoriel de \mathbb{R}^n de dimension $n - \operatorname{rg}(A)$.

On peut appliquer ce résultat pour étudier la compatibilité des systèmes. Je vous propose de démontrer les faits suivants :

1. Le système d'équations

$$\begin{cases} 3x_1 + 2x_2 + x_3 = 1 \\ x_1 + 2x_2 + 3x_3 = 7 \\ x_1 + x_2 + x_3 = 2 \end{cases}$$

est compatible. L'ensemble des solutions du système homogène associé est un sous-espace vectoriel de \mathbb{R}^3 de dimension 2.

2. Le système d'équations

$$\begin{cases} 3x + 5y = 4 \\ 6x + \lambda y = 10 \end{cases}$$

est compatible pour tout $\lambda \in \mathbb{R} \setminus \{10\}$. Il est incompatible pour $\lambda = 10$. L'ensemble des solutions du système homogène correspondant est un sous-espace vectoriel de dimension 0 si $\lambda \neq 10$ et de dimension 1 si $\lambda = 10$.

3. Si $(u_1, u_2)^\sim \in \mathbb{R}^2 \setminus \{0\}$, alors le système d'équations

$$\begin{cases} u_1x = a \\ u_2x = b \end{cases}$$

est compatible si, et seulement si, on a $\det \begin{pmatrix} u_1 & a \\ u_2 & b \end{pmatrix} = 0$.

4. Si $(u_1, u_2, u_3)^\sim$ et $(v_1, v_2, v_3)^\sim$ sont linéairement indépendants dans \mathbb{R}^3 , alors le système d'équations

$$\begin{cases} u_1x + v_1y = a \\ u_2x + v_2y = b \\ u_3x + v_3y = c \end{cases}$$

est compatible si, et seulement si, on a

$$\det \begin{pmatrix} u_1 & v_1 & a \\ u_2 & v_2 & b \\ u_3 & v_3 & c \end{pmatrix} = 0.$$

5. L'ensemble des solutions du système d'équations à 3 inconnues

$$x_1 + x_2 + x_3 = 0$$

est un sous-espace vectoriel de \mathbb{R}^3 de dimension 2. Exprimer les solutions comme combinaisons linéaires de deux vecteurs.

6. L'ensemble des solutions du système d'équations à 4 inconnues

$$\begin{cases} x_1 - x_2 + x_3 - x_4 = 0 \\ 3x_1 - 2x_2 + x_3 - x_4 = 0 \end{cases}$$

est un sous-espace vectoriel de \mathbb{R}^4 , de dimension 2. Exprimer l'ensemble des solutions comme une enveloppe linéaire.

Nous avons jusqu'à présent donné les conditions de compatibilité d'un système linéaire général, et la structure des solutions d'un système linéaire homogène. Il manque évidemment la structure des solutions d'un système linéaire général, quand celui-ci est compatible. Cette structure est déterminée par la proposition suivante.

Proposition 6.3.3. *Soient $(S) : Ax = b$ un système linéaire de p équations à n inconnues, compatible, et x_0 une solution (particulière) de ce système. Pour toute solution y du système homogène associé $(S_0) : Ax = 0$, $x_0 + y$ est une solution de (S) . Réciproquement, toute solution de (S) est de cette forme. L'ensemble des solutions de (S) s'écrit donc*

$$\{x_0 + y : Ay = 0\}.$$

Démonstration. Si on a $Ay = 0$ et $Ax_0 = b$, alors vu la bilinéarité du produit matriciel, on a aussi $A(x_0 + y) = Ax_0 + Ay = b + 0 = b$. Réciproquement, si on a $Ax = b$, alors $x = x_0 + (x - x_0)$ et toujours par bilinéarité, on a $A(x - x_0) = Ax - Ax_0 = b - b = 0$. \square

Nous ne connaissons pas encore le type de structure décrit dans cette proposition. On obtient un tel sous-ensemble de \mathbb{R}^n en *translatant* tous les éléments d'un sous-espace vectoriel. Nous étudierons en détail ce type de structure dans un prochain chapitre. Ce sont des *variétés affines* de \mathbb{R}^n . Nous avons donc montré le théorème de structure suivant.

Proposition 6.3.4. *L'ensemble des solutions d'un système linéaire compatible (S) de p équations à n inconnues est une variété affine de \mathbb{R}^n de dimension $n - \text{rg}(S)$.*

Table des matières

1	Espaces vectoriels réels	2
1.1	Introduction	2
1.2	Premières définitions et exemples	2
1.2.1	Une représentation intuitive et commode	6
1.2.2	Combinaisons linéaires	6
1.3	Sous-espaces vectoriels	7
1.4	Bases et dimension	11
1.5	Composantes	17
1.6	Equations de sous-espaces vectoriels	19
1.6.1	Equations paramétriques	19
1.6.2	Cas particuliers : droites et hyperplans	24
1.6.3	Faisceaux de plans en dimension 3	25
2	Espaces affines	26
2.1	Définitions	26
2.2	Vecteurs liés et vecteurs libres	28
2.3	Combinaisons affines	29
2.4	Variétés affines	32
2.5	Intersection et parallélisme de variétés affines	35
2.5.1	Positions relatives de droites en dimensions 2 et 3	36
2.6	Repères et coordonnées	37
2.7	Equations de variétés affines	39
2.7.1	Le cas des droites en dimension 2	42
2.7.2	Le cas des droites en dimension 3	43
2.7.3	Le cas des plans en dimension 3	44
2.7.4	Deux problèmes classiques	45
3	Espaces vectoriels euclidiens	46
3.1	Les produits scalaires	46
3.2	Normes de vecteurs, inégalités célèbres et angles non orientés	48
3.3	Bases orthonormées	50
3.4	Produit mixte et produit vectoriel, angle orienté	53
3.4.1	Orientation d'un espace vectoriel, changements de bases orthonormées	53
3.4.2	Produit mixte	54
3.4.3	Le produit vectoriel	55
3.4.4	Angle orienté dans le plan	58
3.5	Complément orthogonal et projection orthogonale	59
4	Espaces affines euclidiens	62
4.1	Espaces affines euclidiens et distances	62
4.2	Variétés affines orthogonales	63
4.3	Projections orthogonales	64
4.4	Distances de variétés affines, et cas des dimensions deux et trois	65

4.4.1	Le cas de la dimension 2	67
4.4.2	Le cas de la dimension 3	67
4.5	Angles de droites et plans	68
5	Théorie des courbes	69
5.1	Quelques points techniques	69
5.1.1	Quelques résultats sur la dérivation des fonctions à valeurs vectorielles . .	71
5.2	Arc réguliers de courbes : notions élémentaires	73
5.2.1	Paramétrages : définitions et exemples	73
5.2.2	Paramétrages équivalents	74
5.2.3	Tangente en un point d'un arc régulier	75
5.2.4	Orientation	76
5.2.5	Longueur d'arc et paramétrage naturel	76
5.2.6	Paramétrages naturels	78
5.3	Le trièdre de Frenet	78
5.4	Normale et courbure signées des courbes planes	82
6	Appendice : Calcul matriciel	85
6.1	Éléments de calcul matriciel	85
6.1.1	Sous-matrices	85
6.1.2	Déterminants des matrices carrées : définitions	86
6.1.3	Déterminants des matrices carrées : propriétés	87
6.1.4	Rang d'une matrice	90
6.2	Systèmes linéaires	90
6.3	Compatibilité et structure des solutions	92