

Université Blaise Pascal
U.F.R. Sciences et Technologies
Département de Mathématiques et Informatique

Licence Sciences - Langues
Deuxième année

MATHÉMATIQUES

Notes de cours

2013-2014

FRANÇOIS DUMAS

Ces pages constituent les notes du cours du premier semestre de la deuxième année de la licence sciences-langues, à l'U.F.R Sciences et Technologies de l'Université Blaise Pascal.

Elles sont organisées en sept chapitres, incluant de nombreux exercices d'application, complétés (p. 158 à 173) par des sujets de devoirs.

Elles correspondent à un contenu, un niveau et un mode d'exposition propres au programme et aux objectifs de cette formation.

Le fait de les mettre à la disposition des étudiants comme support à leur travail personnel ne signifie pas qu'il s'agisse d'un document parfaitement finalisé, dans sa conception comme dans sa rédaction. Je suis donc par avance reconnaissant à toutes celles et tous ceux qui me signaleront les erreurs, manques, imperfections, coquilles,... qu'il contient inmanquablement.

`Francois.Dumas@univ-bpclermont.fr`

Table des matières

1 Réduction des endomorphismes	7
1.1 Notions générales sur les valeurs propres	8
1.1.1 Valeurs propres, vecteurs propres, sous-espaces propres	8
1.1.2 Exemples de problèmes de valeurs propres	8
1.1.3 Polynôme caractéristique	9
1.1.4 Multiplicité d'une valeur propre	11
1.1.5 Quelques exemples explicites de différentes situations possibles	12
1.1.6 Exercices	13
1.2 Diagonalisation	15
1.2.1 Endomorphisme diagonalisable, matrice diagonalisable	15
1.2.2 Condition nécessaire et suffisante de diagonalisabilité	16
1.2.3 La question de la trigonalisabilité	17
1.2.4 Exercices	19
1.3 Applications de la diagonalisation	20
1.3.1 Application au calcul des puissances d'une matrice	20
1.3.2 Application à l'étude de suites récurrentes	21
1.3.3 Application à la résolution de systèmes différentiels linéaires	22
1.3.4 Exercices	25
2 Espaces vectoriels euclidiens	27
2.1 Produit scalaire et norme euclidienne	27
2.1.1 Notion de produit scalaire	27
2.1.2 Orthogonalité	29
2.1.3 Représentation des formes linéaires	31
2.1.4 Familles orthogonales de polynômes	32
2.1.5 Exercices	33
2.2 Groupe orthogonal	35
2.2.1 Endomorphismes orthogonaux	35
2.2.2 Matrices orthogonales	36
2.2.3 Orientation et produit mixte	37
2.2.4 Exercices.	39
2.3 Endomorphismes symétriques	41
2.3.1 Transposition	41
2.3.2 Diagonalisation des endomorphismes symétriques	41
2.3.3 Exercices	43

2.4	Applications géométriques	44
2.4.1	Description des isométries vectorielles en dimension 2	44
2.4.2	Angles	46
2.4.3	Description des isométries vectorielles en dimension 3	50
2.4.4	Exercices	51
3	Espaces vectoriels normés	53
3.1	Normes et notions topologiques associées	53
3.1.1	Notion de norme	53
3.1.2	Distance associée à une norme	55
3.1.3	Boules et parties bornées	56
3.1.4	Parties ouvertes	57
3.1.5	Parties fermées	58
3.1.6	Normes équivalentes	59
3.1.7	Exercices	60
3.2	Suites dans un evn	62
3.2.1	Suites convergentes	62
3.2.2	Opérations sur les suites convergentes	63
3.2.3	Suites extraites	65
3.2.4	Suites de Cauchy	66
3.2.5	Exercices	67
3.3	Fonctions d'un evn dans un evn	69
3.3.1	Limites dans les evn	69
3.3.2	Continuité dans les evn	71
3.3.3	Applications lipschitziennes, applications linéaires continues	72
3.3.4	Exercices	74
4	Séries numériques	77
4.1	Notion de série	77
4.1.1	Terminologie des séries	77
4.1.2	Convergence d'une série	77
4.1.3	Exemples classiques	78
4.1.4	Espace vectoriel des séries convergentes	80
4.1.5	Exercices	80
4.2	Séries à termes réels positifs	81
4.2.1	Critère de majoration	81
4.2.2	Séries de Riemann	82
4.2.3	Règle de d'Alembert	83
4.2.4	Règle de Cauchy	84
4.2.5	Exercices	84
4.3	Séries numériques à termes quelconques	85
4.3.1	Convergence absolue	85
4.3.2	Séries (réelles) alternées	86
4.3.3	Exercices	87
4.4	Résultats complémentaires	88
4.4.1	Produit de deux séries	88

4.4.2	Le problème de la sommation par paquets	89
4.4.3	Le problème de l'ordre des termes	90
4.4.4	Comparaison entre séries et intégrales	92
5	Suites et séries de fonctions	95
5.1	Convergence simple et uniforme d'une suite de fonctions	95
5.1.1	Notions de convergence d'une suite de fonctions	95
5.1.2	Convergence uniforme et continuité	97
5.1.3	Convergence uniforme sur tout segment	98
5.1.4	Convergence uniforme, intégration et dérivation.	98
5.1.5	Exercices	101
5.2	Séries de fonctions et convergence normale	102
5.2.1	Convergence simple et uniforme d'une série de fonctions	102
5.2.2	Convergence normale	103
5.2.3	Exercices	106
6	Séries entières	107
6.1	Convergence des séries entières	108
6.1.1	Rayon de convergence d'une série entière	108
6.1.2	Convergence normale d'une série entière	111
6.1.3	Fonctions définies par la somme d'une série entière	112
6.1.4	Cas réel : application aux équations différentielles	114
6.1.5	Exercices	115
6.2	Développement en séries entières	116
6.2.1	Fonction développable en série entière.	116
6.2.2	Exemples classiques.	118
6.2.3	Exercices	120
7	Géométrie affine en dimension 2 et 3	121
7.1	Espaces affines	121
7.1.1	Notion d'espace affine	121
7.1.2	Sous-espaces affines	122
7.1.3	Parallélisme	124
7.1.4	Bases affines	125
7.1.5	Coordonnées et représentations paramétriques des sous-espaces affines	127
7.1.6	Equations cartésiennes des sous-espaces affines	128
7.1.7	Barycentre	131
7.1.8	Cas particulier des espaces affines euclidiens.	134
7.1.9	Exercices	137
7.2	Applications affines	139
7.2.1	Notion d'application affine	139
7.2.2	Applications affines bijectives, groupe affine	142
7.2.3	Homothéties, translations	143
7.2.4	Projections, symétries	146
7.2.5	Le cas des isométries affines.	148
7.2.6	Exercices	152

Chapitre 1

Réduction des endomorphismes

Considérons un \mathbb{K} -espace vectoriel E (où \mathbb{K} est un corps ; pour nous ce sera \mathbb{R} ou \mathbb{C}) et f un endomorphisme de E .

a) L'image par f d'une droite vectorielle de E , lorsqu'elle n'est pas nulle, est une droite vectorielle. Existe-t-il des droites qui soient leur propre image par f ? C'est une des questions que nous allons étudier dans ce chapitre.

b) Supposons que E est de dimension finie. En choisissant une base \mathcal{B} de E , on peut représenter f par sa matrice carrée A dans la base \mathcal{B} . Or il existe des types particuliers de matrices pour lesquelles les règles de calculs sont particulièrement simples, en particulier les matrices diagonales (le produit de deux matrices diagonales est diagonale, le déterminant d'une matrice diagonale est égal au produit de ses coefficients diagonaux,...), ou plus généralement les matrices triangulaires. On comprend donc que, pour représenter f par sa matrice A dans la base \mathcal{B} , on a tout intérêt à choisir \mathcal{B} telle que A soit diagonale (ou à défaut triangulaire). Encore faut-il que cela soit possible, c'est-à-dire qu'une telle base existe. C'est ce problème (pour un endomorphisme donné, déterminer une base dans laquelle sa matrice soit la plus simple possible : diagonale, triangulaire...) que l'on va aborder dans ce chapitre. Cette question est directement liée à la question a) précédente. Les applications (en analyse et en géométrie en particulier) sont extrêmement importantes.

c) On connaît divers *invariants* pour les matrices carrées : il s'agit d'objets mathématiques qui ne changent pas quand on remplace une matrice A par une matrice semblable, c'est-à-dire une matrice carrée $A' = P^{-1}AP$, avec P inversible. On sait que cela signifie que ces objets sont en fait attachés non pas à la matrice A elle-même, mais à l'endomorphisme f dont A ou A' sont les matrices dans deux bases. Citons :

- le rang, qui est un entier naturel : $\text{rg } A = \text{rg}(P^{-1}AP) = \text{rg } f$;
- la trace, qui est un scalaire : $\text{tr } A = \text{tr}(P^{-1}AP) = \text{tr } f$;
- le déterminant, qui est aussi un scalaire : $\det A = \det(P^{-1}AP) = \det f$.

On va dans ce chapitre définir un autre invariant, qui est cette fois un polynôme (en un certain sens, que l'on précisera plus loin, il "contient" les précédents), et qui est un outil fondamental pour résoudre les questions a) et b) formulées précédemment.

1.1 Notions générales sur les valeurs propres

1.1.1 Valeurs propres, vecteurs propres, sous-espaces propres

Soit E un \mathbb{K} -espace vectoriel. Soit f un endomorphisme de E .

a) DÉFINITIONS

1. Une *valeur propre* de f dans \mathbb{K} est un scalaire $\lambda \in \mathbb{K}$ pour lequel il existe un vecteur $v \in E$ non-nul tel que $f(v) = \lambda.v$. On utilisera l'abréviation v.p. pour valeur propre.
2. On appelle *spectre* de f l'ensemble des v.p. de f .
3. Si λ est une v.p. de f , on appelle *vecteur propre* de f associé à λ tout vecteur $v \in E$ non-nul tel que $f(v) = \lambda.v$.
4. Pour toute v.p. λ de f , on appelle *sous-espace propre* de f associé à la v.p. λ le sous-espace vectoriel $E_\lambda = \text{Ker}(f - \lambda.\text{id})$ de E .

b) REMARQUE. Soit $\lambda \in \mathbb{K}$. Considérons le sous-espace vectoriel $E_\lambda = \text{Ker}(f - \lambda.\text{id})$.

- Par définition, E_λ est l'ensemble des vecteurs $v \in E$ tels que $f(v) = \lambda.v$.

- Dire que λ est une v.p. de f signifie que le ss-e.v. E_λ n'est pas réduit à $\{0_E\}$. Si E est de dimension finie n , on a alors : $1 \leq \dim E_\lambda \leq n$.

- Si λ est une v.p. de f , le sous-espace vectoriel E_λ est donc formé des vecteurs propres associés à λ (qui par définition sont tous non-nuls), et du vecteur nul (qui n'est pas un vecteur propre, mais qui appartient quand même au sous-espace propre).

c) REMARQUE. Si v est un vecteur propre associé à une valeur propre λ , alors v ne peut pas être vecteur propre pour une autre valeur propre $\mu \neq \lambda$. En effet, $f(v) = \lambda.v = \mu.v$ implique $(\lambda - \mu).v = 0_E$, ce qui, puisque $v \neq 0_E$, conduit nécessairement à $\lambda = \mu$.

1.1.2 Exemples de problèmes de valeurs propres

a) EXEMPLE (équation différentielle). On prend ici pour E le \mathbb{R} -espace vectoriel des fonctions réelles d'une variable réelle de classe C^∞ sur \mathbb{R} . On considère l'application $f : E \rightarrow E$ qui, à toute fonction $y \in E$ associe sa dérivée $y' \in E$. Il est clair que f est linéaire.

Un réel λ est une v.p. de f si et seulement s'il existe une fonction non identiquement nulle $y \in E$ telle que $y' = \lambda y$. On sait que, pour tout $\lambda \in \mathbb{R}$, les solutions de cette équation différentielle sont les fonctions de la forme $y = ky_\lambda$ avec $k \in \mathbb{R}$ quelconque, où l'on note y_λ la fonction $y_\lambda : t \mapsto e^{\lambda t}$. En d'autres termes, le spectre de f est l'ensemble \mathbb{R} et, pour tout $\lambda \in \mathbb{R}$, le sous-espace propre E_λ associé à la v.p. λ est la droite vectorielle dirigée dans E par y_λ .

b) EXEMPLE (espace de suites). On prend ici pour E le \mathbb{R} -espace vectoriel des suites bornées de réels. On considère l'application $f : E \rightarrow E$ qui, à toute suite $u = (u_n)_{n \geq 0}$, associe la suite $u^+ = (u_{n+1})_{n \geq 0}$. Il est clair que f est linéaire.

Un réel λ est une v.p. de f si et seulement s'il existe une suite non identiquement nulle $u \in E$ telle que $u_{n+1} = \lambda u_n$ pour tout $n \geq 0$. Il est clair que, pour tout $\lambda \in \mathbb{R}$, les suites vérifiant cette condition sont les suites géométriques définies par $u_n = \lambda^n u_0$. Une telle suite est bornée si et seulement si $(u_0 = 0)$ ou $(u_0 \neq 0 \text{ et } |\lambda| \leq 1)$.

En d'autres termes, le spectre de f est l'ensemble $[-1, 1]$ et, pour tout $\lambda \in [-1, 1]$, le sous-espace propre E_λ associé à la v.p. λ est la droite vectorielle dirigée par la suite géométrique $(\lambda^n)_{n \geq 0}$.

c) EXEMPLE (projections et symétries). On prend ici un \mathbb{K} -espace vectoriel quelconque E , et l'on considère deux sous-espaces vectoriels non-nuls F et H tels que $E = F \oplus H$.

Tout vecteur $x \in E$ se décompose de façon unique sous la forme $x = y + z$ avec $y \in F$ et $z \in H$.

La *projection* sur F parallèlement à H est l'application p qui à tout $x \in E$ ainsi décomposé associe $p(x) = y$.

On a : $p \circ p = p$, $\text{Ker } p = H$ et $\text{Im } p = F$.

La *symétrie* par rapport à F parallèlement à H est l'application s qui à tout $x \in E$ ainsi décomposé associe $s(x) = y - z$.

On a : $s \circ s = \text{id}_E$, $\text{Ker } s = \{0_E\}$ et $\text{Im } s = E$



Si un réel λ est une v.p. de p , alors il existe un vecteur non-nul $u \in E$ tel que $p(u) = \lambda u$, d'où $p^2(u) = p(\lambda u) = \lambda p(u) = \lambda^2 u$. Puisque $p \circ p = p$, il vient $\lambda u = \lambda^2 u$. Les seules v.p. possibles sont donc $\lambda = 0$ ou $\lambda = 1$. Il est clair que réciproquement 0 et 1 sont bien des v.p. de p , de sous-espaces propres associés $E_0 = H$ et $E_1 = F$. En particulier, on a : $E = E_0 \oplus E_1$.

Si un réel λ est une v.p. de s , alors il existe un vecteur non-nul $u \in E$ tel que $s(u) = \lambda u$, d'où $s^2(u) = s(\lambda u) = \lambda s(u) = \lambda^2 u$. Puisque $s \circ s = \text{id}_E$, il vient $u = \lambda^2 u$. Les seules v.p. possibles sont donc $\lambda = -1$ ou $\lambda = 1$. Il est clair que réciproquement -1 et 1 sont bien des v.p. de p , de sous-espaces propres associés $E_{-1} = H$ et $E_1 = F$. En particulier, on a : $E = E_{-1} \oplus E_1$.

d) COMMENTAIRES. Le fait que, pour les projections et pour les symétries, E soit somme directe des sous-espaces propres est une propriété très importante, que l'on étudiera dans un contexte plus général plus loin en 1.2.

Les exemples précédents mettent en évidence que, même pour des espaces vectoriels de dimension infinie, on peut dans certaines situations particulières déterminer des v.p. et leurs sous-espaces propres associés simplement en revenant aux définitions. Mais dans le cas des espaces vectoriels de dimension finie, on dispose d'une méthode systématique que l'on va maintenant développer.

1.1.3 Polynôme caractéristique

Soit E un \mathbb{K} -espace vectoriel de dimension finie n .

a) DÉFINITION. Soit f un endomorphisme de E . On appelle *polynôme caractéristique* de f le polynôme $P_f(x) = \det(f - x \cdot \text{id}_E) \in \mathbb{K}[x]$.

b) THÉORÈME. Soit f un endomorphisme de E . Les valeurs propres de f dans \mathbb{K} sont exactement les zéros dans \mathbb{K} de son polynôme caractéristique $P_f(x)$.

Preuve. Soit $\lambda \in \mathbb{K}$. On a vu à la remarque b) du 1.1.1 que λ est une v.p. de f si et seulement si le noyau de l'endomorphisme $(f - \lambda \cdot \text{id}_E)$ n'est pas réduit à $\{0_E\}$, c'est-à-dire si et seulement si $(f - \lambda \cdot \text{id}_E)$ n'est pas injectif. Comme E est de dimension finie, on sait que cela est encore équivalent à dire que $(f - \lambda \cdot \text{id}_E)$ n'est pas bijectif, et donc à $\det(f - \lambda \cdot \text{id}_E) = 0$. En résumé, les v.p. de f sont exactement les scalaires λ tels que $P_f(\lambda) = 0$. \square

c) REMARQUES (passage aux matrices).

- Soit \mathcal{B} une base quelconque de E . Soit f un endomorphisme de E . Soit A la matrice de f par rapport à la base \mathcal{B} . La matrice par rapport à la base \mathcal{B} de l'endomorphisme $(f - x.\text{id}_E)$ est alors $(A - x.I_n)$. On a donc $\det(f - x.\text{id}_E) = \det(A - x.I_n)$, et ceci quelle que soit la base \mathcal{B} choisie.

- Réciproquement, toute matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ peut être considérée comme la matrice d'un certain endomorphisme f de $E = \mathbb{K}^n$ par rapport à la base canonique. On appelle *polynôme caractéristique* de la matrice A le polynôme caractéristique de f , c'est-à-dire le polynôme $P_A(x) = \det(A - x.I_n)$. Ses zéros dans \mathbb{K} seront appelés les *valeurs propres* de la matrice A dans \mathbb{K} ; ce sont bien sûr les valeurs propres de l'endomorphisme f .

d) CONSÉQUENCES PRATIQUES. On suppose toujours que f est un endomorphisme d'un espace vectoriel E de dimension finie n .

(i) Pour déterminer les v.p. de f , on fixe une base quelconque \mathcal{B} , on introduit la matrice A de f par rapport à \mathcal{B} , et on détermine par un calcul de déterminant le polynôme $P_f(x) = P_A(x) = \det(A - x.I_n)$. Les zéros de ce polynôme sont les v.p. de l'endomorphisme f , que l'on appelle aussi les v.p. de la matrice A .

(ii) Comme il est clair que $P_f(x)$ est de degré n , il admet forcément au plus n zéros dans \mathbb{K} . Donc f admet au plus n valeurs propres dans \mathbb{K} .

(iii) *Attention!* Le polynôme $P_f(x)$ n'a pas forcément des zéros dans \mathbb{K} (par exemple s'il est à coefficients réels de degré 2 et de discriminant < 0). Ce qui signifie que l'endomorphisme f ou la matrice A n'a pas forcément de valeurs propres dans \mathbb{K} !

En revanche, un polynôme à coefficients réels $P_f(x)$ a forcément des zéros dans \mathbb{C} , et donc f ou A ont forcément des v.p. dans \mathbb{C} . On verra sur des exemples qu'il est parfois très utile de considérer les v.p. complexes d'une matrice même si celle-ci est au départ une matrice à coefficients réels.

(iv) Le calcul explicite du polynôme caractéristique se ramène le plus souvent à un calcul de déterminant, qui n'est pas forcément simple. Dans le cas évident où $n = 2$, le polynôme caractéristique d'une matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$ s'exprime sous la forme :

$$P_A(x) = \begin{vmatrix} a-x & b \\ c & d-x \end{vmatrix} = x^2 - x(a+d) + (ad-bc).$$

On reconnaît en $(a+d)$ la trace de A et en $(ad-bc)$ son déterminant. C'est le cas le plus simple du résultat général suivant.

e) PROPOSITION. Soit f un endomorphisme d'un espace vectoriel E de dimension finie n . Son polynôme caractéristique $P_f(x) = \det(f - x.\text{id}_E)$ est un polynôme de degré n , à coefficients dans \mathbb{K} , tel que :

- le coefficient de x^n est $(-1)^n$;
- le coefficient de x^{n-1} est $(-1)^{n-1} \text{tr } f$;
- le coefficient constant est $\det f$.

On a donc :
$$P_f(x) = (-1)^n x^n + (-1)^{n-1} (\text{tr } f) x^{n-1} + \dots + \det f.$$

En d'autres termes, pour toute matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$, on a :

$$P_A(x) = (-1)^n x^n + (-1)^{n-1} (\text{tr } A) x^{n-1} + \dots + \det A.$$

Preuve : Le résultat est vrai pour $n = 2$ comme on vient de le voir à la fin du c) ci-dessus. Pour $n \geq 3$, notons $A = (a_{i,j})_{1 \leq i,j \leq n}$ la matrice de f par rapport à une base \mathcal{B} de E , et développons par rapport à la première ligne :

$$\begin{vmatrix} a_{1,1}-x & a_{1,2} & a_{1,3} & \dots & a_{1,n-1} & a_{1,n} \\ a_{2,1} & a_{2,2}-x & a_{2,3} & \dots & a_{2,n-1} & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3}-x & \dots & a_{3,n-1} & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \dots & a_{n,n-1} & a_{n,n}-x \end{vmatrix} = (a_{1,1} - x) \begin{vmatrix} a_{2,2}-x & a_{2,3} & \dots & a_{2,n-1} & a_{2,n} \\ a_{3,2} & a_{3,3}-x & \dots & a_{3,n-1} & a_{3,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n,2} & a_{n,3} & \dots & a_{n,n-1} & a_{n,n}-x \end{vmatrix} + \dots$$

où le reste désigné par les points de suspension est de degré $< n - 1$. En d'autres termes, $P_A(x) = (a_{1,1} - x)P_{A'}(x) + \dots$, où A' est la matrice carrée d'ordre $n - 1$ extraite de A obtenue en supprimant la première ligne et la première colonne.

Par récurrence, supposons que : $P_{A'}(x) = (-1)^{n-1}x^{n-1} + (-1)^n(\text{tr } A')x^{n-2} + \dots$.

$$\begin{aligned} \text{Donc : } P_A(x) &= (-x + a_{1,1})[(-1)^{n-1}x^{n-1} + (-1)^n(\text{tr } A')x^{n-2} + \dots] + \dots \\ &= (-1)^n x^n + [a_{1,1}(-1)^{n-1} - (-1)^n(\text{tr } A')]x^{n-1} + \dots \\ &= (-1)^n x^n + (-1)^{n-1} \underbrace{(a_{1,1} + \text{tr } A')}_{\text{tr } A} x^{n-1} + \dots \end{aligned}$$

ce qui prouve les deux premiers points de la proposition. Le troisième est clair en évaluant le polynôme $P_A(x) = \det(A - x.I_n)$ en $x = 0$. \square

1.1.4 Multiplicité d'une valeur propre

a) RAPPEL (sur les polynômes). Soit $P(x) \in \mathbb{K}[x]$ de degré n . Soit α un zéro de $P(x)$ dans \mathbb{K} , c'est-à-dire un nombre $\alpha \in \mathbb{K}$ tel que $P(\alpha) = 0$. On peut alors mettre $(x - \alpha)$ en facteur dans $P(x)$. Il existe donc un unique entier $1 \leq q \leq n$, appelé la *multiplicité* de α , tel que :

$$P(x) = (x - \alpha)^q Q(x)$$

avec $Q(x) \in \mathbb{K}[x]$ de degré $n - q$ vérifiant $Q(\alpha) \neq 0$.

b) DÉFINITION. Soit λ une v.p. dans \mathbb{K} d'un endomorphisme f d'un \mathbb{K} -espace vectoriel E de dimension finie n (ou d'une matrice A carrée d'ordre n à coefficients dans \mathbb{K}). On appelle *multiplicité algébrique* de λ sa multiplicité en tant que zéro du polynôme caractéristique de f (ou de A).

c) PROPOSITION. Avec les données et notations ci-dessus, et en notant E_λ le sous-espace propre associé à la v.p. λ , on a :

$$1 \leq \dim E_\lambda \leq (\text{multiplicité algébrique de } \lambda) \leq n.$$

Preuve : Notons $p = \dim E_\lambda$ et q la multiplicité algébrique de λ . On sait [remarque b) de 1.1.1] que $1 \leq p \leq n$. Soit $\mathcal{C} = (u_1, \dots, u_p)$ une base de E_λ . Comme c'est une famille libre, on peut la compléter en une base $\mathcal{B} = (u_1, \dots, u_p, v_{p+1}, \dots, v_n)$ de E . Notons M la matrice de f par rapport à \mathcal{B} . Pour tout $1 \leq i \leq p$, le vecteur u_i est un vecteur propre de f associé à λ , donc $f(u_i) = \lambda.u_i$. Donc, par définition même de la matrice d'un endomorphisme par rapport à une base, les p premières colonnes de la matrice M sont :

$$M = \begin{pmatrix} \lambda & 0 & 0 & \dots & 0 & * & \dots & * \\ 0 & \lambda & 0 & \dots & 0 & * & \dots & * \\ 0 & 0 & \lambda & \dots & 0 & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & * & \dots & * \\ 0 & 0 & 0 & \dots & 0 & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & * & \dots & * \end{pmatrix} \quad \text{donc} \quad \det(M - x.I_n) = \begin{vmatrix} \lambda-x & 0 & 0 & \dots & 0 & * & \dots & * \\ 0 & \lambda-x & 0 & \dots & 0 & * & \dots & * \\ 0 & 0 & \lambda-x & \dots & 0 & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda-x & * & \dots & * \\ 0 & 0 & 0 & \dots & 0 & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & * & \dots & * \end{vmatrix}.$$

En développant ce déterminant par rapport aux p premières colonnes, on déduit que $P_f(x)$ est de la forme $P_f(x) = (\lambda - x)^p Q(x)$, avec $Q(x) \in \mathbb{K}[x]$ de degré $n - p$. Par définition de la multiplicité algébrique q de $P_f(x)$, on a donc $p \leq q$. \square

d) TERMINOLOGIE. La dimension du sous-espace propre E_λ est parfois appelée la *multiplicité géométrique* de la v.p. λ . On a donc toujours :

$$1 \leq p = (\text{multiplicité géométrique de } \lambda) \leq q = (\text{multiplicité algébrique de } \lambda) \leq n.$$

Une v.p. est *simple* si $q = 1$; sa multiplicité géométrique p est alors 1.

Une v.p. est *double* si $q = 2$; sa multiplicité géométrique p peut alors être 1 ou 2.

Une v.p. est dite *triple* si $q = 3$; sa multiplicité géométrique p peut alors être 1, 2 ou 3.

1.1.5 Quelques exemples explicites de différentes situations possibles

a) PREMIER EXEMPLE. Soit $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$.

On calcule : $P_A(x) = \begin{vmatrix} 1-x & 2 & 0 \\ 0 & 1-x & 1 \\ 1 & 1 & -1-x \end{vmatrix} = (1-x)(x^2 - 2) + 2 = -x^3 + x^2 + 2x = -x(x-2)(x+1)$.

Donc A admet trois valeurs propres simples qui sont 0, 2 et -1 . Déterminons les sous-espaces propres correspondants.

$$(x, y, z) \in E_2 \Leftrightarrow \begin{pmatrix} 1-2 & 2 & 0 \\ 0 & 1-2 & 1 \\ 1 & 1 & -1-2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} -x + 2y = 0 \\ -y + z = 0 \\ x + y - 3z = 0 \end{cases}$$

Donc $E_2 = \{(2y, y, y) \in \mathbb{R}^3 ; y \in \mathbb{R}\}$ est la droite de base (u_2) , où $u_2 = (2, 1, 1)$.

$$(x, y, z) \in E_0 \Leftrightarrow \begin{pmatrix} 1-0 & 2 & 0 \\ 0 & 1-0 & 1 \\ 1 & 1 & -1-0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \Leftrightarrow \begin{cases} x + 2y = 0 \\ y + z = 0 \\ x + y - z = 0 \end{cases}$$

Donc $E_0 = \{(-2y, y, -y) \in \mathbb{R}^3 ; y \in \mathbb{R}\}$ est la droite de base (u_0) , où $u_0 = (2, -1, 1)$.

$$(x, y, z) \in E_{-1} \Leftrightarrow \begin{pmatrix} 1-(-1) & 2 & 0 \\ 0 & 1-(-1) & 1 \\ 1 & 1 & -1-(-1) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \Leftrightarrow \begin{cases} 2x + 2y = 0 \\ 2y + z = 0 \\ x + y = 0 \end{cases}$$

Donc $E_{-1} = \{(y, -y, 2y) \in \mathbb{R}^3 ; y \in \mathbb{R}\}$ est la droite de base (u_{-1}) , où $u_{-1} = (1, -1, 2)$.

b) SECOND EXEMPLE. Soit $A = \begin{pmatrix} 8 & 2 & -2 \\ 2 & 5 & 4 \\ -2 & 4 & 5 \end{pmatrix}$.

On calcule : $P_A(x) = \begin{vmatrix} 8-x & 2 & -2 \\ 2 & 5-x & 4 \\ -2 & 4 & 5-x \end{vmatrix} = \begin{vmatrix} 8-x & 2 & -2 \\ 2 & 5-x & 4 \\ 0 & 9-x & 9-x \end{vmatrix} = (9-x) \begin{vmatrix} 8-x & 2 & -2 \\ 2 & 5-x & 4 \\ 0 & 1 & 1 \end{vmatrix} = (9-x) \begin{vmatrix} 8-x & 2 & -4 \\ 2 & 5-x & x-1 \\ 0 & 1 & 0 \end{vmatrix}$

$$= (9-x)(-1) \begin{vmatrix} 8-x & -4 \\ 2 & x-1 \end{vmatrix} = (9-x)(x^2 - 9x) = -x(x-9)^2.$$

Donc 0 est v.p. simple et 9 est v.p. double. On sait qu'alors E_0 est forcément une droite, mais E_9 peut être soit une droite, soit un plan. Déterminons-les.

$$(x, y, z) \in E_0 \Leftrightarrow \begin{pmatrix} 8-0 & 2 & -2 \\ 2 & 5-0 & 4 \\ -2 & 4 & 5-0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} 8x + 2y - 2z = 0 \\ 2x + 5y + 4z = 0 \\ -2x + 4y + 5z = 0 \end{cases}$$

Après calculs, on trouve que E_0 est la droite de base (u_0) , où $u_0 = (1, -2, 2)$.

$$(x, y, z) \in E_9 \Leftrightarrow \begin{pmatrix} 8-9 & 2 & -2 \\ 2 & 5-9 & 4 \\ -2 & 4 & 5-9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} -x + 2y - 2z = 0 \\ 2x - 4y + 4z = 0 \\ -2x + 4y - 4z = 0 \end{cases}$$

Donc E_9 est le plan d'équation $x - 2y + 2z = 0$, dont une base est (u_9, v_9) , où $u_9 = (2, 1, 0)$ et $v_9 = (2, 0, -1)$. Donc, sur cet exemple, la multiplicité de la v.p. double 9 est égale à la dimension du sous-espace propre correspondant.

c) TROISIÈME EXEMPLE. Soit $A = \begin{pmatrix} 3 & 2 & 4 \\ -1 & 3 & -1 \\ -2 & -1 & -3 \end{pmatrix}$.

$$P_A(x) = \begin{vmatrix} 3-x & 2 & 4 \\ -1 & 3-x & -1 \\ -2 & -1 & -3-x \end{vmatrix} = \begin{vmatrix} -1-x & 2 & 4 \\ 0 & 3-x & -1 \\ x+1 & -1 & -3-x \end{vmatrix} = (x+1) \begin{vmatrix} -1 & 2 & 4 \\ 0 & 3-x & -1 \\ 1 & -1 & -3-x \end{vmatrix} = (x+1) \begin{vmatrix} -1 & 2 & 4 \\ 0 & 3-x & -1 \\ 0 & 1 & 1-x \end{vmatrix} \\ = -(x+1)(x^2 - 4x + 4) = -(x+1)(x-2)^2.$$

Donc -1 est v.p. simple et 2 est v.p. double. On sait qu'alors E_{-1} est forcément une droite, mais E_2 peut être soit une droite, soit un plan. Déterminons-les.

$$(x, y, z) \in E_{-1} \Leftrightarrow \begin{pmatrix} 3-(-1) & 2 & 4 \\ -1 & 3-(-1) & -1 \\ -2 & -1 & -3-(-1) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} 4x + 2y + 4z = 0 \\ -x + 4y - z = 0 \\ -2x - y - 2z = 0 \end{cases}.$$

Après calculs, on trouve que E_{-1} est la droite de base (u_{-1}) , où $u_{-1} = (1, 0, -1)$.

$$(x, y, z) \in E_2 \Leftrightarrow \begin{pmatrix} 3-2 & 2 & 4 \\ -1 & 3-2 & -1 \\ -2 & -1 & -3-2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \Leftrightarrow \begin{cases} x + 2y + 4z = 0 \\ -x + y - z = 0 \\ -2x - y - 5z = 0 \end{cases}.$$

Après calculs, on trouve que E_2 est la droite de base (u_2) , où $u_2 = (2, 1, -1)$. Donc, sur cet exemple, la multiplicité de la v.p. double 2 est strictement supérieure à la dimension du sous-espace propre correspondant.

d) QUATRIÈME EXEMPLE. Soit $A = \begin{pmatrix} -1 & 1 & 1 & 0 \\ -1 & 2 & 1 & -1 \\ 5 & -3 & -2 & 5 \\ 4 & -2 & -2 & 3 \end{pmatrix}$. Après calculs, $P_A(x) = (x+1)(x-1)^3$.

-1 est v.p. simple. On sait qu'alors E_{-1} est forcément une droite. Après calculs, on trouve que E_{-1} est la droite de base (u_{-1}) , où $u_{-1} = (1, 0, 0, -1)$.

$$1 \text{ est v.p. triple. } (x, y, z) \in E_1 \Leftrightarrow \begin{cases} -2x + y + z = 0 \\ -x + y + z - t = 0 \\ 5x - 3y - 3z + 5t = 0 \\ 4x - 2y - 2z + 2t = 0 \end{cases} \Leftrightarrow \dots \Leftrightarrow \begin{cases} x = t = 0 \\ y + z = 0 \end{cases}$$

Donc E_1 est la droite de base (u_1) , où $u_1 = (0, 1, -1, 0)$. Sur cet exemple, le sous-espace propre associé à la v.p. triple 1 est seulement de dimension 1.

e) CINQUIÈME EXEMPLE. Soit $A = \begin{pmatrix} 1 & -1 \\ 3 & -2 \end{pmatrix}$. Après calculs, $P_A(x) = x^2 + x + 1$. Donc A n'admet pas de v.p. dans \mathbb{R} . En revanche, dans \mathbb{C} , elle admet deux v.p. simples distinctes qui sont j et j^2 . On vérifie aisément (par la même méthode que sur les exemples précédents) que E_j est la droite de base (u_j) , où $u_j = (1, 1-j)$ et E_{j^2} est la droite de base (u_{j^2}) , où $u_{j^2} = (1, 1-j^2)$.

1.1.6 Exercices

EXERCICE 1. Un endomorphisme d'un \mathbb{K} -espace vectoriel E de dimension finie admet-il forcément des v.p. dans \mathbb{K} ? (On pourra distinguer suivant que $\mathbb{K} = \mathbb{C}$ ou $\mathbb{K} = \mathbb{R}$).

Un endomorphisme d'un \mathbb{C} -espace vectoriel E qui n'est pas de dimension finie admet-il forcément des v.p. dans \mathbb{C} ? (On pourra considérer dans $E = \mathbb{C}[x]$ l'endomorphisme f qui envoie tout monôme x^i sur x^{i+1}).

EXERCICE 2. Montrer que, si λ est une valeur propre d'un endomorphisme f d'un \mathbb{K} -espace vectoriel E , alors λ^n est valeur propre de l'endomorphisme f^n pour tout $n \in \mathbb{N}^*$; que peut-on dire des sous-espaces propres correspondants?

EXERCICE 3. Deux matrices semblables dans $\mathcal{M}_n(\mathbb{K})$ ont le même polynôme caractéristique (rappeler la preuve); a-t-on la réciproque?

EXERCICE 4. Soit $A \in \mathcal{M}_n(\mathbb{C})$. Montrer que la somme des valeurs propres de A dans \mathbb{C} (chacune répétée autant de fois que sa multiplicité) est égale à la trace de A .

EXERCICE 5. Soient $A \in \mathcal{M}_n(\mathbb{K})$ et P son polynôme caractéristique. Quelle relation a-t-on entre P et le polynôme caractéristique R de la transposée tA .

On suppose de plus que $A \in \text{GL}(n, \mathbb{K})$, et l'on note Q est le polynôme caractéristique de A^{-1} ; quelle relation a-t-on pour tout $\lambda \in \mathbb{K}^*$ entre $Q(\lambda)$ et $P(\lambda^{-1})$?

EXERCICE 6. Soit f un endomorphisme d'un \mathbb{K} -espace vectoriel de dimension finie. Montrer que, si 0 est v.p. de f^m pour un certain entier $m \geq 1$, alors 0 est v.p. de f . Montrer que f est surjectif si et seulement si 0 n'est pas v.p. de f .

EXERCICE 7. Soit f un endomorphisme d'un \mathbb{K} -espace vectoriel. Montrer que f est une homothétie vectorielle de E (c'est-à-dire un multiple scalaire de id_E) si et seulement si tout vecteur non-nul de E est vecteur propre de f .

EXERCICE 8. Soit f un endomorphisme nilpotent d'un \mathbb{K} -espace vectoriel E de dimension finie. Montrer que 0 est v.p. de f , que c'est la seule v.p. de f , et déterminer le polynôme caractéristique de f .

EXERCICE 9. Soient f et g deux endomorphismes d'un \mathbb{C} -espace vectoriel E de dimension finie. Montrer que f et g n'ont aucune valeur propre commune dans \mathbb{C} si et seulement si leurs polynômes caractéristiques P_f et P_g sont premiers entre eux dans $\mathbb{C}[X]$

EXERCICE 10. Soient A et B deux matrices de $\mathcal{M}_n(\mathbb{K})$. Montrer que, si A ou B est inversible, alors AB et BA ont le même polynôme caractéristique.

On suppose maintenant A et B non-inversibles. Montrer que, pour tout $\lambda \in K$, les polynômes $Q_\lambda(x) = \det[(A - xI_n)B - \lambda I_n]$ et $R_\lambda(x) = \det[B(A - xI_n) - \lambda I_n]$ sont égaux dans $\mathbb{K}[x]$. En déduire que dans ce cas aussi AB et BA ont le même polynôme caractéristique.

EXERCICE 11. Soit u un endomorphisme d'un \mathbb{K} -espace vectoriel E de dimension finie. Soit F un sous-espace vectoriel de E , non-nul et distinct de E . On suppose que F est stable par u , et l'on note v l'endomorphisme de F défini par la restriction de u à F . Montrer que le polynôme caractéristique P_v divise le polynôme caractéristique P_u .

Montrer que si H est un supplémentaire de F dans E stable par u , et si l'on note $w \in \text{End } H$ la restriction de u à H , on a $P_u = P_v P_w$.

EXERCICE 12. On note E le \mathbb{R} -espace vectoriel de matrices carrées d'ordre 2 à coefficients réels. On fixe dans E une matrice $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ et on considère l'application f de E dans E qui, à toute matrice $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ associe $f(M) = AM - MA$. Montrer que f est un endomorphisme de E et déterminer ses v.p. et sous-espaces propres.

EXERCICE 13. On considère dans $\mathcal{M}_n(\mathbb{R})$ la matrice : $A = \begin{pmatrix} 0 & \dots & 0 & 1 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 1 \\ 1 & \dots & 1 & 1 \end{pmatrix}$. Quelle est son rang?

En déduire que 0 est v.p. de A et déterminer la dimension du sous-espace propre E_0 . En écrivant le système d'équations que satisfont les coordonnées d'un vecteur propre de A associé à une valeur propre non-nulle λ , déterminer les valeurs possibles λ . Vérifier réciproquement que ce sont bien des valeurs propres.

1.2 Diagonalisation

1.2.1 Endomorphisme diagonalisable, matrice diagonalisable

On se place dans un \mathbb{K} -espace vectoriel E de dimension finie n .

a) LEMME PRÉLIMINAIRE. Soit f un endomorphisme de E . On suppose que f admet s valeurs propres distinctes $\lambda_1, \lambda_2, \dots, \lambda_s$ dans \mathbb{K} . Alors, le sous-espace somme $F = E_{\lambda_1} + E_{\lambda_2} + \dots + E_{\lambda_s}$ des sous-espaces propres de f est une somme directe : $F = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \dots \oplus E_{\lambda_s}$.

Preuve : Il s'agit de montrer que tout vecteur de F se décompose de façon unique en une somme $u_1 + u_2 + \dots + u_s$ avec $u_1 \in E_{\lambda_1}, u_2 \in E_{\lambda_2}, \dots, u_s \in E_{\lambda_s}$. Par définition de F , l'existence d'une telle décomposition est claire. Le problème est seulement l'unicité. Quitte à faire la différence membre à membre de deux telles décompositions, on est ramené à montrer simplement que, pour tous $u_1 \in E_{\lambda_1}, u_2 \in E_{\lambda_2}, \dots, u_s \in E_{\lambda_s}$, on a :

$$\text{si } u_1 + u_2 + \dots + u_s = 0_E, \text{ alors } u_1 = u_2 = \dots = u_s = 0_E.$$

Supposons donc $u_1 + u_2 + \dots + u_s = 0_E$ avec $u_i \in E_{\lambda_i}, 1 \leq i \leq s$. En appliquant f , on a :

$$\lambda_1.u_1 + \lambda_2.u_2 + \dots + \lambda_s.u_s = f(u_1) + f(u_2) + \dots + f(u_s) = f(u_1 + u_2 + \dots + u_s) = f(0_E) = 0_E.$$

Par ailleurs, en multipliant par λ_1 , on a aussi $\lambda_1 u_1 + \lambda_1 u_2 + \dots + \lambda_1 u_s = 0_E$, d'où par différence membre à membre : $\sum_{i=2}^s (\lambda_i - \lambda_1).u_i = 0_E$. On réitère. En appliquant f , il vient :

$$\sum_{i=2}^s (\lambda_i - \lambda_1)\lambda_i.u_i = \sum_{i=2}^s (\lambda_i - \lambda_1).f(u_i) = f\left(\sum_{i=2}^s (\lambda_i - \lambda_1).u_i\right) = f(0_E) = 0_E.$$

Par ailleurs, en multipliant par λ_2 , on obtient $\sum_{i=2}^s (\lambda_i - \lambda_1)\lambda_2.u_i = 0_E$, d'où par différence membre à membre : $\sum_{i=3}^s (\lambda_i - \lambda_1)(\lambda_i - \lambda_2).u_i = 0_E$. De proche en proche, on parvient ainsi à : $(\lambda_s - \lambda_1)(\lambda_s - \lambda_2) \dots (\lambda_s - \lambda_{s-1}).u_s = 0_E$. Comme les λ_i sont deux à deux distincts, on conclut $u_s = 0_E$. La somme de départ $u_1 + \dots + u_s = 0_E$ se réduit donc à $u_1 + \dots + u_{s-1} = 0_E$. En réitérant le même raisonnement, on obtient successivement $u_{s-1} = 0_E, \dots; u_2 = 0_E, u_1 = 0_E$. Ce qui achève la preuve. \square

b) DÉFINITIONS. Un endomorphisme f de E est dit *diagonalisable* lorsqu'il existe une base \mathcal{C} de E telle que la matrice de f dans la base \mathcal{C} soit une matrice diagonale D . Une matrice carrée d'ordre n est dite *diagonalisable* lorsqu'elle est semblable à une matrice diagonale.

Si A désigne la matrice de f dans une base quelconque \mathcal{B} de E , on a donc :

$$(\text{l'endomorphisme } f \text{ est diagonalisable}) \Leftrightarrow (\text{la matrice } A \text{ est diagonalisable}).$$

Dans ce cas, en notant $A = M_{\mathcal{B}}(f)$, $D = M_{\mathcal{C}}(f)$ diagonale, et $P = \text{Pass}_{\mathcal{B} \rightarrow \mathcal{C}} \in \text{GL}(n, \mathbb{K})$, on a :

$$A = PDP^{-1}$$

Remarquons qu'alors les coefficients diagonaux de la matrice D sont exactement les v.p. de f , chacune apparaissant un nombre de fois égal à sa multiplicité algébrique : il suffit pour le voir de développer $\det(D - x.I_n) = P_D(x) = P_A(x) = P_f(x)$.

En particulier :

- la trace de A est la somme des valeurs propres de A , chacune comptée avec sa multiplicité ;
- le déterminant A est le produit des valeurs propres de A , chacune comptée avec sa multiplicité.

1.2.2 Condition nécessaire et suffisante de diagonalisabilité

On se place dans un \mathbb{K} -espace vectoriel E de dimension finie n .

a) THÉORÈME. Soit f un endomorphisme de E . Les conditions suivantes sont équivalentes :

- (i) f est diagonalisable ;
- (ii) il existe une base \mathcal{C} de E qui est formée de vecteurs propres de f ;
- (iii) E est la somme directe des sous-espaces propres de f ;
- (iv) le polynôme caractéristique $P_f(x)$ admet n zéros (comptés avec leur multiplicité), et la multiplicité algébrique de chaque valeur propre λ de f est égale à la dimension du sous-espace propre E_λ correspondant ;
- (v) le polynôme caractéristique $P_f(x)$ est produit de n facteurs de degré 1, et la multiplicité algébrique de chaque valeur propre λ de f est égale à la dimension du sous-espace propre E_λ correspondant.

Preuve. Supposons que l'on a (i). Il existe donc une base \mathcal{C} de E telle que la matrice D de f par rapport à \mathcal{C} est diagonale. Certains de ses coefficients peuvent être égaux : on note $\lambda_1, \lambda_2, \dots, \lambda_s$ les valeurs distinctes de ces coefficients diagonaux (donc $1 \leq s \leq n$). Quitte à permuter les vecteurs de \mathcal{C} , on peut sans restriction supposer qu'ils apparaissent dans l'ordre :

$$D = M_{\mathcal{C}}(f) = \begin{pmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_1 \end{bmatrix} & & \\ & \begin{bmatrix} \lambda_2 & & \\ & \ddots & \\ & & \lambda_2 \end{bmatrix} & & \\ & & \ddots & & \\ & & & \begin{bmatrix} \lambda_s & & \\ & \ddots & \\ & & \lambda_s \end{bmatrix} \end{pmatrix}$$

(tous les autres coefficients hors de la diagonale étant des zéros).

Pour tout $1 \leq i \leq s$, notons q_i le nombre de fois où apparaît le coefficient λ_i sur la diagonale. Donc la base \mathcal{C} de E est la réunion $\mathcal{C}_1 \cup \dots \cup \mathcal{C}_s$ où chaque \mathcal{C}_i est formé de q_i vecteurs de \mathcal{C} vérifiant $f(u) = \lambda_i \cdot u$. Chaque λ_i est donc une v.p. de f , et la dimension du sous-espace propre associé E_{λ_i} est q_i . Cet entier q_i est aussi la multiplicité algébrique de la v.p. λ_i , puisque qu'il suffit de développer le déterminant $\det(D - x.I_n)$ pour obtenir $P_f(x) = \det(D - x.I_n) = (-1)^n (x - \lambda_1)^{q_1} (x - \lambda_2)^{q_2} \dots (x - \lambda_s)^{q_s}$. Donc (v) est vérifiée.

• L'implication (v) \Rightarrow (iv) est claire.

• Supposons (iv). On a d'une part : $P_f(x) = (-1)^n (x - \lambda_1)^{q_1} (x - \lambda_2)^{q_2} \dots (x - \lambda_s)^{q_s}$, où $\lambda_1, \lambda_2, \dots, \lambda_s$ sont les v.p. distinctes de f , et d'autre part : $\dim E_{\lambda_i} = q_i$ pour tout $1 \leq i \leq s$.

Considérons le ss-e.v. somme $F = E_{\lambda_1} + E_{\lambda_2} + \dots + E_{\lambda_s}$. D'après le lemme a), on a en fait $F = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \dots \oplus E_{\lambda_s}$, donc $\dim F = \dim E_{\lambda_1} + \dim E_{\lambda_2} + \dots + \dim E_{\lambda_s} = q_1 + q_2 + \dots + q_s = \deg P_f(x) = n = \dim E$. Donc $E = F = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \dots \oplus E_{\lambda_s}$, ce qui prouve (iii).

• Il résulte de (iii) que l'on peut former une base \mathcal{C} de E en prenant la réunion disjointe : $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2 \cup \dots \cup \mathcal{C}_s$, où chaque \mathcal{C}_i est une base de E_{λ_i} (donc formée de q_i vecteurs propres associés à la v.p. λ_i) pour tout $1 \leq i \leq s$. Ceci prouve (ii).

- Supposons (ii). Avec les notations précédentes, pour tout vecteur $u \in \mathcal{C}$, il existe un unique indice $1 \leq i \leq s$ tel que $u \in \mathcal{C}_i$, et l'on a alors $f(u) = \lambda_i \cdot u$. La matrice de f par rapport à la base \mathcal{C} est ainsi diagonale par construction.

On a montré que : (i) \Rightarrow (v) \Rightarrow (iv) \Rightarrow (iii) \Rightarrow (ii) \Rightarrow (i), ce qui achève la preuve. \square

b) COROLLAIRE. Soit f un endomorphisme de E . Si f admet n v.p. distinctes (chacune étant nécessairement une v.p. simple), alors f est diagonalisable.

Preuve. On applique tout ce qui précède au cas $n = s$, d'où $q_i = 1$ pour tout $1 \leq i \leq n$. Dans ce cas, chaque sous-espace propre est une droite. \square

c) ATTENTION! Ce corollaire ne donne qu'une condition suffisante pour qu'un endomorphisme soit diagonalisable; elle n'est nullement nécessaire, comme l'exprime le théorème général a) et comme le montre le deuxième des exemples ci-dessous.

d) EXEMPLES. Reprenons les exemples traités en 1.1.5.

- Dans l'exemple (a), la matrice A considérée est diagonalisable, d'après le corollaire ci-dessus. Une base de vecteurs propres est $\mathcal{C} = (u_2, u_{-1}, u_0)$, donc :

$$A = PDP^{-1}, \text{ avec } P = \begin{pmatrix} 2 & 1 & 2 \\ 1 & -1 & -1 \end{pmatrix} \in \text{GL}(3, \mathbb{R}) \text{ et } D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

- Dans l'exemple (b), la matrice A considérée est diagonalisable, d'après le (iv) du théorème c) ci-dessus. Une base de vecteurs propres est $\mathcal{C} = (u_0, u_9, v_9)$, donc :

$$A = PDP^{-1}, \text{ avec } P = \begin{pmatrix} 1 & 2 & 2 \\ -2 & 1 & 0 \\ 2 & 0 & -1 \end{pmatrix} \in \text{GL}(3, \mathbb{R}) \text{ et } D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix}.$$

- Dans l'exemple (c), la matrice A considérée n'est pas diagonalisable, car le sous-espace propre associé à la valeur propre double 2 est seulement de dimension 1.

- De même dans l'exemple (d), la matrice A considérée n'est pas diagonalisable, car le sous-espace propre associé à la valeur propre triple 1 est seulement de dimension 1.

- Enfin, dans l'exemple (e), A n'est pas diagonalisable dans \mathbb{R} , mais est diagonalisable dans \mathbb{C} (car admet deux v.p. simples distinctes complexes). Une base de vecteurs propres est $\mathcal{C} = (u_j, u_{j^2})$, donc : $A = PDP^{-1}$, avec $P = \begin{pmatrix} 1 & 1 \\ 1-j & 1-j^2 \end{pmatrix} \in \text{GL}(2, \mathbb{C})$ et $D = \begin{pmatrix} j & 0 \\ 0 & j^2 \end{pmatrix}$.

1.2.3 La question de la trigonalisabilité

On se place dans un \mathbb{K} -espace vectoriel E de dimension finie n .

On a traité au paragraphe précédent la situation la plus favorable : celle où la matrice carrée A considérée est semblable à une matrice diagonale. Ce n'est pas toujours le cas! On traite donc dans ce paragraphe une situation moins favorable, mais utile quand même pour les applications pratiques, celle où A est seulement semblable à une matrice triangulaire.

a) DÉFINITIONS. Un endomorphisme f de E est dit *trigonalisable* lorsqu'il existe une base \mathcal{C} de E telle que la matrice de f dans la base \mathcal{C} soit une matrice triangulaire (supérieure) T . Une matrice carrée d'ordre n est dite *trigonalisable* lorsqu'elle est semblable à une matrice triangulaire (supérieure).

Si A désigne la matrice de f dans une base quelconque \mathcal{B} de E , on a donc :

$$(\text{l'endomorphisme } f \text{ est trigonalisable}) \Leftrightarrow (\text{la matrice } A \text{ est trigonalisable}).$$

Dans ce cas, avec $A = M_{\mathcal{B}}(f)$, $T = M_{\mathcal{C}}(f)$ triangulaire, et $P = \text{Pass}_{\mathcal{B} \rightarrow \mathcal{C}} \in \text{GL}(n, \mathbb{K})$, on a :

$$A = PTP^{-1}$$

b) THÉORÈME. Soit f un endomorphisme de E . Les conditions suivantes sont équivalentes :

- (i) f est trigonalisable sur \mathbb{K} ;
- (ii) le polynôme caractéristique $P_f(x)$ est produit de n facteurs de degré 1 sur \mathbb{K} .

Preuve. Supposons que l'on a (i). Il existe donc une base \mathcal{C} de E telle que la matrice T de f par rapport à \mathcal{C} est triangulaire supérieure. Donc $P_f(x) = P_T(x)$. Comme T est triangulaire, il est clair que $P_T(x) = (-1)^n(x - \alpha_1)(x - \alpha_2) \dots (x - \alpha_n)$. Donc (ii) est vérifié (et les coefficients diagonaux de T sont les v.p. de f).

• Pour la réciproque, on raisonne par récurrence sur n . Le résultat est trivial si $n = 1$. Supposons-le vrai pour tout endomorphisme de tout \mathbb{K} -e.v. de dimension $n - 1$. Prenons f un endomorphisme d'un \mathbb{K} -e.v. de dimension n tel que $P_f(x)$ est produit de n facteurs de degré 1 sur \mathbb{K} . Il admet donc au moins un zéro dans \mathbb{K} . Soit donc λ un zéro de $P_f(x)$ dans \mathbb{K} , c'est-à-dire une v.p. de f . Soit u un vecteur propre de f associé à λ . On a $u \neq 0$, donc d'après 1.5.4.b, on peut compléter u en une base $\mathcal{B} = (u, v_2, v_3, \dots, v_n)$ de E . Notons $\mathcal{C} = (v_2, v_3, \dots, v_n)$, qui est une famille libre (car sous-famille de \mathcal{B}). Notons $F = \text{Vect } \mathcal{C}$. D'après 1.5.2.b, \mathcal{C} est une base de F , de sorte que $\dim F = n - 1$. D'après 1.5.3.b, F est un supplémentaire dans E de la droite vectorielle Δ de base (u) .

Comme $f(u) = \lambda.u$, la matrice de f par rapport à \mathcal{B} est de la forme :

$$A = \begin{pmatrix} \lambda & \alpha_2 & \alpha_3 \dots & \alpha_n \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}$$

où $B \in \mathcal{M}_{n-1}(\mathbb{K})$ et $\alpha_j \in \mathbb{K}$ pour tout $2 \leq j \leq n$.

Appelons g l'endomorphisme de F dont B est la matrice par rapport à la base \mathcal{C} (attention, ce n'est pas la restriction de f à F , car F n'est a priori pas stable par f). On a donc :

$$f(v_j) = \alpha_j.u + g(v_j) \quad \text{pour tout } 2 \leq j \leq n.$$

On calcule alors à partir de A : $P_A(f) = \det(A - x.I_n) = (\lambda - x) \det(B - x.I_{n-1})$, c'est-à-dire : $P_f(x) = (\lambda - x)P_g(x)$. Comme on a supposé que $P_f(x)$ est produit de n facteurs de degré 1, il en résulte que $P_g(x)$ est produit de $n - 1$ facteurs de degré 1. On peut appliquer à g l'hypothèse de récurrence, il existe une base $\mathcal{C}' = (w_2, w_3, \dots, w_n)$ de F telle que la matrice T de g par rapport à la base \mathcal{C}' est triangulaire supérieure. Comme $E = \Delta \oplus F$, il résulte de 1.5.3.b que $\mathcal{B}' = (u, w_2, w_3, \dots, w_n)$ est une base de E . Chaque w_i (pour $2 \leq i \leq n$) est c.l. de v_2, \dots, v_n , donc :

$$f(w_i) = \beta_i.u + g(w_i) \quad \text{pour tout } 2 \leq i \leq n,$$

où le scalaire β_i est c.l. des α_j . Donc la matrice de f par rapport à \mathcal{C}' est :

$$A' = \begin{pmatrix} \lambda & \beta_2 & \beta_3 \dots & \beta_n \\ 0 & & & \\ \vdots & & T & \\ 0 & & & \end{pmatrix},$$

qui est triangulaire ; ce qui achève la preuve. □

c) COROLLAIRE (cas où $\mathbb{K} = \mathbb{C}$). Tout endomorphisme d'un \mathbb{C} -espace vectoriel de dimension finie est trigonalisable sur \mathbb{C} ; toute matrice carrée à coefficients complexes est trigonalisable.

Preuve. On sait que tout polynôme à coefficients dans \mathbb{C} se décompose en un produit de facteurs de degré 1 sur \mathbb{C} ; on peut donc appliquer le théorème précédent. □

d) REMARQUE IMPORTANTE. Soit A une matrice carrée d'ordre n à coefficients dans \mathbb{R} . On peut toujours la considérer comme un élément de $\mathcal{M}_n(\mathbb{C})$. Elle est donc toujours trigonalisable sur \mathbb{C} , mais cela ne signifie bien sûr pas qu'elle est trigonalisable sur \mathbb{R} .

Si par exemple $P_A(x) = -(x-2)(x^2+1)$, A n'est pas trigonalisable sur \mathbb{R} mais l'est sur \mathbb{C} puisque $P_A(x) = -(x-2)(x-i)(x+i)$.

e) EXEMPLES. Reprenons les exemples vus en 1.1.5. On a déjà vu que les exemples (a) et (b) correspondent à des matrices diagonalisables sur \mathbb{R} , et que l'exemple e) est diagonalisable sur \mathbb{C} (il n'est sur \mathbb{R} ni diagonalisable ni trigonalisable puisque qu'il n'admet pas de v.p. réelles.)

Dans les exemples (c) et (d), les matrices considérées ne sont pas diagonalisables, mais elles sont trigonalisables sur \mathbb{R} puisque le polynôme caractéristique se décompose en produit de facteurs de degré 1. Détaillons :

- Reprenons la matrice A de l'exemple (c) de 1.1.5. Soit f l'endomorphisme de \mathbb{R}^3 dont A est la matrice par rapport à la base canonique. Le sous-espace propre associé à la valeur simple -1 est la droite de base (u_{-1}) avec $u_{-1} = (1, 0, -1)$. Le sous-espace propre associé à la valeur double 2 est la droite de base (u_2) avec $u_2 = (2, 1, -1)$. Quelle que soit la façon de compléter la famille libre (u_{-1}, u_2) en une base \mathcal{C} par l'adjonction d'un vecteur quelconque qui n'est pas combinaison linéaire de u_{-1} et u_2 , la matrice de f par rapport à la base \mathcal{C} est triangulaire. Prenons par exemple $\mathcal{C} = (u_{-1}, u_2, e_1)$ avec $e_1 = (1, 0, 0)$. Notons P la matrice de passage de la base canonique à la base \mathcal{C} . Donc

$$P = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ -1 & -1 & 0 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 0 & -1 & -1 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}, \quad T = \text{Mat}_{\mathcal{C}}(f) = P^{-1}AP = \begin{pmatrix} -1 & 0 & 3 \\ 0 & 2 & -1 \\ 0 & 0 & 2 \end{pmatrix}.$$

- Reprenons la matrice A de l'exemple (d) de 1.1.5. Soit f l'endomorphisme de \mathbb{R}^4 dont A est la matrice par rapport à la base canonique. Le sous-espace propre associé à la valeur simple -1 est la droite de base (u_{-1}) avec $u_{-1} = (1, 0, 0, -1)$. Le sous-espace propre associé à la valeur triple 1 est la droite de base (u_1) avec $u_1 = (0, 1, -1, 0)$. On sait que l'on peut compléter (u_{-1}, u_1) en une base $\mathcal{C} = (u_{-1}, u_1, v, w)$ de \mathbb{R}^4 telle que la matrice de f par rapport à la base \mathcal{C} est triangulaire. Mais ici, le choix des deux vecteurs v, w n'est pas indifférent ; il existe une méthode générale pour les déterminer (réduction de Jordan, hors de notre programme). Contentons-nous de noter ici que, pour $v = (1, 0, 2, 0)$ et $w = (0, 0, 1, 1)$, on a :

$$P = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 2 & 1 \\ -1 & 0 & 0 & 1 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 2 & -1 & -1 & 1 \\ 0 & 1 & 0 & 0 \\ -1 & 1 & 1 & -1 \\ 2 & -1 & -1 & 2 \end{pmatrix}, \quad T = \text{Mat}_{\mathcal{C}}(f) = P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

1.2.4 Exercices

EXERCICE 1 Pour quelles valeurs des paramètres réels les matrices suivantes sont-elles diagonalisables ?

$$A = \begin{pmatrix} 2 & 3 & -6 \\ m-6 & m-7 & -m+12 \\ m-3 & m-3 & -m+5 \end{pmatrix}, \quad B = \begin{pmatrix} a & 0 & b \\ 0 & a+b & 0 \\ b & 0 & a \end{pmatrix}, \quad C = \begin{pmatrix} a^2 & ab & ab & b^2 \\ ab & a^2 & b^2 & ab \\ ab & b^2 & a^2 & ab \\ b^2 & ab & ab & a^2 \end{pmatrix}, \quad E = \begin{pmatrix} q & 1 & 1 \\ 1 & q & 1 \\ 1 & 1 & q \end{pmatrix}.$$

EXERCICE 2 Soient $a, b \in \mathbb{C}$. On considère dans $\mathcal{M}_n(\mathbb{C})$ la matrice $A = (a_{i,j})_{1 \leq i, j \leq n}$ définie par : $a_{i,j} = b$ si $i \neq j$ et $a_{i,i} = a$. Montrer que A est diagonalisable.

EXERCICE 3. Soit u un endomorphisme d'un \mathbb{K} -espace vectoriel E de dimension finie. On suppose que $\text{Im}(u - \text{id}_E) \cap \text{Im}(u + \text{id}_E) = \{0\}_E$. Montrer que u est diagonalisable.

EXERCICE 4. Soient E un \mathbb{K} -espace vectoriel de dimension finie n et $f \in \text{End } E$.

– Montrer que le sous-ensemble de $\text{End } E$ formé des endomorphismes u tels que $f \circ u = 0$ est un sous-espace vectoriel de $\text{End } E$ de dimension $n \times \dim \text{Ker } f$.

– Soit σ l'application de $\text{End } E$ dans $\text{End } E$ définie par $\sigma(u) = f \circ u$ pour tout $u \in \text{End } E$. Montrer que σ est un endomorphisme de $\text{End } E$, qui admet les mêmes v.p. que f , et qui est diagonalisable si et seulement si f l'est.

EXERCICE 5. Soient $n \geq 1$ un entier fixé et $E = \mathbb{R}_n[x]$ le \mathbb{R} -espace vectoriel des polynômes de degré inférieur ou égal à n . Soient a, b deux réels distincts fixés. Soit f l'application $E \rightarrow E$ définie par $f(P(x)) = (x - a)(x - b)P'(x) - nxP(x)$ pour tout $P \in E$, où P' désigne le polynôme dérivé de P .

– Montrer que f est un endomorphisme de E .

– Montrer que, si un polynôme $P \in E$ est un vecteur propre de f , alors il est de la forme $P(x) = (x - a)^k(x - b)^h Q(x)$ avec $k, h \in \mathbb{N}$ et $Q \in E$. Montrer que Q est nécessairement une constante. En déduire les v.p. et les vecteurs propres de f . Conclure que f est diagonalisable.

EXERCICE 6. On fixe un entier $p \geq 2$. On note U_p le groupe des racines p -ièmes de l'unité dans \mathbb{C} . Soit u un endomorphisme d'un \mathbb{C} -espace vectoriel E vérifiant $u^p = \text{id}_E$.

– Montrer que l'ensemble des v.p. de u est inclus dans U_p .

– Montrer que, pour tout $\omega \in U_p$, l'image de l'endomorphisme $v_\omega = \frac{1}{p} \sum_{j=0}^{p-1} \omega^{-j} u^j$ est égale à $\text{Ker}(u - \omega \text{id}_E)$.

– En notant $\omega_k = \exp(2ik\pi/p)$ pour tout entier $1 \leq k \leq p$, montrer que, pour tout entier $1 \leq n \leq p - 1$, on a $\sum_{k=1}^p (\omega_k)^{-n} = 0$.

– Calculer $\sum_{k=1}^p v_{\omega_k}(x)$ pour tout $x \in E$. Conclure que u est diagonalisable.

1.3 Applications de la diagonalisation

On développe dans le corps du texte trois applications classiques de la diagonalisation : le calcul des puissances d'une matrice carrée, le calcul du terme général de suites définies par des relations récurrences linéaires (qui est une conséquence du calcul précédent), et la résolution des systèmes différentiels linéaires d'ordre 1 à coefficients constants.

Plusieurs autres applications en algèbre et en géométrie figurent dans les exercices en fin de paragraphe.

1.3.1 Application au calcul des puissances d'une matrice

a) PRINCIPE. Soit $A \in \mathcal{M}_n(\mathbb{K})$. Supposons que A est diagonalisable. Comme on l'a vu en 1.2.1, on a $A = PDP^{-1}$ avec D diagonale et $P \in \text{GL}(n, \mathbb{K})$. On calcule alors :

$$A^m = PDP^{-1}PDP^{-1} \dots PDP^{-1} = PD^mP^{-1} \text{ pour tout entier } m \geq 1.$$

Puisque D^m est elle-même diagonale, ceci permet de calculer aisément A^m .

b) EXEMPLE. Soient $a \in \mathbb{R}^*$ et A la matrice $\begin{pmatrix} 0 & a & a^2 \\ a^{-1} & 0 & a \\ a^{-2} & a^{-1} & 0 \end{pmatrix}$. On calcule $P_A(x) = -(x+1)^2(x-2)$.

On montre que le sous-espace propre E_{-1} est le plan d'équation $x + ay + a^2z = 0$ et que le sous-espace propre E_2 est la droite d'équations $x = ay = a^2z$. Donc A est diagonalisable et l'on a : $A = PDP^{-1}$ avec

$$D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad P = \begin{pmatrix} -a & -a^2 & a^2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad P^{-1} = \frac{1}{3} \begin{pmatrix} -1/a & 2 & -a \\ -1/a^2 & -1/a & 2 \\ 1/a^2 & 1/a & 1 \end{pmatrix}.$$

On calcule alors :

$$A^n = PD^nP^{-1} = \frac{1}{3} \begin{pmatrix} -a & -a^2 & a^2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} (-1)^n & 0 & 0 \\ 0 & (-1)^n & 0 \\ 0 & 0 & 2^n \end{pmatrix} \begin{pmatrix} -1/a & 2 & -a \\ -1/a^2 & -1/a & 2 \\ 1/a^2 & 1/a & 1 \end{pmatrix},$$

pour obtenir que : $A^n = \alpha_n A + \beta_n I_3$ où l'on a posé : $\alpha_n = \frac{1}{3}[2^n - (-1)^n]$ et $\beta_n = \frac{1}{3}[2(-1)^n + 2^n]$.

On a en particulier : $A^2 = A + 2I_3$,

d'où $\frac{1}{2}(A - I_3)A = I_3$, ce qui prouve que A est inversible et que $A^{-1} = \frac{1}{2}(A - I_3)$.

1.3.2 Application à l'étude de suites récurrentes

a) PRINCIPE. Les contextes peuvent être divers, mais le principe est toujours d'exprimer des relations de récurrence de type linéaire entre les termes successifs d'une ou plusieurs suites sous forme d'une égalité matricielle entre vecteurs, faisant intervenir une matrice carrée A , de sorte que le calcul de A^n (suivant la méthode que l'on vient de voir lorsque A est diagonalisable) permet de déterminer l'expression du terme général de la suite ou des suites considérées.

On verra diverses formes de relations en exercice ; limitons-nous ici à l'exemple suivant :

b) EXEMPLE (suite récurrente avec une relation linéaire d'ordre 2). Soient a et b deux réels fixés. Soit $(u)_{n \geq 0}$ la suite réelle définie par la donnée des valeurs de u_0 et de u_1 , et par la relation :

$$u_{n+2} = au_n + bu_{n+1} \quad \text{pour tout } n \geq 0.$$

Posons $X_n = \begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix}$ dans \mathbb{R}^2 , $A = \begin{pmatrix} 0 & 1 \\ a & b \end{pmatrix}$, de sorte que la relation s'écrit $X_{n+1} = AX_n$ puisque

$$\begin{pmatrix} u_{n+1} \\ u_{n+2} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ a & b \end{pmatrix} \begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix}.$$

On calcule $P_A(x) = x^2 - bx - a$.

- Si $P_A(x)$ admet deux racines réelles distinctes λ_1 et λ_2 dans \mathbb{R} . Alors λ_1 et λ_2 sont les deux v.p. distinctes de A . Donc A est diagonalisable. Il existe une matrice inversible P telle que $A = PDP^{-1}$ où $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$. On peut écrire $X_{n+1} = AX_n$ sous la forme $X_{n+1} = PDP^{-1}X_n$, ou encore, en définissant $Y_n = \begin{pmatrix} v_n \\ w_n \end{pmatrix} = P^{-1}X_n$, sous la forme : $Y_{n+1} = DY_n$.

On en tire : $Y_n = D^n Y_0$ pour tout $n \geq 0$, donc : $v_n = \lambda_1^n v_0$ et $w_n = \lambda_2^n v_1$. Avec $X_n = PY_n$, on conclut que la suite (u_n) est combinaison linéaire des suites géométriques (λ_1^n) et (λ_2^n) .

- Si $P_A(x)$ admet une racine double réelle λ_0 , qui est donc v.p. double de A , on remplace la diagonalisation de A par une trigonalisation de A et l'on montre de même que la suite (u_n) est combinaison linéaire de la suite géométrique (λ_0^n) et de la suite $(n\lambda_0)^n$.

- Si $P_A(x)$ admet deux racines complexes conjuguées $\lambda_1 = \rho e^{i\theta}$ et $\lambda_2 = \rho e^{-i\theta}$, on applique le premier cas avec une diagonalisation dans \mathbb{C} . On obtient que $\mu \rho^n e^{ni\theta} + \eta \rho^n e^{-i\theta n}$ avec μ, η quelconques dans \mathbb{C} . En posant $\alpha = \mu + \eta$ et $\beta = i(\eta - \mu)$, on déduit que les solutions sont les suites $u_n = \rho^n (\alpha \cos(n\theta) + \beta \sin(n\theta))$ avec α, β quelconques dans \mathbb{R} .

A noter que cette méthode permet de traiter de cas de relations de récurrence linéaires d'ordre $p \geq 2$, en réduisant (diagonalisant dans les bons cas) une matrice carrée d'ordre p .

1.3.3 Application à la résolution de systèmes différentiels linéaires

a) PREMIER EXEMPLE (*cas diagonalisable avec v.p. simples*).

On cherche à déterminer tous les triplets de fonctions réelles $x(t), y(t), z(t)$ dérivables sur \mathbb{R} solutions du système différentiel :

$$(S) \begin{cases} x'(t) = -2x(t) - 2y(t) + 2z(t) \\ y'(t) = 3x(t) + 5y(t) \\ z'(t) = x(t) + 2y(t) + 3z(t) \end{cases}.$$

On note matriciellement ce système sous la forme :

$$X'(t) = AX(t), \text{ avec } X(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}, X'(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix}, A = \begin{pmatrix} -2 & -2 & 2 \\ 3 & 5 & 0 \\ 1 & 2 & 3 \end{pmatrix}.$$

Suivant les méthodes vues ci-dessus, on montre que A est diagonalisable et on réalise la diagonalisation :

$$P_A(x) = -(x+1)(x-2)(x-5), \text{ et } A = PDP^{-1} \text{ où } D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix} \text{ et } P = \begin{pmatrix} 2 & 1 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

On pose : $U(t) = \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix} = P^{-1}X(t)$; d'où : $U'(t) = \begin{pmatrix} u'(t) \\ v'(t) \\ w'(t) \end{pmatrix} = P^{-1}X'(t)$.

On a alors :

$$(S) \Leftrightarrow X'(t) = AX(t) \Leftrightarrow PU'(t) = APU(t) \Leftrightarrow U'(t) = P^{-1}APU(t) \Leftrightarrow U'(t) = DU(t).$$

On est donc ramené à résoudre le système (beaucoup plus simple) :

$$\begin{pmatrix} u'(t) \\ v'(t) \\ w'(t) \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix}, \text{ c'est-à-dire : } \begin{cases} u'(t) = -u(t) \\ v'(t) = 2v(t) \\ w'(t) = 5w(t) \end{cases}.$$

D'après le cours d'analyse, on sait que les solutions de ce dernier système sont :

$$u(t) = \alpha e^{-t}, v(t) = \beta e^{2t}, w(t) = \gamma e^{5t}, \text{ pour } \alpha, \beta, \gamma \text{ décrivant } \mathbb{R}.$$

En revenant à $X(t) = PU(t) = \begin{pmatrix} 2 & 1 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \alpha e^{-t} \\ \beta e^{2t} \\ \gamma e^{5t} \end{pmatrix}$, on conclut que les solutions du système différentiel (S) sont :

$$\begin{cases} x(t) = 2\alpha e^{-t} + \beta e^{2t} \\ y(t) = -\alpha e^{-t} - \beta e^{2t} + \gamma e^{5t} \\ z(t) = \beta e^{2t} + \gamma e^{5t} \end{cases}, \text{ avec } \alpha, \beta, \gamma \in \mathbb{R}.$$

Remarque : algébriquement, les trois fonctions vectorielles $\mathbb{R} \rightarrow \mathbb{R}^3$ définies par $t \mapsto \begin{pmatrix} 2e^{-t} \\ -e^{-t} \\ 0 \end{pmatrix}$,

$t \mapsto \begin{pmatrix} e^{2t} \\ -e^{2t} \\ e^{2t} \end{pmatrix}$, $t \mapsto \begin{pmatrix} 0 \\ e^{5t} \\ e^{5t} \end{pmatrix}$, forment une base de l'espace vectoriel des solutions de (S).

b) DEUXIÈME EXEMPLE (*cas diagonalisable avec v.p. multiples*).

On cherche à déterminer tous les triplets de fonctions réelles $x(t), y(t), z(t)$ dérivables sur \mathbb{R} solutions du système différentiel :

$$(S) \begin{cases} x'(t) = 2x(t) + 3y(t) - 6z(t) \\ y'(t) = -6x(t) - 7y(t) + 12z(t) \\ z'(t) = -3x(t) - 3y(t) + 5z(t) \end{cases}.$$

On note matriciellement ce système sous la forme :

$$X'(t) = AX(t), \text{ avec } X(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}, X'(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix}, A = \begin{pmatrix} 2 & 3 & -6 \\ -6 & -7 & 12 \\ -3 & -3 & 5 \end{pmatrix}.$$

Suivant les méthodes vues ci-dessus, on montre que A est diagonalisable et on réalise la diagonalisation :

$$P_A(x) = -(x+1)^2(x-2), \text{ et } A = PDP^{-1} \text{ où } D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \text{ et } P = \begin{pmatrix} -1 & 1 & 1 \\ 2 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

On pose : $U(t) = \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix} = P^{-1}X(t)$; d'où : $U'(t) = \begin{pmatrix} u'(t) \\ v'(t) \\ w'(t) \end{pmatrix} = P^{-1}X'(t)$.

On a alors :

$$(S) \Leftrightarrow X'(t) = AX(t) \Leftrightarrow PU'(t) = APU(t) \Leftrightarrow U'(t) = P^{-1}APU(t) \Leftrightarrow U'(t) = DU(t).$$

On est donc ramené à résoudre le système (beaucoup plus simple) :

$$\begin{pmatrix} u'(t) \\ v'(t) \\ w'(t) \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix}, \quad \text{c'est-à-dire : } \begin{cases} u'(t) = 2u(t) \\ v'(t) = -v(t) \\ w'(t) = -w(t) \end{cases}.$$

D'après le cours d'analyse, on sait que les solutions de ce dernier système sont :

$$u(t) = \alpha e^{2t}, \quad v(t) = \beta e^{-t}, \quad w(t) = \gamma e^{-t}, \quad \text{pour } \alpha, \beta, \gamma \text{ décrivant } \mathbb{R}.$$

En revenant à $X(t) = PU(t) = \begin{pmatrix} -1 & 1 & 1 \\ 2 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha e^{2t} \\ \beta e^{-t} \\ \gamma e^{-t} \end{pmatrix}$, on conclut que les solutions du système différentiel (S) sont :

$$\begin{cases} x(t) = -\alpha e^{2t} + (\beta + \gamma)e^{-t} \\ y(t) = 2\alpha e^{2t} + (-\beta + \gamma)e^{-t} \\ z(t) = \alpha e^{2t} + \gamma e^{-t} \end{cases}, \quad \text{avec } \alpha, \beta, \gamma \in \mathbb{R}.$$

Remarque : algébriquement, les trois fonctions vectorielles $\mathbb{R} \rightarrow \mathbb{R}^3$ définies par $t \mapsto \begin{pmatrix} -e^{2t} \\ 2e^{2t} \\ e^{2t} \end{pmatrix}$, $t \mapsto \begin{pmatrix} e^{-t} \\ -e^{-t} \\ 0 \end{pmatrix}$, $t \mapsto \begin{pmatrix} e^{-t} \\ e^{-t} \\ e^{-t} \end{pmatrix}$, forment une base de l'espace vectoriel des solutions de (S).

c) TROISIÈME EXEMPLE (cas diagonalisable sur \mathbb{C} mais pas sur \mathbb{R}).

On cherche à déterminer tous les triplets de fonctions réelles $x(t), y(t), z(t)$ dérivables sur \mathbb{R} solutions du système différentiel :

$$(S) \begin{cases} x'(t) = 2y(t) - 2z(t) \\ y'(t) = -2x(t) + z(t) \\ z'(t) = 2x(t) - y(t) \end{cases}.$$

On note matriciellement ce système sous la forme :

$$X'(t) = AX(t), \quad \text{avec } X(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}, \quad X'(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 2 & -2 \\ -2 & 0 & 1 \\ 2 & -1 & 0 \end{pmatrix}.$$

On calcule : $P_A(x) = -x(x^2 + 9)$.

• On résoud d'abord sur \mathbb{C} .

$P_A(x) = -x(x - 3i)(x + 3i)$, donc A est diagonalisable. On a :

$$A = PDP^{-1} \text{ où } D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3i & 0 \\ 0 & 0 & -3i \end{pmatrix} \text{ et } P = \begin{pmatrix} 1 & 4 & 4 \\ 2 & -1+3i & -1-3i \\ 2 & -1-3i & -1+3i \end{pmatrix}.$$

On raisonnant comme sur les 2 exemples a) et b) précédents, on montre que les solutions du système différentiel (S) sont :

$$\begin{cases} x(t) = \alpha + 4\beta e^{3it} + 4\gamma e^{-3it} \\ y(t) = 2\alpha + (-1+3i)\beta e^{3it} + (-1-3i)\gamma e^{-3it} \\ z(t) = 2\alpha + (-1-3i)\beta e^{3it} + (-1+3i)\gamma e^{-3it} \end{cases}, \quad \text{avec } \alpha, \beta, \gamma \in \mathbb{C}.$$

• On cherche ensuite les solutions réelles.

D'après ce qui précède, les 3 fonctions vectorielles $\mathbb{R} \rightarrow \mathbb{C}^3$ définies par :

$$u_0 : t \mapsto \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad v : t \mapsto \begin{pmatrix} 4e^{3it} \\ (-1+3i)e^{3it} \\ (-1-3i)e^{3it} \end{pmatrix}, \quad \bar{v} : t \mapsto \begin{pmatrix} 4e^{-3it} \\ (-1-3i)e^{-3it} \\ (-1+3i)e^{-3it} \end{pmatrix},$$

forment une base de l'espace vectoriel des solutions complexes de (S).

On en déduit qu'une base de l'espace vectoriel des solutions réelles de (S) est formée des 3 fonctions vectorielles $\mathbb{R} \rightarrow \mathbb{C}^3$ définies par : $u_0, w_1 = \frac{1}{2}(v + \bar{v}), w_2 = \frac{1}{2i}(v - \bar{v})$. En d'autres termes :

$$w_1 : t \mapsto \begin{pmatrix} 4 \cos 3t \\ -\cos 3t - 3 \sin 3t \\ -\cos 3t + 3 \sin 3t \end{pmatrix}, \quad w_2 : t \mapsto \begin{pmatrix} 4 \sin 3t \\ -\sin 3t + 3 \cos 3t \\ -\sin 3t - 3 \cos 3t \end{pmatrix}.$$

On conclut que les solutions réelles du système différentiel (S) sont :

$$\begin{cases} x(t) = \alpha + 4\beta \cos 3t + 4\gamma \sin 3t \\ y(t) = 2\alpha + (-\beta + 3\gamma) \cos 3t + (-3\beta - \gamma) \sin 3t \\ z(t) = 2\alpha + (-\beta - 3\gamma) \cos 3t + (3\beta - \gamma) \sin 3t \end{cases}, \quad \text{avec } \alpha, \beta, \gamma \in \mathbb{R}.$$

d) QUATRIÈME EXEMPLE (*cas triangularisable non diagonalisable*).

On cherche à déterminer tous les triplets de fonctions réelles $x(t), y(t), z(t)$ dérivables sur \mathbb{R} solutions du système différentiel :

$$(S) \begin{cases} x'(t) = -y(t) - 2z(t) \\ y'(t) = x(t) + z(t) \\ 5z'(t) = -6x(t) + 8y(t) \end{cases}.$$

On note matriciellement ce système sous la forme :

$$X'(t) = AX(t), \quad \text{avec } X(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}, \quad X'(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix}, \quad A = \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & 1 \\ -\frac{6}{5} & \frac{8}{5} & 0 \end{pmatrix}.$$

On calcule : $P_A(x) = -(x-1)^2(x+2)$.

$\lambda = -2$ est v.p. simple ; le sous-espace propre associé est la droite dirigée par $v_0 = \begin{pmatrix} 3 \\ -4 \\ 5 \end{pmatrix}$.

$\lambda = 1$ est v.p. double ; le sous-espace propre associé est la droite dirigée par $v_1 = \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix}$.

Donc A n'est pas diagonalisable. On sait que si l'on complète la famille libre (v_0, v_1) en une base en lui adjoignant un vecteur v_2 , on a :

$$A = PTP^{-1} \quad \text{avec } T = \begin{pmatrix} -2 & 0 & * \\ 0 & 1 & * \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et } P = \begin{pmatrix} 3 & 3 & * \\ -4 & 1 & * \\ 5 & -2 & * \end{pmatrix}.$$

On verra au chapitre suivant une raison théorique pour laquelle on peut toujours choisir v_2 de telle sorte que $T = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. Pour l'instant, contentons-nous de vérifier par le calcul que, si l'on choisit $v_2 = \begin{pmatrix} -1 \\ -2 \\ 0 \end{pmatrix}$, on a $\mathcal{C} = (v_0, v_1, v_2)$ base de \mathbb{R}^3 , et :

$$A = PTP^{-1} \quad \text{avec } T = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et } P = \text{Pass}_{\mathcal{B} \rightarrow \mathcal{C}} = \begin{pmatrix} 3 & 3 & -1 \\ -4 & 1 & -2 \\ 5 & -2 & 0 \end{pmatrix}.$$

On pose : $U(t) = \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix} = P^{-1}X(t)$; d'où : $U'(t) = \begin{pmatrix} u'(t) \\ v'(t) \\ w'(t) \end{pmatrix} = P^{-1}X'(t)$. Donc :

$$(S) \Leftrightarrow X'(t) = AX(t) \Leftrightarrow PU'(t) = APU(t) \Leftrightarrow U'(t) = P^{-1}APU(t) \Leftrightarrow U'(t) = TU(t).$$

On est ainsi ramené à résoudre le système (beaucoup plus simple) :

$$\begin{pmatrix} u'(t) \\ v'(t) \\ w'(t) \end{pmatrix} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix}, \quad \text{c'est-à-dire : } \begin{cases} u'(t) = -2u(t) \\ v'(t) = v(t) + w(t) \\ w'(t) = w(t) \end{cases}.$$

D'après le cours d'analyse, on sait que les solutions des équations (1) et (3) sont $u(t) = \alpha e^{-2t}$ et $w(t) = \gamma e^t$ pour α, γ décrivant \mathbb{R} . La seconde équation devient $v'(t) = v(t) + \gamma e^t$, dont la solution générale est $v(t) = (\gamma t + \beta)e^t$.

En revenant à $X(t) = PU(t) = \begin{pmatrix} 3 & 3 & -1 \\ -4 & 1 & -2 \\ 5 & -2 & 0 \end{pmatrix} \begin{pmatrix} \alpha e^{-2t} \\ (\gamma t + \beta)e^t \\ \gamma e^t \end{pmatrix}$, on conclut que les solutions du système différentiel (S) sont :

$$\begin{cases} x(t) = 3\alpha e^{-2t} + (3\gamma t + 3\beta - \gamma)e^t \\ y(t) = -4\alpha e^{-2t} + (\gamma t + \beta - 2\gamma)e^t \\ z(t) = 5\alpha e^{-2t} + (-2\gamma t - 2\beta)e^t \end{cases}, \quad \text{avec } \alpha, \beta, \gamma \in \mathbb{R}.$$

1.3.4 Exercices

EXERCICE 1. Diagonaliser $A = \begin{pmatrix} -4 & -6 & 0 \\ 3 & 5 & 0 \\ 3 & 6 & 5 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$ et calculer A^n pour tout $n \in \mathbb{N}$.

EXERCICE 2. On considère deux suites de nombres réels $(u_n)_{n \geq 0}$ et $(v_n)_{n \geq 0}$ telles que, pour tout $n \geq 0$, on ait : $u_{n+1} = -10u_n - 28v_n$ et $v_{n+1} = 6u_n + 16v_n$. En diagonalisant une matrice de $\mathcal{M}_2(\mathbb{R})$ adaptée, calculer les termes généraux u_n et v_n en fonction de u_0, v_0 et n .

EXERCICE 3. Soient $(u_n), (v_n), (w_n)$ trois suites de réels définies par leurs premiers termes $u_0 = 1, v_0 = -1, w_0 = 0$, et les relations de récurrence :

$$u_{n+1} = -4u_n - 6v_n, \quad v_{n+1} = 3u_n + 5v_n, \quad w_{n+1} = 3u_n + 6v_n + 5w_n.$$

– En notant $X_n = \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix}$ pour tout $n \in \mathbb{N}$, déterminer la matrice $A \in \mathcal{M}_3(\mathbb{R})$ telle que $X_{n+1} = AX_n$. Diagonaliser A .

– Calculer A^n pour tout $n \in \mathbb{N}$. En déduire l'expression de u_n, v_n et w_n en fonction de n .

– Montrer que A est inversible et calculer A^{-1} ; la formule donnant A^n trouvée à la question précédente reste-t-elle valable pour les entiers $n < 0$?

EXERCICE 4. Déterminer les fonctions réelles $x(t), y(t), z(t)$ dérivables sur \mathbb{R} solutions des systèmes différentiels :

$$\begin{cases} x'(t) = 3x(t) - 5y(t) \\ y'(t) = -2y(t) \\ z'(t) = 5x(t) - 5y(t) - 2z(t) \end{cases} \quad \begin{cases} x'(t) = 3x(t) - 2y(t) \\ y'(t) = y(t) \\ z'(t) = 2x(t) - 2y(t) + z(t) \end{cases} \quad \begin{cases} x'(t) = 9x(t) + 5y(t) + 5z(t) \\ y'(t) = -5x(t) - y(t) - 5z(t) \\ z'(t) = -5x(t) - 5y(t) - z(t) \end{cases}$$

EXERCICE 5. Soit (E) la courbe du plan d'équation : $x^2 - xy + y^2 - \frac{1}{2} = 0$, par rapport à un repère orthonormé. En diagonalisant la matrice $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$, montrer que (E) est une ellipse centrée en l'origine.

EXERCICE 6. Pour tout $\alpha \in \mathbb{R}$, on considère dans $M_3(\mathbb{R})$ la matrice :

$$A_\alpha = \frac{1}{4} \begin{pmatrix} 6\alpha+1 & 2\alpha-1 & -\sqrt{2}(2\alpha-1) \\ 2\alpha-1 & 6\alpha+1 & \sqrt{2}(2\alpha-1) \\ -\sqrt{2}(2\alpha-1) & \sqrt{2}(2\alpha-1) & 4\alpha+2 \end{pmatrix}.$$

On note ϕ_α l'endomorphisme du \mathbb{R} -espace vectoriel \mathbb{R}^3 dont la matrice par rapport à la base canonique $\mathcal{B} = (e_1, e_2, e_3)$ de \mathbb{R}^3 est A_α .

– Déterminer une base \mathcal{B}' de \mathbb{R}^3 formée de vecteurs propres de A_1 . Montrer que, pour tout $\alpha \in \mathbb{R}$, la matrice de ϕ_α dans la base \mathcal{B}' est une matrice diagonale D_α .

– En déduire, par un calcul simple, que $A_\alpha A_\beta = A_{2\alpha\beta}$ pour tous $\alpha, \beta \in \mathbb{R}$, et que $A_{1/2} = I_3$. Pour quelles valeurs de α la matrice A_α est-elle inversible ?

– En déduire que l'ensemble $G = \{A_\alpha ; \alpha \in \mathbb{R}^*\}$ est un sous-groupe de $GL(3, \mathbb{R})$.

EXERCICE 7. Soit A une matrice carrée d'ordre n à coefficients dans \mathbb{C} admettant n valeurs propres distinctes.

– Montrer qu'il existe une matrice inversible $P \in \text{GL}(n, \mathbb{C})$ telle que $P^{-1}AP$ soit une matrice diagonale D , et en déduire que la famille $\mathcal{B} = \{I_n, A, A^2, \dots, A^{n-1}\}$ est libre dans $\mathcal{M}_n(\mathbb{C})$.

– On note \mathcal{C} le centralisateur de A dans $\mathcal{M}_n(\mathbb{C})$, c'est-à-dire l'ensemble des matrices M de $\mathcal{M}_n(\mathbb{C})$ qui vérifient $AM = MA$. Montrer que, pour toute $M \in \mathcal{C}$, la matrice $P^{-1}MP$ est diagonale.

– Dédurre des questions précédentes que \mathcal{C} est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{C})$ dont une base est \mathcal{B} .

Chapitre 2

Espaces vectoriels euclidiens

L'objectif de ce chapitre est d'introduire sur un espace vectoriel E dont le corps des scalaires est \mathbb{R} une structure algébrique complémentaire, celle de produit scalaire. Cette notion permet de disposer dans E de la notion d'orthogonalité (entre vecteurs, entre sous-espaces...) et d'une forme de "longueur" ou de "distance" définie à partir de la norme associée au produit scalaire. Un \mathbb{R} -espace vectoriel ainsi muni d'un produit scalaire s'appelle un espace préhilbertien réel. De tels espaces en dimension infinie ont une grande importance en analyse fonctionnelle. Néanmoins, conformément au programme de ce cours, on se concentre dans ce chapitre sur le cas où E est de dimension finie (on dit alors que E est un espace vectoriel euclidien). On développe dans ce contexte l'étude des endomorphismes de E qui préservent le produit scalaire, à la fois dans son aspect algébrique matriciel (où l'on retrouve des considérations du chapitre précédent sur la diagonalisation) et sur ses applications en géométrie concrète du plan ou de l'espace.

Dans tout ce chapitre, E désigne un espace vectoriel sur \mathbb{R} , de dimension finie.

2.1 Produit scalaire et norme euclidienne

2.1.1 Notion de produit scalaire

a) DÉFINITION. Soit E un \mathbb{R} -espace vectoriel. On appelle *produit scalaire* sur E toute application $\langle \cdot | \cdot \rangle$ de $E \times E$ dans \mathbb{R} qui est bilinéaire, symétrique, définie positive.

Précisons le sens de ces expressions :

Dire que $\langle \cdot | \cdot \rangle$ est *bilinéaire* signifie qu'elle est linéaire par rapport à chacune des deux variables, c'est-à-dire que :

$$(i) \quad \langle \lambda x + \mu y | z \rangle = \lambda \langle x | z \rangle + \mu \langle y | z \rangle \quad \text{pour tous } x, y, z \in E, \lambda, \mu \in \mathbb{R}.$$

$$(i') \quad \langle x | \lambda y + \mu z \rangle = \lambda \langle x | y \rangle + \mu \langle x | z \rangle \quad \text{pour tous } x, y, z \in E, \lambda, \mu \in \mathbb{R}.$$

Dire que $\langle \cdot | \cdot \rangle$ est *symétrique* signifie que :

$$(ii) \quad \langle x | y \rangle = \langle y | x \rangle \quad \text{pour tous } x, y \in E.$$

Dire que $\langle \cdot | \cdot \rangle$ est *définie positive* signifie que :

$$(iii) \quad \langle x | x \rangle \geq 0 \quad \text{pour tout } x \in E,$$

$$(iv) \quad \text{quel que soit } x \in E, \quad (\langle x | x \rangle = 0 \text{ si et seulement si } x = 0_E).$$

Comme il est clair que, si on a (ii), les conditions (i) et (i') sont équivalentes, on peut aussi bien dire qu'un produit scalaire est une application $E \times E \rightarrow \mathbb{R}$ vérifiant (i), (ii), (iii) et (iv).

b) EXEMPLES.

• Soit $E = \mathbb{R}^n$. On définit un produit scalaire sur E , dit produit scalaire *canonique*, en posant :

$$\langle x | y \rangle = \sum_{i=1}^n x_i y_i \quad \text{pour tous } x = (x_1, \dots, x_n) \text{ et } y = (y_1, \dots, y_n) \text{ dans } E.$$

• Soit $E = \mathcal{M}_n(\mathbb{R})$. On définit un produit scalaire sur E en posant :

$$\langle A | B \rangle = \text{trace}(A \times {}^t B) \quad \text{pour toutes } A, B \in E.$$

• Soit $E = \mathcal{C}([0, 1], \mathbb{R})$. On définit un produit scalaire sur E en posant :

$$\langle f | g \rangle = \int_0^1 f(t)g(t) dt \quad \text{pour toutes } f, g \in E.$$

c) DÉFINITIONS. Un \mathbb{R} -espace vectoriel muni d'un produit scalaire et qui est de dimension finie est appelé un *espace vectoriel euclidien*.

C'est le cas des deux premiers exemples ci-dessus. Le troisième exemple n'en est pas un car il n'est pas de dimension finie (on parle alors simplement d'espace préhilbertien réel, mais on ne les étudiera pas dans ce cours).

Soit E un \mathbb{R} -espace vectoriel euclidien. On appelle *norme euclidienne* associée au produit scalaire $\langle \cdot | \cdot \rangle$ l'application $\| \cdot \|$ de E dans \mathbb{R}_+ définie par :

$$\|x\| = \sqrt{\langle x | x \rangle} \quad \text{pour tout } x \in E.$$

d) REMARQUE. Il résulte immédiatement de cette définition et des propriétés du produit scalaire que, pour tout $x \in E$, on a :

$$\|\lambda x\| = |\lambda| \|x\| \quad \text{pour tout } \lambda \in \mathbb{R}, \quad \|x\| \geq 0, \quad \|x\| = 0 \Leftrightarrow x = 0_E$$

e) PROPOSITION (Inégalité de Cauchy-Schwarz). *Dans tout espace vectoriel euclidien E , on a :*

$$|\langle x | y \rangle| \leq \|x\| \|y\| \quad \text{pour tous } x, y \in E,$$

avec égalité si et seulement si x et y sont colinéaires.

Preuve. On fixe $x, y \in E$. L'inégalité étant trivialement vraie si x ou y est nul, on peut supposer dans la suite que x et y sont non-nuls.

On considère l'application $f : \mathbb{R} \rightarrow \mathbb{R}$ défini par $f(t) = \|tx + y\|^2$. Il est clair que $f(t) \geq 0$ pour tout $t \in \mathbb{R}$. Or $f(t) = \langle tx + y | tx + y \rangle = \|x\|^2 t^2 + 2\langle x | y \rangle t + \|y\|^2$ est polynomiale en t de degré 2 (car le coefficient dominant $\|x\|^2$ est non-nul). Le fait que $f(t)$ soit positif pour tout $t \in \mathbb{R}$ implique donc que son discriminant Δ est négatif. Comme $\Delta = 4\langle x | y \rangle^2 - 4\|x\|^2 \|y\|^2$, on obtient l'inégalité voulue.

Si l'on a l'égalité, alors $\Delta = 0$, donc f admet un zéro double $t_0 \in \mathbb{R}$. On a ainsi $\|t_0 x + y\|^2 = 0$, donc $t_0 x + y = 0_E$, donc x et y sont colinéaires. Réciproquement, si x et y sont colinéaires, on a $x = 0_E$ (auquel cas l'égalité est claire) ou $y = tx$ pour un certain réel t ; dans ce cas $\langle x | y \rangle = \langle x | tx \rangle = t\|x\|^2$ et $\|x\| \|y\| = \|x\| \|tx\| = |t| \|x\|^2$; d'où l'égalité voulue. \square

f) PROPOSITION (Inégalité triangulaire). *Dans tout espace vectoriel euclidien E , on a :*

$$\|x + y\| \leq \|x\| + \|y\| \quad \text{pour tous } x, y \in E,$$

avec égalité si et seulement si x et y sont colinéaires.

Preuve. Pour tous $x, y \in E$, on a :

$$\|x + y\|^2 = \langle x + y | x + y \rangle = \|x\|^2 + 2\langle x | y \rangle + \|y\|^2 \leq \|x\|^2 + 2|\langle x | y \rangle| + \|y\|^2,$$

d'où en appliquant l'inégalité de Cauchy-Schwarz :

$$\|x + y\|^2 \leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2,$$

ce qui donne l'inégalité voulue. L'égalité se produit lorsque $\langle x | y \rangle = \|x\| \|y\|$, ce qui achève la preuve en appliquant le cas d'égalité de la proposition précédente. \square

g) LEMME (identités de polarisation). *Dans tout espace vectoriel euclidien E , on a pour tous vecteurs $x, y \in E$:*

$$\langle x | y \rangle = \frac{1}{2} (\|x + y\|^2 - \|x\|^2 - \|y\|^2) = \frac{1}{2} (\|x\|^2 + \|y\|^2 - \|x - y\|^2) = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2).$$

Preuve. Evidente, laissée à titre d'exercice. \square

h) REMARQUES. Les trois propriétés du d) ci-dessus et l'inégalité triangulaire f) se traduisent en disant que $\|\cdot\|$ satisfait les axiomes d'une norme sur E (voir chapitre suivant). On remarquera aussi que les preuves des propriétés d) à g) n'utilisent pas le fait que E est de dimension finie, et restent donc vraies dans le cadre plus général d'espaces vectoriels sur \mathbb{R} munis d'un produit scalaire sans hypothèse de dimension (espaces préhilbertiens au programme du second semestre).

2.1.2 Orthogonalité

a) DÉFINITIONS. Soit E un espace vectoriel euclidien. Deux vecteurs x et y de E sont dits *orthogonaux* lorsque leur produit scalaire est nul. On note alors $x \perp y$.

$$x \perp y \Leftrightarrow \langle x | y \rangle = 0.$$

D'après la propriété (iv) de 2.1.1, le vecteur nul est le seul vecteur qui est orthogonal à lui-même :

$$x \perp x \Leftrightarrow x = 0_E.$$

Deux parties A et B de E sont dites *orthogonales* lorsque tout vecteur de A est orthogonal à tout vecteur de B . On note alors $A \perp B$.

$$A \perp B \Leftrightarrow \forall x \in A, \forall y \in B, \langle x | y \rangle = 0.$$

Pour toute partie non vide A de E , on appelle *orthogonal de A* , noté A^\perp , l'ensemble des vecteurs de E qui sont orthogonaux à tous les vecteurs de A . Il est facile de vérifier (le faire en exercice!) que c'est un sous-espace vectoriel de E .

$$A^\perp = \{x \in E; \forall y \in A, \langle x | y \rangle = 0\} \text{ sous-espace vectoriel de } E.$$

Il résulte immédiatement de cette définition que, pour deux parties non-vides A et B de E :

$$\left[A \perp B \Leftrightarrow A \subseteq B^\perp \Leftrightarrow B \subseteq A^\perp \right] \quad \text{et} \quad \left[A \subseteq B \Rightarrow B^\perp \subseteq A^\perp \right].$$

Enfin, on déduit aisément des propriétés (i) à (iv) de 2.1.1 (le faire en exercice!) que le vecteur nul 0_E est orthogonal à tous les vecteurs de E , et que c'est le seul à avoir cette propriété :

$$\{0_E\}^\perp = E \quad \text{et} \quad E^\perp = \{0_E\}.$$

On considère dans la suite l'orthogonal de parties de E qui sont des sous-espaces vectoriels. L'orthogonal F^\perp d'un tel sous-espace F prend parfois le nom de *supplémentaire orthogonal* de F , ceci en raison du théorème fondamental suivant.

b) THÉORÈME. *Soit E un espace vectoriel euclidien. Pour tout sous-espace vectoriel F de E :*

$$F^\perp \text{ est un sous-espace vectoriel de } E, \quad E = F \oplus F^\perp \quad \text{et} \quad (F^\perp)^\perp = F.$$

Preuve. Le résultat étant clair si $F = \{0_E\}$ ou $F = E$, on suppose que F est un sous-espace non-nul et distinct de E . Il est clair que $F \cap F^\perp = \{0_E\}$ (car on aurait sinon un élément non-nul de F orthogonal à tous les éléments de F , donc à lui-même, ce qui est impossible).

Considérons une base (e_1, \dots, e_p) de F . Construisons l'application $f : E \rightarrow \mathbb{R}^p$ qui, à tous vecteur x de E associe le p -uplet $f(x) = (\langle x | e_1 \rangle, \langle x | e_2 \rangle, \dots, \langle x | e_p \rangle)$. D'après la propriété (i) de 2.1.1, f est linéaire. Son noyau $\text{Ker } f$ est formé des vecteurs $x \in E$ tels que $f(x)$ est nul dans \mathbb{R}^p , c'est-à-dire tels que $\langle x | e_j \rangle = 0$ pour tout $1 \leq j \leq p$. En d'autres termes, un vecteur $x \in E$ est dans $\text{Ker } f$ si et seulement s'il est orthogonal à tous les vecteurs de la base (e_1, e_2, \dots, e_p) de F , et donc finalement par linéarité à tout vecteur de F . On a ainsi prouvé que $\text{Ker } f = F^\perp$. En particulier, F^\perp est un sous-espace vectoriel de E .

Dès lors, en notant $n = \dim E$, on tire de la formule du rang $\dim \text{Im } f + \dim \text{Ker } f = n$ que $\dim F^\perp = n - \dim \text{Im } f$. Comme $\text{Im } f \subseteq \mathbb{R}^p$, on a $\dim \text{Im } f \leq p$, d'où $\dim F^\perp \geq n - p$.

Par ailleurs, $\dim(F + F^\perp) = \dim F + \dim F^\perp - \dim(F \cap F^\perp) = \dim F + \dim F^\perp = p + \dim F^\perp$. Comme évidemment $\dim(F + F^\perp) \leq n$, il vient $\dim F^\perp \leq n - p$. Finalement : $\dim F^\perp = n - p$. Ceci prouve que $E = F \oplus F^\perp$.

En appliquant ce qui précède à F^\perp au lieu de F , on obtient $\dim(F^\perp)^\perp = n - \dim F^\perp = n - (n - p) = p$. Ainsi $\dim(F^\perp)^\perp = \dim F$. Comme il est clair (par définition de l'orthogonal) que $F \subseteq (F^\perp)^\perp$, on conclut à l'égalité. \square

c) VOCABULAIRE. Soit $\mathcal{B} = (e_1, e_2, \dots, e_n)$ une base dans un espace vectoriel euclidien E .

On dit que la base \mathcal{B} est une *base orthogonale* lorsque tous les vecteurs e_i sont orthogonaux deux à deux (c'est-à-dire $\langle e_i | e_j \rangle = 0$ pour tous $1 \leq i \neq j \leq n$).

On dit que la base \mathcal{B} est une *base orthonormale* lorsque c'est une base orthogonale et que, de plus, tous les vecteurs e_i sont de norme égale à 1 (c'est-à-dire $\langle e_i | e_j \rangle = \delta_{i,j}$ pour tous $1 \leq i, j \leq n$).

Il est clair que, si on a une base orthogonale (e_1, e_2, \dots, e_n) , il suffit de définir $e'_j = \frac{1}{\|e_j\|} e_j$ pour tout $1 \leq j \leq n$ pour obtenir une base orthonormale $(e'_1, e'_2, \dots, e'_n)$.

d) THÉORÈME. *Dans tout espace vectoriel euclidien non-nul, il existe des bases orthonormales.*

Preuve. On raisonne par récurrence sur la dimension $n \geq 1$ de E .

Si $n = 1$, il suffit de choisir un vecteur non-nul quelconque x de E et de poser $e_1 = \frac{1}{\|x\|} x$; il est clair que (e_1) est une base orthonormale de E .

Supposons la propriété vraie pour tous les espaces vectoriel euclidiens de dimension $n - 1$ et considérons un espace vectoriel euclidien E de dimension n . Choisissons un vecteur non-nul quelconque x de E et posons $e_n = \frac{1}{\|x\|} x$. Considérons la droite F de base (e_n) . D'après le théorème précédent, on a $E = F \oplus F^\perp$, avec $\dim F^\perp = n - 1$. En appliquant l'hypothèse de récurrence à F^\perp , il existe une base orthonormale $(e_1, e_2, \dots, e_{n-1})$ de F^\perp . Il est clair que $(e_1, e_2, \dots, e_{n-1}, e_n)$ est alors une base orthonormale de E . \square

e) PROPOSITION (expression du produit scalaire dans une base orthonormale). Soit E un espace vectoriel euclidien de dimension $n \geq 1$, muni d'une base orthonormale \mathcal{B} . Pour tous vecteurs x et y de E , de composantes respectives (x_1, \dots, x_n) et (y_1, \dots, y_n) dans la base \mathcal{B} , on a :

$$\langle x | y \rangle = \sum_{j=1}^n x_j y_j \quad \text{et} \quad \|x\| = \sqrt{\sum_{j=1}^n x_j^2}.$$

Preuve. On a : $\langle x | y \rangle = \langle \sum_{i=1}^n x_i e_i | \sum_{j=1}^n y_j e_j \rangle = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \langle e_i | e_j \rangle$, d'où la première égalité puisque $\langle e_i | e_j \rangle = 0$ si $i \neq j$ et $\langle e_i | e_i \rangle = 1$. La seconde découle de la première. \square

f) REMARQUE (*projection orthogonale et symétrie orthogonale*). Il résulte aussi du théorème b) que l'on peut appliquer les notions de projection et de symétrie vues en 1.1.2 du chapitre 1 dans le cas où la direction de projection (ou symétrie) est orthogonale au sous-espace sur lequel on projette (ou par rapport auquel on symétrise).

Plus précisément, si F est un sous-espace vectoriel d'un \mathbb{R} -espace vectoriel euclidien E , avec donc $E = F \oplus F^\perp$, on appelle *projection orthogonale* sur F la projection sur F parallèlement à F^\perp , et *symétrie orthogonale* par rapport à F la symétrie par rapport à F parallèlement à F^\perp .

Ainsi, tout vecteur $x \in E$ se décomposant de façon unique sous la forme $x = y + z$ avec $y \in F$ et $z \in F^\perp$, la projection orthogonale p sur F et la symétrie orthogonale s par rapport à F sont respectivement définies par $p(x) = y$ et $s(x) = y - z$. On a :

$$p \circ p = p, \quad \text{Im } p = F = E_1, \quad \text{Ker } p = F^\perp = E_0.$$

$$s \circ s = \text{id}_E, \quad \text{Im } s = E, \quad \text{Ker } s = \{0_E\}, \quad E_1 = F, \quad E_{-1} = F^\perp.$$

exemple : $\dim E = 2, \dim F = 1$	exemple : $\dim E = 3, \dim F = 2$	exemple : $\dim E = 3, \dim F = 1$

2.1.3 Représentation des formes linéaires

a) RAPPELS. Si E est un espace vectoriel sur \mathbb{R} , les applications linéaires de E dans \mathbb{R} sont appelées les *formes linéaires* sur \mathbb{R} . Elles forment un \mathbb{R} -espace vectoriel que l'on note E^* et que l'on appelle l'*espace dual* de E . Si E est de dimension finie n , alors E^* est de dimension finie n .

b) PROPOSITION. Soit E un espace vectoriel euclidien.

- (i) Pour tout vecteur $a \in E$, l'application $\varphi_a : E \rightarrow \mathbb{R}$ définie par $\varphi_a(x) = \langle a | x \rangle$ pour tout $x \in E$ est une forme linéaire sur E .
- (ii) Réciproquement, pour toute forme linéaire $f \in E^*$, il existe un unique vecteur $a \in E$ tel que $f = \varphi_a$.

Preuve. Soient $x, y \in E$ et $\lambda, \mu \in \mathbb{R}$ quelconques. Par bilinéarité du produit scalaire, on a : $\varphi_a(\lambda x + \mu y) = \langle a | \lambda x + \mu y \rangle = \lambda \langle a | x \rangle + \mu \langle a | y \rangle = \lambda \varphi_a(x) + \mu \varphi_a(y)$, ce qui montre que φ_a est linéaire et prouve (i).

Pour (ii), considérons l'application $\Phi : E \rightarrow E^*$ définie par $\Phi(a) = \varphi_a$ pour tout $a \in E$. Elle est linéaire : en effet, quels que soient $\lambda, \mu \in \mathbb{R}$, $a, b \in E$, on a pour tout $x \in E$: $\varphi_{\lambda a + \mu b}(x) = \langle \lambda a + \mu b | x \rangle = \lambda \langle a | x \rangle + \mu \langle b | x \rangle = \lambda \varphi_a(x) + \mu \varphi_b(x)$, ce qui prouve bien que $\Phi(\lambda a + \mu b) = \lambda \Phi(a) + \mu \Phi(b)$. Déterminons son noyau. Dire qu'un vecteur a est dans $\text{Ker } \Phi$ signifie que $\varphi_a = 0_{E^*}$, ce qui équivaut à dire que, pour tout $x \in E$, on a $\varphi_a(x) = 0$, c'est-à-dire $\langle a | x \rangle = 0$. Parce que $\langle \cdot | \cdot \rangle$ est un produit scalaire, on sait que cela n'est possible (pour tout $x \in E$) que lorsque $a = 0_E$. Ainsi $\text{Ker } \Phi = \{0_E\}$. Donc l'application $\Phi : E \rightarrow E^*$ est injective. Comme $\dim E = \dim E^*$, cela équivaut à la bijectivité de Φ . Donc, pour tout $f \in E^*$, il existe un unique $a \in E$ telle que $f = \Phi(a)$, ce qui est l'assertion (ii) voulue. \square

2.1.4 Familles orthogonales de polynômes

La plupart des applications des produits scalaires que l'on verra dans la suite de ce chapitre sont de nature géométrique. On insère néanmoins ici une application de nature un peu différente, touchant à des espaces vectoriels de polynômes.

a) RAPPELS. On se place dans l'espace vectoriel $\mathbb{R}[x]$ des polynômes à coefficients réels en une indéterminée x . Pour tout entier $n \geq 0$, on note $\mathbb{R}_n[x]$ le sous-espace vectoriel des polynômes de degré inférieur ou égal à n . Comme tout élément p de $\mathbb{R}_n[x]$ s'écrit de façon unique $p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$ avec $a_i \in \mathbb{R}$ pour tout $0 \leq i \leq n$, il est clair que la famille $\mathcal{B} = (1, x, x^2, \dots, x^{n-1}, x^n)$ est une base de $\mathbb{R}_n[x]$. On l'appelle la base canonique de $\mathbb{R}_n[x]$. En particulier, on retient que $\dim \mathbb{R}_n[x] = n + 1$.

Comment reconnaître qu'une famille donnée de $n + 1$ polynômes est une base de $\mathbb{R}_n[x]$? Une condition suffisante pratique et usuelle est la suivante :

Si p_0, p_1, \dots, p_n sont des polynômes tels que chaque p_i soit de degré égal à i (pour tout $0 \leq i \leq n$), alors la famille (p_0, p_1, \dots, p_n) est une base de $\mathbb{R}_n[x]$.

En effet : il suffit pour le démontrer d'observer que la matrice de cette famille \mathcal{C} par rapport à la base canonique \mathcal{B} est triangulaire supérieure, avec sur la diagonale des coefficients tous non-nuls (ce sont les coefficients du terme de plus haut degré de chaque p_i). Il en résulte que cette matrice est inversible, ce qui équivaut à la propriété pour \mathcal{C} d'être une base de $\mathbb{R}_n[x]$.

b) LEMME. Soit $n \geq 1$ un entier fixé. L'application $\langle \cdot | \cdot \rangle : \mathbb{R}_n[x] \times \mathbb{R}_n[x] \rightarrow \mathbb{R}$ définie par : $\langle p | q \rangle = \int_{-1}^{+1} p(t)q(t) dt$ pour tous $p, q \in \mathbb{R}_n[x]$ est un produit scalaire sur $\mathbb{R}_n[x]$.

Preuve. Analogie au troisième exemple du 2.1.1 ci-dessus ; le détail des calculs est laissé à titre d'exercice. \square

Une question naturelle est alors de voir si la base canonique est une base orthogonale pour ce produit scalaire. Ce n'est pas le cas car un calcul simple montre que $\langle x^m | x^k \rangle = 0$ si $m + k$ est impair, mais $\frac{2}{m+k+1} \neq 0$ si $m + k$ est pair.

c) PROPOSITION. Une base orthogonale de $\mathbb{R}_n[x]$ est donnée par la famille $\mathcal{L} = (\ell_0, \ell_1, \ell_2, \dots, \ell_n)$, où, pour tout entier $m \geq 0$, on désigne par ℓ_m le polynôme de degré m défini par :

$$\ell_m(x) = \frac{d^m}{dx^m} [(x^2 - 1)^m].$$

Preuve. Il est clair que ℓ_m est de degré m pour tout $m \geq 0$. Il résulte donc du rappel a) ci-dessus que \mathcal{L} est une base de $\mathbb{R}_n[x]$. Par ailleurs, pour tout $0 \leq m \leq n$ et tout $0 \leq k \leq m-1$, on peut calculer $I_{k,m} := \langle \ell_m | x^k \rangle$ par intégration par parties :

$$I_{k,m} = \int_{-1}^{+1} x^k \frac{d^m}{dx^m} [(x^2-1)^m] dx = \left[x^k \frac{d^{m-1}}{dx^{m-1}} [(x^2-1)^m] \right]_{-1}^{+1} - \int_{-1}^{+1} kx^{k-1} \frac{d^{m-1}}{dx^{m-1}} [(x^2-1)^m] dx$$

Or comme ± 1 est zéro d'ordre m de $(x^2-1)^m$, il est aussi zéro d'ordre $m-j$ de $\frac{d^j}{dx^j} [(x^2-1)^m]$ pour tout $j \leq m-1$. En particulier le polynôme $\frac{d^{m-1}}{dx^{m-1}} [(x^2-1)^m]$ s'annule en $+1$ et en -1 , de sorte que l'expression précédente se réduit à

$$I_{k,m} = -k \int_{-1}^{+1} x^{k-1} \frac{d^{m-1}}{dx^{m-1}} [(x^2-1)^m] dx.$$

En itérant, il vient de même :

$$I_{k,m} = +k(k-1) \int_{-1}^{+1} x^{k-2} \frac{d^{m-2}}{dx^{m-2}} [(x^2-1)^m] dx,$$

et après k itérations :

$$I_{k,m} = (-1)^k k! \int_{-1}^{+1} \frac{d^{m-k}}{dx^{m-k}} [(x^2-1)^m] dx = (-1)^k k! \left[\frac{d^{m-k-1}}{dx^{m-k-1}} [(x^2-1)^m] \right]_{-1}^{+1}.$$

Cette dernière expression est nulle par le même argument que celui utilisé précédemment (± 1 est zéro du polynôme $\frac{d^{m-k-1}}{dx^{m-k-1}} [(x^2-1)^m]$ car $m-k-1 \leq m-1$).

On a ainsi montré que $\langle \ell_m | x^k \rangle = 0$ pour tous $0 \leq k \leq m-1$, donc par bilinéarité $\langle \ell_m | \ell_k \rangle = 0$ pour tous $0 \leq k \leq m-1$, et finalement par symétrie $\langle \ell_m | \ell_k \rangle = 0$ pour tous $k \neq m$. \square

d) REMARQUES ET DÉFINITION. On peut évidemment rendre orthonormale la base orthogonale \mathcal{L} en multipliant chaque polynôme ℓ_m par l'inverse de sa norme.

Le calcul (utilisant les intégrales de Wallis) permet d'observer que : $\|\ell_m\|^2 = \frac{2^{2m+1}(m!)^2}{2m+1}$.

Ceci explique que l'on remplace souvent les polynômes ℓ_m par leurs multiples $L_m = \frac{1}{2^m m!} \ell_m$, qui restent évidemment deux à deux orthogonaux, mais vérifient plus simplement : $\|L_m\| = \sqrt{\frac{2}{2m+1}}$.

Ces polynômes s'appellent les *polynômes de Legendre*. Les premiers termes sont :

$$\begin{aligned} L_0(x) &= 1 & L_2(x) &= \frac{3}{2}x^2 - \frac{1}{2} & L_4(x) &= \frac{35}{8}x^4 - \frac{15}{4}x^2 + \frac{3}{8} \\ L_1(x) &= x & L_3(x) &= \frac{5}{2}x^3 - \frac{3}{2}x & L_5(x) &= \frac{63}{8}x^5 - \frac{35}{4}x^3 + \frac{15}{8}x \end{aligned}$$

Ils interviennent dans la résolution de certains types d'équations différentielles.

e) REMARQUE. Nous avons donné ici à titre d'exercice d'illustration un exemple de famille orthogonale de polynômes, parmi les plus simples. A côté des polynômes de Legendre, d'autres familles classiques font l'objet d'innombrables problèmes liant des questions d'algèbre et d'analyse : polynômes de Tchébychev, de Laguerre, d'Hermite, etc...

2.1.5 Exercices

D'autres propriétés importantes liées aux produits scalaires sont regroupées dans les exercices suivants. Elles doivent être considérées comme des résultats de cours à connaître.

EXERCICE 1 (*procédé d'orthogonalisation de Gram-Schmidt*). Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base d'un espace vectoriel euclidien E . Montrer que l'on peut construire une base *orthogonale* $\mathcal{C} = (\varepsilon_1, \dots, \varepsilon_n)$ de E qui vérifie $\text{Vect}(\varepsilon_1, \dots, \varepsilon_p) = \text{Vect}(e_1, \dots, e_p)$ pour tout $1 \leq p \leq n$.

Indication : considérer par récurrence la suite des $\varepsilon_i = e_i - \sum_{j=1}^{i-1} \lambda_{j,i} \varepsilon_j$ où $\lambda_{j,i} = \frac{1}{\|\varepsilon_j\|^2} \langle \varepsilon_j | e_i \rangle$.

EXERCICE 2 (*propriété de Pythagore*). Soit E un espace vectoriel euclidien. Montrer que deux vecteurs x et y sont orthogonaux si et seulement si $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.

Montrer que, si (x_1, \dots, x_p) est une famille de p vecteurs deux à deux orthogonaux, alors $\|\sum_{j=1}^p x_j\|^2 = \sum_{j=1}^p \|x_j\|^2$.

EXERCICE 3 (*équations d'hyperplan vectoriel et vecteur normal*). Soit E un espace vectoriel euclidien de dimension n . On rappelle que, par définition, un hyperplan de E est un sous-espace vectoriel de dimension $n - 1$. Soit \mathcal{B} une base orthonormale de E .

– Montrer (on pourra utiliser 2.1.3) que, si H est un hyperplan de E , alors il existe des réels $\alpha_1, \dots, \alpha_n$ non tous nuls (et non uniques) tels que H est l'ensemble des vecteurs de E dont les composantes (x_1, \dots, x_n) dans la base \mathcal{B} vérifient la relation $\alpha_1 x_1 + \dots + \alpha_n x_n = 0$. On dit que cette relation est une équation de H dans \mathcal{B} .

– Réciproquement montrer que, si $\alpha_1, \dots, \alpha_n$ sont des réels non tous nuls, alors l'ensemble des vecteurs de E dont les coordonnées (x_1, \dots, x_n) dans \mathcal{B} vérifient $\alpha_1 x_1 + \dots + \alpha_n x_n = 0$ est un hyperplan vectoriel H .

– Avec les notations ci-dessus, montrer que H^\perp est la droite vectorielle dirigée par le vecteur non-nul a dont les composantes dans la base \mathcal{B} sont $(\alpha_1, \dots, \alpha_n)$. On dit que a est un vecteur normal à H .

– Soient H et H' deux hyperplans de E d'équations respectives $\alpha_1 x_1 + \dots + \alpha_n x_n = 0$ et $\alpha'_1 x_1 + \dots + \alpha'_n x_n = 0$ dans une base orthonormale. Comment traduire sur les coefficients α_i et α'_i le fait que $H = H'$? que $H^\perp \subset H'$? Peut-on avoir $H \perp H'$?

– Que deviennent ces résultats dans les cas particulier où $\dim E = 2$? où $\dim E = 3$?

EXERCICE 4 (*projection orthogonale et minimalisation de la distance*). Soit E un espace vectoriel euclidien. Pour toute partie non-vide A de E , et tout vecteur $x \in E$, on définit la distance entre x et A par :

$$d(x, A) = \inf_{y \in A} \|x - y\|.$$

– Montrer que, pour tout sous-espace vectoriel F de E , et pour tout vecteur $x \in E$, on a $d(x, F) = \|x - p(x)\|$ où p désigne la projection orthogonale sur F (en d'autres termes, la distance minimale entre x et les vecteurs de F est atteinte en l'unique vecteur $p(x)$).

– On suppose ici que F est un hyperplan H , de vecteur normal a de composantes $(\alpha_1, \dots, \alpha_n)$ dans une base orthonormale de E . Montrer que, pour tout vecteur $x \in E$, on a :

$$d(x, H) = \frac{|(x|a)|}{\|a\|}.$$

Que devient cette égalité dans les cas particulier où $\dim E = 2$? où $\dim E = 3$?

EXERCICE 5 (*réflexion échangeant deux vecteurs*). Soit E un espace vectoriel euclidien. On appelle réflexion toute symétrie orthogonale par rapport à un hyperplan de E . Montrer que, si x et x' sont deux vecteurs non-nuls distincts de même norme, alors il existe une unique réflexion qui les échange. (Indication : considérer l'hyperplan de vecteur normal $x - x'$).

2.2 Groupe orthogonal

2.2.1 Endomorphismes orthogonaux

On se place dans un espace vectoriel euclidien E . On note n sa dimension.

a) LEMME FONDAMENTAL. *Pour tout endomorphisme f de E , les conditions suivantes sont équivalentes :*

- (i) $\langle f(x) | f(y) \rangle = \langle x | y \rangle$ pour tous $x, y \in E$ (on dit que f conserve le produit scalaire);
- (ii) $\|f(x)\| = \|x\|$ pour tout $x \in E$ (on dit que f conserve la norme euclidienne);
- (iii) pour toute base orthonormale $\mathcal{B} = (e_1, \dots, e_n)$ de E , la famille $(f(e_1), \dots, f(e_n))$ est une base orthonormale de f (on dit que f transforme toute base orthonormale en une base orthonormale);
- (iv) il existe une base orthonormale $\mathcal{B} = (e_1, \dots, e_n)$ de E telle que la famille $(f(e_1), \dots, f(e_n))$ soit une base orthonormale de f .

Preuve. Il est clair que (i) implique (ii), et, puisque $f(x+y) = f(x) + f(y)$, on a aussi que (ii) implique (i) d'après la première identité de polarisation vue au g) de 2.1.1.

Supposons que l'on a (i). On a en particulier $\|f(0_E)\| = \|0_E\|$, donc $\|f(0_E)\| = 0$, donc $f(0_E) = 0_E$. En d'autres termes $\text{Ker } f = \{0_E\}$. Ainsi f est injective, c'est-à-dire bijective puisque f est un endomorphisme en dimension finie. Dès lors, considérons une base orthonormale $\mathcal{B} = (e_1, \dots, e_n)$ de E . La famille $\mathcal{C} = (f(e_1), \dots, f(e_n))$ est une base de E puisque f est bijective. De plus, pour tous $1 \leq i, j \leq n$, on a $\langle f(e_i) | f(e_j) \rangle = \langle e_i | e_j \rangle = \delta_{i,j}$, ce qui prouve que la base \mathcal{C} est orthonormale. Ainsi (iii) est satisfaite. Il est trivial que (iii) implique (iv).

Supposons pour finir qu'il existe une base orthonormale $\mathcal{B} = (e_1, \dots, e_n)$ de E telle que la famille $\mathcal{C} = (f(e_1), \dots, f(e_n))$ soit une base orthonormale de E . Quels que soient x, y dans E , de composantes respectives (x_1, \dots, x_n) et (y_1, \dots, y_n) dans la base \mathcal{B} , on a $\langle x | y \rangle = \sum_{i=1}^n x_i y_i$ d'après la proposition e) de 2.1.2. Mais comme f est linéaire, (x_1, \dots, x_n) et (y_1, \dots, y_n) sont aussi les composantes respectives de $f(x)$ et $f(y)$ dans la base \mathcal{C} . De sorte que $\sum_{i=1}^n x_i y_i = \langle f(x) | f(y) \rangle$, ce qui prouve (i). \square

b) DÉFINITION. On appelle *endomorphisme orthogonal* tout endomorphisme de E satisfaisant l'une des conditions équivalentes du lemme précédent.

Comme on l'a vu dans la preuve, il résulte de la condition (ii) qu'un endomorphisme orthogonal est nécessairement injectif, donc bijectif (car E est de dimension finie). On emploie aussi le mot d'*isométrie vectorielle* comme synonyme d'endomorphisme orthogonal.

c) EXEMPLE. Pour tout sous-espace vectoriel F de E , la symétrie orthogonale s par rapport à F est un endomorphisme orthogonal.

En effet. Soit $x \in E$ quelconque, décomposé en $x = y + z$ avec $y \in F$ et $z \in F^\perp$. On a $s(x) = y - z$, donc $\|s(x)\|^2 = \langle y - z | y - z \rangle = \|y\|^2 + \|z\|^2 - 2\langle y | z \rangle$. Mais $y \perp z$ donc cette expression se réduit à $\|y\|^2 + \|z\|^2$, qui est de même égale à $\langle y + z | y + z \rangle$. Ainsi $\|s(x)\| = \|x\|$.

d) THÉORÈME ET DÉFINITION. *L'ensemble des endomorphismes orthogonaux d'un espace vectoriel euclidien E est un sous-groupe de $\text{GL}(E)$.*

On l'appelle le groupe orthogonal de E , noté $\text{O}(E)$.

Preuve. Rappelons d'abord que $GL(E)$ désigne le groupe linéaire de E , c'est-à-dire le groupe (pour la loi \circ) des automorphismes de l'espace vectoriel E , c'est-à-dire des endomorphismes de E qui sont bijectifs. On a donc bien $O(E) \subset GL(E)$. Il est clair aussi que $\text{id}_E \in O(E)$.

Soient f et g deux éléments de $O(E)$. Pour tout $x \in E$, on a $\|(g \circ f)(x)\| = \|g(f(x))\| = \|f(x)\| = \|x\|$ en appliquant d'abord le fait que g conserve la norme, puis que f conserve la norme. Ceci montre que $g \circ f \in O(E)$. Donc $O(E)$ est stable pour la loi \circ .

Soit $f \in O(E)$. Notons f^{-1} sa bijection réciproque. On a $\|x\| = \|f(f^{-1}(x))\|$ pour tout $x \in E$. Mais comme f conserve la norme, $\|f(f^{-1}(x))\| = \|f^{-1}(x)\|$. Ainsi $\|x\| = \|f^{-1}(x)\|$. Ceci montre que $f^{-1} \in O(E)$. Donc $O(E)$ est stable par passage à la réciproque. \square

2.2.2 Matrices orthogonales

a) **RAPPEL** (*transposition des matrices*). Pour toute matrice $A \in \mathcal{M}_n(\mathbb{R})$, la transposée de A est la matrice ${}^tA \in \mathcal{M}_n(\mathbb{R})$ obtenue en échangeant les lignes et les colonnes de A :

$$\text{si } A = (a_{ij})_{1 \leq i, j \leq n}, \text{ alors } {}^tA = (a_{ji})_{1 \leq i, j \leq n}.$$

Pour toutes $A, B \in \mathcal{M}_n(\mathbb{R})$, et tout $\lambda \in \mathbb{R}$, on a :

$${}^t(A + B) = {}^tA + {}^tB, \quad {}^t(\lambda A) = \lambda {}^tA, \quad {}^t(AB) = {}^tB {}^tA, \quad {}^t({}^tA) = A.$$

b) **DÉFINITION**. On appelle *matrice orthogonale* d'ordre n toute matrice $A \in \mathcal{M}_n(\mathbb{R})$ telle que :

$$A {}^tA = {}^tAA = I_n.$$

Une matrice orthogonale est donc nécessairement inversible et vérifie $A^{-1} = {}^tA$.

c) **THÉORÈME**. *L'ensemble des matrices orthogonales d'ordre n est un sous-groupe de $GL(n, \mathbb{R})$. On le note $O(n, \mathbb{R})$.*

Preuve. Rappelons d'abord que $GL(n, \mathbb{R})$ désigne le groupe des matrices carrées inversibles dans $\mathcal{M}_n(\mathbb{R})$. On a donc bien $O(n, \mathbb{R}) \subset GL(n, \mathbb{R})$. Il est clair aussi que $I_n \in O(n, \mathbb{R})$.

Soient A et B deux matrices de $O(n, \mathbb{R})$. On a : ${}^t(AB) = {}^tB {}^tA = B^{-1}A^{-1} = (AB)^{-1}$, ce qui montre que $O(n, \mathbb{R})$ est stable par produit. De plus, ${}^t(A^{-1}) = {}^t({}^tA) = A = (A^{-1})^{-1}$. Donc $O(n, \mathbb{R})$ est stable par passage à l'inverse, ce qui achève la preuve. \square

Le fait que l'on utilise la même terminologie et des notations voisines pour désigner à la fois les endomorphismes orthogonaux de E et les matrices orthogonales d'ordre n est justifié par le théorème suivant.

d) **THÉORÈME**. *Pour tout endomorphisme f d'un espace vectoriel euclidien E , les conditions suivantes sont équivalentes :*

1. f est un endomorphisme orthogonal de E ;
2. pour toute base orthonormale \mathcal{B} de E , la matrice de f dans la base \mathcal{B} est une matrice orthogonale ;
3. il existe une base orthonormale \mathcal{B} de E telle que la matrice de f dans la base \mathcal{B} soit une matrice orthogonale.

Preuve. Soit f un endomorphisme de E , et $A = (a_{ij})_{1 \leq i, j \leq n}$ la matrice de f dans une base orthonormale $\mathcal{B} = (e_1, \dots, e_n)$ de E .

Par définition de A , on a $f(e_i) = \sum_{k=1}^n a_{ki} e_k$ pour tout $1 \leq i \leq n$. Donc pour $1 \leq i, j \leq n$, on calcule :

$$\langle f(e_i) | f(e_j) \rangle = \left\langle \sum_{k=1}^n a_{ki} e_k \mid \sum_{\ell=1}^n a_{\ell j} e_\ell \right\rangle = \sum_{k=1}^n \sum_{\ell=1}^n a_{ki} a_{\ell j} \langle e_k | e_\ell \rangle = \sum_{k=1}^n a_{ki} a_{kj}.$$

La famille $(f(e_1), \dots, f(e_n))$ est une base orthonormale si et seulement si $\langle f(e_i) | f(e_j) \rangle = \delta_{i,j}$ pour tous $1 \leq i, j \leq n$, c'est-à-dire si et seulement si $\sum_{k=1}^n a_{ki} a_{kj} = \delta_{i,j}$ pour tous $1 \leq i, j \leq n$, égalités entre coefficients qui traduisent exactement l'égalité matricielle ${}^t A A = I_n$. \square

e) COROLLAIRE. Soit \mathcal{B} une base orthonormale de E . L'application qui, à tout endomorphisme de E , associe sa matrice dans la base \mathcal{B} définit un isomorphisme de groupes de $O(E)$ sur $O(n, \mathbb{R})$.

Preuve. Considérons l'application Φ qui, à tout endomorphisme f de E associe sa matrice $A = \Phi(f)$ par rapport à la base \mathcal{B} . On sait déjà que f est bijective si et seulement si A est inversible, et donc Φ définit une application du groupe $GL(E)$ des automorphismes de E dans le groupe $GL(n, \mathbb{R})$ des matrices carrées d'ordre n inversibles. On sait aussi que, si g est un autre automorphisme de E de matrice B dans la base \mathcal{B} , alors la matrice de $g \circ f$ est BA . En d'autres termes, $\Phi(g \circ f) = \Phi(g)\Phi(f)$ pour tous $f, g \in GL(E)$, ce qui s'exprime en disant que Φ détermine un isomorphisme de groupe de $GL(E)$ sur $GL(n, \mathbb{R})$.

Le théorème ci-dessus prouve de plus que $f \in O(E)$ si et seulement si $\Phi(f) \in O(n, \mathbb{R})$, et donc la restriction de Φ au sous-groupe $O(E)$ détermine un isomorphisme de groupes de $O(E)$ sur $O(n, \mathbb{R})$. \square

2.2.3 Orientation et produit mixte

a) THÉORÈME. Soit n un entier ≥ 1 fixé. Soit E un espace vectoriel euclidien de dimension n .

- (i) Si A est une matrice orthogonale, alors son déterminant vaut $+1$ ou -1 . L'ensemble des matrices orthogonales dont le déterminant vaut $+1$ est un sous-groupe de $O(n, \mathbb{R})$ appelé sous-groupe spécial orthogonal, noté $SO(n, \mathbb{R})$.
- (ii) Si f est un endomorphisme orthogonal de E , alors son déterminant vaut $+1$ ou -1 . L'ensemble des endomorphismes orthogonaux dont le déterminant vaut $+1$ est un sous-groupe de $O(E)$ appelé le sous-groupe spécial orthogonal de E , noté $SO(E)$.
- (iii) L'application qui, à tout endomorphisme de E , associe sa matrice dans une base orthonormale définit un isomorphisme de groupes de $SO(E)$ sur $SO(n, \mathbb{R})$.

Preuve. Si $A \in O(n, \mathbb{R})$, on a ${}^t A A = I_n$, donc $\det({}^t A A) = 1$, c'est-à-dire $\det({}^t A) \det A = 1$. Mais $\det({}^t A) = \det A$, d'où $(\det A)^2 = 1$ et finalement $\det A = \pm 1$. Il est clair que I_n est de déterminant $+1$, que le produit de deux matrices de déterminant $+1$ est de déterminant $+1$, et que l'inverse d'une matrice de déterminant $+1$ est de déterminant $+1$; on conclut que $SO(n, \mathbb{R})$ est un sous-groupe (c'est le noyau du morphisme de groupe $\det : O(n, \mathbb{R}) \rightarrow \{\pm 1\}$).

Le point (i) étant ainsi montré, le (ii) en découle de façon immédiate avec le théorème d) de 2.2.2 puisque $\det f = \det A$ où $A = \Phi(f)$, et (iii) résulte alors du corollaire e) de 2.2.2. \square

b) REMARQUES. On a par définition $SO(E) = O(E) \cap SL(E)$. Le groupe $SO(E)$ est parfois noté $O^+(E)$, et ses éléments sont parfois appelés *isométries vectorielles directes*, ou *isométries vectorielles positives*.

En notant $O^-(E)$ l'ensemble des endomorphismes orthogonaux dans $O(E)$ qui sont de déterminant -1 (que l'on appelle parfois *isométries vectorielles indirectes* ou *isométries vectorielles négatives*) on a : $O(E) = O^+(E) \cup O^-(E)$, mais attention, $O^-(E)$ n'est pas un sous-groupe de $O(E)$!

c) PROPOSITION (exemple fondamental des symétries orthogonales). Soit F un sous-espace vectoriel de dimension p dans un espace vectoriel euclidien E de dimension n . Soit s la symétrie orthogonale par rapport à F . Alors :

$$s \in O(E), \quad \det s = (-1)^{n-p}, \quad s \in SO(E) \Leftrightarrow (n-p \text{ est pair}).$$

Preuve. Soit (e_1, \dots, e_p) une base orthonormale de F . Comme $E = F \oplus F^\perp$, il existe une base orthonormale (e_{p+1}, \dots, e_n) de F^\perp telle que $(e_1, \dots, e_p, e_{p+1}, \dots, e_n)$ soit une base orthonormale \mathcal{B} de E . On a $s(e_i) = e_i$ pour tous les $1 \leq i \leq p$ et $s(e_j) = -e_j$ pour tous les $p+1 \leq j \leq n$. La matrice A de s dans la base \mathcal{B} est donc diagonale, avec p termes diagonaux égaux à 1 et $n-p$ termes diagonaux égaux à -1 . Il en résulte d'une part que ${}^tAA = I_n$, et d'autre part que $\det A = (-1)^{n-p}$, ce qui montre les résultats voulus. \square

Ainsi, si E est de dimension 2, toute symétrie orthogonale par rapport à une droite vectorielle est une isométrie indirecte. Si E est de dimension 3, toute symétrie orthogonale par rapport à un plan vectoriel est une isométrie indirecte, mais toute symétrie orthogonale par rapport à une droite vectorielle est une isométrie directe. D'une façon générale, si E est de dimension n , toute symétrie orthogonale par rapport à un hyperplan vectoriel est une isométrie indirecte.

d) REMARQUES ET DÉFINITIONS (*orientation*). Soit E un espace vectoriel euclidien de dimension n . Soient \mathcal{B} et \mathcal{B}' deux bases orthonormales de E . On sait qu'il existe un unique automorphisme $f \in GL(E)$ qui transforme \mathcal{B} en \mathcal{B}' . Parce que \mathcal{B} et \mathcal{B}' sont des bases orthonormales, il résulte du lemme a) de 2.2.1 que l'on a nécessairement $f \in O(E)$. On peut alors définir :

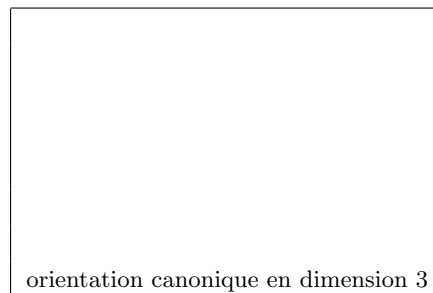
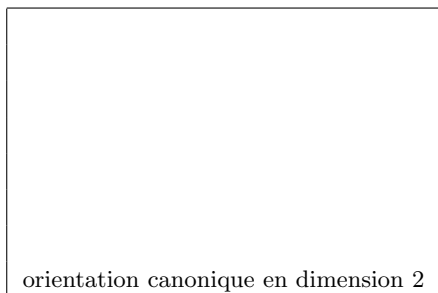
Deux bases orthonormales \mathcal{B} et \mathcal{B}' sont dites *de même orientation* lorsque l'unique isométrie vectorielle $f \in O(E)$ qui transforme \mathcal{B} en \mathcal{B}' est une isométrie directe [c'est-à-dire $f \in O^+(E)$].

Sinon [c'est-à-dire $f \in O^-(E)$], on dira que \mathcal{B} et \mathcal{B}' sont d'orientations opposées.

Orienter l'espace vectoriel euclidien E consiste à choisir arbitrairement une base orthonormale de référence \mathcal{B} ; toute base orthonormale \mathcal{B}' ayant la même orientation que \mathcal{B} sera dite *directe*, toute base orthonormale \mathcal{B}' ayant l'orientation opposée à celle de \mathcal{B} sera dite *indirecte*.

Il résulte immédiatement de ces définitions que :

- Deux bases orthonormales sont de même orientation si et seulement si la matrice de passage de l'une à l'autre est de déterminant $+1$.
- Une isométrie est directe si et seulement si elle conserve l'orientation (ie. elle transforme toute base orthonormale directe en une base orthonormale directe).



e) PROPOSITION ET DÉFINITION (*produit mixte*). Soit E un espace vectoriel euclidien orienté de dimension n . Soient u_1, u_2, \dots, u_n des vecteurs de E . Pour toute base orthonormale directe \mathcal{B} de E , le réel

$$[u_1, u_2, \dots, u_n] := \det \text{Mat}_{\mathcal{B}}(u_1, u_2, \dots, u_n)$$

est indépendant de la base orthonormale \mathcal{B} choisie. On l'appelle le *produit mixte* des vecteurs u_1, u_2, \dots, u_n .

Preuve. Soient \mathcal{B} et \mathcal{B}' deux bases orthonormales directes de E . La matrice de passage P de \mathcal{B} à \mathcal{B}' est donc orthogonale, et directe ($\det P = 1$). Soit $X = (u_1, u_2, \dots, u_n)$ une famille de n vecteurs de E ; notons M la matrice de la famille X dans la base \mathcal{B} , et M' sa matrice dans la base \mathcal{B}' . Alors $M = PM'$, d'où $\det M = \det P \det M' = \det M'$. \square

f) PROPOSITION ET DÉFINITION (*produit vectoriel en dimension 3*). Soit E un espace vectoriel euclidien orienté de dimension 3. Quels que soient deux vecteurs u et v de E , il existe un unique vecteur, noté $u \wedge v$ et appelé le *produit vectoriel* de u par v , tel que :

$$[u, v, w] = \langle u \wedge v | w \rangle \quad \text{pour tout } w \in E.$$

Preuve. Fixons $u, v \in E$ quelconques. Il est clair, parce que le déterminant est linéaire par rapport à sa troisième variable, que l'application $f : w \mapsto [u, v, w]$ définit une application $E \rightarrow \mathbb{R}$ qui est linéaire. En d'autres termes, f est une forme linéaire. En appliquant la proposition b) de 2.1.3, il existe donc un unique vecteur $a \in E$ tel que $f = \varphi_a$, c'est-à-dire tel que $\langle a | w \rangle = [u, v, w]$ pour tout $w \in E$. Ce vecteur unique a dépend bien sûr des vecteurs u et v de départ, et l'on note donc $a = u \wedge v$. \square

On trouvera ci-dessous à l'exercice 7 des propriétés classiques du produit vectoriel.

2.2.4 Exercices.

EXERCICE 1 (*un résultat souvent utile*). Soient E un espace vectoriel euclidien et $f \in O(E)$. Montrer que, pour tout sous-espace vectoriel F de E , on a $f(F)^\perp = f(F^\perp)$, en déduire qu'en particulier, si F est un sous-espace vectoriel stable par f , alors F^\perp aussi.

EXERCICE 2 (*valeurs propres d'un endomorphisme orthogonal*). Soient E un espace vectoriel euclidien et $f \in O(E)$. Montrer que les sous-espaces $E_1 = \text{Ker}(f - \text{id}_E)$ et $E_{-1} = \text{Ker}(f + \text{id}_E)$ sont orthogonaux. Montrer que les seules v.p. réelles possibles pour f sont 1 et -1 .

EXERCICE 3 (*caractérisations des symétries orthogonales*). Soient E un espace vectoriel euclidien et $f \in O(E)$. Montrer que les conditions suivantes sont équivalentes :

- (i) f est une symétrie orthogonale; (ii) $f \circ f = \text{id}_E$; (iii) f est diagonalisable.

EXERCICE 4 (*un exemple concret en dimension 3*). Soit E un espace vectoriel euclidien de dimension 3, muni d'une base orthonormale \mathcal{B} . Soit f l'endomorphisme de E dont la matrice dans la base \mathcal{B} est :

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Montrer que f est une symétrie orthogonale par rapport à un sous-espace que l'on déterminera.

EXERCICE 5 (*un exemple concret en dimension 3*). Soit E un espace vectoriel euclidien de dimension 3, muni d'une base orthonormale \mathcal{B} . Pour tous réels a, b, c , on considère l'endomorphisme $f_{a,b,c}$ de E dont la matrice dans la base \mathcal{B} est :

$$M_{a,b,c} = \begin{pmatrix} 0 & 0 & a \\ b & 0 & 0 \\ 0 & c & 0 \end{pmatrix}.$$

Déterminer les valeurs de a, b, c pour lesquelles $f_{a,b,c} \in \text{O}(E)$. On suppose désormais que $f_{a,b,c} \in \text{O}(E)$. Déterminer les valeurs propres réelles de $f_{a,b,c}$ et les sous-espaces propres associés. Est-ce que $f_{a,b,c}$ est diagonalisable sur \mathbb{R} ?

EXERCICE 6 (*un exemple concret en dimension 3*). Soit E un espace vectoriel euclidien de dimension 3, muni d'une base orthonormale \mathcal{B} . Pour tout couple de réel a, b , on considère l'endomorphisme $f_{a,b}$ de E dont la matrice dans la base \mathcal{B} est :

$$M_{a,b} = \begin{pmatrix} 3a+b & -4a & 0 \\ -4a & -3a+b & 0 \\ 0 & b & 1 \end{pmatrix}.$$

Déterminer les valeurs de a et b pour lesquelles $f_{a,b} \in \text{GL}(E)$. Déterminer les valeurs de a et b pour lesquelles $f_{a,b} \in \text{O}(E)$. Si $a = \frac{1}{5}$ et $b = 0$, montrer que $f_{1/5,0}$ est une symétrie orthogonale par rapport à un sous-espace vectoriel F de E que l'on déterminera.

EXERCICE 7. (*produit vectoriel en dimension 3*). Soit E un espace vectoriel euclidien orienté de dimension 3.

– En utilisant les propriétés du déterminant, montrer que, pour tous $u, v, w \in E$, $\alpha, \beta \in \mathbb{R}$:

$$u \wedge u = 0_E, \quad v \wedge u = -u \wedge v, \quad (\alpha u + \beta v) \wedge w = \alpha u \wedge w + \beta v \wedge w.$$

$$u \perp (u \wedge v) \quad \text{et} \quad v \perp (u \wedge v).$$

$$u \text{ et } v \text{ colinéaires si et seulement si } u \wedge v = 0_E.$$

$$\text{si } u \text{ et } v \text{ non colinéaires, } (u, v, u \wedge v) \text{ est libre.}$$

– Montrer que, si (x, y, z) et (x', y', z') sont les composantes respectives de deux vecteurs $u, v \in E$ dans une base orthonormale directe, alors les composantes dans cette même base de $u \wedge v$ sont : $(yz' - zy', zx' - xz', xy' - yx')$.

– Montrer que, si (u, v) est une famille orthonormale de deux vecteurs de E , alors $(u, v, u \wedge v)$ est une base orthonormale directe de E .

– Montrer que, quels que soient $u, v, w \in E$, on a : $u \wedge (v \wedge w) = \langle u | w \rangle v - \langle u | v \rangle w$.

EXERCICE 8 (*matrices de permutation*). On fixe un entier $n \geq 2$, et on considère l'espace euclidien \mathbb{R}^n muni du produit scalaire canonique. Pour toute permutation σ dans le groupe symétrique S_n , on désigne par f_σ l'application $\mathbb{R}^n \rightarrow \mathbb{R}^n$ transformant un vecteur quelconque $x = (x_1, x_2, \dots, x_n)$ en $f_\sigma(x) = (x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)})$. On note $M_\sigma \in M_n(\mathbb{R})$ la matrice de f_σ par rapport à la base canonique.

Montrer que f_σ appartient au groupe orthogonal $\text{O}(E)$. Montrer qu'il existe une valeur propre réelle commune à tous f_σ , pour σ décrivant S_n . Montrer qu'il existe un hyperplan H de \mathbb{R}^n stable par f_σ pour tout $\sigma \in S_n$.

EXERCICE 9 (*matrices de Householder et symétries orthogonales*). On fixe un entier $n \geq 2$, et on considère l'espace euclidien \mathbb{R}^n muni du produit scalaire canonique. On note \mathcal{B} la base canonique de \mathbb{R}^n . Pour tout vecteur non-nul u de \mathbb{R}^n , on note U la matrice colonne des composantes de u dans la base \mathcal{B} et on introduit l'endomorphisme h_u de \mathbb{R}^n dont la matrice par rapport à \mathcal{B} est : $H_u = I_n - \frac{2}{\|u\|^2} U^t U$.

Montrer que la matrice H_u est symétrique et orthogonale. Montrer que h_u est la symétrie orthogonale par rapport à l'hyperplan $(\mathbb{R}u)^\perp$.

2.3 Endomorphismes symétriques

2.3.1 Transposition

On se place dans un espace vectoriel euclidien E .

a) LEMME. Soit \mathcal{B} une base orthonormale de E . Une base \mathcal{B}' de E est orthonormale si et seulement la matrice de passage de \mathcal{B} à \mathcal{B}' est une matrice orthogonale. En particulier, toute matrice de passage entre deux bases orthonormales de E est une matrice orthogonale.

Preuve. La matrice de passage P de \mathcal{B} à \mathcal{B}' donne en colonnes les composantes des vecteurs de \mathcal{B}' quand on les exprime dans la base \mathcal{B} . En d'autres termes, c'est aussi la matrice dans la base \mathcal{B} de l'endomorphisme f de E qui envoie la base \mathcal{B} sur la base \mathcal{B}' . Le lemme résulte donc de l'application immédiate du théorème d) de 2.2.2. \square

b) COROLLAIRE. Soit $A \in O(n, \mathbb{R})$ une matrice orthogonale. La famille des vecteurs colonnes de A constitue une base orthonormale de l'espace vectoriel euclidien \mathbb{R}^n muni de sa base canonique.

Preuve. Evident. \square

c) PROPOSITION ET DÉFINITION. Pour tout endomorphisme f de E , il existe un unique endomorphisme de E , noté ${}^t f$, appelé le transposé de f , vérifiant :

$$\langle f(x) | y \rangle = \langle x | {}^t f(y) \rangle \quad \text{pour tous } x, y \in E.$$

Pour toute base orthonormale \mathcal{B} de E , la matrice de ${}^t f$ dans la base \mathcal{B} est la transposée de la matrice de f dans cette même base \mathcal{B} :

$$\text{Mat}_{\mathcal{B}}({}^t f) = {}^t \text{Mat}_{\mathcal{B}}(f)$$

Preuve. Soit \mathcal{B} une base orthonormale de E . Soient f un endomorphisme de E et $A = (a_{ij})_{1 \leq i, j \leq n}$ la matrice de f dans la base \mathcal{B} . Appelons g l'endomorphisme de E dont la matrice dans la base \mathcal{B} est ${}^t A = (a_{ji})_{1 \leq i, j \leq n}$.

Prenons deux vecteurs quelconques x et y de E . Notons X et Y les matrices colonnes de leurs composantes respectives dans la base \mathcal{B} . Notons X' et Y' les matrices colonnes des composantes respectives des vecteurs $f(x)$ et $g(y)$ dans la base \mathcal{B} .

Parce que \mathcal{B} est orthonormale, l'expression du produit scalaire donne $\langle f(x) | y \rangle = {}^t X' Y$. Comme $X' = AX$ et $Y' = {}^t AY$, on en déduit que :

$$\langle f(x) | y \rangle = {}^t (AX) Y = {}^t X {}^t A Y = {}^t X Y' = \langle x | g(y) \rangle.$$

L'endomorphisme g choisi est donc une solution du problème considéré; les mêmes calculs montrent que c'est l'unique solution, et l'on peut donc poser $g = {}^t f$.

Enfin, si \mathcal{B}' est une autre base orthonormale, et si l'on note A' la matrice de f dans cette nouvelle base \mathcal{B}' , on sait que $A' = P^{-1} A P$ où P est la matrice de passage de \mathcal{B} à \mathcal{B}' . Comme on l'a vu au lemme a) ci-dessus, on a ${}^t P = P^{-1}$. La matrice A'' de l'endomorphisme ${}^t f$ dans la base \mathcal{B}' est de même $A'' = P^{-1} ({}^t A) P$. On a $A'' = ({}^t P) ({}^t A) ({}^t P^{-1}) = {}^t (P^{-1} A P) = {}^t A'$, ce qui achève la preuve. \square

2.3.2 Diagonalisation des endomorphismes symétriques

On se place dans un espace vectoriel euclidien E . On note n sa dimension.

a) DÉFINITION. Un endomorphisme f de E est dit *symétrique* lorsque ${}^t f = f$ dans $\text{End } E$, ce qui est équivalent à :

$$\langle f(x) | y \rangle = \langle x | f(y) \rangle \quad \text{pour tous } x, y \in E.$$

Une matrice A dans $\mathcal{M}_n(\mathbb{R})$ est dite *symétrique* lorsque $A = {}^tA$.

Il est clair d'après la proposition c) de 2.3.1 que f est symétrique si et seulement si sa matrice dans une base orthonormale de E est une matrice symétrique.

b) EXEMPLES. Les projections orthogonales et les symétries orthogonales sont des endomorphismes symétriques.

En effet. Soient F un sous-espace vectoriel de E , p la projection orthogonale sur F et s la symétrie orthogonale par rapport à F . Soient x, y deux vecteurs quelconques de E , décomposés en $x = x' + x''$ et $y = y' + y''$ avec $x', y' \in F$ et $x'', y'' \in F^\perp$.

$$\langle p(x) | y \rangle = \langle x' | y' + y'' \rangle = \langle x' | y' \rangle + \langle x' | y'' \rangle = \langle x' | y' \rangle,$$

$$\langle x | p(y) \rangle = \langle x' + x'' | y' \rangle = \langle x' | y' \rangle + \langle x'' | y' \rangle = \langle x' | y' \rangle.$$

$$\langle s(x) | y \rangle = \langle x' - x'' | y' + y'' \rangle = \langle x' | y' \rangle - \langle x'' | y' \rangle + \langle x' | y'' \rangle - \langle x'' | y'' \rangle = \langle x' | y' \rangle - \langle x'' | y'' \rangle,$$

$$\langle x | s(y) \rangle = \langle x' + x'' | y' - y'' \rangle = \langle x' | y' \rangle + \langle x'' | y' \rangle - \langle x' | y'' \rangle - \langle x'' | y'' \rangle = \langle x' | y' \rangle - \langle x'' | y'' \rangle,$$

ce qui montre bien que p et s sont symétriques.

Mais on a mieux : on a vu à l'exemple c) de 1.1.2 que p et s sont diagonalisables. Dans le cas de p , on a $E = E_0 \oplus E_1$ avec $E_0 = \text{Ker } p = F^\perp$ et $E_1 = \text{Im } p = F$. Dans le cas de s , on a $E = E_{-1} \oplus E_1$ avec $E_{-1} = F^\perp$ et $E_1 = F$. Dans les deux cas, les sous-espaces propres sont donc orthogonaux !

Cette dernière observation est un cas particulier du théorème suivant, qui est l'un des plus importants de la théorie. On établit d'abord quelques résultats intermédiaires.

c) LEMME. *Si λ et μ sont deux valeurs propres réelles distinctes d'un endomorphisme symétrique, alors les sous-espaces propres associés E_λ et E_μ sont orthogonaux.*

Preuve. Soit f un endomorphisme symétrique de E . Supposons que λ et μ sont deux v.p. distinctes de f . Soient $x \in E_\lambda$ et $y \in E_\mu$. Donc $f(x) = \lambda x$ et $f(y) = \mu y$. L'hypothèse f symétrique se traduit par $\langle f(x) | y \rangle = \langle x | f(y) \rangle$, donc $\langle \lambda x | y \rangle = \langle x | \mu y \rangle$, ou encore $\lambda \langle x | y \rangle = \mu \langle x | y \rangle$. Comme $\lambda \neq \mu$, on déduit que $\langle x | y \rangle = 0$, c'est-à-dire $x \perp y$. \square

d) LEMME. *Le polynôme caractéristique d'un endomorphisme symétrique se décompose complètement dans \mathbb{R} en produit de polynômes de degré 1.*

Preuve. Notons $\dim E = n$. Soient f un endomorphisme symétrique de E , et A sa matrice dans une base orthonormale de E . En particulier A est une matrice symétrique dans $\mathcal{M}_n(\mathbb{R})$. Soient $P_f = P_A$ le polynôme caractéristique de f ou A . C'est un polynôme à coefficients réels, et on peut donc le considérer aussi comme polynôme à coefficients complexes. A ce titre, on sait que, dans $\mathbb{C}[x]$, il se décompose sous la forme $P_A(x) = (-1)^n (x - \lambda_1)(x - \lambda_2) \dots (x - \lambda_n)$ avec les λ_i dans \mathbb{C} (pas forcément deux à deux distincts). Notre but est de montrer que $\lambda_i \in \mathbb{R}$ pour tout $1 \leq i \leq n$.

La matrice A peut être considérée comme la matrice d'un endomorphisme g de \mathbb{C}^n par rapport à la base canonique de \mathbb{C}^n . Prenons l'un des λ_i en le notant simplement λ ; c'est un zéro dans \mathbb{C} de P_A , donc une v.p. de g , et l'on peut considérer un vecteur propre $x \in \mathbb{C}^n$ associé à la v.p. λ . On a $x \neq 0_{\mathbb{C}^n}$ et $g(x) = \lambda x$. Notons X la matrice colonne des composantes de x dans la base \mathcal{B} , de sorte que : $AX = \lambda X$. En prenant les conjugués de tous les coefficients dans cette égalité, on a $\overline{A} \overline{X} = \overline{\lambda} \overline{X}$. Mais $\overline{A} = A$ puisque A est à coefficients réels. Donc

$A\bar{X} = \bar{\lambda}\bar{X}$. Si l'on écrit en ligne cette égalité de colonnes (en transposant), on obtient ${}^t\bar{X}A = \bar{\lambda}{}^t\bar{X}$, c'est-à-dire ${}^t\bar{X}A = \bar{\lambda}{}^t\bar{X}$, puisque A est symétrique.

En combinant les égalités $\bar{\lambda}{}^t\bar{X} = {}^t\bar{X}A$ et $AX = \lambda X$, on obtient $\bar{\lambda}{}^t\bar{X}X = {}^t\bar{X}AX = \lambda{}^t\bar{X}X$. Or ${}^t\bar{X}X$ est un réel strictement positif (c'est la somme des carrés des modules des composantes dans \mathbb{C} du vecteur non-nul x). On conclut que $\bar{\lambda} = \lambda$, c'est-à-dire $\lambda \in \mathbb{R}$. \square

e) THÉORÈME FONDAMENTAL. *Tout endomorphisme symétrique f d'un espace vectoriel euclidien E est diagonalisable. Plus précisément, il existe des bases orthogonales de E constituées de vecteurs propres de f .*

Preuve. On raisonne par récurrence sur $n = \dim E$. Pour $n = 1$, le résultat est clair. Prenons $n \geq 2$, supposons par hypothèse de récurrence le résultat vrai pour tout espace vectoriel euclidien de dimension $n - 1$, et considérons un espace vectoriel euclidien E de dimension n . Soit f un endomorphisme symétrique de E . D'après le lemme d), il admet des v.p. réelles ; soit λ l'une d'elle. Considérons u_n un vecteur propre associé à λ . Quitte à multiplier u_n par le réel $\|u_n\|^{-1}$, on peut sans restriction choisir u_n de norme 1. Soit H l'hyperplan vectoriel orthogonal à la droite vectorielle $F = \mathbb{R}u_n$ engendrée par u_n (voir exercice 3 de 2.1.5). Comme on l'a vu à l'exercice 1 de 2.2.4, le fait que F soit stable par f (ce qui est la définition même du fait que u_n est vecteur propre) implique que $H = F^\perp$ est aussi stable par f . Donc la restriction de f à H détermine un endomorphisme f' de H . Par hypothèse de récurrence appliquée à f' , il existe une base orthonormale $\mathcal{B}' = (u_1, \dots, u_{n-1})$ de H constituée de vecteurs propres de f' donc de f . En adjoignant u_n , on obtient une famille $\mathcal{B} = (u_1, \dots, u_{n-1}, u_n)$ qui est une base orthonormale de E (car $F \oplus H = E$ avec $F \perp H$) et qui est constituée de vecteurs propres de f . \square

f) COROLLAIRE FONDAMENTAL. *Toute matrice symétrique A dans $\mathcal{M}_n(\mathbb{R})$ est diagonalisable. Plus précisément, il existe une matrice diagonale D à coefficients réels et une matrice orthogonale $P \in O(n, \mathbb{R})$ telles que $A = PDP^{-1}$.*

Preuve. Soit A une matrice symétrique dans $\mathcal{M}_n(\mathbb{R})$. Soit f l'endomorphisme de l'espace vectoriel euclidien $E = \mathbb{R}^n$ (muni du produit scalaire canonique) telle que A soit la matrice de f dans la base canonique \mathcal{B} (qui est bien sûr orthonormale). Comme on l'a vu à la définition a), f est un endomorphisme symétrique. En appliquant le théorème e) ci-dessus, il existe une base orthonormale \mathcal{B}' de E telle que la matrice de f dans la base \mathcal{B}' soit une matrice diagonale D . On a donc $A = PDP^{-1}$ où P est la matrice de passage de \mathcal{B} à \mathcal{B}' . Il résulte enfin du lemme a) de 2.3.1 que P est orthogonale, ce qui achève la preuve. \square

2.3.3 Exercices

EXERCICE 1. Montrer que le sous-ensemble \mathcal{S} de $\mathcal{M}_n(\mathbb{R})$ formé des matrices symétriques est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$. En définissant une matrice antisymétrique comme une matrice $A \in \mathcal{M}_n(\mathbb{R})$ telle que ${}^tA = -A$, montrer que le sous-ensemble \mathcal{A} des matrices symétriques est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$. Montrer que $\mathcal{M}_n(\mathbb{R}) = \mathcal{S} \oplus \mathcal{A}$. Déterminer les dimensions respectives \mathcal{S} et \mathcal{A} .

EXERCICE 2. Soit E un espace euclidien. Soit u un endomorphisme de E .

– Montrer que $\text{Ker } u = (\text{Im } {}^tu)^\perp$ et $\text{Ker } {}^tu = (\text{Im } u)^\perp$. En déduire que, si u est symétrique, le noyau et l'image de u sont supplémentaires et orthogonaux.

– Montrer que si u est symétrique et $u(x) \perp x$ pour tout $x \in E$, alors u est nul.

– On suppose qu'il existe un sous-espace vectoriel M de E tel que $\|u(x)\| = \|x\|$ pour tout $x \in M$, et $v(x) = 0_E$ pour tout $x \in M^\perp$. Montrer que $\text{Ker } v = M^\perp$.

EXERCICE 3 (*exemples de diagonalisation de matrices symétriques*). Reprendre l'exemple b) du paragraphe 1.1.5 et vérifier que les sous-espaces propres sont orthogonaux (ou encore que la matrice de passage est orthogonale).

Faire de même pour les matrices B, C, E de l'exercice 1 de 1.2.4, et pour la matrice A de l'exercice 2. Faire de même pour les matrices A_α de l'exercice 6 de 1.3.4.

EXERCICE 4 (*exemple de diagonalisation d'une matrice symétrique*). On considère dans l'espace vectoriel euclidien \mathbb{R}^3 l'endomorphisme u dont la matrice dans la base canonique est

$$A = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}.$$

– Montrer que u est diagonalisable de quatre façons différentes : (1) sans aucun calcul ; (2) en calculant le polynôme caractéristique et les sous-espaces propres ; (3) en utilisant le théorème du rang ; (4) en calculant A^2 .

– Déterminer une base orthonormale de E dans laquelle la matrice de u est diagonale.

EXERCICE 5 (*exemple de diagonalisation d'une matrice symétrique*). On considère dans l'espace vectoriel euclidien \mathbb{R}^4 l'endomorphisme u dont la matrice dans la base canonique est

$$A = \begin{pmatrix} a^2 & ab & ac & ad \\ ab & b^2 & bc & bd \\ ac & bc & c^2 & dc \\ ad & bd & cd & d^2 \end{pmatrix}, \quad \text{où } a, b, c, d \text{ sont des réels non tous nuls.}$$

Montrer que u est symétrique. Montrer que 0 est valeur propre triple de u et déterminer le sous-espace propre associé. Soit λ l'autre valeur propre de u ; déterminer le sous-espace propre associé, et calculer λ .

2.4 Applications géométriques

2.4.1 Description des isométries vectorielles en dimension 2

On fixe un \mathbb{R} -espace vectoriel euclidien E de dimension 2. On va ci-dessous décrire explicitement le groupe orthogonal $O(E)$, c'est-à-dire tous les endomorphismes orthogonaux de E . Conformément à l'usage le plus répandu en géométrie, on dira ici "isométrie vectorielle" plutôt que "endomorphisme orthogonal" (rappelons que les deux termes sont strictement synonymes).

a) RÉFLEXION. Soit D est une droite vectorielle dans E ; rappelons que la symétrie orthogonale par rapport à D est l'application $s_D : E \rightarrow E$ qui, à tout vecteur $u \in E$ décomposé de façon unique en $v+w$ avec $v \in D$ et $w \in D^\perp$, associe $s_D(u) = v - w$. Une telle symétrie orthogonale par rapport à une droite D est appelée *réflexion d'axe D* (voir aussi exercice 5 de 2.1.5). Au vu des propriétés démontrées précédemment sur les symétries, on a :

Réflexion

PROPOSITION. *Toute réflexion vectorielle est une isométrie vectorielle. La bijection réciproque de la réflexion s_D est $s_D^{-1} = s_D$. L'ensemble E_1 des vecteurs fixés par s_D est égal à la droite D .*

b) ROTATION. On appelle *rotation* vectorielle de E toute composée de deux réflexions vectorielles.

Une rotation r est donc de la forme $r = s_D \circ s_{D'}$. Attention, il n'y a pas unicité de D et D' (cf. exercice 2 de 2.4.4).

Remarquons que $r = \text{id}_E$ si et seulement si $D = D'$.

On a pour les rotations un résultat comparable à celui formulé ci-dessus pour les réflexions :

Rotation

PROPOSITION. Toute rotation vectorielle est une isométrie vectorielle. La bijection réciproque de la rotation $r = s_{D'} \circ s_D$ est la rotation $r^{-1} = s_D \circ s_{D'}$. L'ensemble E_1 des vecteurs fixés par r est réduit à $\{0_E\}$ dès lors que $D \neq D'$.

Preuve. Soient D et D' deux droites de E , et r la rotation $s_D \circ s_{D'}$. Il est clair que r est une isométrie comme composée de deux isométries, et que $r^{-1} = (s_D \circ s_{D'})^{-1} = s_{D'}^{-1} \circ s_D^{-1} = s_{D'} \circ s_D$. Soit maintenant $u \in E$ tel que $r(u) = u$. On a donc $s_D(u) = s_{D'}(u)$. En décomposant $u = v + w = v' + w'$ avec $v \in D$, $w \in D^\perp$, $v' \in D'$, $w' \in D'^\perp$ on a $v - w = v' - w'$. En sommant membre à membre les deux égalités, il vient $v = v'$, qui est donc un vecteur de D et de D' . Comme on a supposé que $D \neq D'$, on a $D \cap D' = \{0_E\}$. Il s'ensuit que $v = v' = 0_E$. Donc $u = w = w'$. Mais alors $u \in D^\perp \cap D'^\perp$, et comme on a aussi $D^\perp \cap D'^\perp = \{0_E\}$, on conclut que $u = 0_E$. \square

c) PROPOSITION (*matrice d'une réflexion ou d'une rotation dans une base orthonormale*). Soit \mathcal{B} une base orthonormale de E . Soit f une application linéaire de E dans E . Alors :

- (i) f est une réflexion vectorielle si et seulement s'il existe deux réels a et b vérifiant $a^2 + b^2 = 1$ tels que la matrice de f relativement à la base \mathcal{B} soit égale à $\begin{pmatrix} a & b \\ b & -a \end{pmatrix}$
- (ii) f est une rotation vectorielle si et seulement s'il existe deux réels a et b vérifiant $a^2 + b^2 = 1$ tels que la matrice de f relativement à la base \mathcal{B} soit égale à $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$

Preuve. Soit f une réflexion. Soit D son axe. Soit v un vecteur unitaire de D ; notons (α, β) les composantes de v dans la base \mathcal{B} . Comme f est une isométrie, on a $\|f(v)\| = \|v\| = 1$, donc $\alpha^2 + \beta^2 = 1$. Soit w le vecteur de composantes $(-\beta, \alpha)$. Il est unitaire et orthogonal à v , donc $w \in D^\perp$. Ainsi $\mathcal{B}' = (v, w)$ est une base orthonormale de E , et il est clair que la matrice de f dans la base \mathcal{B}' est $S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. La matrice de passage de \mathcal{B} à \mathcal{B}' est $P = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$. Donc :

$$\text{Mat}_{\mathcal{B}}(f) = PSP^{-1} = \begin{pmatrix} \alpha^2 - \beta^2 & 2\alpha\beta \\ 2\alpha\beta & -\alpha^2 + \beta^2 \end{pmatrix}$$

de la forme voulue avec $a := \alpha^2 - \beta^2$ et $b := 2\alpha\beta$, qui vérifient bien $a^2 + b^2 = (\alpha^2 + \beta^2)^2 = 1$.

- Supposons maintenant qu'il existe deux réels a et b vérifiant $a^2 + b^2 = 1$ tels que la matrice de f dans \mathcal{B} soit $A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$. Elle est orthogonale (il suffit de calculer tAA et d'utiliser le fait que $a^2 + b^2 = 1$). Elle est symétrique, donc diagonalisable d'après les résultats de 2.3.2. L'exercice 3 de 2.2.4 implique alors que f est une symétrie orthogonale. Comme il est clair que $A \neq I_2$ et $A \neq -I_2$, on conclut que f est une réflexion par rapport à une droite.

- Supposons que f est une rotation. C'est la composée de deux réflexions. Il en résulte avec le point (i) qu'il existe des réels a, b, c, d vérifiant $a^2 + b^2 = c^2 + d^2 = 1$ tels que :

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \begin{pmatrix} c & d \\ d & -c \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$$

avec $\alpha = ac + bd$ et $\beta = ad - bc$, qui vérifient bien $\alpha^2 + \beta^2 = (a^2 + b^2)(c^2 + d^2) = 1$.

- Supposons enfin qu'il existe deux réels a et b vérifiant $a^2 + b^2 = 1$ tels que la matrice de f dans \mathcal{B} soit $M = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. On a $M = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ ce qui, avec le point (i), montre que f est la composée de deux réflexions, et donc que f est une rotation. \square

d) THÉORÈME. Lorsque E est un espace vectoriel euclidien de dimension 2, le groupe orthogonal $O(E)$ est constitué des rotations et des réflexions.

Plus précisément, les isométries directes du plan vectoriel E sont les rotations vectorielles, et les isométries indirectes du plan vectoriel E sont les réflexions vectorielles.

Preuve. Il résulte de la proposition précédente que, dans une base orthonormale quelconque, toute rotation est de déterminant 1 et toute réflexion est de déterminant -1 . Réciproquement, soit f une isométrie quelconque de E . Soit $M = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ sa matrice dans une base orthonormale \mathcal{B} de E . D'après le théorème d) de 2.2.2, M est une matrice orthogonale. Donc $\det M = ad - bc = \varepsilon \in \{-1, 1\}$ et $\varepsilon \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = M^{-1} = {}^t M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, d'où $d = \varepsilon a$, $c = -\varepsilon b$, et $a^2 + b^2 = 1$. Il suffit donc d'appliquer la proposition précédente pour conclure que f est une rotation si $\varepsilon = 1$ et une réflexion si $\varepsilon = -1$. \square

e) REMARQUE. Il résulte du théorème que la composée d'un nombre pair de réflexions est une rotation, la composée d'un nombre impair de réflexions est une réflexion, la composée d'une rotation et d'une réflexion est une réflexion, et la composée d'un nombre quelconque de rotations est une rotation.

f) REMARQUE. Le sous-groupe $O^+(E)$ est abélien, ce qui signifie que, quelles que soient f, g deux rotations de E , on a : $g \circ f = f \circ g$.

En effet. Il est clair que, pour $a, b, c, d \in \mathbb{R}$, $\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} c & -d \\ d & c \end{pmatrix} = \begin{pmatrix} ac-bd & -ad-bc \\ ad+bc & ac-bd \end{pmatrix} = \begin{pmatrix} c & -d \\ d & c \end{pmatrix} \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

2.4.2 Angles

E désigne ici un \mathbb{R} -espace vectoriel euclidien de dimension deux orienté.

a) THÉORÈME. Soit $r \in O^+(E)$ une rotation de E .

(i) Il existe $\theta \in \mathbb{R}$ tel que la matrice de r dans toute base orthonormale directe de E soit

$$R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

(ii) La matrice de r dans toute base orthonormale indirecte de E est alors égale à $R_{-\theta}$.

Preuve. Soit \mathcal{B} une base orthonormale directe de E . Comme on l'a vu à la proposition c) de 2.4.1, il existe $(a, b) \in \mathbb{R}^2$ satisfaisant $a^2 + b^2 = 1$ tel que la matrice de r dans \mathcal{B} soit $M = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. D'après un résultat d'analyse bien connu (mais non trivial !) l'égalité $a^2 + b^2 = 1$ implique qu'il existe $\theta \in \mathbb{R}$, non unique (défini modulo $2\pi\mathbb{Z}$), tel que $a = \cos \theta$ et $b = \sin \theta$. Donc la matrice de r par rapport à la base \mathcal{B} est R_θ .

Considérons maintenant une autre base orthonormale \mathcal{B}' de E , notons f l'isométrie transformant \mathcal{B} en \mathcal{B}' , P la matrice de passage de \mathcal{B} à \mathcal{B}' (qui n'est autre que la matrice de f dans la base \mathcal{B}), et M' la matrice de r par rapport à \mathcal{B}' . Donc $M' = P^{-1}R_\theta P$.

Supposons d'abord que \mathcal{B}' est directe; alors $f \in O^+(E)$, et comme $O^+(E)$ est abélien [remarque f) de 2.4.1], l'égalité $M' = P^{-1}R_\theta P$ devient $M' = R_\theta$. Ceci achève de prouver (i).

Pour (ii), supposons maintenant que \mathcal{B}' est indirecte. On a $f \in O^-(E)$, donc $r \circ f \in O^-(E)$, d'où $r \circ f \circ r \circ f = \text{id}$ et $f^2 = \text{id}$ puisque $r \circ f$ et f sont des réflexions vectorielles [voir théorème d) de 2.4.1]. Il en résulte que $f^{-1} \circ r \circ f = f \circ r \circ f = r^{-1}$, ce qui se traduit matriciellement par $M' = P^{-1}R_\theta P = R_\theta^{-1}$. Il est évident de vérifier que $R_\theta^{-1} = R_{-\theta}$. \square

b) DÉFINITION. Pour tout réel θ , notons r_θ l'unique rotation dont la matrice par rapport à toute base orthonormale directe est la matrice R_θ . On l'appelle la *rotation d'angle θ* .

Remarque : on devrait plus rigoureusement, conformément à ce qui suit, l'appeler la rotation dont l'angle a pour mesure θ .

c) PROPOSITION. Avec les notations ci-dessus, on a :

(i) Pour tous $\theta, \theta' \in \mathbb{R}$, on a : $(r_\theta = r_{\theta'}) \Leftrightarrow (R_\theta = R_{\theta'}) \Leftrightarrow (\exists k \in \mathbb{Z}, \theta' = \theta + 2k\pi)$.

(ii) En particulier : $r_0 = \text{id}_E$ et $r_\pi = r_{-\pi} = -\text{id}_E$.

(iii) Pour tous $\theta, \theta' \in \mathbb{R}$, on a : $R_{\theta+\theta'} = R_\theta R_{\theta'}$ et $r_{\theta+\theta'} = r_\theta \circ r_{\theta'}$

Preuve. Les points (i) et (ii) sont clairs. Pour (iii), considérons deux réels quelconque θ et θ' , et les deux rotations r_θ et $r_{\theta'}$ de matrices respectives R_θ et $R_{\theta'}$ par rapport à une même base orthonormale. La matrice de $r_\theta \circ r_{\theta'}$ dans cette base est :

$$\begin{aligned} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \theta' & -\sin \theta' \\ \sin \theta' & \cos \theta' \end{pmatrix} &= \begin{pmatrix} \cos \theta \cos \theta' - \sin \theta \sin \theta' & -\sin \theta \cos \theta' - \cos \theta \sin \theta' \\ \sin \theta \cos \theta' + \cos \theta \sin \theta' & \cos \theta \cos \theta' - \sin \theta \sin \theta' \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta + \theta') & -\sin(\theta + \theta') \\ \sin(\theta + \theta') & \cos(\theta + \theta') \end{pmatrix}, \text{ d'où le résultat. } \quad \square \end{aligned}$$

Algébriquement, on traduit ces propriétés en disant que l'application $r : \mathbb{R} \rightarrow O^+(E)$ définie par $\theta \mapsto r_\theta$ est un morphisme de groupes, surjectif, de noyau $2\pi\mathbb{Z}$.

A noter que, comme $\theta + \theta' = \theta' + \theta$, on retrouve le fait déjà observé à la remarque f) de 2.4.1 que les rotations commutent entre elles pour la loi \circ .

Géométriquement, le lemme suivant va permettre de passer de la notion d'angle d'une rotation à celle d'angle de vecteurs.

d) LEMME. Soient u et v deux vecteurs non-nuls de même norme. Il existe une unique rotation vectorielle r de E telle que $r(u) = v$.

1. *Preuve géométrique.* Parce que $\|u\| = \|v\|$, les vecteurs $u+v$ et $u-v$ sont orthogonaux. Il en résulte que, si l'on appelle s la réflexion vectorielle d'axe dirigé par $u+v$, on a $s(u) = v$. Donc, en notant s' la réflexion d'axe dirigé par u , la rotation $r = s \circ s'$ vérifie $r(u) = s(s'(u)) = s(u) = v$.

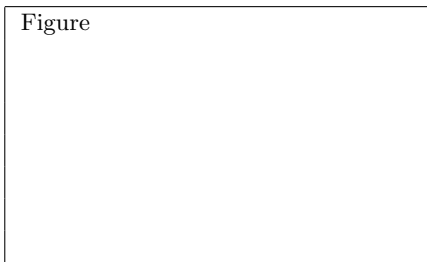
Pour l'unicité, supposons qu'il existe deux rotations r et r' telles que $r(u) = v = r'(u)$. Posons $r'' = r^{-1} \circ r'$.

C'est une rotation (composée de deux rotations), qui vérifie $r''(u) = u$ par construction et fixe donc un vecteur non-nul de E . On en déduit (proposition b de 2.4.1) que $r'' = \text{id}_E$, donc $r = r'$. \square

2. *Preuve analytique.* Plaçons-nous dans une base orthonormale \mathcal{B} . Soient (x, y) les composantes de u , et (x', y') celles de v . On cherche à déterminer les couples $(a, b) \in \mathbb{R}^2$ tels que : $\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix}$ et $a^2 + b^2 = 1$. Or le système $\begin{cases} xa - yb = x' \\ ya + xb = y' \end{cases}$ a pour déterminant $x^2 + y^2 \neq 0$, donc il admet une unique solution donnée par les formules de Cramer :

$$a = \frac{xx' + yy'}{x^2 + y^2} \quad \text{et} \quad b = \frac{xy' - yx'}{x^2 + y^2}, \quad \text{d'où} \quad a^2 + b^2 = \frac{(x^2 + y^2)(x'^2 + y'^2)}{(x^2 + y^2)^2} = 1.$$

D'où l'existence et l'unicité de la solution cherchée. \square



e) DÉFINITIONS. On se place toujours dans un plan vectoriel euclidien orienté E .

• Soient deux vecteurs non-nuls u et v dans E . On peut toujours se ramener à des vecteurs de même norme en considérant les vecteurs *unitaires*

$$u_1 = \frac{1}{\|u\|}u \quad \text{et} \quad v_1 = \frac{1}{\|v\|}v,$$

et il existe d'après le lemme ci-dessus une unique rotation r telle que $r(u_1) = v_1$.

• Soient quatre vecteurs non-nuls u, v, u', v' dans E . On dit que (u, v) et (u', v') déterminent le même *angle* lorsque l'unique rotation r telle que $r(u_1) = v_1$ vérifie aussi $r(u'_1) = v'_1$. On note alors :

$$\widehat{(u, v)} = \widehat{(u', v')}$$

Plus formellement, on définit une relation entre couples de vecteurs non-nuls traduisant le fait que l'unique rotation r telle que $r(u_1) = v_1$ est celle qui vérifie aussi $r(u'_1) = v'_1$; on montre que cette relation est une relation d'équivalence dans l'ensemble des couples de vecteurs non-nuls de E , et un angle est une classe d'équivalence pour cette relation.

A ce niveau, il ne serait pas nécessaire de supposer E orienté pour définir la notion d'angle de vecteurs; cela est en revanche nécessaire pour parler de mesure d'angle.



• On appelle *mesure* d'un angle de vecteurs $\widehat{(u, v)}$ tout réel θ tel que l'unique rotation qui envoie u_1 sur v_1 soit la rotation r_θ définie en c) ci-dessus.

$$[\theta \text{ est une mesure de } \widehat{(u, v)}] \Leftrightarrow [r_\theta(u_1) = v_1].$$

La mesure d'un angle n'est pas unique, mais définie modulo 2π : si θ est une mesure de $\widehat{(u, v)}$, ses autres mesures sont les réels $\theta + 2k\pi$ avec $k \in \mathbb{Z}$.

Exemples :

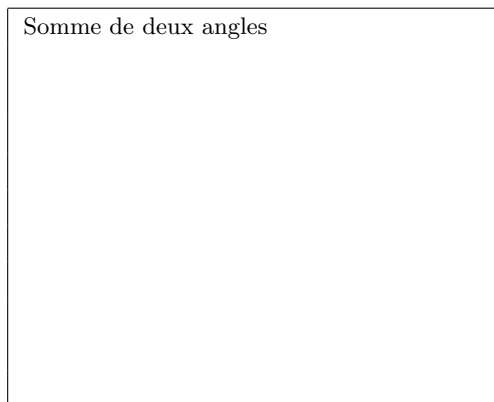
1. l'angle nul est l'angle égal à $\widehat{(u, u)}$ pour tout vecteur non-nul $u \in E$; la rotation correspondante est $\text{id}_E = r_0$; ses mesures sont tous les $2k\pi$ avec $k \in \mathbb{Z}$.
2. l'angle plat est l'angle égal à $\widehat{(u, -u)}$ pour tout vecteur non-nul $u \in E$; la rotation correspondante est $-\text{id}_E = r_\pi$; ses mesures sont tous les $\pi + 2k\pi$ avec $k \in \mathbb{Z}$.
3. deux vecteurs non-nuls $u, v \in E$ forment une base orthonormale directe de E si et seulement si une mesure de $\widehat{(u, v)}$ est $\frac{\pi}{2}$

• Soient u, v, u', v' quatre vecteurs non-nuls de E . On peut toujours choisir un vecteur w tel que $\widehat{(u', v')} = \widehat{(v, w)}$, et l'on définit alors par une forme de relation de Chasles la *somme des deux angles* :

$$\widehat{(u, v)} + \widehat{(u', v')} = \widehat{(u, v)} + \widehat{(v, w)} = \widehat{(u, w)}.$$

La somme de deux angles étant ainsi définie par la composition des deux rotations correspondantes, il résulte directement de cette définition que,

si θ est une mesure de $\widehat{(u, v)}$ et θ' est une mesure de $\widehat{(u', v')}$, alors une mesure de $\widehat{(u, v)} + \widehat{(u', v')}$ est $\theta + \theta'$.



REMARQUE. C'est ce principe qui est à la base des systèmes pratiques de mesure des angles. Comme il établit que la mesure de la somme de deux angles est la somme de leurs mesures respectives, il permet de définir à partir d'une valeur arbitraire d'un angle donné (généralement l'angle plat) de calculer la mesure des différentes fractions de cet angle (suivant l'unité choisie : π en radians, 180 en degrés, 200 en grades).

- L'opposé d'un angle $\widehat{(u, v)}$ est l'angle $\widehat{(v, u)}$, et ceci indépendamment du couple de vecteurs non-nuls choisis.

La somme d'un angle et de son opposé est l'angle nul puisque $\widehat{(u, v)} + \widehat{(v, u)} = \widehat{(u, u)}$.

Opposé d'un angle

- Comme les différentes mesure d'un même angle diffèrent d'un multiple entier de 2π , on peut sans ambiguïté définir *le cosinus et le sinus d'un angle* comme le cosinus et le sinus de l'une quelconque de ses mesures. On note :

$$\cos \widehat{(u, v)} = \cos \theta \quad \text{et} \quad \sin \widehat{(u, v)} = \sin \theta, \quad \text{où } \theta \text{ est une mesure de l'angle } \widehat{(u, v)}.$$

f) THÉORÈME. Pour tout couple (u, v) de vecteurs non-nuls d'un plan vectoriel euclidien orienté E , on a :

$$\cos \widehat{(u, v)} = \frac{\langle u | v \rangle}{\|u\| \|v\|} \quad \text{et} \quad \sin \widehat{(u, v)} = \frac{[u, v]}{\|u\| \|v\|}.$$

Preuve. Définissons u_1, v_1 unitaires par : $u_1 = \frac{1}{\|u\|}u$ et $v_1 = \frac{1}{\|v\|}v$. Soit θ une mesure de $\widehat{(u_1, v_1)} = \widehat{(u, v)}$.

Ainsi : $r_\theta(u_1) = v_1$.

Soit w_1 le vecteur unitaire orthogonal à u_1 tel que la base orthonormale $\mathcal{B}_1 = \{u_1, w_1\}$ soit directe.

Dans la base \mathcal{B}_1 , les composantes de u_1 sont $(1, 0)$.

Dans la base \mathcal{B}_1 , les composantes de v_1 forment la première colonne de R_θ (puisque $v_1 = r_\theta(u_1)$) et $\text{Mat}_{\mathcal{B}_1}(r_\theta) = R_\theta$ et sont donc égales à $(\cos \theta, \sin \theta)$.

Donc $\langle u_1 | v_1 \rangle = \cos \theta$ et $[u_1, v_1] = \det \text{Mat}_{\mathcal{B}_1}(u_1, v_1) = \sin \theta$. D'où le résultat. \square

Figure

Dans la preuve ci-dessus, on travaille dans la base orthonormale directe \mathcal{B}_1 adaptée à la donné de u et v . Si maintenant on prend une base orthonormale directe \mathcal{B} quelconque, et si l'on note (x, y) et (x', y') les composantes respectives de u et v dans la base \mathcal{B} , on obtient :

$$\cos \widehat{(u, v)} = \frac{xx' + yy'}{x^2 + y^2} \quad \text{et} \quad \sin \widehat{(u, v)} = \frac{xy' - yx'}{x^2 + y^2},$$

ce que l'on avait déjà observé dans la preuve du lemme c).

2.4.3 Description des isométries vectorielles en dimension 3

Dans toute ce paragraphe, E est un espace vectoriel euclidien orienté de dimension 3.

a) PROPOSITION ET DÉFINITION. Soit $u \in E$ un vecteur non-nul. Soit θ un réel. Il existe une isométrie directe f de E telle que la matrice de f dans toute base orthonormale directe \mathcal{B} de E admettant $u_1 = \frac{1}{\|u\|}u$ comme premier vecteur est :

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}.$$

Cette isométrie est appelé la rotation vectorielle d'axe dirigé et orienté par u , et d'angle θ .

Preuve. Soit $\theta \in \mathbb{R}$. Soient $u \in E$ non-nul, et $u_1 = \frac{1}{\|u\|}u$. Soient u_2, u_3 deux vecteurs tels que (u_1, u_2, u_3) constitue une base orthonormale directe \mathcal{B} de E .

Soit f l'endomorphisme de E dont la matrice dans la base \mathcal{B} est $A_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$.

On vérifie aisément que : ${}^t A_\theta A_\theta = I_3$ et $\det A_\theta = +1$, et donc $f \in O^+(E)$. Si \mathcal{B}' est une autre base orthonormale directe de la forme (u_1, v_2, v_3) , la matrice de passage P de \mathcal{B} à \mathcal{B}' est de la forme $\begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix}$ avec $R = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in O^+(2, \mathbb{R})$. Parce que $O^+(2, \mathbb{R})$ est abélien (on l'a vu à la remarque f) de 2.4.1), on a $\text{Mat}_{\mathcal{B}'}(f) = P^{-1}A_\theta P = A_\theta$. \square

b) DESCRIPTION GÉOMÉTRIQUE.

Considérons D la droite de base u_1 et H le plan de base (u_2, u_3) , qui sont deux sous-espaces orthogonaux.

La restriction de f à la droite D est id_D .

La restriction de f au plan $H = D^\perp$ est la rotation vectorielle plane d'angle θ , de matrice $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ dans la base (u_2, u_3) .

Rotation en dimension 3

c) THÉORÈME. Les isométries directes de E sont les rotations vectorielles.

Preuve. Soit $f \in O^+(E)$. En particulier, $\det f = 1$. Son polynôme caractéristique détermine une fonction polynomiale $\mathbb{R} \rightarrow \mathbb{R}$ de la forme $P_f(x) = -x^3 + \dots + \det f = -x^3 + \dots + 1$. Cette fonction est continue sur \mathbb{R} , tend vers $-\infty$ quand x tend vers $+\infty$, et prend la valeur positive $+1$ en $x = 0$: il résulte alors du théorème des valeurs intermédiaire qu'il existe $\lambda \in]0, +\infty[$ tel que $P_f(\lambda) = 0$. Donc λ est une valeur propre de f , qui vérifie $\lambda > 0$. Or on a vu à l'ex. 2 de 2.2.4 que les seules v.p. possibles d'une isométrie sont 1 et -1 . Donc ici $\lambda = 1$.

On vient d'établir qu'il existe $u \in E$ non-nul tel que $f(u) = u$. Normalisons-le en $u_1 = \frac{1}{\|u\|}u$, et choisissons deux vecteurs u_2, u_3 de E tel que $\mathcal{B} = (u_1, u_2, u_3)$ soit une base orthonormale directe de E . Appelons D la droite de base u_1 dans E , et H le plan D^\perp , dont une base orthonormale est (u_2, u_3) . Parce que D est évidemment stable par f , il en est de même de son orthogonal H (d'après l'ex. 1 de 2.2.4). Donc la matrice de f dans la base \mathcal{B} est de la forme $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & c \\ 0 & b & d \end{pmatrix}$. On conclut en écrivant que A est orthogonale et de déterminant $+1$. \square

d) REMARQUE. On peut de même obtenir une description concrète des isométries indirectes en dimension 3. Sans développer ici la démonstration, citons pour mémoire le résultat : les isométries indirectes de E sont : $-\text{id}_E$, les symétries orthogonales par rapport à un plan H (appelées réflexions par rapport à H), les composées d'une réflexion par rapport à un plan H avec une rotation d'axe H^\perp et d'angle dans $]0, \pi[$.

2.4.4 Exercices

EXERCICE 1. Dans chacun des cas suivants, on considère une matrice $M \in \mathcal{M}_2(\mathbb{R})$ et l'endomorphisme f d'un plan vectoriel euclidien orienté E qui admet M pour matrice dans une base orthonormale directe de \mathcal{E} . Dans chaque cas, montrer que M est une matrice orthogonale, préciser si elle est directe ou indirecte, et décrire géométriquement f :

$$A = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}, \quad B = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad C = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{pmatrix}.$$

EXERCICE 2 (*non-unicité de l'écriture d'une rotation comme composée de deux réflexions*). Soit E un plan vectoriel euclidien. Soit r est une rotation vectorielle de E . Montrer que, pour toute droite D , il existe une droite D' telle que $r = s_D \circ s_{D'}$ et une droite D'' telle que $r = s_{D''} \circ s_D$ [indication : utiliser la remarque f) de 2.4.1 pour introduire D' comme axe de la réflexion $s_D \circ r$]. Faire une figure.

EXERCICE 3. Montrer que, de façon similaire au lemme d) de 2.4.2, on a aussi : si u et v sont deux vecteurs non-nuls de même norme d'un plan vectoriel euclidien E , il existe une unique réflexion s de E telle que $s(u) = v$.

EXERCICE 4. Montrer que, dans un plan vectoriel euclidien orienté E , toute réflexion transforme un angle orienté en son opposé [indication : pour une réflexion s_D d'axe D , et deux vecteurs non-nuls quelconque $u, v \in E$, considérer la rotation $s_D \circ s_{D'}$ où $D' = (u - v)^\perp$].

EXERCICE 5. Soient u, v, u', v' quatre vecteurs de même normes dans un plan vectoriel euclidien orienté E . Montrer que les assertions suivantes sont équivalentes : (i) l'unique rotation qui envoie u sur u' est égale à l'unique rotation qui envoie v sur v' ; (ii) l'unique rotation qui envoie u sur v est égale à l'unique rotation qui envoie u' sur v' . En déduire que :

$$\widehat{(u, v)} = \widehat{(u', v')} \Leftrightarrow \widehat{(u, u')} = \widehat{(v, v')}.$$

EXERCICE 6 (*Isométries vectorielles en dimension 3*). Soit $\mathcal{B} = (e_1, e_2, e_3)$ une base orthonormale directe d'un espace vectoriel euclidien E de dimension 3. Donner la matrice dans la base \mathcal{B} de la rotation d'axe dirigé en orienté par $e_1 + e_2 + e_3$, et d'angle $\frac{2\pi}{3}$.

EXERCICE 7 (*Isométries vectorielles en dimension 3*). Déterminer la nature géométrique de l'endomorphisme de E dont la matrice dans une base orthonormale d'un espace vectoriel euclidien de dimension 3 est $A = \frac{1}{3} \begin{pmatrix} 2 & 1 & 2 \\ -2 & 2 & 1 \\ -1 & -2 & 2 \end{pmatrix}$.

EXERCICE 8 (*Isométries vectorielles en dimension 3*). Soit E un espace vectoriel euclidien de dimension 3, rapporté à une base orthonormale \mathcal{B} . Pour tous réel a, b, c , on note $f_{a,b,c}$ l'endomorphisme de E dont la matrice par rapport à la base \mathcal{B} est : $M_{a,b,c} = \begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}$.

– Montrer que $f_{a,b,c} \in O(E)$ si et seulement si $|a + b + c| = 1$ et $ab + bc + ca = 0$.

– Montrer qu'un endomorphisme $f_{a,b,c}$ qui appartient à $O(E)$ est involutif si et seulement si $b = c$.

– Montrer que l'ensemble G des endomorphismes $f_{a,b,c}$ qui appartiennent à $O(E)$ et qui sont involutifs est un groupe fini, dont on déterminera explicitement tous les éléments, et dont on dressera la table pour la loi \circ .

– Décrire géométriquement chacun des éléments de G , en indiquant s'il s'agit d'un élément de $O^+(E)$ ou de $O^-(E)$, et en précisant le sous-espace des vecteurs invariants.

EXERCICE 9 (*Engendrement du groupe orthogonal par les symétries hyperplanes*). Démontrer que toute isométrie vectorielle f d'un espace vectoriel euclidien E de dimension n (autre que

l'identité) s'écrit sous la forme $f = s_1 \circ s_2 \circ \dots \circ s_p$ où $1 \leq p \leq n$ et s_1, s_2, \dots, s_p sont des symétries hyperplanes.

[Indication : s'il existe $u \in E$ non-nul tel que $f(u) = u$, considérer la droite vectorielle D dirigée par u , l'hyperplan $H = D^\perp$, et achever par récurrence; sinon, raisonner de même avec D la droite vectorielle dirigée par $u - f(u)$ pour u un vecteur non-nul de E .]

EXERCICE 10 (*Centre du groupe orthogonal*). Soit E un espace vectoriel euclidien de dimension $n \geq 2$. On appelle centre de $O(E)$ l'ensemble Z de tous les $f \in O(E)$ qui commutent pour la loi \circ avec tous les éléments de $O(E)$:

$$Z = \{f \in O(E); g \circ f = f \circ g \text{ pour tout } g \in O(E)\}.$$

On définit de même le centre Z^+ de $O^+(E)$. Le but de l'exercice est de déterminer Z et Z^+ .

– Soit H un sous-espace vectoriel de E . On note s_H la symétrie orthogonale de E par rapport à H . Soit $f \in O(E)$ quelconque; posons $h = f \circ s_H \circ f^{-1}$. Montrer que $h(u) = u$ pour tout $u \in f(H)$ et que $h(u) = -u$ pour tout $u \in f(H^\perp)$. En déduire que $h = s_{f(H)}$. Conclure que :

$$\text{pour tout } f \in O(E), [f \circ s_H = s_H \circ f] \Leftrightarrow [f(H) = H]. \quad (*)$$

– Montrer que si $f \in O(E)$ vérifie $f(\Delta) = \Delta$ pour toute droite Δ de E , alors $f = \text{id}_E$ ou $f = -\text{id}_E$. En déduire que si $\dim E \geq 3$ et si $f \in O(E)$ vérifie $f(\Pi) = \Pi$ pour tout plan Π de E , alors $f = \text{id}_E$ ou $f = -\text{id}_E$.

– A l'aide de (*) et de la question précédente, déterminer quels sont les éléments de Z . Déterminer de même quels sont les éléments de Z^+ (on pourra distinguer trois cas suivant que $\dim E$ est impaire, ou paire et au moins égale à 4, ou égale à 2).

Chapitre 3

Espaces vectoriels normés

L'objectif de ce chapitre est d'introduire des notions d'analyse bien connues dans \mathbb{R} ou \mathbb{C} (convergence de suites, limites de fonctions,...) dans un contexte plus général où les points ne sont plus des nombres réels ou complexes mais des éléments d'un espace vectoriel E . Pour cela, la "proximité" entre deux points (qui s'exprime dans \mathbb{R} avec la valeur absolue de leur différence, et dans \mathbb{C} avec le module de leur différence) fait intervenir la notion plus générale de norme dans E .

Dans toute la suite, la lettre \mathbb{K} désigne \mathbb{R} ou \mathbb{C} . On travaille dans ce chapitre dans un \mathbb{K} -espace vectoriel E . Suivant les applications et exemples considérés, E pourra être de dimension finie ($E = \mathbb{K}^n$ pour un certain entier $n \geq 1$), ou de dimension infinie (par exemple $E = \mathbb{K}[X]$, ou $E = \mathcal{C}^0(I, \mathbb{K})$, ou d'autres espaces de fonctions...)

3.1 Normes et notions topologiques associées

3.1.1 Notion de norme

a) DÉFINITION. On appelle *norme* sur un \mathbb{K} -espace vectoriel E toute application $N : E \rightarrow \mathbb{R}_+$ vérifiant les trois propriétés suivantes :

- (i) pour tout $\lambda \in \mathbb{K}$ et tout $x \in E$, on a $N(\lambda x) = |\lambda|N(x)$,
- (ii) pour tout $x \in E$, $N(x) = 0$ si et seulement si $x = 0_E$,
- (iii) pour tous $x, y \in E$, $N(x + y) \leq N(x) + N(y)$.

La troisième propriété est appelée l'inégalité triangulaire.

Commentaire : on peut montrer que les conditions (i) et (iii) impliquent que $N(x) \geq 0$ pour tout $x \in E$, de sorte que l'on pourrait se contenter de prendre $N : E \rightarrow \mathbb{R}$ au départ.

Un espace vectoriel E ainsi muni d'une norme N est appelé un *espace vectoriel normé*, ce que l'on abrège traditionnellement en evn.

Une autre notation usuelle pour une norme est $\|x\|$ au lieu de $N(x)$.

b) EXEMPLES.

- La valeur absolue définit une norme sur \mathbb{R} (vu comme \mathbb{R} -espace vectoriel de dimension 1).

- Le module définit une norme sur \mathbb{C} (vu comme un \mathbb{C} -espace vectoriel de dimension 1, ou comme un \mathbb{R} -espace vectoriel de dimension 2).
- La norme associée à un produit scalaire sur un espace vectoriel euclidien E est une norme [voir points c), d) et f) de 2.1.1].

c) EXEMPLES : LES TROIS NORMES USUELLES SUR \mathbb{K}^n . Pour $E = \mathbb{K}^n$ et pour tout vecteur $x = (x_1, \dots, x_n)$ de E , on pose :

$$\|x\|_1 = \sum_{i=1}^n |x_i|, \quad \|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2}, \quad \|x\|_\infty = \sup_{1 \leq i \leq n} |x_i| = \max_{1 \leq i \leq n} |x_i|.$$

On définit ainsi trois normes sur \mathbb{K}^n (preuve en exercice). La seconde est appelée norme euclidienne lorsque $\mathbb{K} = \mathbb{R}$ et norme hermitienne lorsque $\mathbb{K} = \mathbb{C}$.

d) EXEMPLE : NORME DE LA CONVERGENCE UNIFORME. On prend ici $E = \mathcal{B}(X, \mathbb{K})$ le \mathbb{K} -espace vectoriel des fonctions $X \rightarrow \mathbb{K}$ qui sont bornées sur X (où X est un ensemble quelconque non vide). Pour toute fonction $f \in E$, on pose :

$$\|f\|_\infty = \sup_{x \in X} |f(x)|.$$

On définit ainsi une norme sur $\mathcal{B}(X, \mathbb{K})$ (preuve en exercice). Elle est appelée (on verra plus loin pourquoi ce nom) norme de la convergence uniforme.

e) EXEMPLE : NORMES 1 ET 2 POUR LES FONCTIONS INTÉGRABLES. On prend ici $E = \mathcal{C}^0([a, b], \mathbb{K})$ le \mathbb{K} -espace vectoriel des fonctions continues $[a, b] \rightarrow \mathbb{K}$ (où $[a, b]$ est un intervalle fermé borné non vide de \mathbb{R}). Pour toute fonction $f \in E$, on pose :

$$\|f\|_1 = \int_a^b |f(t)| dt \quad \text{et} \quad \|f\|_2 = \sqrt{\int_a^b |f(t)|^2 dt}.$$

On définit ainsi deux normes sur $\mathcal{C}^0([a, b], \mathbb{K})$ (preuve en exercice). Elles sont appelées respectivement la norme 1 et la norme 2 sur E , qui sont des cas particuliers de l'exemple suivant.

f) EXEMPLE : NORMES DE HÖLDER. On fixe un réel $p \in]1, +\infty[$.

- On définit une norme sur $E = \mathbb{K}^n$ (où $n \geq 1$ est un entier fixé) en posant, pour tout $x = (x_1, \dots, x_n) \in E$:

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}.$$

L'exemple c) ci-dessus correspond aux cas $p = 1$ et $p = 2$, et l'on a $\lim_{p \rightarrow \infty} \|x\|_p = \|x\|_\infty$ pour tout $x \in E$.

- On définit une norme sur $E = \mathcal{C}^0([a, b], \mathbb{K})$ (où $[a, b]$ est un intervalle fermé borné non vide de \mathbb{R}) en posant, pour toute $f \in E$:

$$\|f\|_p = \left(\int_a^b |f(t)|^p dt \right)^{1/p}.$$

L'exemple e) ci-dessus correspond aux cas $p = 1$ et $p = 2$, et l'on a $\lim_{p \rightarrow \infty} \|f\|_p = \|f\|_\infty$ pour toute $f \in E$.

- Les preuves de ces résultats pourront être établies en exercice.

3.1.2 Distance associée à une norme

a) DÉFINITION. Soit E un evn. On note $\|\cdot\|$ sa norme et l'on pose :

$$d(x, y) = \|x - y\| \quad \text{pour tous } x, y \in E.$$

On définit ainsi une application $d : E^2 \rightarrow \mathbb{R}_+$ que l'on appelle la *distance* associée à la norme.

Remarque : en particulier la norme d'un vecteur est la distance entre ce vecteur et le vecteur nul : $\|x\| = d(x, 0_E)$ pour tout $x \in E$.

b) PROPOSITION. Soient E un evn, et d la distance associée à sa norme $\|\cdot\|$. Quels que soient les vecteurs x, y, z de E , on a :

- $d(x, y) = d(y, x)$, ← symétrie
- $d(x, y) = 0$ si et seulement si $x = y$, ← séparation
- $d(x, z) \leq d(x, y) + d(y, z)$, ← inégalité triangulaire
- pour tout $\lambda \in \mathbb{K}$, $d(\lambda x, \lambda y) = |\lambda|d(x, y)$, ← homogénéité
- $d(x + z, y + z) = d(x, y)$. ← invariance par translation

Preuve. Calculs élémentaires, laissés en exercice. □

c) REMARQUES. Les trois premières propriétés sont celles qui, de façon générale, définissent une distance sur un ensemble E . On peut montrer qu'une distance sur un espace vectoriel E qui de plus satisfait les quatrième et cinquième conditions est nécessairement associée à une norme sur E . On pourra à titre d'exercice, dans le cas où $E = \mathbb{R}^2$ muni de la norme euclidienne, illustrer par des dessins les trois dernières propriétés.

inégalité triangulaire	homogénéité	invariance par translation

d) LEMME PRATIQUE. Soient E un evn, et d la distance associée à sa norme $\|\cdot\|$. Quels que soient les vecteurs x, y, z de E , on a :

$$\|x - y\| \geq \left| \|x\| - \|y\| \right| \quad \text{et} \quad d(x, y) \geq |d(x, z) - d(y, z)|.$$

Preuve. $\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\|$ et donc $\|x - y\| \geq \|x\| - \|y\|$. Comme x et y jouent des rôles symétriques, on a de même $\|x - y\| \geq \|y\| - \|x\|$. Ainsi $\|x - y\|$ est supérieur ou égal au plus grand des nombres $\|x\| - \|y\|$ et $\|y\| - \|x\|$, ce qui prouve la première inégalité. Pour la seconde, il suffit de remarquer que $d(x, y) = \|x - y\| = \|(x - z) - (y - z)\| \geq \left| \|x - z\| - \|y - z\| \right| = |d(x, z) - d(y, z)|$. □

3.1.3 Boules et parties bornées

a) DÉFINITIONS. Soient E un evn, et d la distance associée à sa norme $\|\cdot\|$. Pour tout $a \in E$ et tout réel $r > 0$, on appelle *boule ouverte de centre a et de rayon r* l'ensemble noté $B(a, r)$ des vecteurs de E situés à une distance de a strictement inférieure à r :

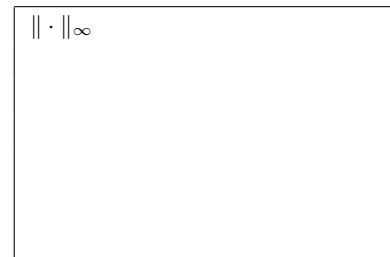
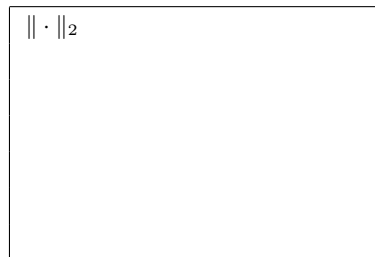
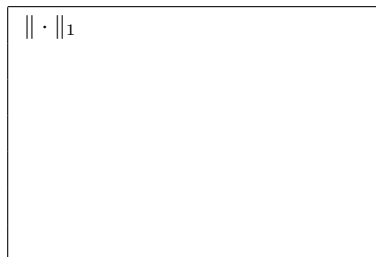
$$B(a, r) = \{x \in E; d(a, x) < r\} = \{x \in E; \|x - a\| < r\}.$$

De façon naturelle, on définit de même la boule fermée $\overline{B}(a, r)$ et la sphère $S(a, r)$ par :

$$\overline{B}(a, r) = \{x \in E; d(a, x) \leq r\} \quad \text{et} \quad S(a, r) = \{x \in E; d(a, x) = r\}.$$

Une boule n'a qu'un centre et qu'un rayon, qui la déterminent complètement ; en d'autres termes, quels que soient $a, b \in E$ et r, s des réels strictement positifs, $B(a, r) = B(b, s)$ si et seulement si $a = b$ et $r = s$ (voir exercice 6 de 3.1.7).

b) ILLUSTRATION. Dans le cas où $E = \mathbb{R}^2$, dessiner les boules $B(0, 1)$ pour chacune des trois normes $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$.



c) PROPOSITION ET DÉFINITION. Soient E un evn, et d la distance associée à sa norme $\|\cdot\|$. Soit A une partie de E . Les assertions suivantes sont équivalentes :

- (i) il existe un réel positif M tel que $d(x, y) \leq M$ pour tous $x, y \in A$;
- (ii) il existe un réel positif r tel que $\|x\| \leq r$ pour tout $x \in A$.

Lorsque ces conditions sont satisfaites, on dit que A est *bornée*.

Preuve On peut supposer A non-vide. Supposons (i) vérifié. Fixons un élément $a \in A$. Alors pour tout $x \in A$, on a : $\|x\| = \|x - a + a\| \leq \|x - a\| + \|a\| \leq M + \|a\|$, d'où l'assertion (ii) avec $r = M + \|a\|$. Supposons (ii) vérifié. Alors pour tous $x, y \in A$, on a : $d(x, y) = \|x - y\| \leq \|x\| + \|y\| \leq r + r$, d'où l'assertion (i) avec $M = 2r$. \square

En d'autres termes, une partie est bornée lorsqu'il existe une boule (ouverte ou fermée) qui la contient (faire un dessin!). Les propriétés suivantes sont immédiates (preuves en exercice) :

- toute partie incluse dans une partie bornée est bornée,
- toute réunion d'un nombre fini de parties bornées est bornée,
- toute partie finie est bornée.

Il est parfois utile, pour une partie bornée A de E , de considérer son *diamètre*, défini par :

$$\text{diam } A = \sup_{x, y \in A} d(x, y).$$

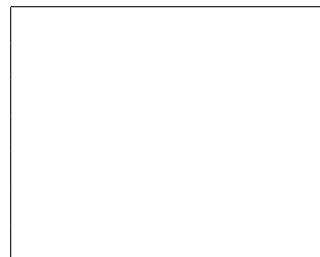
3.1.4 Parties ouvertes

a) **EXEMPLE INTRODUCTIF.** Fixons dans un evn E une boule ouverte $A = B(a, r)$, avec $a \in E$ et $r > 0$ fixés.

Soit x un élément quelconque de A . On a donc $d(a, x) < r$. Posons alors $\delta = r - d(a, x) > 0$. Pour tout $y \in B(x, \delta)$, on a : $d(a, y) \leq d(a, x) + d(x, y) < d(a, x) + \delta = r$, et donc $y \in A$.

On a ainsi prouvé que :

pour tout $x \in A$, il existe un réel $\delta > 0$ tel que $B(x, \delta) \subseteq A$.

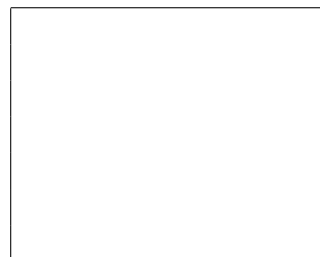


b) **DÉFINITION.** Soit E un evn. Une partie A de E est dite *ouverte* lorsqu'elle vérifie la propriété suivante :

pour tout $x \in A$, il existe un réel $\delta > 0$ tel que $B(x, \delta) \subseteq A$.

- On traduit parfois cette définition en disant qu'une partie est ouverte lorsqu'elle est un voisinage de chacun de ses points.

- Par convention, l'ensemble vide \emptyset est considéré comme ouvert.



c) **EXEMPLES.**

- E est une partie ouverte de E (immédiat)
- Toute boule ouverte dans E est une partie ouverte de E (voir exemple introductif).
- Toute réunion d'une famille quelconque (finie ou infinie) de parties ouvertes de E est encore une partie ouverte de E (preuve immédiate, à rédiger).
- Toute intersection d'une famille *finie* de parties ouvertes de E est encore une partie ouverte de E (preuve immédiate, à rédiger).

Attention, l'intersection d'une famille infinie de parties ouvertes peut ne pas être ouverte (contre-exemple : prendre $E = \mathbb{R}$ muni de la norme de la valeur absolue, et $A_n =]-\frac{1}{n}, \frac{1}{n}[$ pour tout entier $n \geq 1$; chaque A_n est une partie ouverte et l'intersection $\bigcap_{n \geq 1} A_n$ est le singleton $\{0\}$, qui n'est pas une partie ouverte).

d) **PROPOSITION** (une définition équivalente d'une partie ouverte). *Une partie A d'un evn E est ouverte si et seulement si elle est une réunion de boules ouvertes de E .*

Preuve. La condition est clairement suffisante puisque chaque boule ouverte est une partie ouverte et qu'une réunion de parties ouvertes est encore une partie ouverte.

Pour montrer qu'elle est nécessaire, supposons A ouverte. Pour tout $a \in A$, il existe une boule ouverte $B(a, r_a)$ incluse dans A . Donc $\bigcup_{a \in A} B(a, r_a) \subseteq A$. Mais par ailleurs $\bigcup_{a \in A} B(a, r_a) \supseteq A$ puisque chaque $a \in A$ appartient à la boule $B(a, r_a)$. D'où $A = \bigcup_{a \in A} B(a, r_a)$. \square

e) **DÉFINITION.** Soit A une partie d'un evn E . On appelle *intérieur* de A , noté $\overset{\circ}{A}$, la plus grande (au sens de l'inclusion) partie ouverte de E incluse dans A , c'est-à-dire encore la réunion de toute les parties ouvertes de E incluses dans A . En particulier :

[on a toujours : $\overset{\circ}{A} \subseteq A$] et [$A = \overset{\circ}{A}$ si et seulement si A est ouverte].

f) **PROPOSITION.** *Soit A une partie d'un evn E . Pour tout $x \in A$, on a :*

$x \in \overset{\circ}{A}$ si et seulement s'il existe un réel $\delta > 0$ tel que $B(x, \delta) \subseteq A$.

Preuve. Supposons $x \in \overset{\circ}{A}$. Comme $\overset{\circ}{A}$ est ouverte, il existe une boule $B(x, \delta)$ incluse dans $\overset{\circ}{A}$, donc incluse dans A . Supposons réciproquement que $x \in A$ soit tel qu'il existe une boule $B(x, \delta)$ incluse dans A . Comme $B(x, \delta)$ est une partie ouverte incluse dans A , il résulte de la définition de $\overset{\circ}{A}$ que $B(x, \delta) \subseteq \overset{\circ}{A}$, et en particulier $x \in \overset{\circ}{A}$. \square

3.1.5 Parties fermées

a) DÉFINITION. Soit E un evn. Une partie A de E est dite *fermée* lorsque son complémentaire dans E est une partie ouverte.

b) EXEMPLES.

- E et \emptyset sont parties fermées de E (immédiat)
- Toute boule fermée dans E est une partie fermée de E .

En effet, soit $A = \overline{B}(a, r)$ une boule fermée de E , avec $a \in E$ et $r > 0$. Soit C le complémentaire de A dans E . Soit x un élément quelconque de C . On a donc $d(a, x) > r$. Posons alors $\delta = d(a, x) - r > 0$. Pour tout $y \in B(x, \delta)$, on a : $d(x, y) < \delta$ donc $d(a, x) - d(x, y) > d(a, x) - \delta = r$, de sorte que $d(a, y) \geq |d(a, x) - d(x, y)| > r$, et donc $y \in C$. Ceci prouve que C est ouverte, c'est-à-dire que A est fermée.



- Toute intersection d'une famille quelconque (finie ou infinie) de parties fermées de E est encore une partie fermée de E . Toute réunion d'une famille *finie* de parties fermées de E est encore une partie fermée de E (les preuves se déduisent immédiatement, par passage au complémentaire, des propriétés correspondantes sur les parties ouvertes).

- Il en résulte par exemple que :

- Toute sphère est fermée (car $S(a, r) = (E \setminus B(a, r)) \cap \overline{B}(a, r)$ est l'intersection de deux parties fermées).
- Tout singleton est fermé (car $\{a\} = \bigcap_{r>0} \overline{B}(a, r)$ est une intersection de parties fermées).
- Toute partie finie de E est fermée (car égale à la réunion finie des singletons de ses éléments).

c) REMARQUES. Une partie peut être à la fois ouverte et fermée (par exemple \emptyset ou E); et une partie peut n'être ni ouverte ni fermée.

Exemple : prendre $E = \mathbb{R}$ muni de la norme de la valeur absolue, et considérer $A =]0, 1]$. Ce n'est pas une partie ouverte de E car il n'existe pas de boule ouverte (ici intervalle ouvert) centrée en 1 incluse dans A . Et son complémentaire $C =]-\infty, 0] \cup]1, +\infty[$ n'est pas non plus une partie ouverte car il n'existe pas de boule ouverte (ici intervalle ouvert) centrée en 0 incluse dans C .

Attention, la réunion d'une famille infinie de parties fermées peut ne pas être fermée

Exemple : prendre $E = \mathbb{R}$ muni de la norme de la valeur absolue, et $A_x = \{x\}$ pour tout réel $x \in]0, 1[$; chaque A_x est fermée mais $\bigcup_{x \in]0, 1[} A_x$ est $]0, 1[$, qui n'est pas une partie fermée.

d) DÉFINITION. Soit A une partie d'un evn E . On appelle *adhérence* de A , noté \bar{A} , la plus petite (au sens de l'inclusion) partie fermée de E contenant A , c'est-à-dire encore l'intersection de toute les parties fermées de E contenant A . En particulier :

$$[\text{ on a toujours : } A \subseteq \bar{A}] \quad \text{et} \quad [A = \bar{A} \text{ si et seulement si } A \text{ est fermée}].$$

La caractérisation suivante de l'adhérence est duale de la caractérisation de l'intérieur vue à la fin de 3.1.4.

e) PROPOSITION. Soit A une partie d'un evn E . Pour tout $x \in E$, on a :

$$x \in \bar{A} \text{ si et seulement si, pour tout réel } \delta > 0, \text{ on a } B(x, \delta) \cap A \neq \emptyset.$$

Preuve. Posons $C = E \setminus A$ le complémentaire de A dans E . Il résulte de façon évidente des définitions de l'intérieur et de l'adhérence que $E \setminus \bar{A} = \overset{\circ}{C}$. Dès lors, $x \in \bar{A}$ équivaut à $x \notin \overset{\circ}{C}$.

Or d'après la proposition achevant 3.1.4, $x \in \overset{\circ}{C}$ équivaut à l'existence d'un réel $\delta > 0$ tel que $B(x, \delta) \subseteq C$. Sa négation équivaut donc à : pour réel $\delta > 0$, il existe $y \in B(x, \delta)$ tel que $y \notin C$, ce qui donne l'assertion voulue. \square

3.1.6 Normes équivalentes

La donnée d'une norme sur un espace vectoriel E permet ainsi de définir les notions topologiques de parties ouvertes et de parties fermées, et d'étudier comme on va le voir ci-dessous la convergence des suites d'éléments de E ou la continuité des applications à valeurs dans E . Il importe donc d'identifier ce qui, dans ces notions, dépend ou non du choix de la norme choisie.

a) DÉFINITION. Soit N et N' deux normes sur un \mathbb{K} -espace vectoriel E . On dit que N est *équivalente* à N' , ce que l'on note $N \sim N'$, lorsqu'il existe deux réels $\alpha, \beta > 0$ tels que :

$$\text{pour tout } x \in E, \text{ on a : } \alpha N(x) \leq N'(x) \leq \beta N(x).$$

Il est facile de vérifier (laissé en exercice) que \sim est effectivement une relation d'équivalence (c'est-à-dire qu'elle est réflexive, symétrique et transitive).

b) EXEMPLE. Pour tout entier $n \geq 1$, les trois normes $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ sur l'espace vectoriel \mathbb{K}^n sont équivalentes.

En effet, on obtient facilement pour tout $x = (x_1, \dots, x_n) \in \mathbb{K}^n$ les inégalités :

$$\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty \quad \text{et} \quad \|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty.$$

L'observation ci-dessus n'est qu'une illustration directe du théorème général suivant que, conformément au programme de ce cours, nous ne démontrerons pas ici :

c) THÉORÈME (admis). Sur un \mathbb{K} -espace vectoriel de dimension finie, toutes les normes sont équivalentes.

d) CONTRE-EXEMPLE. Donnons un exemple de deux normes non-équivalentes sur un même espace vectoriel E (bien sûr de dimension infinie). Prenons $E = \mathcal{C}^0([0, 1], \mathbb{R})$ le \mathbb{R} -espace vectoriel des fonctions continues de l'intervalle $[0, 1]$ dans \mathbb{R} . Pour tout entier $n \geq 1$, définissons la fonction affine par intervalles f_n donnée par $f(x) = n(1 - nx)$ pour $x \in [0, \frac{1}{n}]$ et $x = 0$ pour $x \in [\frac{1}{n}, 1]$ (faire un dessin!). On calcule $\|f_n\|_\infty = \sup_{x \in [0, 1]} |f_n(x)| = n$ et $\|f_n\|_1 = \int_0^1 |f_n(t)| dt = \frac{1}{2}$. Si les deux normes étaient équivalentes, il existerait $\alpha, \beta > 0$ tel que $\alpha \leq \frac{\|f_n\|_\infty}{\|f_n\|_1} \leq \beta$ pour tout $n \geq 1$, ce qui n'est pas le cas puisque $\frac{\|f_n\|_\infty}{\|f_n\|_1} = 2n$ tend vers $+\infty$ quand n tend vers l'infini. On conclut que ces deux normes ne sont pas équivalentes.

e) PROPOSITION. Soient N et N' deux normes d'un \mathbb{K} -espace vectoriel E . Les conditions suivantes sont équivalentes :

- (i) les normes N et N' sont équivalentes,
- (ii) les parties ouvertes de E pour la norme N sont exactement les parties ouvertes de E pour la norme N' ,
- (iii) les parties fermées de E pour la norme N sont exactement les parties fermées de E pour la norme N' ,

Preuve. Supposons (i). Il existe $\alpha, \beta > 0$ tel que $\alpha N(x) \leq N'(x) \leq \beta N(x)$. Soit A une partie ouverte de l'evn E muni de la norme N . Quel que soit $a \in A$, il existe $r > 0$ tel que $B_N(a, r) \subseteq A$, avec la notation $B_N(a, r) := \{x \in E; N(x - a) < r\}$. Il est clair que $B_N(a, r)$ contient $B_{N'}(a, \alpha r) := \{x \in E; N'(x - a) < \alpha r\}$. Ainsi $B_{N'}(a, \alpha r) \subseteq A$, ce qui prouve que A est une partie ouverte de l'evn E muni de la norme N' . Par symétrie de l'équivalence des normes, ceci montre que la propriété (ii). L'équivalence de (ii) et (iii) est claire par passage au complémentaire. Le fait que (ii) implique (i), qui est beaucoup moins utile dans la pratique, est laissé en exercice. \square

3.1.7 Exercices

EXERCICE 1. Soit E un \mathbb{K} -espace vectoriel muni de deux normes N_1 et N_2 . On note B_1 et B_2 les boules ouvertes de centre 0_E et de rayon 1 associées. Montre que $B_1 = B_2$ implique $N_1 = N_2$.

EXERCICE 2. Soient a_1, \dots, a_n des réels. Soit $N : \mathbb{K}^n \rightarrow \mathbb{R}$ définie par $N(x_1, \dots, x_n) = a_1|x_1| + \dots + a_n|x_n|$. A quelles conditions sur les paramètres a_1, \dots, a_n l'application N définit-elle une norme sur \mathbb{K}^n ?

EXERCICE 3. Montrer que l'application $N : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $N(x, y) = \sup_{t \in [0, 1]} |x + ty|$ est une norme sur \mathbb{R}^2 . Représenter la boule unité pour N . Comparer N avec $\|\cdot\|_\infty$.

EXERCICE 4. Soient f_1, \dots, f_n des fonctions $[0, 1] \rightarrow \mathbb{R}$ continues sur $[0, 1]$. Soit $N : \mathbb{R}^n \rightarrow \mathbb{R}$ définie par $N(x_1, \dots, x_n) = \|x_1 f_1 + \dots + x_n f_n\|_\infty$. A quelles conditions sur les fonctions f_1, \dots, f_n l'application N est-elle une norme sur \mathbb{R}^n ?

EXERCICE 5. On note $E = \mathcal{M}_n(\mathbb{R})$ l'espace vectoriel des matrices carrées réelles d'ordre n .

- On pose $\|A\|_2 = (\sum_{i,j=1}^n a_{ij}^2)^{\frac{1}{2}}$ pour toute $A = (a_{ij})_{1 \leq i, j \leq n} \in E$.

Montrer que $\|\cdot\|_2$ est une norme sur E , et vérifie de plus $\|AB\|_2 \leq \|A\|_2 \|B\|_2$ pour tous $A, B \in E$.

- Mêmes questions pour $\|A\| = \sup_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$; vérifier de plus que si λ est une valeur propre de A , alors $|\lambda| \leq \|A\|$.

EXERCICE 6. Soit E un \mathbb{K} -espace vectoriel non-nul; en notant $B(a, r)$ la boule ouverte de centre a et de rayon r pour tous $a \in E, r \in \mathbb{R}_+^*$, montrer que, pour tous $a, b \in E$, pour tous $r, s \in \mathbb{R}_+^*$, et pour tout $\lambda \in \mathbb{R}^*$, on a :

- (i) $B(a + b, r + s) = B(a, r) + B(b, s)$, (ii) $\lambda B(a, r) = B(\lambda a, |\lambda|r)$,
- (iii) $B(a, r) \cap B(b, s) \neq \emptyset \Leftrightarrow \|a - b\| < r + s$, (iv) $B(a, r) = B(b, s) \Leftrightarrow (a = b \text{ et } r = s)$.

EXERCICE 7. Montrer que si F un sous-espace vectoriel d'un evn E est ouvert, alors $F = E$.

EXERCICE 8. Soient A, B deux parties d'un evn E telles que $d(A, B) > 0$, avec la notation $d(A, B) = \inf_{x \in A, y \in B} d(x, y)$. Montrer qu'il existe deux parties ouvertes disjointes U, V de E tels que $A \subseteq U$ et $B \subseteq V$.

EXERCICE 9. Montrer que si F un sous-espace vectoriel d'un evn E d'intérieur non-vidé, alors $F = E$.

EXERCICE 10. Soient A, B deux parties d'un evn E .

- On suppose $A \subset B$. Montrer que $A^\circ \subset B^\circ$ et $\overline{A} \subset \overline{B}$.
- Comparer $(A \cap B)^\circ$ avec $A^\circ \cap B^\circ$, puis $(A \cup B)^\circ$ avec $A^\circ \cup B^\circ$.
- Comparer $\overline{A \cup B}$ avec $\overline{A} \cup \overline{B}$, puis $\overline{A \cap B}$ avec $\overline{A} \cap \overline{B}$.

EXERCICE 11. Soit E un \mathbb{R} -espace vectoriel normé. Rappelons qu'une partie A de E est dite convexe lorsque, pour tous $x, y \in A$, le segment $[x, y]$ est inclus dans A , ce qui équivaut à dire que $\lambda x + (1 - \lambda)y \in A$ quel que soit $\lambda \in [0, 1]$. Montrer que si A est une partie convexe de E , alors son adhérence et son intérieur sont aussi convexes. Montrer qu'une partie ouverte U de E est convexe si et seulement si $U + U = 2U$.

EXERCICE 12. Soient E le \mathbb{R} -espace vectoriel des fonctions réelles d'une variable réelle qui sont bornées sur $[0, 1]$ et F le sous-espace vectoriel de celles qui continues sur $[0, 1]$, munis de la norme $\|\cdot\|_\infty$. Soit $f_0 \in E$ définie par $f_0(x) = 1$ si $0 \leq x \leq \frac{1}{2}$, et $f_0(x) = 2$ si $\frac{1}{2} < x \leq 1$. Calculer la distance entre cet élément f_0 et le sous-espace F [ie. $d(f_0, F) = \inf_{g \in F} \|f_0 - g\|_\infty$].

EXERCICE 13. Soit N une norme sur l'espace vectoriel $E = \mathcal{M}_n(\mathbb{R})$ des matrices carrées d'ordre n à coefficients réels. Montrer qu'il existe un réel $c > 0$ tel que $N(AB) \leq cN(A)N(B)$ pour toutes $A, B \in E$.

EXERCICE 14. Soit E le \mathbb{R} -espace vectoriel des fonctions $[0, 1] \rightarrow \mathbb{R}$ de classe C^1 . On pose :

$$\|x\| = \sup_{t \in [0, 1]} (|x(t)| + |x'(t)|) \quad \text{et} \quad N(x) = \sup_{t \in [0, 1]} |x(t)| + \sup_{t \in [0, 1]} |x'(t)|. \quad \text{pour tout } x \in E.$$

Montrer que l'on définit ainsi deux normes sur E , et qu'elles sont équivalentes.

EXERCICE 15. Soit $p \geq 1$ un entier fixé. On note E le \mathbb{C} -espace vectoriel $\mathbb{C}_p[X]$ des polynômes de degré $\leq p$. Pour tout $Z = (z_0, z_1, \dots, z_p) \in \mathbb{C}^{p+1}$, on définit une application $N_Z : E \rightarrow \mathbb{R}_+$ en posant : $N_Z(P) = \sum_{0 \leq i \leq p} |P(z_i)|$ pour tout $P \in E$. Donner une condition nécessaire et suffisante portant sur Z pour que N_Z soit une norme sur E . Deux normes N_Z et $N_{Z'}$ de ce type sont-elles équivalentes ?

EXERCICE 16. Soient E le \mathbb{R} -espace vectoriel des fonctions réelles d'une variable réelle qui sont continues sur $[0, 1]$, et E^+ le sous-ensemble des fonctions de E qui sont positives sur $[0, 1]$ et ne s'annulent qu'en un nombre fini de points de $[0, 1]$. On pose

$$\|f\|_\varphi = \int_0^1 |f(t)|\varphi(t) dt \quad \text{pour toutes } f \in E, \varphi \in E^+.$$

- Montrer que $\|\cdot\|_\varphi$ est une norme sur E pour toute $\varphi \in E^+$.
- Montrer que si φ_1, φ_2 sont deux fonctions dans E^+ supposées de plus strictement positives sur $[0, 1]$, alors les normes $\|\cdot\|_{\varphi_1}$ et $\|\cdot\|_{\varphi_2}$ sont équivalentes.
- les normes $\|\cdot\|_x$ et $\|\cdot\|_{x^2}$ sont elles équivalentes ? [Indication : calculer ces deux normes en la fonction $f_n(x) = (1 - x)^n$ avec $n \in \mathbb{N}$].

3.2 Suites dans un evn

3.2.1 Suites convergentes

a) DÉFINITIONS. Soit E un evn. On appelle *suite* d'éléments de E , ou suite dans E , toute application $u : \mathbb{N} \rightarrow E$. En posant $u_n \in E$ l'image de n pour tout $n \in \mathbb{N}$, il est d'usage de noter :

$$u = (u_n)_{n \in \mathbb{N}}, \quad \text{ou encore } u = (u_n)_{n \geq 0}, \quad \text{ou plus simplement } u = (u_n).$$

L'expression du vecteur u_n en fonction de n est appelé le *terme général* de la suite.

L'ensemble des suites d'éléments de E se note $E^{\mathbb{N}}$. Il est fréquent que l'on ne considère le terme général u_n qu'à partir d'un certain rang n_0 , et que l'on étudie donc en fait la suite $(u_n)_{n \geq n_0}$.

b) DÉFINITION. Soient E un evn, et $\|\cdot\|$ sa norme. Une suite $u = (u_n)_{n \in \mathbb{N}}$ est dite *bornée* lorsque l'ensemble $\{u_n; n \in \mathbb{N}\}$ est une partie bornée de E (au sens de 3.1.3), c'est-à-dire lorsqu'il existe un réel positif r tel que $\|u_n\| \leq r$ pour tout $n \in \mathbb{N}$.

c) DÉFINITION. Soient E un evn, et d la distance associée à sa norme $\|\cdot\|$. Soit (u_n) une suite d'éléments de E . Soit ℓ un élément quelconque de E . On dit que (u_n) *converge vers* ℓ lorsque, pour toute boule ouverte $B := B(\ell, \varepsilon)$ centrée en ℓ , il existe un rang N à partir duquel tous les termes de la suite (u_n) appartiennent à la boule B .

Ce qui peut se traduire par : $\forall \varepsilon \in \mathbb{R}_+^*, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq N \Rightarrow u_n \in B(\ell, \varepsilon)]$,

ou encore : $\forall \varepsilon \in \mathbb{R}_+^*, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq N \Rightarrow d(u_n, \ell) < \varepsilon]$

ou encore : $\forall \varepsilon \in \mathbb{R}_+^*, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq N \Rightarrow \|u_n - \ell\| < \varepsilon]$

d) LEMME (unicité de la limite). *Supposons que, dans un evn E , une suite (u_n) converge vers un élément ℓ_1 et converge vers un élément ℓ_2 , alors $\ell_1 = \ell_2$.*

Preuve. Fixons un réel $\varepsilon > 0$ quelconque. D'une part il existe un entier $N_1 \geq 0$ tel que $u_n \in B(\ell_1, \varepsilon)$ pour tout $n \geq N_1$, d'autre part il existe un entier $N_2 \geq 0$ tel que $u_n \in B(\ell_2, \varepsilon)$ pour tout $n \geq N_2$. Notons N le plus grand des deux entiers N_1 et N_2 . Pour tout $n \geq N$, on a à la fois $d(u_n, \ell_1) < \varepsilon$ et $d(u_n, \ell_2) < \varepsilon$, donc par inégalité triangulaire $d(\ell_1, \ell_2) < 2\varepsilon$. Ainsi le réel $d(\ell_1, \ell_2)$ est strictement inférieur à tout réel strictement positif, ce qui équivaut à dire qu'il est nul, et donc $\ell_1 = \ell_2$. \square

Ce lemme justifie la terminologie suivante.

e) DÉFINITIONS. Soit E un evn. Soit (u_n) une suite d'éléments de E . On dit que (u_n) est *convergente* dans E (ou *converge* dans E) lorsqu'il existe un élément ℓ de E , qui est alors nécessairement unique, tel que (u_n) converge vers ℓ .

On dit alors que ℓ est la *limite* de (u_n) dans E .

On note : $\ell = \lim_{n \rightarrow +\infty} u_n$, ou $\ell = \lim_n u_n$, ou $\ell = \lim u_n$, ou $u_n \longrightarrow \ell$.

Une suite que ne converge pas dans E est dite *divergente* dans E .

Le lemme suivant est élémentaire, mais d'un usage extrêmement fréquent et pratique.

f) LEMME. Soit E un evn. Une suite (u_n) d'éléments de E converge vers une limite $\ell \in E$ si et seulement si la suite $(\|u_n - \ell\|)$ converge dans \mathbb{R} vers 0.

Preuve. Laissée en exercice (voir plus loin exercice 1 de 3.2.5). \square

g) PROPOSITION. Dans un evn, toute suite convergente est bornée.

Preuve. Soit (u_n) une suite dans un evn E convergente vers une limite $\ell \in E$. En appliquant la définition de la limite par exemple pour $\varepsilon = 1$, il existe un entier $N \geq 0$ tel que, pour tout $n \geq N$, on a $\|u_n - \ell\| < 1$ et donc $\|u_n\| \leq \|u_n - \ell\| + \|\ell\| \leq 1 + \|\ell\|$. En notant r le plus grand des réels positifs $\|u_0\|, \|u_1\|, \dots, \|u_N\|, 1 + \|\ell\|$, on a $\|u_n\| \leq r$ pour tout $n \geq 0$. \square

h) THÉORÈME (caractérisation séquentielle de l'adhérence). Soit A une partie d'un evn E .

- (i) Un élément x de E appartient à \overline{A} si et seulement s'il existe une suite d'éléments de A qui converge vers x .
- (ii) En particulier A est une partie fermée si et seulement si toute suite d'éléments de A convergeant dans E a sa limite qui est dans A .

Preuve. Supposons $x \in \overline{A}$. Il résulte de la dernière proposition de 3.1.5 que, pour tout entier $n \geq 1$, on peut choisir un élément u_n dans l'intersection $B(x, \frac{1}{n}) \cap A$. On construit ainsi une suite (u_n) d'éléments de A , et comme $\|u_n - x\| < \frac{1}{n}$ pour tout $n \geq 1$, il est clair que cette suite (u_n) converge vers x . Supposons réciproquement qu'il existe une suite (u_n) d'éléments de A qui converge vers x . Prenons un réel $\delta > 0$ quelconque; il existe un rang $N \geq 0$ tel que, pour tout $n \geq N$, on a $\|u_n - x\| < \delta$, et donc $u_n \in B(x, \delta)$. En particulier $B(x, \delta) \cap A \neq \emptyset$, et ceci pour tout $\delta > 0$, donc $x \in \overline{A}$. Ceci achève la preuve de (i), et (ii) en découle puisque A est fermée si et seulement si $\overline{A} = A$. \square

i) OBSERVATION IMPORTANTE. Les notions de suite convergente et de suite bornée sont invariantes par l'équivalence des normes.

Plus explicitement, si une suite (u_n) d'éléments d'un \mathbb{K} -espace vectoriel E converge au sens d'une norme $\|\cdot\|$ de E vers une limite $\ell \in E$, alors (u_n) converge vers ℓ aussi au sens de toute norme $\|\cdot\|' \sim \|\cdot\|$. De même si (u_n) est bornée.

3.2.2 Opérations sur les suites convergentes

a) DÉFINITIONS. Puisque l'on considère des suites à valeurs dans un evn E , les lois d'espaces vectoriel sur E (addition et produit externe par un scalaire) permettent de définir des lois analogues sur les suites : si $u = (u_n)_{n \geq 0}$ et $v = (v_n)_{n \geq 0}$ sont deux suites d'éléments de E , et si $\lambda \in \mathbb{K}$ est un scalaire quelconque, $u + v$ est la suite dont le terme général est $u_n + v_n$ et λu est la suite dont le terme général est λu_n , ceci donc pour tout $n \geq 0$.

$$(u_n)_{n \geq 0} + (v_n)_{n \geq 0} = (u_n + v_n)_{n \geq 0} \quad \text{et} \quad \lambda(u_n)_{n \geq 0} = (\lambda u_n)_{n \geq 0}.$$

Il est alors facile de vérifier que, du fait que les axiomes de la structure d'espace vectoriel sont vérifiés pour les lois de E , ils le sont aussi pour les lois que l'on en a déduit sur l'ensemble $E^{\mathbb{N}}$ des suites d'éléments de E . On retiendra donc :

l'ensemble $E^{\mathbb{N}}$ des suites d'éléments d'un evn E sur \mathbb{K} est un \mathbb{K} -espace vectoriel.

b) PROPOSITION. Soit E un evn. Soient u et v deux suites dans $E^{\mathbb{N}}$. Soit $\lambda \in \mathbb{K}$ un scalaire.

- (i) Si u converge dans E vers une limite ℓ et v converge dans E vers une limite ℓ' , alors la suite $u + v$ converge dans E vers $\ell + \ell'$.
- (ii) Si u converge dans E vers une limite ℓ , alors la suite λu converge dans E vers $\lambda \ell$.

En particulier, les suites convergentes forment un sous-espace vectoriel de $E^{\mathbb{N}}$.

Preuve. Les résultats voulus se déduisent des propriétés analogues déjà connues pour les suites de réels, via l'utilisation du lemme f) de 3.2.1.

Pour tout $n \in \mathbb{N}$, on a $\|(u_n + v_n) - (\ell + \ell')\| = \|u_n - \ell + v_n - \ell'\| \leq \|u_n - \ell\| + \|v_n - \ell'\|$. Chacune des suites réelles ($\|u_n - \ell\|$) et ($\|v_n - \ell'\|$) converge vers 0 dans \mathbb{R} (d'après le lemme f) de 3.2.1. On sait qu'alors la somme de ces deux suites converge vers $0 + 0 = 0$. Ainsi ($\|(u_n + v_n) - (\ell + \ell')\|$) converge vers 0 dans \mathbb{R} , ce qui prouve le résultat de (i) en réappliquant le lemme f) de 3.2.1.

Pour tout $n \in \mathbb{N}$, on a $\|\lambda u_n - \lambda \ell\| = |\lambda| \|u_n - \ell\|$. Comme la suite (u_n) converge vers ℓ dans E , la suite ($\|u_n - \ell\|$) converge vers 0 dans \mathbb{R} (d'après le lemme f) de 3.2.1. Alors $\lim |\lambda| \|u_n - \ell\| = 0$. Ainsi la suite réelle ($\|\lambda u_n - \lambda \ell\|$) converge vers 0, ce qui prouve le résultat de (ii) en réappliquant le lemme f) de 3.2.1. \square

Le point (ii) ci-dessus peut être étendu sous la forme suivante, concernant le produit d'une suite de $E^{\mathbb{N}}$ par une suite de scalaires (preuves en exercice).

Soient $(u_n)_{n \geq 0}$ une suite d'éléments de E et $(\lambda_n)_{n \geq 0}$ une suite d'éléments de \mathbb{K} .

Si $(\lambda_n)_{n \geq 0}$ converge vers 0 dans \mathbb{K} , et si $(u_n)_{n \geq 0}$ est bornée dans E , alors $(\lambda_n u_n)_{n \geq 0}$ converge vers 0_E dans E .

Si $(\lambda_n)_{n \geq 0}$ est bornée dans \mathbb{K} , et si $(u_n)_{n \geq 0}$ converge vers 0_E dans E , alors $(\lambda_n u_n)_{n \geq 0}$ converge vers 0_E dans E .

Si $(\lambda_n)_{n \geq 0}$ converge dans \mathbb{K} vers une limite λ , et si $(u_n)_{n \geq 0}$ converge dans E vers une limite ℓ , alors $(\lambda_n u_n)_{n \geq 0}$ converge dans E vers $\lambda \ell$.

c) REMARQUE (Produit de deux suites convergentes dans une algèbre normée). Soit E un evn. On suppose de plus que E est muni d'une multiplication interne. Rappelons qu'on dit que E est une algèbre lorsque la multiplication est associative et cohérente avec les lois d'espace vectoriel de E au sens où : $x(y + z) = xy + xz$ et $(y + z)x = yx + zx$ pour tous $x, y, z \in E$ et $(\lambda x)y = x(\lambda y) = \lambda(xy)$ pour tous $x, y \in E$ et $\lambda \in \mathbb{R}$.

On dit que E est une algèbre normée lorsque la norme sur l'evn E vérifie de plus la condition (dite de sous-multiplicativité) :

$$\|xy\| \leq \|x\| \|y\| \quad \text{pour tous } x, y \in E.$$

On a alors la proposition suivante (qui généralise le résultat bien connu pour les suites réelles ou complexes) :

PROPOSITION. Soit E une algèbre normée. Soient $u = (u_n)$ et $v = (v_n)$ deux suites dans $E^{\mathbb{N}}$. Si u converge dans E vers une limite ℓ et v converge dans E vers une limite ℓ' , alors la suite $(u_n v_n)$ converge dans E vers $\ell \ell'$.

Preuve. Pour tout $n \in \mathbb{N}$, on a :

$$\|u_n v_n - \ell \ell'\| = \|u_n v_n - u_n \ell' + u_n \ell' - \ell \ell'\| \leq \|u_n\| \|v_n - \ell'\| + \|u_n - \ell\| \|\ell'\|.$$

D'après le lemme f) de 3.2.1, on a d'une part $\lim \|u_n - \ell\| = 0$ donc $\lim \|u_n - \ell\| \|\ell'\| = 0$, et d'autre part $\lim \|v_n - \ell'\| = 0$ donc $\lim \|u_n\| \|v_n - \ell'\| = 0$ car la suite (u_n) est bornée ; l'inégalité ci-dessus implique alors que $\lim \|u_n v_n - \ell \ell'\| = 0$, ce qui montre le résultat voulu grâce une nouvelle fois au lemme f) de 3.2.1. \square

3.2.3 Suites extraites

a) DÉFINITION. Soit E un evn. Soit (u_n) une suite dans $E^{\mathbb{N}}$; on appelle *suite extraite* de (u_n) toute suite $(u_{\sigma(n)})$ où σ est une application $\mathbb{N} \rightarrow \mathbb{N}$ qui est strictement croissante.

Par exemple (u_{n^2}) est extraite de (u_n) , correspondant à $\sigma(n) = n^2$, de premiers termes $(u_0, u_1, u_4, u_9, \dots)$.

Remarque : on utilisera souvent dans les preuves le fait (évident par récurrence) que, si $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ est strictement croissante, alors $\sigma(m) \geq m$ pour tout $m \in \mathbb{N}$.

b) DÉFINITION. Soit E un evn. Soit (u_n) une suite dans $E^{\mathbb{N}}$; on appelle *valeur d'adhérence* de la suite (u_n) tout élément $a \in E$ tel qu'il existe une suite extraite de (u_n) qui converge vers a .

Par exemple : la suite de réels $(-1)^n$ admet 1 et -1 pour valeurs d'adhérence.

c) LEMME. Soient E un evn et $\|\cdot\|$ sa norme. Soit (u_n) une suite dans $E^{\mathbb{N}}$. Soit a un élément quelconque de E . Les conditions suivantes sont équivalentes.

- (i) a est une valeur d'adhérence de la suite (u_n) ;
- (ii) pour tout réel $\varepsilon > 0$ et tout entier $k \geq 0$, il existe un entier $n \geq k$ tel que $\|u_n - a\| < \varepsilon$;

Preuve. Supposons d'abord qu'il existe une suite extraite $(u_{\sigma(n)})$ telle que $\lim_n u_{\sigma(n)} = a$. Soit $\varepsilon > 0$; il existe $N \in \mathbb{N}$ tel que, pour tout $n \geq N$, $\|u_{\sigma(n)} - a\| < \varepsilon$. Soit $k \in \mathbb{N}$ quelconque. Posons $m = \max\{k, N\}$ et $n = \sigma(m)$. D'après la remarque faite à la fin du a), on a : $n \geq m$ et donc $n \geq k$. Par ailleurs, $m \geq N$ donc $\|u_{\sigma(m)} - a\| < \varepsilon$, c'est-à-dire $\|u_n - a\| < \varepsilon$.

On a ainsi montré que : $\forall \varepsilon > 0, \forall k \in \mathbb{N}, \exists n \geq k, \|u_n - a\| < \varepsilon$.

Réciproquement, supposons que a vérifie la condition (ii). On construit par récurrence une suite extraite $(u_{\sigma(n)})$ de la façon suivante : choisissons arbitrairement $\sigma(0)$; supposons $\sigma(n-1)$ connu pour un $n \geq 1$. Traduisons la condition (ii) avec $\varepsilon = \frac{1}{n}$ et $k = \sigma(n-1) + 1$: on peut trouver un entier, notons-le $\sigma(n)$, tel que $\sigma(n) \geq k > \sigma(n-1)$ et $\|u_{\sigma(n)} - a\| < \frac{1}{n}$. On en déduit que la suite de réels positifs $(\|u_{\sigma(n)} - a\|)$ converge vers 0. Cela prouve (voir exercices 3.2.5) que $(u_{\sigma(n)})$ converge dans E vers a . C'est une suite extraite de (u_n) puisque σ est strictement croissante par construction.

d) PROPOSITION. Si une suite (u_n) converge dans E vers une limite ℓ , alors toute suite extraite de (u_n) converge aussi vers ℓ . En d'autres termes, ℓ est l'unique valeur d'adhérence de (u_n) .

Preuve. On suppose que la suite (u_n) converge vers ℓ . Fixons $\varepsilon > 0$ quelconque; il existe $N \in \mathbb{N}$ tel que, pour tout $n \in \mathbb{N}$, on a $[n \geq N \implies \|u_n - \ell\| < \varepsilon]$.

Soit $(u_{\sigma(n)})$ une suite extraite de (u_n) . Tout $n \geq N$ vérifie $\sigma(n) \geq \sigma(N)$ puisque σ est croissante. Mais $\sigma(N) \geq N$ d'après la remarque de la fin du a) ci-dessus. Finalement, tout $n \geq N$ vérifie $\sigma(n) \geq N$, et donc $\|u_{\sigma(n)} - \ell\| < \varepsilon$. On conclut que $(u_{\sigma(n)})$ converge vers ℓ . \square

e) REMARQUE. La propriété ci-dessus permet de montrer que certaines suites divergent; par exemple la suite de réels $((-1)^n)$ diverge puisqu'elle admet deux suites extraites $((-1)^{2n})$ et $((-1)^{2n+1})$ qui convergent vers des limites différentes.

Réciproquement, une suite peut admettre une unique valeur d'adhérence sans converger. Par exemple la suite de réels $(\frac{1}{2}, 2, \frac{1}{3}, 3, \dots, \frac{1}{n}, n, \dots)$ admet 0 pour unique valeur d'adhérence, mais ne converge pas dans \mathbb{R} (en revanche on peut montrer qu'une suite bornée qui admet une unique valeur d'adhérence converge vers cette valeur).

f) REMARQUE (lien entre valeur d'adhérence d'une suite et adhérence d'une partie). Soit $(u_n)_{n \in \mathbb{N}}$ une suite d'un evn E . Notons $A = \{u_p; p \in \mathbb{N}\}$ le sous-ensemble de E formé des valeurs prises par la suite. Il est faux de penser que les valeurs d'adhérence de la suite (u_n) sont les éléments de l'adhérence \overline{A} de A [prendre par exemple dans $E = \mathbb{R}$ la suite $(\frac{1}{n+1})$]. En revanche, on peut montrer (voir exercice 4 en 3.2.5 ci-dessous) que, si l'on note $\text{VA}(u_n)$ l'ensemble des valeurs d'adhérence de la suite (u_n) , alors on a :

$$\text{VA}(u_n) = \bigcap_{n \in \mathbb{N}} \overline{\{u_p; p \geq n\}}.$$

Il en résulte que $\text{VA}(u_n) \subset \overline{A}$, mais l'inclusion peut être stricte.

g) COMPLÉMENT. Parmi les résultats classiques sur les valeurs d'adhérence et les suites extraites, citons le résultat suivant qui donne, en dimension finie, une condition suffisante pour l'existence de valeurs d'adhérence :

THÉORÈME. Si E est un evn de dimension finie, toute suite bornée d'éléments de E admet au moins une valeur d'adhérence.

Ce résultat, dit théorème de Bolzano-Weierstrass, n'est pas explicitement au programme de ce cours. Mais il a été montré en première année dans le cas où $E = \mathbb{R}$, et on verra à l'exercice 8 en 3.2.5 comment on peut en déduire le résultat général.

3.2.4 Suites de Cauchy

a) DÉFINITION. Soient E un evn, et $\|\cdot\|$ sa norme. Une suite $u = (u_n)_{n \in \mathbb{N}}$ est appelée une *suite de Cauchy* lorsqu'elle vérifie la condition :

$$\forall \varepsilon \in \mathbb{R}_+^*, \exists N \in \mathbb{N}, \forall p, q \in \mathbb{N}, (p \geq N \text{ et } q \geq N) \implies \|u_p - u_q\| < \varepsilon.$$

b) PROPOSITION. Soient E un evn.

- (i) Toute suite convergente dans E est de Cauchy.
- (ii) Toute suite de Cauchy dans E est bornée.

Preuve. Soit (u_n) une suite d'éléments de E . Supposons que (u_n) converge vers une limite $\ell \in E$. Soit $\varepsilon > 0$. Il existe $N \in \mathbb{N}$ tel que : $|u_n - \ell| < \frac{\varepsilon}{2}$ pour tout $n \geq N$. Si $p, q \in \mathbb{N}$ tels que $p \geq N$ et $q \geq N$, alors :

$$\|u_p - u_q\| = \|(u_p - \ell) + (\ell - u_q)\| \leq \|u_p - \ell\| + \|u_q - \ell\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

ce qui prouve (i). Pour (ii), supposons que (u_n) est de Cauchy dans E . En choisissant $\varepsilon = 1$, il existe $N \in \mathbb{N}$ tel que : $\|u_n - u_N\| < 1$ pour tout $n \geq N$. Pour un tel $n \geq N$, on a :

$$\|u_n\| = \|u_n - u_N + u_N\| \leq \|u_n - u_N\| + \|u_N\| < 1 + \|u_N\|.$$

Posons $M = \max\{\|u_0\|, \dots, \|u_{N-1}\|, 1 + \|u_N\|\} \in \mathbb{R}_+$; on a $\|u_n\| \leq M$ pour tout $n \in \mathbb{N}$. \square

c) REMARQUE. Par contraposition du point (i) ci-dessus, on peut montrer qu'une suite est divergente en vérifiant qu'elle n'est pas de Cauchy.

Exemple : dans $E = \mathbb{R}$, considérons la suite $(u_n)_{n \geq 1}$ définie par $u_n = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n}$.

Pour tout $n \in \mathbb{N}$, on a $|u_{2n} - u_n| = u_{2n} - u_n = \frac{1}{n+1} + \frac{1}{n+2} + \dots + \frac{1}{2n}$.

Chaque terme est $\geq \frac{1}{2n}$, donc $|u_{2n} - u_n| \geq \frac{n}{2n} = \frac{1}{2}$.

Ainsi il existe $\varepsilon = \frac{1}{2}$ tel que, pour tout $N \in \mathbb{N}$, les entiers $p = N$ et $q = 2N$ vérifient $p \geq N, q \geq N$ et $|u_p - u_q| = |u_N - u_{2N}| \geq \varepsilon$.

On conclut que (u_n) n'est pas de Cauchy dans \mathbb{R} , donc ne converge pas dans \mathbb{R} .

d) THÉORÈME. Soient E un evn de dimension finie. Une suite d'éléments de E est convergente dans E si et seulement si elle est de Cauchy dans E .

Preuve. On a vu à la proposition b) que toute suite convergente est de Cauchy. C'est la réciproque qu'il faut maintenant établir dans le cas particulier où E est de dimension finie.

- On le montre d'abord pour $E = \mathbb{R}$, muni de la valeur absolue. Soit $(u_n) \in \mathbb{R}^{\mathbb{N}}$ une suite de Cauchy. D'après la proposition b), elle est bornée. Donc d'après le théorème de Bolzano-Weierstrass dans \mathbb{R} vu en première année [voir aussi théorème g) du 3.2.3 et exercice 8 de 3.2.5 pour une généralisation], elle admet une valeur d'adhérence $a \in \mathbb{R}$. Montrons qu'elle converge vers a .

Pour cela, fixons $\varepsilon > 0$. Parce que (u_n) est de Cauchy, il existe $N_0 \in \mathbb{N}$ tel que :

$$\forall p, q \in \mathbb{N}, (p \geq N_0 \text{ et } q \geq N_0) \implies |u_p - u_q| < \frac{\varepsilon}{2}. \quad (1)$$

Parce que a est valeur d'adhérence de (u_n) , il résulte du lemme c) de 3.2.3 que :

$$\forall k \in \mathbb{N}, \exists m \geq k, |u_m - a| < \frac{\varepsilon}{2}. \quad (2)$$

Ecrivons (2) pour $k = N_0$: on déduit qu'il existe $m_0 \geq N_0$ tel que : $|u_{m_0} - a| < \frac{\varepsilon}{2}$.

En appliquant (1) à $p \geq N_0$ quelconque, et à $q = m_0$, on obtient $|u_p - u_{m_0}| < \frac{\varepsilon}{2}$.

Donc : $|u_p - a| = |(u_p - u_{m_0}) + (u_{m_0} - a)| \leq |u_p - u_{m_0}| + |u_{m_0} - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$.

Bilan : on a montré qu'il existe $N_0 \in \mathbb{N}$ tel que, pour tout $p \geq N_0$, on a $|u_p - a| < \varepsilon$.

Ceci étant vrai quel que soit $\varepsilon > 0$, on conclut que (u_n) converge a .

- On en déduit le résultat pour $E = \mathbb{C}$, muni du module. Soit (z_n) une suite de nombre complexe. Pour tout $n \geq 0$, notons x_n la partie réelle de z_n et y_n la partie imaginaire de z_n . On définit ainsi deux suites (x_n) et (y_n) de réels vérifiant $z_n = x_n + iy_n$ pour tout $n \geq 0$. Supposons que (z_n) est de Cauchy dans \mathbb{C} . On sait qu'alors (voir exercice 6 de 3.2.5, première question) les suites (x_n) et (y_n) sont de Cauchy dans \mathbb{R} . D'après ce que l'on vient de démontrer dans \mathbb{R} , on en déduit que (x_n) converge vers un réel a et que (y_n) converge vers un réel b . On conclut avec la seconde question du même exercice que (z_n) converge dans \mathbb{C} vers le complexe $a + ib$.

- On le montre ensuite pour $E = \mathbb{K}^p$, quel que soit un entier $p \geq 1$. C'est vrai pour $p = 1$ d'après ce qu'on vient de voir. Pour passer à un p quelconque, on raisonne exactement comme ci-dessus en utilisant l'exercice 7 de 3.2.5 au lieu de l'exercice 6.

- On conclut en utilisant le fait que tout \mathbb{K} -espace vectoriel de dimension finie p est isomorphe à \mathbb{K}^p (résultat classique et simple d'algèbre linéaire), et que toutes les normes sur un tel espace sont équivalentes (ce que l'on a admis au c) de 3.1.6. \square

Un espace E qui a la propriété que toute suite de Cauchy dans E converge dans E (et où donc les notions de suites convergentes et de suites de Cauchy sont équivalentes) est dit *complet*.

D'après ce qui précède : \mathbb{R} est complet, \mathbb{C} est complet, \mathbb{K}^n est complet (pour tout $n \geq 1$), tout evn de dimension finie est complet...

Dans un tel evn complet, le *critère de Cauchy* (c'est-à-dire le fait qu'être de Cauchy équivaut à être convergente) permet de montrer qu'une suite converge sans connaître sa limite.

3.2.5 Exercices

EXERCICE 1. Soit (u_n) une suite d'éléments d'un evn E . Montrer que (u_n) converge vers ℓ dans E si et seulement si la suite $(\|u_n - \ell\|)$ converge vers 0 dans \mathbb{R} . Montrer que (u_n) converge

vers 0_E dans E si et seulement la suite $(\|u_n\|)$ converge vers 0 dans \mathbb{R} . Montrer que, si (u_n) converge vers ℓ dans E , alors la suite $(\|u_n\|)$ converge vers $\|\ell\|$ dans \mathbb{R} .

EXERCICE 2. Démontrer les propriétés énoncées à la fin du b) de 3.2.2

EXERCICE 3. Soient (u_n) une suite dans $\mathbb{K}^{\mathbb{N}}$ et $\ell \in \mathbb{K}$. Montrer que (u_n) converge vers ℓ si et seulement si les suites extraites $(u_{2p})_{p \geq 0}$ et $(u_{2p+1})_{p \geq 0}$ convergent toutes les deux vers ℓ .

EXERCICE 4. Démontrer les propriétés de la remarques f) de 3.2.3

EXERCICE 5. Soient $(z_n) \in \mathbb{C}^{\mathbb{N}}$, $(x_n) \in \mathbb{R}^{\mathbb{N}}$, $(y_n) \in \mathbb{R}^{\mathbb{N}}$ telles que $z_n = x_n + iy_n$ pour tout $n \in \mathbb{N}$. Montrer que la suite complexe (z_n) est de Cauchy si et seulement si les suites réelles (x_n) et (y_n) sont de Cauchy. Montrer que la suite (z_n) converge dans \mathbb{C} si et seulement si les suites (x_n) et (y_n) convergent dans \mathbb{R} , et que l'on a alors $\lim z_n = \lim x_n + i \lim y_n$.

EXERCICE 6. Soient $(z_n) \in \mathbb{C}^{\mathbb{N}}$, $(x_n) \in \mathbb{R}^{\mathbb{N}}$, $(y_n) \in \mathbb{R}^{\mathbb{N}}$ telles que $z_n = x_n + iy_n$ pour tout $n \in \mathbb{N}$. Montrer que si (z_n) admet dans \mathbb{C} une valeur d'adhérence $z = a + ib$ avec $a, b \in \mathbb{R}$, alors a est valeur d'adhérence de (x_n) et b valeur d'adhérence de (y_n) .

Montrer que la réciproque est fausse. [Prendre $z_n = i^n$; vérifier que 1 est valeur d'adhérence des suites réelles (x_n) et (y_n) , mais $1 + i$ n'est pas valeur d'adhérence dans \mathbb{C} de (z_n)].

EXERCICE 7. Considérons p suites $U_1 = (u_{1,n})_{n \geq 0}$, $U_2 = (u_{2,n})_{n \geq 0}$, \dots , $U_p = (u_{p,n})_{n \geq 0}$ dans $\mathbb{K}^{\mathbb{N}}$, où p est un entier ≥ 1 . On définit une suite $(W_n)_{n \geq 0}$ d'éléments de \mathbb{K}^p en posant $W_n = (u_{1,n}, u_{2,n}, \dots, u_{p,n})$ pour tout $n \geq 0$.

Montrer que W est de Cauchy dans \mathbb{K}^p si et seulement si chaque U_i (pour $1 \leq i \leq p$) est de Cauchy dans E .

Montrer que W converge dans \mathbb{K}^p si et seulement si chaque U_i (pour $1 \leq i \leq p$) est convergente dans E ; vérifier qu'alors $\lim W = (\lim U_1, \lim U_2, \dots, \lim U_p)$.

EXERCICE 8 (*Autour du théorème de Bolzano Weirstrass*). Réviser une preuve du théorème vu en première année : de toute suite bornée d'éléments de \mathbb{R} , on peut extraire une suite convergente. En déduire, en utilisant l'exercice 6, que : de toute suite bornée d'éléments de \mathbb{C} , on peut extraire une suite convergente. Montrer qu'une preuve identique, utilisant l'exercice 7, permet de montrer que : de toute suite bornée d'éléments de \mathbb{K}^p , on peut extraire une suite convergente. Conclure.

EXERCICE 9 (*Algèbre normée des matrices carrées*). On se place dans $E = \mathcal{M}_p(\mathbb{R})$. Montrer que, pour toute norme N sur E , il existe un réel $c > 0$ tel que E soit une algèbre normée pour la norme cN (voir exercice 13 de 3.1.7).

- Soient A et B deux éléments de E . Montrer que, si la suite $((AB)^n)$ converge dans E vers la matrice nulle O_p , alors il en est de même de la suite $((BA)^n)$.

- Soient A et B deux éléments de E telles que $AB = BA$. Montrer que, si la suite (A^n) converge dans E vers une matrice P , et si la suite (B^n) converge dans E vers une matrice Q , alors $PQ = QP$.

- Soient (A_n) une suite de matrices inversibles dans E . Montrer que, si les suites (A_n) et (A_n^{-1}) convergent dans E vers des matrices respectivement notées de A et B , alors A est inversible et $B = A^{-1}$.

- Soit A une matrice de E . Montrer qu'il existe une matrice $M \in E$ telle que la suite (M^n) converge vers A si et seulement si $A^2 = A$.

- Soit (A_n) une suite de matrices de E convergeant dans E vers une matrice A . Montrer qu'il existe un entier k tel que $\text{rg } A_n \geq \text{rg } A$ pour tout $n \geq k$. Montrer que si toutes les matrices A_n sont de même rang p , alors $\text{rg } A \leq p$.

EXERCICE 10. Montrer que toute partie fermée d'un evn E peut s'écrire comme intersection d'une suite décroissante de parties ouvertes de E .

EXERCICE 11. On se place dans l'evn $E = \mathbb{R}^2$. On désigne par p_1 et p_2 les applications coordonnées $E \rightarrow \mathbb{R}$ définie $p_1(x, y) = x$ et $p_2(x, y) = y$.

- Montrer que si U est une partie ouverte de E , alors $p_1(U)$ et $p_2(U)$ sont des parties ouvertes de \mathbb{R} .

- Montrer que la partie $H = \{(x, y) \in E; xy = 1\}$ est fermée dans E , mais que $p_1(H)$ et $p_2(H)$ ne sont pas fermées dans \mathbb{R} .

- Montrer que, si F est une partie fermée de E et si $p_2(F)$ est bornée dans \mathbb{R} , alors $p_1(F)$ est fermée dans \mathbb{R} .

EXERCICE 12. On se place dans l'evn E des suites de nombres réels bornées muni de la norme $\|\cdot\|_\infty$. Soit F l'ensemble des suites de nombres réels qui sont identiquement nulles à partir d'un certain rang. Montrer que F est un sous-espace vectoriel de E . Est-ce une partie ouverte de E ? Est-ce une partie fermée de E ?

EXERCICE 13. Soit A une partie non-vide de \mathbb{R} telle que, pour tout $x \in \mathbb{R}$, il existe un unique réel $y \in A$ tel que $\|x - y\| = d(x, A)$. Montrer que A est un intervalle fermé.

EXERCICE 14. Montrer que si F est un sous-espace vectoriel d'un evn E , alors son adhérence \overline{F} est aussi un sous-espace vectoriel de E .

EXERCICE 15. (*Moyenne des itérés d'une isométrie vectorielle*) Soient E un espace euclidien et $u \in O(E)$. On pose $v = \text{id}_E - u$. Montrer que $\text{Ker } v = (\text{Im } v)^\perp$.

Soit $x \in E$; on pose $p_n(x) = \frac{1}{n} \sum_{i=0}^{n-1} u^i(x)$ (pour $n \in \mathbb{N}^*$). Montrer que la suite $(p_n(x))_{n \geq 1}$ converge dans E vers le projeté orthogonal y de x sur $\text{Ker } v$, c'est-à-dire que :

$$\lim_{n \rightarrow +\infty} \|p_n(x) - y\| = 0.$$

3.3 Fonctions d'un evn dans un evn

Dans cette partie, on s'intéresse aux applications $f : E \rightarrow F$ où E et F sont deux evn sur \mathbb{K} . On note $\|\cdot\|_E$ la norme de E et $\|\cdot\|_F$ la norme de F . Ce que l'on va faire ici généralise donc les notions de limites et de continuité déjà bien connues lorsque f est une fonction d'une variable réelle (ie. $E = \mathbb{R}$) à valeurs réelles ou complexes (ie. $F = \mathbb{R}$ ou \mathbb{C}).

3.3.1 Limites dans les evn

a) DÉFINITION. Soient E et F deux evn. Soit X une partie de E . Soit $f : X \rightarrow F$ une application. Soit $a \in \overline{X}$ un élément de E qui est adhérent à X . Soit ℓ un élément de F . On dit que f admet ℓ pour *limite* en a lorsque :

pour tout $\varepsilon \in \mathbb{R}_+^*$, il existe $\delta \in \mathbb{R}_+^*$ tel que, pour tout $x \in X$ vérifiant $\|x - a\|_E < \delta$, on a $\|f(x) - \ell\|_F < \varepsilon$.

ce que l'on écrit encore :

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in X, \|x - a\|_E < \delta \Rightarrow \|f(x) - \ell\|_F < \varepsilon.$$

Il est facile de vérifier (adapter la preuve faite en 3.2.1.d pour les limites de suites) que la limite, si elle existe, est unique, ce qui justifie la notation :

$$\lim_{x \rightarrow a} f(x) = \ell, \quad \text{ou encore} \quad \lim_a f = \ell.$$

Remarquons qu'il est crucial ici que $a \in \overline{X}$ pour qu'il existe effectivement des $x \in X$ qui vérifient $\|x - a\|_E < \delta$ (puisque $B(a, \delta) \cap X \neq \emptyset$ par définition de l'adhérence de X).

Il est clair enfin que l'existence d'une limite et sa valeur ne changent pas si l'on remplace $\|\cdot\|_E$ par une norme sur E équivalente à $\|\cdot\|_E$, et $\|\cdot\|_F$ par une norme sur F équivalente à $\|\cdot\|_F$. En particulier, si E et F sont de dimensions finies, l'existence et la valeur de la limite sont indépendantes des normes choisies.

b) LEMME (limites et fonctions coordonnées). Soit $f : X \rightarrow F$ où X est une partie d'un evn E , et F est un evn de dimension finie n . Soit \mathcal{B} une base de F et $f_i : X \rightarrow \mathbb{K}$ les fonctions coordonnées de f relativement à la base \mathcal{B} . Soit $\ell \in F$; on note (ℓ_1, \dots, ℓ_n) les composantes de ℓ dans la base \mathcal{B} . Soit $a \in \overline{X}$. Alors :

$$\left[\lim_{x \rightarrow a} f(x) = \ell \text{ dans } F \right] \Leftrightarrow \left[\lim_{x \rightarrow a} f_i(x) = \ell_i \text{ dans } \mathbb{K} \text{ pour tout } 1 \leq i \leq n \right].$$

Preuve. Laissée à titre d'exercice; s'inspirer des exercices 6 et 7 de 3.2.5. □

En particulier, lorsque f est à valeurs dans \mathbb{C} , on a :

$$\left[\lim_{x \rightarrow a} f(x) = \ell \in \mathbb{C} \right] \Leftrightarrow \left[\lim_{x \rightarrow a} (\operatorname{Re} f)(x) = \operatorname{Re}(\ell) \in \mathbb{R} \text{ et } \lim_{x \rightarrow a} (\operatorname{Im} f)(x) = \operatorname{Im}(\ell) \in \mathbb{R} \right].$$

c) PROPOSITION (limites et opérations). Soient E et F deux espaces vectoriels normés sur \mathbb{K} . Soit f et g deux applications $X \rightarrow F$ définies sur une partie X de E . Soit $a \in \overline{X}$ un point adhérent à X . Soient $\ell, \ell' \in F$. On suppose que $\lim_{x \rightarrow a} f(x) = \ell$ et $\lim_{x \rightarrow a} g(x) = \ell'$. Alors :

(i) $\lim_{x \rightarrow a} (f + g)(x) = \ell + \ell'$;

(ii) $\lim_{x \rightarrow a} \lambda f(x) = \lambda \ell$ pour tout $\lambda \in \mathbb{K}$;

(iii) plus généralement, $\lim_{x \rightarrow a} \lambda(x) f(x) = \lambda \ell$ pour toute application $\lambda : X \rightarrow \mathbb{K}$ admettant une limite $\lambda \in \mathbb{K}$ en a ;

(iv) si on suppose de plus que F est une algèbre normée, $\lim_{x \rightarrow a} (fg)(x) = \ell \ell'$.

Preuve. Laissée à titre d'exercice; s'inspirer des preuves analogues pour les suites dans un evn (voir 3.2.2) ou pour les limites des fonctions dans la cas des fonctions d'une variable réelle à valeurs dans \mathbb{K} (voir cours de première année). □

d) THÉORÈME (caractérisation séquentielle de la limite). Soient E et F deux evn. Soit X une partie de E . Soit $f : X \rightarrow F$ une application. Soit $a \in \overline{X}$ un élément de E qui est adhérent à X . Soit ℓ un élément de F . Alors f admet ℓ pour limite en a si et seulement si, pour toute suite (u_n) d'éléments de X convergeant vers a dans E , la suite $f(u_n)$ converge vers ℓ dans F .

Preuve. Supposons d'abord que $\lim_{x \rightarrow a} f(x) = \ell$. Considérons une suite (u_n) d'éléments de X convergeant vers a (ce qui, d'après le théorème h de 3.2.1, est cohérent avec le fait que $a \in \overline{X}$). Soit $\varepsilon > 0$ fixé. Il existe $\delta > 0$ tel que, pour tout $x \in X$ vérifiant $\|x - a\|_E < \delta$, on a $\|f(x) - \ell\|_F < \varepsilon$. Mais par ailleurs, pour ce $\delta > 0$, il existe un rang $N \in \mathbb{N}$ tel que, pour tout $n \geq N$, on a $\|u_n - a\|_E < \delta$, et donc d'après ce qui précède, $\|f(u_n) - \ell\|_F < \varepsilon$. On a ainsi prouvé que $\lim_{n \rightarrow +\infty} f(u_n) = \ell$ dans F .

Pour la réciproque, raisonnons par contraposée en supposant que f n'admette pas ℓ pour limite en a . Il existe donc un réel $\varepsilon > 0$ tel que :

$$\forall \delta > 0, \exists x \in X, \|x - a\|_E < \delta \text{ et } \|f(x) - \ell\|_F \geq \varepsilon.$$

Prenons $n \in \mathbb{N}$ quelconque et appliquons ce qui précède à $\delta = \frac{1}{n+1}$. Il existe donc $x_n \in X$ tel que $\|x_n - a\|_E < \frac{1}{n+1}$ et $\|f(x_n) - \ell\|_F \geq \varepsilon$. On définit ainsi une suite (x_n) d'éléments de X , qui converge vers a dans E (puisque $\|x_n - a\|_E < \frac{1}{n+1}$ pour tout $n \in \mathbb{N}$ avec $\lim_n \frac{1}{n+1} = 0$) mais telle que $f(x_n)$ ne converge pas vers ℓ dans F (puisque $\|f(x_n) - \ell\|_F \geq \varepsilon$ pour tout $n \in \mathbb{N}$ avec $\varepsilon > 0$). Ce qui achève la preuve. \square

e) PROPOSITION (critère de Cauchy pour l'existence d'une limite). Les hypothèses et données sont celles du théorème précédent. On suppose de plus que F est de dimension finie. Alors f admet une limite en a si et seulement si on a :

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x, y \in X, [\|x - a\|_E < \delta \text{ et } \|y - a\|_E < \delta] \Rightarrow \|f(x) - f(y)\|_F < \varepsilon.$$

Preuve. Laissée à titre d'exercice : faire une preuve directe dans le sens où la propriété est évidente (voir aussi proposition b.(i) de 3.2.4) ; puis utiliser le théorème d) ci-dessus et le théorème d de 3.2.4 pour le sens réciproque. \square

3.3.2 Continuité dans les evn

a) DÉFINITIONS. Soient E et F deux evn. Soit X une partie de E . Soit $f : X \rightarrow F$ une application. Soit $a \in X$ un élément de E (dans X et pas seulement dans \overline{X}). On dit que f est *continue en a* lorsque f admet $f(a)$ pour limite en a , ce qui équivaut à :

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in X, \|x - a\|_E < \delta \Rightarrow \|f(x) - f(a)\|_F < \varepsilon.$$

On dit que f est *continue sur X* lorsque f est continue en a pour tout $a \in X$. On note $\mathcal{C}^0(X, F)$ ou parfois simplement $\mathcal{C}(X, F)$ l'ensemble des applications continues de X dans F . Il est clair par exemple que toute application constante sur E est continue sur E . L'application id_E est continue de E dans E .

Ces notions restent inchangées si on remplace une norme par une norme équivalente, ce qui est toujours vérifié si E et F sont de dimension finie.

b) PROPOSITION (caractérisation séquentielle de la continuité). Soient E et F deux evn. Soit X une partie de E . Soit $f : X \rightarrow F$ une application. Soit $a \in X$. Alors f est continue en a si et seulement si, pour toute suite (u_n) d'éléments de X convergeant vers a dans E , la suite $f(u_n)$ converge vers $f(a)$ dans F .

Preuve. Résulte immédiatement de la définition a) ci-dessus et du théorème d) de 3.3.2. \square

c) PROPOSITION (opérations sur les fonctions continues). Soient E et F deux espaces vectoriels normés sur \mathbb{K} . Soit f et g deux applications $X \rightarrow F$ définies sur une partie X de E . Soit $a \in X$ un élément de X . On suppose que f et g sont continues en a . Alors :

- (i) $f + g$ est continue en a ;
- (ii) λf est continue en a pour tout $\lambda \in \mathbb{K}$;
- (iii) plus généralement, λf est continue en a pour toute application $\lambda : X \rightarrow \mathbb{K}$ continue en a ;
- (iv) si on suppose de plus que F est une algèbre normée, fg est continue en a .

Preuve. Laissée à titre d'exercice ; résulte immédiatement de la définition a) ci-dessus et de la proposition c) de 3.3.2. □

Il est clair que les assertions ci-dessus restent vraies si l'on remplace "continue en a " par "continue sur X ". En particulier, on déduit de (i) et (ii), puis de (iv) le corollaire suivant :

d) COROLLAIRE. Soient E et F deux espaces vectoriels normés sur \mathbb{K} . Soit X une partie E . Alors :

- (i) $\mathcal{C}(X, F)$ est un espace vectoriel sur \mathbb{K} .
- (ii) Si de plus F est une algèbre normée, $\mathcal{C}(X, F)$ est une algèbre.

On laisse au lecteur le soin d'écrire le détail des preuves (toutes évidentes), et de rédiger à titre d'exercice une preuve du résultat complémentaire suivant :

e) PROPOSITION (composée d'applications continues). Soient E, F, G trois sous-espaces vectoriels normés sur \mathbb{K} . Soient X une partie de E et X' une partie de F . Soit $f : X \rightarrow F$ telle que $f(X) \subseteq X'$ et $g : X' \rightarrow G$. On peut donc considérer $g \circ f : X \rightarrow G$.

- (i) Si f est continue en un point a de X et si g est continue en $b := f(a)$, alors $g \circ f$ est continue en a .
- (ii) Si f est continue sur X et si g est continue sur X' , alors $g \circ f$ est continue sur X .

3.3.3 Applications lipschitziennes, applications linéaires continus

a) DÉFINITION Soient E et F deux evn. Soit X une partie de E . Soit $f : X \rightarrow F$ une application. Cette application est dite *lipschitzienne* lorsqu'il existe un réel $\lambda > 0$ tel que :

$$\forall x, y \in X, \|f(x) - f(y)\|_F \leq \lambda \|x - y\|_E.$$

On dit alors que f est lipschitzienne de *rapport* λ . Toute isométrie est lipschitzienne avec $\lambda = 1$. Une application λ -lipschitzienne avec $\lambda < 1$ est dite contractante. L'un des intérêts de cette notion est qu'elle fournit une classe facile d'applications continues, en vertu de la proposition suivante :

b) PROPOSITION. Soient E et F deux evn. Soit X une partie de E . Toute application $f : X \rightarrow F$ qui est lipschitzienne est continue sur X .

Preuve. Soit $a \in X$ quelconque. Pour tout $x \in X$, on a $0 \leq \|f(x) - f(a)\|_F \leq k \|x - a\|_E$. Comme $\lim_{x \rightarrow a} \|x - a\|_E = 0$ dans \mathbb{R} , on déduit de cette inégalité que $\lim_{x \rightarrow a} \|f(x) - f(a)\|_F = 0$, donc $\lim_{x \rightarrow a} f(x) = f(a)$, c'est-à-dire que f est continue en a . Ceci étant vrai pour tout $a \in X$, le résultat voulu est démontré. □

c) **EXEMPLES ET REMARQUES.** Il est clair que, si une application est lipschitzienne de rapport $\lambda > 0$, alors elle est lipschitzienne de rapport λ' pour tout $\lambda' > \lambda$.

Un exemple canonique important est celui de l'application norme de E dans \mathbb{R} . Plus précisément, si on note $f : E \rightarrow \mathbb{R}$ l'application qui, à tout vecteur $x \in E$ associe $f(x) := \|x\|_E$, alors f est lipschitzienne donc continue.

En effet : cela résulte de l'inégalité $|\|x\|_E - \|y\|_E| \leq \|x - y\|_E$ pour tous $x, y \in E$.

Un autre exemple général utile d'application lipschitzienne donc continue est celui des fonctions coordonnées lorsque E est de dimension finie.

Plus précisément : si E est de dimension finie n et si \mathcal{B} est une base de E , on peut considérer les applications $p_k : E \rightarrow \mathbb{K}$ pour tout $1 \leq k \leq n$ qui, à un tout vecteur x de composantes (x_1, \dots, x_n) dans la base \mathcal{B} , associent la coordonnée x_k . Si l'on considère dans E la norme $\|\cdot\|_\infty$ définie par $\|x\|_\infty = \max_{1 \leq k \leq n} |x_k|$, il est clair que $|p_k(x) - p_k(y)| = |x_k - y_k| \leq \|x - y\|_\infty$, ce qui prouve le résultat voulu.

d) **THÉORÈME (cas des applications linéaires).** Soient E et F deux evn.

(i) Une application linéaire $f : E \rightarrow F$ est continue (sur E) si et seulement si elle vérifie la condition :

il existe $k \in \mathbb{R}_+$ tel que, pour tout $x \in E$, on a $\|f(x)\|_F \leq k\|x\|_E$.

(ii) Si E est de dimension finie, toute application linéaire $f : E \rightarrow F$ est continue sur E .

Preuve. Pour montrer (i), supposons que f est continue. En particulier f est continue en 0_E . Ecrivons la définition avec $\varepsilon = 1$: il existe $\delta > 0$ tel que, pour tout $y \in E$ vérifiant $\|y\|_E < \delta$, on a $\|f(y)\|_F < 1$. Prenons alors $x \in E$, $x \neq 0_E$, quelconque. Posons $y = \frac{\delta}{2\|x\|_E}x$, de sorte $\|y\|_E = \frac{\delta}{2} < \delta$, donc $\|f(y)\|_F < 1$. Or $f(y) = \frac{\delta}{2\|x\|_E}f(x)$, de sorte que $\|f(x)\|_F < \frac{2}{\delta}\|x\|_E$. D'où le résultat voulu avec $k = \frac{2}{\delta}$.

Réciproquement, supposons l'existence d'un réel $k > 0$ vérifiant $\|f(x)\|_F \leq k\|x\|_E$ pour tout $x \in E$. Alors, quels que soient $x, y \in E$, on a $\|f(x) - f(y)\|_F = \|f(x - y)\|_F \leq k\|x - y\|_E$. Ainsi f est lipschitzienne, et donc continue sur E d'après la proposition b).

Pour montrer (ii), on suppose maintenant que E est de dimension finie n , et on considère une application $f : E \rightarrow F$ quelconque. On fixe une base $\mathcal{B} = (e_1, \dots, e_n)$ de E et la norme $\|\cdot\|_\infty$ sur E définie par $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$ où $x = \sum_{i=1}^n x_i e_i$. Parce que E est de dimension finie, on sait que toute norme $\|\cdot\|_E$ sur E est équivalente à $\|\cdot\|_\infty$. Il existe $\alpha \in \mathbb{R}_+^*$ tel que $\|x\|_\infty \leq \alpha\|x\|_E$ pour tout $x \in E$. Alors, pour tout $x = \sum_{i=1}^n x_i e_i$ de E , on a :

$$\|f(x)\|_F = \left\| \sum_{i=1}^n x_i f(e_i) \right\|_F \leq \sum_{i=1}^n |x_i| \|f(e_i)\|_F \leq \left(\sum_{i=1}^n \|f(e_i)\|_F \right) \|x\|_\infty \leq k\|x\|_E,$$

avec $k := \alpha \sum_{i=1}^n \|f(e_i)\|_F > 0$, ce qui d'après le point (i) prouve que f est continue. \square

e) **COMPLÈMENT (norme d'une application linéaire).** Soient E et F deux evn. On note :

$\mathcal{L}(E, F)$ l'espace vectoriel des applications linéaires $E \rightarrow F$.

$\mathcal{LC}(E, F)$ l'espace vectoriel des applications linéaires et continues $E \rightarrow F$.

Il résulte du point (i) du théorème que, pour toute application $f \in \mathcal{LC}(E, F)$, l'ensemble de tous les réels $\frac{\|f(x)\|_F}{\|x\|_E}$ pour x décrivant $E \setminus \{0_E\}$ est une partie majorée (et non-vide) de \mathbb{R}_+ ; il admet donc une borne supérieure. On note :

$$\|f\| = \sup_{x \in E, x \neq 0_E} \frac{\|f(x)\|_F}{\|x\|_E} \text{ pour toute } f \in \mathcal{L}\mathcal{C}(E, F).$$

Par définition, on a donc :

$$\|f(x)\|_F \leq \|f\| \|x\|_E \text{ pour tous } f \in \mathcal{L}\mathcal{C}(E, F), x \in E.$$

Il est facile de vérifier que l'on a aussi, pour toute $f \in \mathcal{L}\mathcal{C}(E, F)$:

$$\|f\| = \sup_{x \in E, x \neq 0_E} \frac{\|f(x)\|_F}{\|x\|_E} = \sup_{x \in E, \|x\|_E \leq 1} \|f(x)\|_F = \sup_{x \in E, \|x\|_E = 1} \|f(x)\|_F.$$

On peut montrer facilement que, pour toutes $f, g \in \mathcal{L}\mathcal{C}(E, F)$ et tout $\lambda \in \mathbb{K}$, on a :

$$\begin{aligned} f + g \in \mathcal{L}\mathcal{C}(E, F) & \quad \text{et} \quad \|f + g\| \leq \|f\| + \|g\|, \\ \lambda f \in \mathcal{L}\mathcal{C}(E, F) & \quad \text{et} \quad \|\lambda f\| = |\lambda| \|f\|, \end{aligned}$$

ce qui traduit le fait que :

$$\mathcal{L}\mathcal{C}(E, F) \text{ est un espace vectoriel normé pour la norme } \|\cdot\|.$$

On peut enfin établir que, si E, F, G sont trois evn, pour toutes $f \in \mathcal{L}\mathcal{C}(E, F)$ et $g \in \mathcal{L}\mathcal{C}(F, G)$ on a :

$$g \circ f \in \mathcal{L}\mathcal{C}(E, G) \quad \text{et} \quad \|g \circ f\| \leq \|g\| \times \|f\|,$$

ce qui implique en particulier lorsque $E = F = G$ que :

$$\mathcal{L}\mathcal{C}(E, E) \text{ est une algèbre normée pour la norme } \|\cdot\|.$$

3.3.4 Exercices

EXERCICE 1. Soit X une partie de \mathbb{R} et f une application $X \rightarrow \mathbb{R}$. On note A l'ensemble des réels $\frac{f(x)-f(y)}{x-y}$ pour x, y décrivant \mathbb{R} avec $x \neq y$. Montrer que f est lipschitzienne si et seulement si A est borné. Montrer que l'application $g : \mathbb{R} \rightarrow \mathbb{R}$ définie par $g(x) = \frac{1}{1+|x|}$ est lipschitzienne. Montrer que l'application $h : [0, 1] \rightarrow \mathbb{R}$ définie par $h(x) = \sqrt{x}$ n'est pas lipschitzienne.

EXERCICE 2. Soient E un evn et $f : E \rightarrow E$ définie par $f(x) = x$ si $\|x\| \leq 1$ et $f(x) = \frac{1}{\|x\|}x$ si $\|x\| > 1$. Montrer que f est lipschitzienne de rapport 2.

EXERCICE 3. Soient E un evn et $f : E \rightarrow E$ définie par $f(x) = \frac{1}{\max(1, \|x\|)}x$. Montrer que f est lipschitzienne de rapport 2. On suppose de plus que $\|\cdot\|$ est la norme associée à un produit scalaire sur E . Montrer que f est lipschitzienne de rapport 1.

EXERCICE 4. On considère les trois applications $f, g, h : \mathbb{R}^2 \rightarrow \mathbb{R}$ définies par :

$$f(x, y) = \frac{xy}{x^4+y^4}, \quad g(x, y) = \frac{x^2y}{x^4+y^2}, \quad h(x, y) = \frac{xy}{x-y}.$$

Déterminer quatre suites de réels $(x_n)_{n \geq 1}$, $(y_n)_{n \geq 1}$, $(x'_n)_{n \geq 1}$ et $(y'_n)_{n \geq 1}$ convergeant toutes les quatre vers 0 dans \mathbb{R} , telles que les suites de réels $(f(x_n, y_n))_{n \geq 1}$ et $(f(x'_n, y'_n))_{n \geq 1}$ ne convergent pas dans \mathbb{R} vers la même limite. En déduire que f n'admet pas de limite en $(0, 0)$.

Même question pour g et h .

EXERCICE 5. Montrer que $\lim_{(x,y) \rightarrow (0,0)} \frac{x^3 + y^3}{x^2 + y^2} = 0$.

EXERCICE 6. Soit A une partie non-vide d'un evn E . Montrer que, pour tout $x \in E$, le nombre réel positif $d(x, A) := \inf\{\|x - a\|; a \in A\}$ est bien défini. Montrer que l'application $f_A : E \rightarrow \mathbb{R}$ définie par $f(x) = d(x, A)$ est continue car lipschitzienne.

EXERCICE 7. On considère dans l'evn $E = \mathbb{R}^2$ les parties $X = \mathbb{R}_+^* \times \mathbb{R}_+^*$, $Y = \mathbb{R}_+ \times \mathbb{R}_+^*$ et $Z = \mathbb{R}_+ \times \mathbb{R}_+$. On définit $f : X \rightarrow \mathbb{R}$ par $f(x, y) = x^y$ pour tout $(x, y) \in X$. Montrer que f est continue sur X . On prolonge f à Y en posant de plus $f(0, y) = 0$ pour tout $y > 0$. Montrer que ce prolongement de f est continue sur Y . Montrer que l'on ne peut pas prolonger par continuité à Z .

EXERCICE 8. Soit E un evn et $f \in \mathcal{L}\mathcal{C}(E, E)$. Montrer que toute valeur propre λ de f vérifie $|\lambda| \leq \|f\|$.

EXERCICE 9. Soient N_1 et N_2 deux normes sur un espace vectoriel E . Montrer qu'elles sont équivalentes si et seulement si l'application id_E est continue en tant qu'application de l'evn E muni de N_1 dans l'evn E muni de N_2 , et aussi continue en tant qu'application de l'evn E muni de N_2 dans l'evn E muni de N_1 .

EXERCICE 10. Soient E et F deux evn et $f : E \rightarrow F$ linéaire. Montrer que, si la suite $(f(x_n))$ est bornée dans F pour toute suite (x_n) convergeant vers 0_E dans E , alors f est continue.

EXERCICE 11. Soient E et F deux evn. Soit (x_n) une suite d'éléments de E convergeant dans E vers une limite ℓ . Soit (f_n) une suite d'application de $\mathcal{L}\mathcal{C}(E, F)$ convergeant dans $\mathcal{L}\mathcal{C}(E, F)$ vers une application f (au sens de la norme $\|\cdot\|$ définie à partir des normes $\|\cdot\|_E$ et $\|\cdot\|_F$). Montrer que la suite $(f(x_n))$ converge dans F vers $f(a)$.

EXERCICE 12 (*plus petite et plus grande valeur propre d'une matrice symétrique*). On fixe une matrice $A \in \mathcal{M}_n(\mathbb{R})$ que l'on suppose symétrique. On note u l'endomorphisme de $E = \mathbb{R}^n$ dont A est la matrice par rapport à la base canonique \mathcal{B} . Pour tout $x \in E$, en désignant par X la matrice colonne des composantes de x par rapport à \mathcal{B} , on a donc (pour le produit scalaire canonique) : $\langle u(x) | x \rangle = {}^t(AX)X = {}^tXAX$.

Soit $r : E \setminus \{0_E\} \rightarrow \mathbb{R}$ l'application (dite de Rayleigh) définie par $r(x) = \frac{1}{\|x\|^2} \langle u(x) | x \rangle$. On note $I = \{r(x); x \in E, x \neq 0_E\}$. Montrer que $I = r(S_n)$ où $S_n = \{x \in E; \|x\| = 1\}$ est la sphère unité de E . En déduire qu'il existe $x_m, x_M \in E$ non-nuls tels que $r(x_m) = \inf I$ et $r(x_M) = \sup I$. Montrer que la plus petite valeur propre λ_m de A et la plus grande λ_M sont données par :

$$\lambda_m = \inf_{X \neq 0} \frac{{}^tXAX}{{}^tXX} \quad \text{et} \quad \lambda_M = \sup_{X \neq 0} \frac{{}^tXAX}{{}^tXX}, \quad \text{avec} \quad r(E \setminus \{0_E\}) = [\lambda_m, \lambda_M].$$

EXERCICE 13. On considère l'evn $E = \mathcal{C}([0, 1], \mathbb{R})$ muni de la norme $\|\cdot\|_\infty$, et l'evn $F = \mathbb{R}$ muni de la valeur absolue. Montrer que l'application $\varphi : E \rightarrow F$ définie par $\varphi(f) = f(1) - f(0)$ pour toute $f \in E$ est linéaire continue et que $\|\varphi\| = 2$.

EXERCICE 14. On considère l'evn $E = \mathcal{C}([0, 1], \mathbb{R})$ muni de la norme $\|\cdot\|_2$, et l'evn $F = \mathbb{R}$ muni de la valeur absolue. Pour toute application $\varphi \in E$, on définit l'application $T_\varphi : E \rightarrow F$ par $T_\varphi(f) = \int_0^1 f(t)\varphi(t) dt$ pour toute $f \in E$. Montrer que T_φ est linéaire continue et que $\|T_\varphi\| = \|\varphi\|_2$.

EXERCICE 15. On considère l'evn $E = \mathcal{C}([0, 1], \mathbb{R})$ pour la norme $N_1 = \|\cdot\|_\infty$. On considère l'evn $F = \mathcal{C}^1([0, 1], \mathbb{R})$ pour la norme N_2 définie par $N_2(f) = \|f\|_\infty + \|f'\|_\infty$. Soit $T : E \rightarrow F$ où, pour toute $f \in E$, l'application $T(f) \in F$ est définie par $T(f)(x) = \int_0^x f(t) dt$ pour tout $x \in [0, 1]$. Montrer que $T \in \mathcal{L}\mathcal{C}(E, F)$ et que $\|T_\varphi\| = 2$.

EXERCICE 16. On considère l'espace vectoriel $E = \mathcal{C}([0, 1], \mathbb{R})$ et l'endomorphisme $u \in \mathcal{L}(E, E)$ où, pour toute $f \in E$, l'application $u(f) \in E$ est définie par $u(f)(x) = f(x) - f(0)$ pour tout $x \in [0, 1]$.

Montrer que, si E est muni de la norme $\|\cdot\|_\infty$, alors u est continu et $\|u\| = 2$.

Montrer que, si E est muni de la norme $\|\cdot\|_1$, alors u n'est pas continu [*indication* : considérer dans E les fonctions $f_n : x \mapsto (n+1)(1-x)^n$, vérifier que l'on a $\|f_n\|_1 = 1$ et $\|u(f_n)\|_1 = n$, et en déduire le résultat.

EXERCICE 17. On considère l'evn $E = \ell^\infty(\mathbb{R})$ des suites de réels bornées, muni de la norme $\|\cdot\|_\infty$. Pour toute suite $x = (x_n)$ de E , on définit la suite $F(x) = (y_n)$ par $y_n = u_{n+1}$ pour tout $n \in \mathbb{N}$, et la suite $G(x) = (z_n)$ par $z_n = u_{n+1} - u_n$ pour tout $n \in \mathbb{N}$. Montrer que les applications $F : E \rightarrow E$ et $G : E \rightarrow E$ ainsi définies sont linéaires et continues. Montrer que $\|F\| = 1$ et $\|G\| = 2$.

EXERCICE 18. On considère l'evn $E = \mathcal{C}([0, 1], \mathbb{R})$ muni de la norme $\|\cdot\|_1$, et l'evn $F = \mathbb{R}$ muni de la valeur absolue. Montrer que l'application $\varphi : E \rightarrow F$ définie par $\varphi(f) = \int_0^1 tf(t) dt$ pour toute $f \in E$ est continue et que $\|\varphi\| = 1$.

Chapitre 4

Séries numériques

Conformément à ce que prévoit le programme, ce chapitre est plus précisément consacré à l'étude des séries à *termes réels ou complexes*. Dans tout le chapitre, \mathbb{K} désigne \mathbb{R} ou \mathbb{C} , et $|\cdot|$ désigne respectivement la valeur absolue ou le module.

4.1 Notion de série

4.1.1 Terminologie des séries

Soit $(u_n)_{n \geq 0}$ une suite d'éléments de \mathbb{K} . Posons pour tout $N \in \mathbb{N}$:

$$S_N = u_0 + u_1 + u_2 + \cdots + u_N = \sum_{n=0}^N u_n \in \mathbb{K}.$$

S_N s'appelle la N -ième *somme partielle* associée à la suite $(u_n)_{n \geq 0}$. On introduit ainsi une nouvelle suite $(S_N)_{N \geq 0}$ dans $\mathbb{K}^{\mathbb{N}}$, dite *suite des sommes partielles* associée à $(u_n)_{n \geq 0}$.

On appelle *série numérique* la donnée d'un couple formé par une suite $(u_n)_{n \geq 0}$ d'éléments de \mathbb{K} et la suite de ses sommes partielles. On note cette série : $\sum_{n \geq 0} u_n$, ou plus simplement $\sum u_n$. L'élément u_n s'appelle n -ième terme, ou *terme général* de la série. Une série dont le terme général est dans \mathbb{R} s'appelle une *série réelle*.

Ce vocabulaire reste valable pour une suite $(u_n)_{n \geq p}$ définie à partir d'un certain rang p ; on note $\sum_{n \geq p} u_n$ la série associée.

4.1.2 Convergence d'une série

a) DÉFINITIONS. On dit qu'une série numérique $\sum_{n \geq 0} u_n$ *converge* lorsque la suite (S_N) de ses sommes partielles converge dans \mathbb{K} . Dans ce cas, la limite de la suite (S_N) s'appelle la *somme* de la série $\sum_{n \geq 0} u_n$. On la note : $\sum_{n=0}^{+\infty} u_n$. Donc :

$$\text{pour une série convergente } \sum_{n \geq 0} u_n, \text{ on a : } \sum_{n=0}^{+\infty} u_n = \lim_{N \rightarrow +\infty} \sum_{n=0}^N u_n \in \mathbb{K}.$$

On dit qu'une série numérique *diverge* lorsqu'elle ne converge pas. On dit que deux séries numériques sont *de même nature* lorsqu'elles sont toutes les deux convergentes ou toutes les deux divergentes.

Le lemme suivant exprime que : la nature d'une série n'est pas modifiée lorsque l'on change l'indice de départ ; mais quand il y a convergence, la valeur de la somme peut être modifiée.

b) LEMME (Changement d'indice de départ). Soient $\sum_{n \geq 0} u_n$ une série numérique, et $p \in \mathbb{N}$. Alors, les séries $\sum_{n \geq 0} u_n$ et $\sum_{n \geq p} u_n$ sont de même nature. De plus, si elles convergent, on a :

$$\sum_{n=0}^{+\infty} u_n = \sum_{n=p}^{+\infty} u_n + \sum_{n=0}^{p-1} u_n.$$

PREUVE. Cela résulte directement de l'égalité (pour $N \geq p$) : $\sum_{n=0}^N u_n = \sum_{n=p}^N u_n + \sum_{n=0}^{p-1} u_n$ \square

c) PROPOSITION (Condition nécessaire de convergence). Si une série numérique converge, alors son terme général tend vers 0. Plus précisément, si la série $\sum u_n$ converge, alors la suite (u_n) converge vers 0.

PREUVE. Soit $\sum u_n$ une série convergente ; notons S sa somme. Si (S_n) désigne la suite des sommes partielles, on a $u_n = S_n - S_{n-1}$ pour tout $n \geq 1$. Comme (S_n) converge vers S , il est clair que $(S_n - S_{n-1})$ converge vers 0. \square

Cette proposition indique qu'une série numérique $\sum u_n$ telle que la suite (u_n) ne converge pas vers 0 est nécessairement divergente. On dit alors qu'elle est *grossièrement divergente*.

ATTENTION ! Cette condition du c) est NÉCESSAIRE mais NON SUFFISANTE. Il existe des suites (u_n) tendant vers 0 telles que la série $\sum u_n$ diverge. (cf. exemples ci-dessous)

d) REMARQUE (critère de Cauchy). On sait que, dans \mathbb{R} ou \mathbb{C} , toute suite de Cauchy est convergente. Donc une série $\sum u_n$ numérique converge si et seulement si la suite des sommes partielles est de Cauchy, c'est-à-dire :

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, \forall p \geq 1, \left| \sum_{k=1}^p u_{n+k} \right| < \varepsilon,$$

en remarquant que $S_{n+p} - S_n = \sum_{k=0}^{n+p} u_k - \sum_{k=0}^n u_k = \sum_{k=n+1}^{n+p} u_k$.

d) REMARQUE (partie réelle et partie imaginaire). On sait qu'une suite de nombres complexes (u_n) converge dans \mathbb{C} si et seulement si les suites réelles $(\operatorname{Re} u_n)$ et $(\operatorname{Im} u_n)$ convergent dans \mathbb{R} , et que dans ce cas $\lim u_n = \lim \operatorname{Re} u_n + i \lim \operatorname{Im} u_n$. En appliquant cela à la suite des sommes partielles d'une série de nombres complexes, on déduit que :

Soit $\sum u_n$ une série complexe. Pour tout $n \geq 0$, notons $u_n = x_n + iy_n$ avec $x_n, y_n \in \mathbb{R}$. La série complexe $\sum u_n$ converge si et seulement si les séries réelles $\sum x_n$ et $\sum y_n$. Dans ce cas, on a :

$$\sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{+\infty} x_n + i \sum_{n=0}^{+\infty} y_n.$$

4.1.3 Exemples classiques

a) SÉRIE GÉOMÉTRIQUE. Fixons $z \in \mathbb{K}$, et posons $u_n = z^n$ pour tout $n \in \mathbb{N}$. La série $\sum_{n \geq 0} u_n$ ainsi définie s'appelle la *série géométrique* de raison z .

Elle est convergente si et seulement si $|z| < 1$, et dans ce cas sa somme est $\sum_{n=0}^{+\infty} z^n = \frac{1}{1-z}$.

En effet : si $|z| \geq 1$, le terme général $u_n = z^n$ ne tend pas vers 0, donc la série est grossièrement divergente. Si $|z| < 1$, alors $S_N = \sum_{n=0}^N z^n = \frac{1-z^{N+1}}{1-z}$, donc $\lim_N S_N = \frac{1}{1-z}$, donc $\sum_{n \geq 0} z^n$ est convergente de somme $\frac{1}{1-z}$.

b) SÉRIE HARMONIQUE. Posons $u_n = \frac{1}{n}$ pour tout $n \in \mathbb{N}^*$. La série $\sum_{n \geq 1} u_n$ ainsi définie s'appelle la *série harmonique*. Elle est divergente (bien que son terme général tende vers 0).

En effet : Pour tout $N \geq 1$, on a la minoration $S_N = \sum_{k=1}^N \frac{1}{k} \geq \sum_{k=1}^N \ln(1 + \frac{1}{k})$.

Or, $\sum_{k=1}^N \ln(1 + \frac{1}{k}) = \sum_{k=1}^N \ln(\frac{k+1}{k}) = \sum_{k=1}^N (\ln(k+1) - \ln k) = \ln(1+N)$. Donc $S_N \geq \ln(1+N)$.

Comme $\ln(1+N)$ tend vers $+\infty$ quand N tend vers $+\infty$, on en déduit que la suite $(S_N)_N$ diverge vers $+\infty$, et donc la série de terme général $\frac{1}{n}$ est divergente.

c) SÉRIE HARMONIQUE ALTERNÉE. Posons $u_n = \frac{(-1)^n}{n}$ pour tout $n \in \mathbb{N}^*$. La série $\sum_{n \geq 1} u_n$ ainsi définie s'appelle la *série harmonique alternée*. Elle est convergente de somme : $\sum_{n=1}^{+\infty} \frac{(-1)^n}{n} = -\ln 2$.

En effet : $-\sum_{k=1}^n \frac{(-1)^k}{k} = -\sum_{k=1}^n (-1)^k \int_0^1 t^{k-1} dt = \int_0^1 \sum_{k=1}^n (-t)^{k-1} dt = \int_0^1 \frac{1-(-t)^n}{1+t} dt$.

On scinde cette dernière intégrale en : $\int_0^1 \frac{1-(-t)^n}{1+t} dt = \int_0^1 \frac{dt}{1+t} + (-1)^{n+1} \int_0^1 \frac{t^n}{1+t} dt$.

Or $\int_0^1 \frac{dt}{1+t} = \ln 2$, et par ailleurs $0 \leq \int_0^1 \frac{t^n}{1+t} dt \leq \int_0^1 t^n dt = \frac{1}{n+1}$ donc $\lim_{n \rightarrow +\infty} \int_0^1 \frac{t^n}{1+t} dt = 0$.

On conclut que la suite (S_n) des sommes partielles converge vers $\ln 2$. Donc la série de terme général $\frac{(-1)^n}{n}$ est convergente et sa somme est $\sum_{n=1}^{+\infty} \frac{(-1)^n}{n} = -\ln 2$.

d) SÉRIES DITES "TÉLÉSCOPIQUES". On désigne par ce terme familier des séries dont le terme général u_n peut s'exprimer sous la forme $u_n = v_{n+1} - v_n$ pour une certaine suite (v_n) . On calcule alors les sommes partielles sous la forme :

$$S_N = \sum_{n=0}^N (v_{n+1} - v_n) = v_1 - v_0 + v_2 - v_1 + v_3 - v_2 + \dots + v_N - v_{N-1} + v_{N+1} - v_N = v_{N+1} - v_0.$$

On en déduit que, si la suite (v_n) converge vers une limite ℓ , alors la suite (S_N) converge vers $\ell - v_0$, et donc la série de terme général u_n est convergente et sa somme est $\ell - v_0$.

Par exemple, la série $\sum_{n \geq 1} \frac{1}{n(n+1)}$ est convergente, de somme $\sum_{n=1}^{+\infty} \frac{1}{n(n+1)} = 1$.

En effet : Posons $u_n = \frac{1}{n(n+1)}$ pour tout $n \geq 1$. On remarque que $u_n = \frac{1}{n} - \frac{1}{n+1}$. Donc, pour $N \geq 1$, on a : $S_N = 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \dots + \frac{1}{N} - \frac{1}{N+1} = 1 - \frac{1}{N+1}$. La suite $(S_N)_{N \geq 1}$ converge donc vers 1.

e) REMARQUE. Il est rare dans la pratique que l'on puisse ainsi faire un calcul direct de la somme d'une série convergente. La plupart des résultats que l'on va voir dans la suite permettent seulement de déterminer la nature d'une série.

4.1.4 Espace vectoriel des séries convergentes

a) DÉFINITIONS (*somme de deux séries, produit d'une série par un scalaire*). Soient $\sum u_n$ et $\sum v_n$ deux séries numériques. Leur somme est la série de terme général $u_n + v_n$. Si de plus $\lambda \in \mathbb{K}$, le produit externe de $\sum u_n$ par λ est la série de terme général λu_n .

b) PROPOSITION. Soient $\sum u_n$ et $\sum v_n$ deux séries numériques convergentes. Alors, pour tous $\lambda, \mu \in \mathbb{K}$, la série $\sum(\lambda u_n + \mu v_n)$ converge, et sa somme est :

$$\sum_{n=0}^{+\infty} (\lambda u_n + \mu v_n) = \lambda \sum_{n=0}^{+\infty} u_n + \mu \sum_{n=0}^{+\infty} v_n.$$

PREUVE. On applique aux sommes partielles les résultats correspondants sur les limites de suites. \square

c) COROLAIRE. Les séries numériques dans \mathbb{K} forment un \mathbb{K} -espace vectoriel, et les séries convergentes en forment un sous-espace.

PREUVE. Résulte de la structure d'espace vectoriel de $\mathbb{K}^{\mathbb{N}}$, et du b) ci-dessus. \square

4.1.5 Exercices

EXERCICE 1. Montrer que la somme d'une série convergente et d'une série divergente est nécessairement divergente. Montrer par deux exemples bien choisis que la somme de deux séries divergentes peut être convergente ou divergente.

EXERCICE 2. Dans chacun des deux cas suivants, déterminer si la série de terme général u_n converge, et si oui calculer sa somme.

a) $u_n = \sin n$, b) $u_n = \ln(1 - \frac{1}{n^2})$, c) $u_n = \arctan \frac{1}{1+n+n^2}$.

EXERCICE 3 (*Série de Tannery*). Soit x un réel différent de 1 et de -1 . On considère la série de terme général :

$$u_n = \frac{2x^{2^n}}{x^{2^{n+1}} - 1}.$$

A partir de l'égalité $\frac{t}{t^2-1} = \frac{1}{t-1} - \frac{1}{t+1}$ pour tout réel $t \neq \pm 1$, montrer que les sommes partielles $S_p = \sum_{n=0}^p u_n$ vérifient $S_p = \frac{2}{x-1} - \frac{2}{x^{2^{p+1}}-1}$. Conclure que la série $\sum u_n$ est convergente, et calculer sa somme en distinguant suivant que $|x| > 1$ ou que $|x| < 1$

EXERCICE 4 (*Une autre preuve de la divergence de la série harmonique*). Considérons les sommes partielles $S_m = \sum_{n=1}^m \frac{1}{n}$, avec $m \geq 1$. Montrer que :

$$|S_{2m} - S_m| = \frac{1}{m+1} + \frac{1}{m+2} + \dots + \frac{1}{2m} \geq \frac{1}{2}.$$

En déduire que la suite (S_n) n'est pas de Cauchy. Conclure que la série harmonique est divergente.

EXERCICE 5 (*Une autre preuve de la divergence de la série harmonique*). Considérons les sommes partielles $S_n = \sum_{k=1}^n \frac{1}{k}$, avec $n \geq 1$. Montrer que l'entier $m = E(\frac{\ln n}{\ln 2})$ vérifie $n \geq 2^m$ et $S_n \geq S_{2^m}$.

Montrer que :

$$S_{2^m} = \sum_{j=1}^m (\frac{1}{2^{j-1}+1} + \frac{1}{2^{j-1}+2} + \dots + \frac{1}{2^j}) \geq 1 + \frac{m}{2}.$$

Déduire que $\lim S_{2^m} = +\infty$. Conclure que la série harmonique est divergente.

4.2 Séries à termes réels positifs

4.2.1 Critère de majoration

a) LEMME. Soient $\sum u_n$ une série réelle telle que $u_n \geq 0$ à partir d'un certain rang p . Pour tout $N \geq p$, notons $S_N = \sum_{n=p}^N u_n$. La série $\sum u_n$ converge si et seulement si la suite (S_N) est majorée.

PREUVE. Comme $u_n \geq 0$ pour tout $n \geq p$, la suite $(S_N)_{N \geq p}$ est croissante. D'après les résultats connus sur les suites réelles croissantes, ou bien (S_N) est majorée, et alors elle converge, ou bien elle n'est pas majorée, et alors elle tend vers $+\infty$. D'où le résultat. \square

b) THÉORÈME (fondamental). Soient $\sum u_n$ et $\sum v_n$ deux séries réelles telles que, à partir d'un certain rang p , on ait : $0 \leq u_n \leq v_n$ pour tout $n \geq p$.

(i) Si $\sum v_n$ converge, alors $\sum u_n$ converge, et dans ce cas : $0 \leq \sum_{n=p}^{+\infty} u_n \leq \sum_{n=p}^{+\infty} v_n$.

(ii) Si $\sum u_n$ diverge, alors $\sum v_n$ diverge.

PREUVE. Pour tout $N \geq p$, notons $S_N = \sum_{n=p}^N u_n$ et $T_N = \sum_{n=p}^N v_n$.

D'après l'hypothèse, les suites $(S_N)_{N \geq p}$ et $(T_N)_{N \geq p}$ sont croissantes, et vérifient $S_N \leq T_N$ pour tout $N \geq p$. Supposons que $\sum v_n$ converge; d'après le lemme a), la suite (T_N) est majorée, donc (S_N) l'est, et en réappliquant le lemme, on conclut que la série $\sum u_n$ converge. En appliquant le théorème de passage à la limite dans les inégalités, on a $\lim S_N \leq \lim T_N$, ce qui achève de prouver (i). Le point (ii) s'en déduit par contraposition. \square

c) COROLLAIRE (critère de domination). Soient $\sum u_n$ et $\sum v_n$ deux séries à termes réels positifs à partir d'un certain rang. On suppose que $u_n = O(v_n)$ au voisinage de l'infini [rappelons que cela signifie qu'il existe $p \in \mathbb{N}$ et $\alpha \in \mathbb{R}_+^*$ tel que l'on ait $0 \leq u_n \leq \alpha v_n$ pour tout entier $n \geq p$].

(i) Si $\sum v_n$ converge, alors $\sum u_n$ converge.

(ii) Si $\sum u_n$ diverge, alors $\sum v_n$ diverge.

PREUVE. Comme le produit d'une série convergente par un scalaire est une série convergente (cf. 4.1.4), on applique le théorème précédent aux séries $\sum u_n$ et $\sum \alpha v_n$. \square

d) COROLLAIRE (critère d'équivalence). Soient $\sum u_n$ et $\sum v_n$ deux séries à termes réels positifs à partir d'un certain rang. Si u_n et v_n sont équivalents au voisinage de ∞ , alors les séries $\sum u_n$ et $\sum v_n$ sont de même nature.

PREUVE. Par hypothèse, on a à partir d'un certain rang : $u_n = v_n(1 + r_n)$, avec (r_n) suite tendant vers 0. Traduisons cette limite pour $\varepsilon = \frac{1}{2}$: il existe $N \in \mathbb{N}$ tel que pour tout $n \geq N$, on a : $-\frac{1}{2} < r_n < \frac{1}{2}$ donc $\frac{1}{2} < 1 + r_n < \frac{3}{2}$. Donc, à partir d'un certain rang, on a : $0 \leq \frac{1}{2}v_n \leq u_n = v_n(1 + r_n) \leq \frac{3}{2}v_n$. Ceci prouve que $u_n = O(v_n)$ et $v_n = O(u_n)$. On applique alors le corollaire précédent. \square

Donnons tout de suite un exemple très important d'application de ces théorèmes.

4.2.2 Séries de Riemann

a) DÉFINITION. On appelle série de Riemann une série de la forme $\sum_{n \geq 1} \frac{1}{n^\alpha}$, où α est un réel fixé.

b) THÉORÈME. La série de Riemann $\sum_{n \geq 1} \frac{1}{n^\alpha}$ converge si et seulement si $\alpha > 1$.

PREUVE. On raisonne en deux cas :

• Si $\alpha \leq 1$, alors $n^\alpha \leq n$, donc $\frac{1}{n^\alpha} \geq \frac{1}{n} > 0$. Comme la série harmonique diverge comme on l'a vu au b) de 4.1.3, on applique 4.2.1.b.(ii) pour conclure que $\sum \frac{1}{n^\alpha}$ diverge.

• Si $\alpha > 1$, posons $\beta = \alpha - 1 > 0$; définissons $a_n = \frac{1}{n^\beta}$ et $b_n = a_n - a_{n+1}$ pour tout $n \geq 1$. Comme $\beta > 0$, on a : $a_{n+1} < a_n$ donc $b_n > 0$. Formons pour $N \geq 1$ la somme partielle $S_N = \sum_{n=1}^N b_n = a_1 - a_2 + a_2 - a_3 + \dots + a_N - a_{N+1} = 1 - \frac{1}{(N+1)^\beta}$. Comme $\beta > 0$, la suite (S_N) converge vers 1. Ainsi, la série $\sum b_n$ converge (et sa somme est 1).

Or $\frac{1}{n^{\beta+1}}$ est équivalent à $\frac{1}{\beta} b_n$ au voisinage de ∞ .

$$\text{en effet : } b_n = \frac{1}{n^\beta} - \frac{1}{(n+1)^\beta} = \frac{1}{n^\beta} \left[1 - \left(\frac{n}{n+1} \right)^\beta \right] = \frac{1}{n^\beta} \left[1 - \left(1 + \frac{1}{n} \right)^{-\beta} \right]$$

$$\text{On utilise le D.L. : } \left(1 + \frac{1}{n} \right)^{-\beta} = 1 - \beta \frac{1}{n} + \frac{1}{n} \varepsilon_n \text{ avec } \lim_n \varepsilon_n = 0.$$

$$\text{Donc : } b_n = \frac{1}{n^\beta} \left[\beta \frac{1}{n} - \frac{1}{n} \varepsilon_n \right] = \frac{\beta}{n^{\beta+1}} \left[1 - \frac{1}{\beta} \varepsilon_n \right]$$

Ainsi $\frac{1}{n^\alpha} = \frac{1}{n^{\beta+1}}$ est équivalent à $\beta^{-1} b_n$. Comme la série $\sum b_n$ converge, il en est de même de $\sum \beta^{-1} b_n$; il s'agit d'une série à terme réels positifs. On applique le d) de 4.2.1 pour conclure que $\sum \frac{1}{n^\alpha}$ converge dans ce cas. \square

c) COROLLAIRE (Règle de comparaison avec une série de Riemann). Soit $\sum u_n$ une série à termes réels positifs.

(i) S'il existe un réel $\alpha > 1$ tel que $\lim_n n^\alpha u_n = 0$, alors la série $\sum u_n$ converge.

(ii) S'il existe un réel $0 < \alpha \leq 1$ tel que $\lim_n n^\alpha u_n = +\infty$, alors la série $\sum u_n$ diverge.

PREUVE. Traduisons d'abord $\lim_n n^\alpha u_n = 0$ pour $\varepsilon = 1$; il existe $N_1 \in \mathbb{N}$ tel que :

$$\text{pour tout } n \geq N_1, \text{ on a : } 0 \leq n^\alpha u_n < 1, \text{ donc } 0 \leq u_n < \frac{1}{n^\alpha}.$$

Comme $\alpha > 1$, $\sum \frac{1}{n^\alpha}$ converge d'après le b) ci-dessus, d'où $\sum u_n$ converge d'après 4.2.1.b.

Traduisons maintenant $\lim_n n^\alpha u_n = +\infty$ pour $\varepsilon = 1$; il existe $N_2 \in \mathbb{N}$ tel que :

$$\text{pour tout } n \geq N_2, \text{ on a : } n^\alpha u_n > 1, \text{ donc } u_n > \frac{1}{n^\alpha}.$$

Comme $0 < \alpha < 1$, $\sum \frac{1}{n^\alpha}$ diverge d'après le b) ci-dessus, d'où $\sum u_n$ diverge d'après 4.2.1.b. \square

d) EXEMPLE. Soit $a \in \mathbb{R}$ fixé. La série $\sum \exp[-(\ln n)^a]$ converge si et seulement si $a > 1$.

En effet : En effet, posons $u_n = \exp[-(\ln n)^a]$. Si $a = 1$, alors $u_n = \frac{1}{n}$, donc $\sum u_n$ est la série harmonique; on sait qu'elle diverge. On suppose donc maintenant $a \neq 1$. En vue d'utiliser le corollaire ci-dessus, on forme : $n^\alpha u_n = \exp(\alpha \ln n - (\ln n)^a)$ pour un $\alpha > 0$. Or :

$$\lim_{n \rightarrow +\infty} (\alpha \ln n - (\ln n)^a) = \lim_{n \rightarrow +\infty} (\ln n)(\alpha - (\ln n)^{a-1}) = \begin{cases} -\infty & \text{si } a > 1 \\ +\infty & \text{si } a < 1 \end{cases}$$

Si $a > 1$, alors $\lim_n n^\alpha u_n = 0$ pour tout $\alpha > 0$. On applique à $\alpha = 2$ pour conclure avec le point (i) du corollaire précédent que $\sum u_n$ converge.

Si $a < 1$, alors $\lim_n n^\alpha u_n = +\infty$ pour tout $\alpha > 0$. On applique à $\alpha = 1$ pour conclure avec le point (ii) du corollaire précédent que $\sum u_n$ diverge.

e) REMARQUE. Il est clair que, dans le point (i) de l'énoncé c), on peut remplacer la condition $\lim_n n^\alpha u_n = 0$ pour un $\alpha > 1$ par la condition plus faible : il existe $\alpha > 1$ telle que la suite de réels positifs $(n^\alpha u_n)$ soit majorée.

4.2.3 Règle de d'Alembert

a) PROPOSITION. On considère une série réelle $\sum u_n$ telle que $u_n > 0$ à partir d'un certain rang. On suppose que la suite $(\frac{u_{n+1}}{u_n})$ admet une limite $\ell \in \mathbb{R}_+$.

(i) Si $\ell < 1$, alors la série $\sum u_n$ converge.

(ii) Si $\ell > 1$, alors la série $\sum u_n$ diverge.

PREUVE. Supposons d'abord $\ell < 1$. Choisissons un réel λ tel que $\ell < \lambda < 1$. Traduisons $\lim(\frac{u_{n+1}}{u_n}) = \ell$ pour $\varepsilon = \lambda - \ell > 0$; il existe $N \in \mathbb{N}$ tel que :

$$\text{pour tout } n \geq N, \text{ on a } u_n > 0 \text{ et } \ell - \varepsilon < \frac{u_{n+1}}{u_n} < \ell + \varepsilon = \lambda.$$

Ainsi, pour $n \geq N$, on a : $0 < \frac{u_{n+1}}{u_n} < \lambda$. On déduit : $\frac{u_{N+1}}{u_N} \times \frac{u_{N+2}}{u_{N+1}} \times \dots \times \frac{u_{n-1}}{u_{n-2}} \times \frac{u_n}{u_{n-1}} < \lambda^{n-N}$, d'où $\frac{u_n}{u_N} < \lambda^{n-N}$. Par suite, pour tout $n \geq N$, on a $u_n < (u_N \lambda^{-N}) \lambda^n$, et donc $u_n = O(\lambda^n)$. Or, comme $0 < \lambda < 1$, la série géométrique $\sum \lambda^n$ converge (cf. a) du 4.1.3); on applique d) de 4.2.1 pour conclure que $\sum u_n$ converge.

• Pour montrer (ii), supposons $\ell > 1$. Choisissons un réel ε tel que $0 < \varepsilon < \ell - 1$. Traduisons $\lim(\frac{u_{n+1}}{u_n}) = \ell$ pour ce ε ; il existe $N_0 \in \mathbb{N}$ tel que :

$$\text{pour tout } n \geq N_0, \text{ on a : } u_n > 0 \text{ et } 1 < \ell - \varepsilon < \frac{u_{n+1}}{u_n} < \ell + \varepsilon,$$

et donc $u_{n+1} > u_n$. Ainsi, la suite (u_n) est croissante à partir du rang N_0 . Elle est donc minorée par $u_{N_0} > 0$. Elle ne peut donc pas converger vers 0. D'après la proposition c) de 4.1.2, la série $\sum u_n$ est grossièrement divergente. \square

b) EXEMPLE : La série de terme général $u_n = \frac{n!}{n^n}$ converge.

En effet, on a $u_n > 0$ pour tout $n \geq 1$, et : $\frac{u_{n+1}}{u_n} = (\frac{n}{n+1})^n = (1 + \frac{1}{n})^{-n} = \exp(-n \ln(1 + \frac{1}{n}))$ qui tend vers $\exp(-1)$ quand $n \rightarrow +\infty$. Comme $e^{-1} < 1$, on conclut que la série $\sum u_n$ converge.

c) REMARQUE. Si $\lim_n(\frac{u_{n+1}}{u_n}) = 1$, il se peut que $\sum u_n$ converge ou qu'elle diverge (prendre par exemple : $u_n = \frac{1}{n}$ et $u_n = \frac{1}{n^2}$). De même lorsque $(\frac{u_{n+1}}{u_n})$ n'admet pas de limite.

Si $\lim_n(\frac{u_{n+1}}{u_n}) = +\infty$, on a $u_{n+1} > u_n > 0$ à partir d'un certain rang, donc $\sum u_n$ est grossièrement divergente (comme à la fin de la preuve du cas (ii) du théorème a).

4.2.4 Règle de Cauchy

a) PROPOSITION. On considère une série réelle $\sum u_n$ telle que $u_n > 0$ à partir d'un certain rang. On suppose que la suite $(u_n^{\frac{1}{n}})$ admet une limite $\ell \in \mathbb{R}_+$.

- (i) Si $\ell < 1$, alors la série $\sum u_n$ converge.
- (ii) Si $\ell > 1$, alors la série $\sum u_n$ diverge.

PREUVE. On raisonne comme dans la preuve de 4.2.3

- Supposons d'abord $\ell < 1$. Choisissons un réel λ tel que $\ell < \lambda < 1$. Comme dans la preuve du a) de 4.2.3, il existe $N \in \mathbb{N}$ tel que, pour tout $n \geq N$, on a : $0 < (u_n)^{\frac{1}{n}} < \lambda$, donc $0 < u_n < \lambda^n$. La série géométrique $\sum \lambda^n$ converge puisque $\lambda < 1$, donc $\sum u_n$ converge d'après 4.2.1.b.(i).

- Supposons maintenant $\ell > 1$. Comme dans la preuve du a) de 4.2.3, il existe $N_0 \in \mathbb{N}$ tel que, pour tout $n \geq N_0$, on a : $(u_n)^{\frac{1}{n}} > 1$ donc $u_n > 1$. Dès lors, la suite (u_n) ne peut pas converger vers 0. D'après la proposition c) de 4.1.2, la série $\sum u_n$ est grossièrement divergente. \square

b) EXEMPLE : la série $\sum \left(\frac{n+1}{2n-1}\right)^n$ converge, car $\lim_n \left(\frac{n+1}{2n-1}\right)^{\frac{1}{n}} = \frac{1}{2} < 1$.

c) REMARQUES. On peut montrer que si la règle de d'Alembert s'applique, alors la règle de Cauchy aussi, c'est-à-dire que si $\lim_n \frac{u_{n+1}}{u_n} = \ell \in \mathbb{R}_+$, alors $\lim_n u_n^{1/n} = \ell$. Néanmoins, il est souvent plus commode d'étudier la suite $\left(\frac{u_{n+1}}{u_n}\right)$ que la suite $(u_n^{1/n})$.

Réciproquement, on peut avoir $\lim_n (u_n)^{\frac{1}{n}} = l \in \mathbb{R}_+$ sans que $\lim_n \left(\frac{u_{n+1}}{u_n}\right)$ existe.

Contre-exemple : soit (u_n) définie par $u_{2p} = u_{2p+1} = 2^p$ pour tout $p \in \mathbb{N}$.

D'une part $\lim_p (u_{2p})^{\frac{1}{2p}} = \lim_p (u_{2p+1})^{\frac{1}{2p+1}} = \sqrt{2}$, ce qui est suffisant pour conclure que $\lim_n (u_n)^{\frac{1}{n}} = \sqrt{2}$ (résultat classique d'analyse réelle ou complexe, voir aussi chapitre 3 de ce cours). D'autre part, la suite $\left(\frac{u_{n+1}}{u_n}\right)$ vaut alternativement 1 et 2, donc ne converge pas.

Pour conclure : on dispose ainsi de nombreux résultats techniques pour étudier les séries à termes réels positifs. L'un de leurs intérêts réside dans le théorème 4.3.1.b de la section suivante.

4.2.5 Exercices

EXERCICE 1. Déterminer la nature des séries de termes général u_n défini par :

$$u_n = \frac{1}{\sqrt{n(n+1)}}, \quad u_n = \frac{1}{\sqrt{n(n^2+1)}}, \quad u_n = \frac{2n}{n+2^n}, \quad u_n = \frac{n!}{n^n}, \quad u_n = \frac{1}{n \times n^{1/n}}$$

EXERCICE 2. Déterminer la nature des séries de termes général u_n défini par :

$$u_n = \frac{2 + \cos n}{n^p} \quad \text{avec } p \in \mathbb{Z} \text{ fixé}, \quad u_n = \frac{5^n + 1}{2^n - 1}, \quad u_n = \frac{(n^4 + 1)^{1/3}}{n(n-1)^{1/2}}.$$

EXERCICE 3. Déterminer la nature des séries de termes général u_n défini par :

$$u_n = \frac{2^n}{n^2} \sin^{2n} \alpha \quad \text{avec } 0 < \alpha \leq \frac{\pi}{2} \quad \text{et} \quad u_n = a^n \sin^2(n\alpha) \quad \text{avec } \alpha \in \mathbb{R}, 0 < a < 1.$$

EXERCICE 4. Soient $a, b \in \mathbb{R}_+^*$ et $\alpha \in \mathbb{R}$. Déterminer la nature de la série $\sum \frac{1}{(an+b)^\alpha}$.

EXERCICE 5. Soit $a \in \mathbb{R}_+^*$. Déterminer la nature des séries de termes général u_n défini par :

$$u_n = \frac{a^n}{n!}, \quad u_n = n! a^n, \quad u_n = n a^n, \quad u_n = \frac{a^n}{n^n}.$$

EXERCICE 6. Déterminer la nature des séries de termes général u_n défini par :

$$u_n = n^{(1/n)^\alpha} - 1 \quad \text{avec } \alpha \in \mathbb{R}, \quad u_n = \frac{1}{2^{\sqrt{n}}}, \quad u_n = \left(1 + \frac{1}{n}\right)^{-n\sqrt{n}}.$$

EXERCICE 7. Soient $\alpha \in \mathbb{R}$, $x > 0$, $a \geq b \geq 0$. Déterminer la nature des séries de termes général :

$$u_n = \frac{n^\alpha}{(1+x)(1+x^2)\cdots(1+x^n)} \quad \text{et} \quad u_n = (\sqrt{n^2 + an + 2} - \sqrt{n^2 + bn + 1})^n.$$

EXERCICE 8 (*Règle de Raabe et Duhamel*). Soit (u_n) une suite de réels positifs à partir d'un certain rang. On suppose que $\lim_{n \rightarrow +\infty} n\left(\frac{u_n}{u_{n+1}} - 1\right) = \lambda$ existe (finie ou non).

Montrer que, si $\lambda > 1$, alors la série $\sum u_n$ converge, et que si $\lambda < 1$, alors la série $\sum u_n$ diverge.

Appliquer au séries de terme général :

$$u_n = \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdot 6 \cdots (2n)}, \quad u_n = \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdot 6 \cdots (2n+2)},$$

$$u_n = \frac{(n!)^2 \cdot 2^{2n}}{(2n+1)!}, \quad u_n = (n!) \prod_{k=1}^n \ln\left(1 + \frac{1}{k}\right).$$

EXERCICE 9. Soient α, β deux réels tels que $0 < \beta < 1 < \alpha$. On définit une suite (u_n) par $u_0 = 0$, $u_n = \frac{1}{n^\beta}$ si n est une puissance de 2, et $u_n = \frac{1}{n^\alpha}$ sinon. Montrer que la série $\sum u_n$ converge, mais que la suite $\left(\frac{u_{n+1}}{u_n}\right)$ n'est pas bornée et n'admet pas de limite.

EXERCICE 10. Déterminer la nature des séries de termes général u_n défini par :

$$u_n = \frac{\pi}{2} - \arcsin\left(\frac{n}{n+1}\right) \quad \text{et} \quad u_n = \frac{1}{\sin \frac{1}{n}} \left[1 - \cos\left(\frac{1}{n\sqrt{\ln n}}\right)\right].$$

4.3 Séries numériques à termes quelconques

On se place de nouveau dans le cadre général de séries de terme général dans $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} (pas nécessairement dans \mathbb{R}_+).

4.3.1 Convergence absolue

a) DÉFINITION. Une série numérique $\sum_{n \geq 0} u_n$ est dite *absolument convergente* lorsque la série à termes réels positifs $\sum_{n \geq 0} |u_n|$ est convergente.

b) THÉORÈME. Soit $\sum_{n \geq 0} u_n$ une série numérique. Si elle est absolument convergente, alors elle

est convergente. De plus, on a dans ce cas $\left| \sum_{n=0}^{+\infty} u_n \right| \leq \sum_{n=0}^{+\infty} |u_n|$

PREUVE. Pour tout $p \in \mathbb{N}$, notons $S_p = \sum_{n=0}^p u_n$ et $T_p = \sum_{n=0}^p |u_n|$. Fixons $\varepsilon > 0$ quelconque. Par hypothèse, $\sum_{n \geq 0} |u_n|$ converge, donc la suite $(T_p)_{p \geq 0}$ converge. Il en résulte en particulier qu'elle est de Cauchy : il existe donc $N_0 \in \mathbb{N}$ tel que :

$$\text{pour tous } q > p \geq N_0, \text{ on a } |T_q - T_p| < \varepsilon, \text{ c'est-à-dire } \sum_{n=p+1}^q |u_n| < \varepsilon.$$

Or, par inégalité triangulaire, $|\sum_{n=p+1}^q u_n| \leq \sum_{n=p+1}^q |u_n| < \varepsilon$, c'est-à-dire $|S_q - S_p| < \varepsilon$. Ceci prouve que la suite $(S_p)_{p \geq 0}$ est une suite de Cauchy dans \mathbb{K} . Or on sait que, dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$, toute suite de Cauchy est convergente (\mathbb{R} et \mathbb{C} sont complets). Donc la suite $(S_p)_{p \geq 0}$ converge, ce qui signifie que la série $\sum_{n \geq 0} u_n$ est convergente.

Dès lors, on passe à la limite en q dans l'inégalité (triangulaire) $|\sum_{n=0}^q u_n| \leq \sum_{n=0}^q |u_n|$ pour obtenir l'inégalité voulue sur les sommes. \square

c) REMARQUE ET DÉFINITION. La réciproque est fautive !

Contre-exemple. La série harmonique alternée $\sum_{n \geq 1} u_n$ avec $u_n = \frac{(-1)^n}{n}$ est convergente (voir 4.1.3). Comme $|u_n| = \frac{1}{n}$, la série $\sum_{n \geq 1} |u_n|$ est la série harmonique, qui elle est divergente.

Donc la série harmonique alternée est convergente mais non absolument convergente.

Une série numérique qui est convergente mais qui n'est pas absolument convergente est dite parfois *semi-convergente*.

d) PROPOSITION. *Les séries absolument convergentes forment un sous-espace de l'espace vectoriel des séries convergentes.*

PREUVE. L'inclusion découle du b) ci-dessus. Pour la stabilité par combinaison linéaire, considérons $\sum u_n$ et $\sum v_n$ deux séries absolument convergentes, et deux scalaires $\lambda, \mu \in \mathbb{K}$. On a : $|\lambda u_n + \mu v_n| \leq |\lambda| |u_n| + |\mu| |v_n|$. Or par hypothèse, $\sum |u_n|$ et $\sum |v_n|$ convergent, donc $|\lambda| \sum |u_n| + |\mu| \sum |v_n|$ aussi d'après 4.1.4.b. L'inégalité $|\lambda u_n + \mu v_n| \leq |\lambda| |u_n| + |\mu| |v_n|$ implique donc, avec le théorème de majoration 4.2.1.b, que la série $\sum |\lambda u_n + \mu v_n|$ est convergente. Ceci signifie que la série $\sum (\lambda u_n + \mu v_n)$ est absolument convergente, i.e. $\lambda \sum u_n + \mu \sum v_n$ est absolument convergente. \square

Le théorème suivant est un résultat pratique donnant des conditions suffisantes de convergence pour certains types particuliers de séries dont le terme général est un réel changeant de signe suivant la parité de n . Ce résultat, appelé règle des séries alternées, et en fait un cas particulier d'un résultat plus général (qui n'est pas au programme de ce cours) appelé la règle d'Abel.

4.3.2 Séries (réelles) alternées

a) DÉFINITION. Une série réelle $\sum u_n$ est dite *alternée* lorsque son terme général est de la forme $u_n = (-1)^n \alpha_n$ avec $\alpha_n \in \mathbb{R}_+$, ou bien de la forme $u_n = (-1)^{n+1} \alpha_n$ avec $\alpha_n \in \mathbb{R}_+$.

b) THÉORÈME (*règle des séries alternées*). *Soit (α_n) une suite de réels positifs. Si (α_n) est décroissante et convergente vers 0, alors la série alternée $\sum (-1)^n \alpha_n$ est convergente.*

PREUVE. Posons $S_N = \sum_{n=0}^N (-1)^n \alpha_n$ pour tout $N \in \mathbb{N}$. L'hypothèse que la suite (α_n) est décroissante implique que, pour tout entier $p \geq 0$, on a $S_{2p+2} - S_{2p} = \alpha_{2p+2} - \alpha_{2p+1} \leq 0$

et $S_{2p+3} - S_{2p+1} = -\alpha_{2p+3} + \alpha_{2p+2} \geq 0$. Donc la suite $(S_{2p})_{p \geq 0}$ est décroissante et la suite $(S_{2p+1})_{p \geq 0}$ est croissante. De plus $S_{2p+1} - S_{2p} = -\alpha_{2p+1}$, et l'hypothèse que la suite (α_n) converge vers 0 implique alors que $\lim(S_{2p+1} - S_{2p}) = 0$. En d'autres termes, les suites $(S_{2p})_{p \geq 0}$ et $(S_{2p+1})_{p \geq 0}$ sont adjacentes. Elles convergent donc vers une même limite S .

Ainsi, la suite (S_n) est telle que les deux suites extraites $(S_{2p})_{p \geq 0}$ et $(S_{2p+1})_{p \geq 0}$ convergent vers la même limite S . On sait (voir résultats de première année, et aussi chapitre 3) qu'alors la suite (S_n) converge vers S . \square

c) EXEMPLE (*série de Riemann alternée*). Il s'agit de $\sum u_n$ où $u_n = \frac{(-1)^n}{n^\alpha}$, avec $\alpha \in \mathbb{R}$ fixé.

- Si $\alpha \leq 0$, elle diverge grossièrement.
- Si $\alpha > 1$, elle est absolument convergente car $\sum |u_n| = \sum \frac{1}{n^\alpha}$ converge (cf. 4.2.2.b).
- Dans le cas $0 < \alpha \leq 1$, $\sum |u_n|$ diverge (cf. 4.2.2.b) donc $\sum u_n$ n'est pas absolument convergente. Néanmoins, $\sum u_n$ converge par application du théorème b) ci-dessus (car la suite $\frac{1}{n^\alpha}$ décroît vers 0.) Dans ce cas, $\sum u_n$ est semi-convergente.

La *série harmonique alternée* $\sum \frac{(-1)^n}{n}$, qui est semi-convergente comme on l'a vu au c) de 4.3.1, correspond au cas $\alpha = 1$.

4.3.3 Exercices

EXERCICE 1. Soit $\theta \in \mathbb{R}$. Pour tout $n \in \mathbb{N}^*$, on pose $u_n = \frac{\sqrt{n} \ln n}{n^2 + 1} \sin n\theta$. Déterminer la nature de la série $\sum |u_n|$, puis de la série $\sum u_n$.

EXERCICE 2. Soit a un réel positif fixé. On pose $u_n = \sin(\pi\sqrt{n^2 + a^2})$ et $v_n = \frac{a^2 \pi}{\sqrt{n^2 + a^2} + n}$. Montrer que $u_n = (-1)^n \sin v_n$. En déduire la nature de la série $\sum u_n$.

EXERCICE 3. Déterminer la nature de la série $\sum (-1)^n n^{\frac{1}{n}} \sin \frac{1}{n}$.

EXERCICE 4. Déterminer la nature de la série $\sum \frac{1 + (-1)^n \sqrt{n}}{n}$.

EXERCICE 5. Soit $\alpha \in \mathbb{R}_+^*$. On pose $u_n = \frac{(-1)^n}{(-1)^n + n^\alpha}$.

Montrer que $u_n = \frac{(-1)^n}{n^\alpha} + \frac{-1}{n^\alpha(n^\alpha + (-1)^n)}$. En déduire la nature de la série $\sum u_n$.

EXERCICE 6. En utilisant des développements limités appropriés, étudier la nature de la série $\sum u_n$ dans chacun des cas suivants :

$$u_n = \ln n \ln \left(1 + \frac{(-1)^n}{n} \right), \quad u_n = (n^3 + 1)^{\frac{1}{3}} - (n^2 + 1)^{\frac{1}{2}}, \quad u_n = (-1)^n \left[\left(1 + \frac{1}{n} \right)^{-n} - \frac{1}{e} \right].$$

$$u_n = \sqrt{n^4 + 2n + 1} - \sqrt{n^4 + an} \quad \text{avec } a \in \mathbb{R}_+ \quad \text{et} \quad u_n = e^{-\frac{c}{n}} - a - \frac{b}{n} \quad \text{avec } a, b, c \in \mathbb{R}.$$

4.4 Résultats complémentaires

4.4.1 Produit de deux séries

a) DÉFINITION. Soient $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$ deux séries numériques. On appelle *série produit* (ou *produit de Cauchy*) de ces deux séries la série $\sum_{n \geq 0} w_n$ de terme général :

$$w_n = u_0 v_n + u_1 v_{n-1} + \cdots + u_n v_0 = \sum_{p=0}^n u_p v_{n-p} = \sum_{p+q=n} u_p v_q.$$

b) LEMME. Soient $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$ deux séries réelles à termes positifs. Si les deux convergent, alors leur série produit $\sum_{n \geq 0} w_n$ converge, et sa somme vaut :

$$\sum_{n=0}^{+\infty} w_n = \left(\sum_{n=0}^{+\infty} u_n \right) \left(\sum_{n=0}^{+\infty} v_n \right).$$

PREUVE. Notons dans \mathbb{R}_+ les sommes partielles : $U_n = \sum_{k=0}^n u_k$, $V_n = \sum_{k=0}^n v_k$, $W_n = \sum_{k=0}^n w_k$. Par hypothèse, la suite (U_n) , (qui est croissante puisque chaque u_k est positif), converge vers une limite $U = \sum_{n=0}^{+\infty} u_n \in \mathbb{R}_+$. De même la suite (V_n) converge en croissant vers $V = \sum_{n=0}^{+\infty} v_n \in \mathbb{R}_+$. Il en résulte que la suite $(U_n V_n)$ converge en croissant vers UV .

Remarquons que $U_n V_n = \sum_{0 \leq p, q \leq n} u_p v_q$. Par ailleurs, $W_n = \sum_{k=0}^n \left(\sum_{p+q=k} u_p v_q \right) = \sum_{p+q \leq n} u_p v_q$.

D'une part : $(p+q \leq n) \implies (p \leq n \text{ et } q \leq n)$, donc chaque terme de la somme W_n figure dans la somme $U_n V_n$; comme tous les termes sont positifs, on conclut que $W_n \leq U_n V_n$.

D'autre part : $(p \leq n \text{ et } q \leq n) \implies (p+q \leq 2n)$, donc de la même façon : $U_n V_n \leq W_{2n}$.

En résumé : $W_n \leq U_n V_n \leq W_{2n}$ (*)

Puisque la suite $(U_n V_n)$ est majorée par UV , il résulte de (*) que la suite (W_n) est majorée. Mais la suite (W_n) est aussi croissante (car chaque w_k est positif). On conclut que (W_n) converge. Soit W sa limite dans \mathbb{R}_+ . La suite (W_{2n}) , qui est une suite extraite de (W_n) converge donc aussi vers W . On déduit alors de (*) que $W = UV$. \square

c) THÉORÈME. Soient $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$ deux séries numériques. Si les deux sont absolument convergentes, alors leur série produit $\sum_{n \geq 0} w_n$ est absolument convergente, et sa somme vaut :

$$\sum_{n=0}^{+\infty} w_n = \left(\sum_{n=0}^{+\infty} u_n \right) \left(\sum_{n=0}^{+\infty} v_n \right).$$

PREUVE. On raisonne en deux étapes.

• Notons $\sum_{n \geq 0} w_n$ la série produit de $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$, définie par $w_n = \sum_{p+q=n} u_p v_q$. Notons $\sum_{n \geq 0} x_n$ la série produit de $\sum_{n \geq 0} |u_n|$ et $\sum_{n \geq 0} |v_n|$, définie par $x_n = \sum_{p+q=n} |u_p v_q|$.

Par l'inégalité triangulaire : $|w_n| = \left| \sum_{p+q=n} u_p v_q \right| \leq \sum_{p+q=n} |u_p| |v_q| = x_n$, pour tout $n \in \mathbb{N}$.

Or, comme par hypothèse les séries positives $\sum_{n \geq 0} |u_n|$ et $\sum_{n \geq 0} |v_n|$ convergent, le lemme précédent implique que la série $\sum_{n \geq 0} x_n$ converge. On peut donc appliquer le critère de majoration 4.2.1.b pour conclure que la série $\sum_{n \geq 0} |w_n|$ converge, c'est-à-dire que la série $\sum_{n \geq 0} w_n$ est absolument convergente.

• On introduit pour les sommes partielles les notations suivantes :

$$U_n = \sum_{k=0}^n u_k, \quad V_n = \sum_{k=0}^n v_k, \quad W_n = \sum_{k=0}^n w_k = \sum_{k=0}^n \left(\sum_{p+q=k} u_p v_q \right) = \sum_{p+q \leq n} u_p v_q.$$

$$A_n = \sum_{k=0}^n |u_k|, \quad B_n = \sum_{k=0}^n |v_k|, \quad C_n = \sum_{k=0}^n x_k = \sum_{p+q \leq n} |u_p| |v_q|.$$

En particulier le lemme b) se traduit par : $\lim(A_n B_n - C_n) = 0$. On calcule :

$$U_n V_n - W_n = \sum_{0 \leq p, q \leq n} u_p v_q - \sum_{p+q \leq n} u_p v_q = \sum_{(p,q) \in \Delta_n} u_p v_q$$

où l'on a noté $\Delta_n = \{(p, q) \in \mathbb{N}^2; 0 \leq p \leq n, 0 \leq q \leq n, p + q > n\}$. De la même façon :

$$A_n B_n - C_n = \sum_{(p,q) \in \Delta_n} |u_p| |v_q|.$$

Par inégalité triangulaire : $|U_n V_n - W_n| = \left| \sum_{(p,q) \in \Delta_n} u_p v_q \right| \leq \sum_{(p,q) \in \Delta_n} |u_p| |v_q| = A_n B_n - C_n$.

Or $\lim(A_n B_n - C_n) = 0$, donc $\lim(U_n V_n - W_n) = 0$, donc $\lim W_n = (\lim U_n)(\lim V_n)$,
c'est-à-dire : $\sum_{n=0}^{+\infty} w_n = \left(\sum_{n=0}^{+\infty} u_n \right) \left(\sum_{n=0}^{+\infty} v_n \right)$. \square

d) EXEMPLE IMPORTANT (exponentielle). Pour tout $x \in \mathbb{C}$, la série $\sum_{n \geq 0} \frac{x^n}{n!}$ est absolument convergente (c'est clair par la règle de d'Alembert). Donc elle converge ; notons sa somme :

$$\exp x = \sum_{n=0}^{+\infty} \frac{x^n}{n!}.$$

Pour deux nombres complexes x et y , la série produit de $\sum_{n \geq 0} \frac{x^n}{n!}$ et $\sum_{n \geq 0} \frac{y^n}{n!}$ a pour terme général :

$$w_n = \sum_{p=0}^n \frac{x^p}{p!} \frac{y^{n-p}}{(n-p)!} = \frac{1}{n!} \sum_{p=0}^n \binom{n}{p} x^p y^{n-p} = \frac{1}{n!} (x + y)^n.$$

D'après le théorème précédent, cette série produit est absolument convergente, et on a :

$$\exp(x + y) = \sum_{n=0}^{+\infty} \frac{(x+y)^n}{n!} = \left(\sum_{n=0}^{+\infty} \frac{x^n}{n!} \right) \left(\sum_{n=0}^{+\infty} \frac{y^n}{n!} \right) = (\exp x)(\exp y)$$

e) REMARQUE. On peut améliorer le théorème c) en montrant que la série produit reste convergente si l'on suppose seulement que l'une des deux séries données est absolument convergente, l'autre étant simplement supposée convergente (théorème de Mertens).

4.4.2 Le problème de la sommation par paquets

a) EXEMPLE INTRODUCTIF. La série $\sum_{n \geq 0} (-1)^n$ est divergente (grossièrement). Mais en regroupant les termes deux par deux :

$$\underbrace{1 + (-1)}_0 + \underbrace{1 + (-1)}_0 + \underbrace{1 + (-1)}_0 + \dots + \underbrace{(-1)^{2p} + (-1)^{2p+1}}_0 + \dots$$

la série $\sum_{p \geq 0} ((-1)^{2p} + (-1)^{2p+1})$ est de terme général nul, donc convergente.

b) NOTATIONS. Soit $\sum_{n \geq 0} u_n$ une série numérique. Regroupons ses termes par paquets sous la forme : $v_0 = u_0 + u_1 + \dots + u_{\phi(0)}$, $v_1 = u_{\phi(0)+1} + u_{\phi(0)+2} + \dots + u_{\phi(1)}$, et plus généralement :

$$v_n = \sum_{k=\phi(n-1)+1}^{\phi(n)} u_k, \quad \text{avec } \phi : \mathbb{N} \rightarrow \mathbb{N} \text{ strictement croissante.} \quad \text{On a alors :}$$

c) PROPOSITION. Avec les données et notations ci-dessus :

- (i) Si $\sum u_n$ converge, alors $\sum v_n$ converge ; dans ce cas, les deux séries ont la même somme.
- (ii) Si de plus tous les u_n sont des réels positifs, la réciproque est vraie aussi.

PREUVE. Considérons les sommes partielles :

$$U_n = \sum_{j=0}^n u_j \quad \text{et} \quad V_n = \sum_{j=0}^n v_j = \sum_{j=0}^n \left(\sum_{k=\phi(j-1)+1}^{\phi(j)} u_k \right) = \sum_{j=0}^{\phi(n)} u_j = U_{\phi(n)}.$$

Si l'on suppose que la série $\sum_{n \geq 0} u_n$ converge, en notant U sa somme, on a $\lim U_n = U$. Comme ϕ est strictement croissante, $\lim \phi(n) = +\infty$, donc $\lim V_n = \lim U_{\phi(n)} = \lim U_n = U$, ce qui prouve le point (i).

Pour prouver (ii), supposons que chaque u_n est un réel positif. Il est clair que les suites réelles (U_n) et (V_n) sont alors croissantes. Si on suppose que la série $\sum_{n \geq 0} v_n$ converge, alors la suite (V_n) est majorée (car toute suite réelle croissante et non majorée tend vers $+\infty$). Mais pour tout $n \in \mathbb{N}$, on a : $\phi(n) \geq n$ (c'est une conséquence facile du fait que ϕ est strictement croissante) donc : $U_n = \sum_{k=0}^n u_k \leq \sum_{k=0}^{\phi(n)} u_k = U_{\phi(n)} = V_n$. On déduit que la suite (U_n) est majorée. Comme elle est croissante, elle converge. \square

4.4.3 Le problème de l'ordre des termes

a) EXEMPLE INTRODUCTIF. Considérons la série réelle $\sum_{n \geq 1} u_n$, pour $u_n = \frac{(-1)^{n+1}}{\sqrt{n}}$. D'après 4.3.2.c, elle est convergente (mais non absolument convergente)

Changeons l'ordre des termes en renumérotant :

$$\sum_{n \geq 1} u_{\sigma(n)} = 1 + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{5}} + \frac{1}{\sqrt{7}} - \frac{1}{\sqrt{4}} + \frac{1}{\sqrt{9}} + \frac{1}{\sqrt{11}} - \frac{1}{\sqrt{6}} + \dots$$

où σ est la bijection de \mathbb{N}^* sur \mathbb{N}^* définie par :

$$\sigma(3p) = 2p, \quad \sigma(3p-1) = 4p-1, \quad \sigma(3p-2) = 4p-3 \quad \text{pour tout } p \geq 1.$$

Dans $\sum_{n \geq 1} u_{\sigma(n)}$, regroupons ensuite par paquets en : $\sum_{p \geq 1} (u_{\sigma(3p-2)} + u_{\sigma(3p-1)} + u_{\sigma(3p)})$.

Posons $a_n = u_{\sigma(3n-2)} + u_{\sigma(3n-1)} + u_{\sigma(3n)} = \frac{1}{\sqrt{4n-3}} + \frac{1}{\sqrt{4n-1}} - \frac{1}{\sqrt{2n}}$, qui est positif.

On vérifie que a_n est équivalente à $\left(\frac{\sqrt{2}-1}{\sqrt{2}}\right) \frac{1}{\sqrt{n}}$ au voisinage de $+\infty$.

$$\begin{aligned} \text{En effet : } a_n &= \frac{1}{\sqrt{2n}} \left(\frac{1}{\sqrt{2}} \left(1 - \frac{3}{4n}\right)^{-1/2} + \frac{1}{\sqrt{2}} \left(1 - \frac{1}{4n}\right)^{-1/2} - 1 \right) \\ &= \frac{1}{\sqrt{2n}} \left(\frac{1}{\sqrt{2}} \left(1 + \frac{3}{8n} + \frac{1}{n} \varepsilon\left(\frac{1}{n}\right)\right) + \frac{1}{\sqrt{2}} \left(1 + \frac{1}{8n} + \frac{1}{n} \varepsilon\left(\frac{1}{n}\right)\right) - 1 \right) = \frac{1}{\sqrt{2n}} \left(\frac{2-\sqrt{2}}{\sqrt{2}} \left(1 + \varepsilon'\left(\frac{1}{n}\right)\right) \right). \end{aligned}$$

D'après 4.2.2, la série $\sum \frac{1}{\sqrt{2n}}$ diverge, donc d'après 4.2.1.d, il est de même de $\sum a_n$. En appliquant 4.4.3.c.(i), on conclut que $\sum_{n \geq 1} u_{\sigma(n)}$ diverge.

b) DÉFINITION. Une série numérique $\sum_{n \geq 0} u_n$ est dite *commutativement convergente* lorsque, pour tout bijection σ de \mathbb{N} sur \mathbb{N} , la série $\sum_{n \geq 0} u_{\sigma(n)}$ est convergente.

c) THÉORÈME. Toute série numérique absolument convergente est commutativement convergente, et sa somme reste inchangée quand on modifie l'ordre des termes.

PREUVE. Soit $\sum_{n \geq 0} u_n$ absolument convergente, et σ bijection de \mathbb{N} sur \mathbb{N} .

• *Première étape* : on montre que $\sum_{n \geq 0} u_{\sigma(n)}$ est absolument convergente. Pour cela, pour tout $n \in \mathbb{N}$, introduisons $\phi(n) = \max\{\sigma(0), \sigma(1), \dots, \sigma(n)\}$. Comme $\{\sigma(0), \sigma(1), \dots, \sigma(n)\}$ est formé de $n + 1$ entiers ≥ 0 qui sont deux à deux distincts, on a :

$$\phi(n) \geq n. \quad (1)$$

De plus, $\phi(n)$ majore $\{\sigma(0), \sigma(1), \dots, \sigma(n)\}$, donc :

$$\{\sigma(0), \sigma(1), \dots, \sigma(n)\} \subset \{0, 1, 2, \dots, \phi(n)\}. \quad (2)$$

Notons $X_n = \sum_{k=0}^n |u_{\sigma(k)}|$. On tire de (2) que $X_n = |u_{\sigma(0)}| + |u_{\sigma(1)}| + \dots + |u_{\sigma(n)}| \leq \sum_{k=0}^{\phi(n)} |u_k|$.

Comme par hypothèse, la série $\sum_{k \geq 0} |u_k|$ converge, on en déduit que : $X_n \leq \sum_{k=0}^{+\infty} |u_k|$. La suite (X_n) , qui est croissante, est donc majorée ; ceci prouve que la suite (X_n) converge, c'est-à-dire que la série $\sum u_{\sigma(n)}$ est absolument convergente. On a en outre montré que : $\sum_{n=0}^{+\infty} |u_{\sigma(n)}| \leq \sum_{n=0}^{+\infty} |u_n|$. En échangeant le rôle des deux séries, (c'est-à-dire en remplaçant σ par la bijection réciproque σ^{-1}), on conclut que :

$$\sum_{n=0}^{+\infty} |u_{\sigma(n)}| = \sum_{n=0}^{+\infty} |u_n| \quad (3).$$

• *Seconde étape* : les séries $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} u_{\sigma(n)}$ étant convergentes (car absolument convergentes), il reste à montrer que leurs sommes respectives S et S' sont égales.

Pour cela, posons $L_n = \{0, 1, 2, \dots, \phi(n)\} \setminus \{\sigma(0), \sigma(1), \dots, \sigma(n)\}$, de sorte que :

$$\left| \sum_{k=0}^{\phi(n)} u_k - \sum_{k=0}^n u_{\sigma(k)} \right| = \left| \sum_{k \in L_n} u_k \right| \leq \sum_{k \in L_n} |u_k|.$$

En remodifiant le second membre, on obtient finalement :

$$\left| \sum_{k=0}^{\phi(n)} u_k - \sum_{k=0}^n u_{\sigma(k)} \right| \leq \sum_{k=0}^{\phi(n)} |u_k| - \sum_{k=0}^n |u_{\sigma(k)}|. \quad (4)$$

D'après (1), $\lim \phi(n) = +\infty$, donc (3) implique que le second membre de (4) tend vers 0 quand $n \rightarrow +\infty$. Comme le premier membre de (4) tend vers $|S' - S|$, on conclut que $S' = S$. \square

d) REMARQUES. Soit $\sum u_n$ une série convergente mais non absolument convergente. On a vu à l'exemple a) que l'on peut avoir $\sum u_{\sigma(n)}$ qui est divergente. Il se peut aussi que $\sum u_{\sigma(n)}$ converge, mais que sa somme soit différente de celle de $\sum u_n$. (cf. exercice ci-dessous).

On peut montrer (on ne le fera pas ici) que, la réciproque du théorème c) est en fait vraie aussi, c'est-à-dire qu'une série numérique semi-convergente n'est jamais commutativement convergente.

EXERCICE. Considérons la série harmonique alternée $\sum_{n \geq 1} u_n$ avec $u_n = \frac{(-1)^{n+1}}{n}$. Elle est convergente et non absolument convergente (d'après 4.3.1.c ou 4.3.2.c). Vérifier qu'en posant :

$$\sigma(3q) = 4q, \quad \sigma(3q - 1) = 2(2q - 1), \quad \sigma(3q - 2) = 2q - 1 \quad \text{pour tout } q \geq 1,$$

on définit une bijection σ de \mathbb{N} sur \mathbb{N} . Montrer qu'alors la série :

$$\sum_{n \geq 1} u_{\sigma(n)} = 1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \frac{1}{12} + \dots + \frac{1}{2q-1} - \frac{1}{2(2q-1)} - \frac{1}{4q} + \dots$$

converge, mais que : $\sum_{n=1}^{+\infty} u_n = 2 \sum_{n=1}^{+\infty} u_{\sigma(n)}$.

(Indication : faire des sommations par paquets de 3, puis de 2.)

4.4.4 Comparaison entre séries et intégrales

a) PROPOSITION. Soit $\sum_{n \geq 0} a_n$ une série de terme général $a_n \in \mathbb{K}$. Soit $f : [0, +\infty[\rightarrow \mathbb{K}$ définie par $f(t) = a_n$ pour tous $n \in \mathbb{N}$ et $t \in [n, n+1[$. Alors :

(i) La série $\sum_{n \geq 0} a_n$ et l'intégrale $\int_0^{+\infty} f(t) dt$ sont de même nature.

(ii) De plus, dans le cas de convergence, on a : $\sum_{n=0}^{+\infty} a_n = \int_0^{+\infty} f(t) dt$.

Preuve. f est bien définie car tout $t \geq 0$ appartient à un intervalle $[n, n+1[$ et un seul. Par construction, f est constante sur chaque $[n, n+1[$, donc continue par morceaux sur tout segment $[a, b]$ inclus dans $I = [0, +\infty[$. Par définition, cela signifie que $f \in \mathcal{CM}(I, \mathbb{K})$.

Pour tout $N \in \mathbb{N}$, notons :

$$S_N = \sum_{n=0}^N a_n = \int_0^{N+1} f(t) dt.$$



• Supposons que $\int_0^{+\infty} f(t) dt$ converge, c'est-à-dire $\lim_{x \rightarrow +\infty} \int_0^x f(t) dt = L \in \mathbb{K}$. Alors on a :

$$\lim_{N \rightarrow +\infty} S_N = \lim_{N \rightarrow +\infty} \int_0^{N+1} f(t) dt = L \text{ ce qui prouve que la série } \sum_{n \geq 0} a_n \text{ converge.}$$

• Réciproquement, supposons que $\sum_{n \geq 0} a_n$ converge, c'est-à-dire $\lim_{N \rightarrow +\infty} S_N = L \in \mathbb{K}$. Pour tout réel $x \geq 0$, notons $E(x)$ sa partie entière ; on a :

$$\int_0^x f(t) dt = \int_0^{E(x)} f(t) dt + \int_{E(x)}^x f(t) dt = \sum_{n=0}^{E(x)-1} \int_n^{n+1} f(t) dt + \int_{E(x)}^x f(t) dt$$

$$= \sum_{n=0}^{E(x)-1} a_n + (x - E(x))a_{E(x)} = S_{E(x)-1} + (x - E(x))a_{E(x)}$$

Or $\lim_{x \rightarrow +\infty} E(x) = +\infty$. Donc $\lim_{x \rightarrow +\infty} S_{E(x)-1} = L$. On a aussi : $\lim_{x \rightarrow +\infty} a_{E(x)} = \lim_n a_n = 0$ (puisque a_n est le terme général d'une série convergente). Comme $0 \leq x - E(x) < 1$, on déduit $\lim_{x \rightarrow +\infty} (x - E(x))a_{E(x)} = 0$, d'où $\lim_{x \rightarrow +\infty} \int_0^x f(t) dt = L$. \square

b) THÉORÈME (Cas des fonctions réelles positives décroissantes) Notons $I = [0, +\infty[$. Soit $f : I \rightarrow \mathbb{R}$ continue par morceaux sur I , que l'on suppose positive et décroissante sur I . Alors :

la série $\sum_{n \geq 0} f(n)$ et l'intégrale $\int_0^{+\infty} f(t) dt$ sont de même nature.

Preuve. On raisonne en deux étapes. Définissons d'abord pour tout $n \in \mathbb{N}$:

$$x_n = \sum_{k=0}^n f(k) - \int_0^{n+1} f(t) dt.$$

• On va montrer que la suite $(x_n)_{n \geq 0}$ est convergente. Pour cela remarquons d'abord que : pour tous $k \in \mathbb{N}$ et $t \in [k, k+1]$, on a (car f est décroissante) :

$$0 \leq f(k+1) \leq f(t) \leq f(k)$$

donc :

$$0 \leq f(k+1) \leq \int_k^{k+1} f(t) dt \leq f(k). \quad (*)$$

$$\begin{aligned} \text{Or } x_n &= \sum_{k=0}^n f(k) - \sum_{k=0}^n \left(\int_k^{k+1} f(t) dt \right) \\ &= \sum_{k=0}^n \left(f(k) - \int_k^{k+1} f(t) dt \right) \end{aligned}$$

Chaque différence $f(k) - \int_k^{k+1} f(t) dt \geq 0$ étant positive d'après (*), on a $x_{n+1} \geq x_n$ pour tout $n \in \mathbb{N}$. Donc la suite (x_n) est croissante.

Par ailleurs :

$$x_n = \sum_{k=0}^n \left(f(k) - \int_k^{k+1} f(t) dt \right) = f(0) + \sum_{k=0}^{n-1} \left(- \int_k^{k+1} f(t) dt + f(k+1) \right) - \int_n^{n+1} f(t) dt.$$

Chaque différence $\left(- \int_k^{k+1} f(t) dt + f(k+1) \right)$ étant négative d'après (*), il vient :

$x_n \leq f(0) - \int_n^{n+1} f(t) dt$. Toujours d'après (*), on a aussi : $-\int_n^{n+1} f(t) dt \leq -f(n+1)$. Finalement : $x_n \leq f(0) - f(n+1) \leq f(0)$. Ceci étant vrai pour tout $n \in \mathbb{N}$, la suite (x_n) est majorée ; on a déjà vu qu'elle est croissante : on conclut qu'elle converge.

Bilan : en posant $S_n = \sum_{k=0}^n f(k)$, on a $S_n = x_n + \int_0^{n+1} f(t) dt$ avec la suite (x_n) qui converge.

Donc la suite (S_n) converge si et seulement la suite $(\int_0^{n+1} f(t) dt)_{n \geq 0}$ converge.

• Par ailleurs, pour tout réel $x > 0$, on a (parce que f est positive) :

$$\int_0^{E(x)} f(t) dt \leq \int_0^x f(t) dt \leq \int_0^{E(x)+1} f(t) dt.$$

Puisque $\lim_{x \rightarrow +\infty} E(x) = +\infty$, on en déduit que : la suite $(\int_0^n f(t) dt)_{n \geq 0}$ converge dans \mathbb{R} vers une limite finie si et seulement si $\lim_{x \rightarrow +\infty} \int_0^x f(t) dt$ existe dans \mathbb{R} , c'est-à-dire si et seulement si l'intégrale $\int_0^{+\infty} f(t) dt$ est convergente. En rappelant le bilan ci-dessus, on conclut que l'intégrale $\int_0^{+\infty} f(t) dt$ est convergente si et seulement si la suite (S_n) converge dans \mathbb{R} , ce qui équivaut à dire que la série $\sum_{k \geq 0} f(k)$ est convergente. \square

c) COROLLAIRE. Soient un réel $a \geq 0$ et $I = [a, +\infty[$. Si $f : I \rightarrow \mathbb{R}$ est continue par morceaux, positive et décroissante sur I , alors la série $\sum_{n \geq 0} f(a+n)$ et l'intégrale $\int_a^{+\infty} f(t) dt$ sont de même nature.

Preuve. Il suffit d'appliquer le théorème à $f(a+t)$ au lieu de $f(t)$. \square

d) REMARQUE. En particulier, pour tout $\alpha \geq 0$, l'application $t \rightarrow \frac{1}{t^\alpha}$ étant positive, continue par morceaux et décroissante sur $I = [a, +\infty[$ (avec $a > 0$ quelconque), on déduit que l'intégrale $\int_a^{+\infty} \frac{1}{t^\alpha} dt$ converge si et seulement si la série $\sum_{n \geq 0} \frac{1}{(n+a)^\alpha}$ est convergente, c'est-à-dire si et

seulement si $\alpha > 1$. On retrouve ainsi via la comparaison à ce type d'intégrale les conditions de convergence des séries de Riemann de 4.2.2.

e) UNE APPLICATION INTÉRESSANTE. La suite $(\sum_{k=1}^n \frac{1}{k} - \ln n)$ est convergente; sa limite γ est appelée la constante d'Euler.

En effet : soit $I = [1, +\infty[$ et $f : I \rightarrow \mathbb{R}$ continue positive décroissante sur I définie par $f(t) = \frac{1}{t}$. Comme dans la preuve du théorème b), considérons $x_n = \sum_{k=0}^n g(k) - \int_0^{n+1} g(t) dt$, où l'on a noté g l'application $[0, +\infty[\rightarrow \mathbb{R}_+$ définie par $g(t) = f(t+1)$. On a donc :

$$x_n = \sum_{k=0}^n \frac{1}{k+1} - \int_0^{n+1} \frac{1}{t+1} dt = \sum_{k=1}^{n+1} \frac{1}{k} - \int_1^{n+2} \frac{1}{t} dt,$$

Où encore :

$$x_{n-1} = \sum_{k=1}^n \frac{1}{k} - \int_1^{n+1} \frac{1}{t} dt = \sum_{k=1}^n \frac{1}{k} - \int_1^n \frac{1}{t} dt - \int_n^{n+1} \frac{1}{t} dt = \sum_{k=1}^n \frac{1}{k} - \ln n - (\ln(n+1) - \ln n).$$

Comme $\lim_n (\ln(n+1) - \ln n) = \lim_n \ln(1 + \frac{1}{n}) = 0$, et comme on a vu dans la preuve du théorème

b) que la suite (x_n) converge, on conclut que la suite $(\sum_{k=1}^n \frac{1}{k} - \ln n)$ est convergente. \square

La constante d'Euler :

$$\gamma = \lim_{n \rightarrow +\infty} (1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots + \frac{1}{n} - \ln n) = \lim_{n \rightarrow +\infty} (\sum_{k=1}^n \frac{1}{k} - \ln n)$$

joue un rôle très important dans de nombreux problèmes mathématiques. Une valeur approchée de γ est 0,57722.... La question de savoir si γ est ou non rationnel est encore ouverte aujourd'hui !

f) EXERCICE (séries de Bertrand). On fixe $\alpha, \beta \in \mathbb{R}$. On considère la série de terme général :

$$u_n = \frac{1}{n^\alpha (\ln n)^\beta}.$$

Montrer que :

- (i) si $\alpha > 1$, la série $\sum \frac{1}{n^\alpha (\ln n)^\beta}$ converge pour tout β ,
- (ii) si $\alpha < 1$, la série $\sum \frac{1}{n^\alpha (\ln n)^\beta}$ diverge pour tout β ,
- (iii) si $\alpha = 1$, la série $\sum \frac{1}{n (\ln n)^\beta}$ converge si et seulement si $\beta > 1$.

Chapitre 5

Suites et séries de fonctions

Conformément à ce que prévoit le programme, ce chapitre est plus précisément consacré à l'étude des suites et séries de fonctions *d'une variable réelle, à valeurs réelles ou complexes*.

On observera néanmoins (voir la troisième remarque de 5.1.2.c pour plus de détails) que les définitions et résultats des sections 5.1.1 et 5.1.2 restent vrais dans le contexte plus général de fonctions d'une partie d'un evn à valeurs dans un evn.

Dans tout le chapitre, \mathbb{K} désigne \mathbb{R} ou \mathbb{C} , et $|\cdot|$ désigne respectivement la valeur absolue ou le module.

5.1 Convergence simple et uniforme d'une suite de fonctions

5.1.1 Notions de convergence d'une suite de fonctions

Fixons d'abord les notations. On considère X une partie non vide de \mathbb{R} . Soit $(f_n)_{n \geq 0}$ une suite d'applications $X \rightarrow \mathbb{K}$. Donc :

pour tout $n \in \mathbb{N}$, $f_n \in \mathcal{F}(X, \mathbb{K})$ et pour tous $n \in \mathbb{N}$ et $x \in X$, $f_n(x) \in \mathbb{K}$.

a) DÉFINITION (convergence simple). On dit que la suite (f_n) converge *simplement* sur X vers une application $f \in \mathcal{F}(X, \mathbb{K})$ lorsque, pour tout $x \in X$, la suite $(f_n(x))_{n \geq 0}$ converge dans \mathbb{K} vers $f(x)$. Ceci équivaut à :

$$\forall x \in X, \lim_{n \rightarrow +\infty} f_n(x) = f(x),$$

c'est-à-dire à :

$$\forall x \in X, \forall \varepsilon > 0, \exists N_{\varepsilon, x} \in \mathbb{N}, \forall n \geq N_{\varepsilon, x}, |f_n(x) - f(x)| < \varepsilon.$$

b) DÉFINITION (convergence uniforme). On dit que la suite (f_n) converge uniformément sur X vers une application $f \in \mathcal{F}(X, \mathbb{K})$ lorsque :

$$\forall \varepsilon > 0, \exists N_\varepsilon \in \mathbb{N}, \forall n \geq N_\varepsilon, \forall x \in X, |f_n(x) - f(x)| < \varepsilon,$$

(on observera avec soin la place des quantificateurs en comparaison de la définition précédente).

Ceci équivaut encore à :

$$\left\{ \begin{array}{l} \text{les applications } (f_n - f) \text{ sont bornées sur } X \text{ pour } n \text{ assez grand,} \\ \text{et } \lim_{n \rightarrow +\infty} (\|f_n - f\|_\infty) = 0, \end{array} \right.$$

de sorte que la convergence uniforme sur X de la suite (f_n) s'interprète comme la convergence de la suite (f_n) dans l'evn $\mathcal{B}(X, \mathbb{K})$ muni de la norme $\|\cdot\|_\infty$ [voir exemple d) de 3.1.1].

c) EXEMPLES. • Soit $I =]0, +\infty[$; pour tout $n \in \mathbb{N}^*$ et tout $x \in I$, on pose : $f_n(x) = \frac{1}{nx}$. Il est clair que la suite (f_n) converge simplement vers la fonction nulle sur I , mais ne converge pas uniformément sur I car : $\exists \varepsilon = \frac{1}{2}, \forall N \in \mathbb{N}^*, \exists n = N, \exists x = \frac{1}{n}, |f_n(x) - 0| = f_n(\frac{1}{n}) = 1 > \varepsilon$.

• Soit $I = [0, 1]$; pour tout $n \in \mathbb{N}^*$ et tout $x \in I$, on pose (représenter graphiquement f et g) :

$$f_n(x) = \begin{cases} x & \text{si } x \in [0, \frac{1}{2n}] \\ -x + \frac{1}{n} & \text{si } x \in [\frac{1}{2n}, \frac{1}{n}] \\ 0 & \text{si } x \in [\frac{1}{n}, 1] \end{cases} \quad g_n(x) = \begin{cases} 2nx & \text{si } x \in [0, \frac{1}{2n}] \\ 2 - 2nx & \text{si } x \in [\frac{1}{2n}, \frac{1}{n}] \\ 0 & \text{si } x \in [\frac{1}{n}, 1] \end{cases}$$

Les suites (f_n) et (g_n) convergent simplement vers la fonction nulle sur I .

En effet : pour tout $x \in]0, 1]$, on a pour n assez grand $x > \frac{1}{n}$ donc $f_n(x) = g_n(x) = 0$. De plus $f_n(0) = g_n(0) = 0$. Donc $\lim_{n \rightarrow +\infty} f_n(x) = \lim_{n \rightarrow +\infty} g_n(x) = 0$ pour tout $x \in I$.

La suite (f_n) converge uniformément sur I .

En effet : $\|f_n - 0\|_\infty = \sup_{x \in I} |f_n(x)| = \frac{1}{2n}$ qui tend vers 0 quand $n \rightarrow \infty$.

La suite (g_n) ne converge pas uniformément sur I ,

En effet : $\|g_n - 0\|_\infty = \sup_{x \in I} |g_n(x)| = 1$ qui ne tend pas vers 0 quand $n \rightarrow \infty$.

d) PROPOSITION (*relation entre convergences simple et uniforme*). Soit (f_n) une suite d'applications de $\mathcal{F}(X, \mathbb{K})$. Si elle converge uniformément sur X vers une application $f \in \mathcal{F}(X, \mathbb{K})$, alors elle converge simplement sur X vers f . La réciproque est fautive en général.

Preuve. L'implication est claire d'après les définitions a) et b) elles-mêmes. On a vu en c) des contre-exemples à la réciproque. \square

e) PROPOSITION (*condition de Cauchy pour la convergence uniforme*). Une suite (f_n) d'applications de $\mathcal{F}(X, \mathbb{K})$ converge uniformément sur X si et seulement si elle vérifie la condition de Cauchy uniforme sur X :

$$\forall \varepsilon > 0, \exists N_\varepsilon \in \mathbb{N}, \forall p, q \geq N_\varepsilon, \sup_{x \in X} |f_p(x) - f_q(x)| < \varepsilon$$

Preuve. Supposons que (f_n) converge uniformément sur X vers une fonction $f \in \mathcal{F}(X, \mathbb{K})$. Fixons $\varepsilon > 0$; il existe $N \in \mathbb{N}$ tel que : $\forall n \geq N, \forall x \in X, |f_n(x) - f(x)| < \frac{\varepsilon}{2}$. Si $p, q \geq N$, alors pour tout $x \in X$, on a : $|f_p(x) - f_q(x)| \leq |f_p(x) - f(x)| + |f(x) - f_q(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. Donc (f_n) vérifie la condition de Cauchy uniforme sur X .

• Supposons réciproquement que (f_n) vérifie la condition de Cauchy uniforme sur X . Fixons $\varepsilon > 0$. Il existe $N_\varepsilon \in \mathbb{N}$ tel que $\forall p, q \geq N_\varepsilon, \sup_{x \in X} |f_p(x) - f_q(x)| < \varepsilon$.

Donc pour tout $x \in X$, la suite $(f_n(x))$ est de Cauchy dans \mathbb{K} ; comme \mathbb{K} est complet, elle converge dans \mathbb{K} . En d'autres termes, la suite (f_n) converge simplement; notons $f : X \rightarrow \mathbb{K}$ l'application limite. Si $p \geq N_\varepsilon$ et $x \in X$, on passe à la limite pour $q \rightarrow +\infty$ dans l'inégalité $|f_p(x) - f_q(x)| < \varepsilon$ vraie pour tout $q \geq N_\varepsilon$: on obtient $|f_p(x) - f(x)| \leq \varepsilon$.

On a ainsi montré qu'il existe $N_\varepsilon \in \mathbb{N}$ tel que : $\forall p \geq N_\varepsilon, \forall x \in X, |f_p(x) - f(x)| \leq \varepsilon$. Ceci étant pour tout $\varepsilon > 0$, on conclut : (f_n) converge vers f uniformément sur X . \square

A noter que, là encore, la condition de Cauchy uniforme s'interprète comme le fait que la suite (f_n) est une suite de Cauchy dans l'evn $\mathcal{B}(X, \mathbb{K})$ muni de la norme $\|\cdot\|_\infty$.

f) EXERCICE (*des propriétés à connaître*). Montrer que si (f_n) et (g_n) convergent uniformément sur X vers f et g resp., alors $(\lambda f_n + \mu g_n)$ converge uniformément sur X vers $\lambda f + \mu g$, quels que soient $\lambda, \mu \in \mathbb{K}$. Montrer que si (f_n) converge uniformément sur X vers f , alors $(|f_n|)$ converge uniformément sur X vers $|f|$.

5.1.2 Convergence uniforme et continuité

a) EXEMPLES PRÉLIMINAIRES

- Soient $I = [0, 1]$ et $f_n(x) = x^n$ pour tout $n \in \mathbb{N}$ et tout $x \in I$. La suite (f_n) converge simplement sur I vers f définie par $f(x) = 0$ pour tout $x \in [0, 1[$ et $f(1) = 1$. Chaque f_n est continue sur I mais la fonction limite f n'est pas continue sur I (non continue en 1).

- Soient $I = [0, +\infty[$ et $f_n(x) = \frac{n}{x+n}$ pour tout $n \in \mathbb{N}^*$ et tout $x \in I$. Pour tout $x \in I$ fixé, $\lim_{n \rightarrow +\infty} f_n(x) = 1$ donc la suite (f_n) converge simplement sur I vers la fonction constante valant 1. Par ailleurs, pour tout $n \in \mathbb{N}^*$ fixé, $\lim_{x \rightarrow +\infty} f_n(x) = 0$. Ainsi :

$$\lim_{x \rightarrow +\infty} \lim_{n \rightarrow +\infty} f_n(x) = \lim_{x \rightarrow +\infty} f(x) = 1 \neq 0 = \lim_{n \rightarrow +\infty} \lim_{x \rightarrow +\infty} f_n(x).$$

b) THÉORÈME. Soient X une partie non vide de \mathbb{R} , et (f_n) une suite d'applications de $\mathcal{F}(X, \mathbb{K})$. On suppose que (f_n) converge uniformément sur X vers une application $f \in \mathcal{F}(X, \mathbb{K})$.

(i) Soit $a \in X$ quelconque; si chaque f_n est continue en a , alors f est continue en a .

(ii) Si chaque f_n est continue sur X , alors f est continue sur X .

Preuve. Soit $a \in X$. Fixons un réel $\varepsilon > 0$ quelconque. Puisque la suite (f_n) converge vers f uniformément sur X , il existe un entier $N \geq 0$ tel que : $\forall n \geq N, \forall x \in X, |f_n(x) - f(x)| < \varepsilon$. Pour un tel entier $n \geq N$, la continuité de f_n en a se traduit (pour le réel $\varepsilon > 0$ considéré) par l'existence d'un réel $\eta > 0$ tel que : $\forall x \in X, |x - a| < \eta \Rightarrow |f_n(x) - f_n(a)| < \varepsilon$. Pour un tel $x \in X$ vérifiant $|x - a| < \eta$, on a donc en combinant les deux assertions et en utilisant l'inégalité triangulaire : $|f(x) - f(a)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(a)| + |f_n(a) - f(a)| < 3\varepsilon$, ce qui prouve que f est continue en a . Le point (i) étant établi, le point (ii) s'en déduit de façon évidente. \square

c) REMARQUES.

- Par contraposée, ce théorème peut permettre de montrer qu'une convergence n'est pas uniforme : par exemple, dans le premier exemple préliminaire de a) on peut conclure que la suite (f_n) ne converge pas uniformément sur I .

- La convergence uniforme est une condition suffisante mais non nécessaire a priori pour la continuité de f . Il se peut que f soit continue sur I sans que la convergence soit uniforme sur I (comme dans le premier exemple de 5.1.1.c).

- On vérifie sans difficulté que la notion de convergence uniforme et tout le contenu du paragraphe 5.1.1 restent valables si l'on considère des fonctions $X \rightarrow F$, où X est un ensemble non vide et F est un evn [il suffit de remplacer dans chaque définition les expressions de la forme $|f_n(x) - f(x)|$ par $\|f_n(x) - f(x)\|_F$].

De plus, le théorème 5.1.2.b reste vrai si l'on considère des fonctions $X \rightarrow F$, où X est une partie d'un evn E et F est un evn [il suffit de remplacer dans la preuve les expressions de la forme $|f_n(x) - f(x)|$ ou $|f_n(x) - f_n(a)|$ par $\|f_n(x) - f(x)\|_F$ ou $\|f_n(x) - f_n(a)\|_F$, et les expressions de la forme $|x - a|$ par $\|x - a\|_E$].

d) COMPLÉMENT IMPORTANT SUR L'INTERVERSION DES LIMITES. Le théorème b) ci-dessus consiste à montrer que $\lim_{x \rightarrow a} f(x) = f(a)$, c'est-à-dire que :

$$\lim_{x \rightarrow a} \lim_{n \rightarrow +\infty} f_n(x) = \lim_{n \rightarrow +\infty} \lim_{x \rightarrow a} f_n(x).$$

En cela, c'est en fait une conséquence du théorème plus général, dont la preuve est laissée en exercice.

THÉORÈME. Soient X une partie non vide de \mathbb{R} et $a \in \overline{X}$ un point adhérent à X . Soit (f_n) une suite d'applications de $\mathcal{F}(X, \mathbb{K})$. On suppose que d'une part, pour tout $n \in \mathbb{N}$, f_n admet une limite ℓ_n en a , et d'autre part que la suite (f_n) converge uniformément sur X vers une application $f \in \mathcal{F}(X, \mathbb{K})$. Alors :

(1) la suite (ℓ_n) converge dans \mathbb{K} , (2) f admet une limite en a , (3) $\lim_a f = \lim_n \ell_n$,

c'est-à-dire que l'on a bien : $\lim_{x \rightarrow a} \lim_{n \rightarrow +\infty} f_n(x) = \lim_{n \rightarrow +\infty} \lim_{x \rightarrow a} f_n(x)$;

La même question se pose bien sûr pour des limites du type $x \rightarrow +\infty$ ou $x \rightarrow -\infty$ [voir le second exemple préliminaire du a) ci-dessus]. On peut vérifier que le théorème d'interversion des limites reste vrai :

- lorsque X est un intervalle du type $]\alpha, +\infty[$ et lorsque $a = +\infty$.

- lorsque X est un intervalle du type $]-\infty, \beta[$ et lorsque $a = -\infty$.

5.1.3 Convergence uniforme sur tout segment

a) EXEMPLE. Reprenons le premier exemple de 5.1.1.c, avec $I =]0; +\infty[$ et $f_n : I \rightarrow \mathbb{R}$ définie par $f_n(x) = \frac{1}{nx}$. La suite (f_n) converge simplement mais non uniformément sur I vers la fonction nulle. Quel que soit $[a, b] \subset I$, on a $\sup_{x \in [a, b]} |f_n(x)| = \frac{1}{na}$ qui tend vers 0 lorsque $n \rightarrow +\infty$. Donc la suite (f_n) converge vers 0 uniformément sur tout intervalle de la forme $[a, b]$ inclus dans I .

b) REMARQUE. Soit I un intervalle non vide de \mathbb{R} . Soit f une application $I \rightarrow \mathbb{K}$. Soit (f_n) une suite d'applications $I \rightarrow \mathbb{K}$. Lorsque (f_n) converge vers f uniformément sur tout segment $[a, b]$ inclus dans I , on dit parfois que (f_n) converge vers f *localement uniformément*. Dans la pratique, cette propriété peut être suffisante pour obtenir des énoncés intéressants sur la fonction limite, comme dans la situation suivante.

c) PROPOSITION. Soit (f_n) une suite d'applications $I \rightarrow \mathbb{K}$, où I est un intervalle non vide de \mathbb{R} . On suppose qu'il existe $f : I \rightarrow \mathbb{K}$ telle que (f_n) converge vers f uniformément sur tout segment $[a, b]$ inclus dans I . Si chaque f_n est continue sur I , alors f est continue sur I .

Preuve. Soit $x \in I$. Il existe $b > a \in I$ tels que $x \in [a, b] \subseteq I$. La convergence étant uniforme sur $[a, b]$, il résulte du théorème 5.1.2.b que f est continue sur $[a, b]$, donc en x . Et ceci pour tout $x \in I$. Donc f est continue sur I . \square

5.1.4 Convergence uniforme, intégration et dérivation.

a) THÉORÈME (*convergence uniforme et intégration sur un segment*). Soient $I = [a, b]$ un intervalle fermé borné de \mathbb{R} , pour des réels fixés $a \leq b$, et (f_n) une suite d'applications de $I \rightarrow \mathbb{K}$. On suppose que chaque f_n est continue sur I et que la suite (f_n) converge uniformément sur I vers une application $f : I \rightarrow \mathbb{K}$. Alors la suite $(\int_a^b f_n(t) dt)$ converge dans \mathbb{K} et l'on a :

$$\lim_{n \rightarrow +\infty} \int_a^b f_n(t) dt = \int_a^b f(t) dt$$

Preuve. Observons d'abord que, d'après le théorème 5.1.2.b, on sait que f est continue sur $[a, b]$, ce qui justifie l'existence de l'intégrale $\int_a^b f(t) dt$. Il résulte des propriétés connues de l'intégrale que, pour tout entier $n \geq 0$:

$$\left| \int_a^b f_n(t) dt - \int_a^b f(t) dt \right| = \left| \int_a^b (f_n - f)(t) dt \right| \leq \int_a^b |f_n(t) - f(t)| dt \leq (b-a) \sup_{a \leq t \leq b} |f_n(t) - f(t)|.$$

Puisque $\lim_{n \rightarrow +\infty} \|f_n - f\|_\infty = 0$ par l'hypothèse de convergence uniforme, on en déduit que

$$\lim_{n \rightarrow +\infty} \left(\int_a^b f_n(t) dt - \int_a^b f(t) dt \right) = 0, \text{ d'où le résultat. } \square$$

b) REMARQUES.

- Si (f_n) est une suite de fonctions continues sur $[a, b]$ convergeant simplement sur $[a, b]$ vers une fonction f continue sur $[a, b]$, il se peut que la suite $(\int_a^b f_n(t) dt)$ ne converge pas vers $\int_a^b f(t) dt$ dans \mathbb{K} . Dans ce cas bien sûr, la convergence n'est pas uniforme.

Par exemple : en posant $f_n(x) = \begin{cases} (-1)^n n^3 x & \text{si } x \in [0, \frac{1}{n}] \\ (-1)^{n+1} n^3 (x - \frac{2}{n}) & \text{si } x \in [\frac{1}{n}, \frac{2}{n}] \\ 0 & \text{si } x \in [\frac{2}{n}, 1] \end{cases}$ on définit une suite

$(f_n)_{n \geq 2}$ d'applications $[0, 1] \rightarrow \mathbb{R}$ qui converge simplement vers la fonction nulle, mais telle que la suite de réels $(\int_0^1 f_n(t) dt)_{n \geq 2}$ diverge puisque son terme général vaut $(-1)^n n$.

- Si (f_n) est une suite de fonctions continues sur $[a, b]$ convergeant simplement mais non uniformément sur $[a, b]$ vers une fonction f continue sur $[a, b]$, il se peut cependant que la suite $(\int_a^b f_n(t) dt)$ converge vers $\int_a^b f(t) dt$ dans \mathbb{K} .

Par exemple : en posant $f_n(x) = nx^n(1-x)$ pour tout $x \in [0, 1]$, on définit une suite $(f_n)_{n \geq 0}$ d'applications $[0, 1] \rightarrow \mathbb{R}$ qui converge simplement vers la fonction nulle, telle que la suite de réels $(\int_a^b f_n(t) dt)_{n \geq 0}$ converge puisque son terme général vaut $\frac{n}{(n+1)(n+2)}$, et telle que $(f_n)_{n \geq 0}$ ne converge pas vers la fonction nulle uniformément sur $[0, 1]$ (car on vérifie par un calcul direct du maximum de f_n sur $[0, 1]$ que $\lim_{n \rightarrow +\infty} \|f_n\|_\infty = 1/e$).

- Dans le théorème, on peut supposer f et les f_n seulement continues par morceaux sur I .

c) COROLLAIRE (convergence uniforme et primitives). Soit I un intervalle quelconque de \mathbb{R} (non vide, non réduit à un point). Soit (g_n) une suite d'applications $I \rightarrow \mathbb{K}$. On suppose que chaque g_n est continue sur I , et que la suite (g_n) converge uniformément sur tout segment inclus dans I vers une application $g : I \rightarrow \mathbb{K}$. Fixons un élément $a \in I$ quelconque et, pour tout $n \geq 0$, notons $h_n : I \rightarrow \mathbb{K}$ la primitive de g_n sur I telle que $h_n(a) = 0$.

Alors la suite (h_n) converge uniformément sur tout segment inclus dans I vers une application $h : I \rightarrow \mathbb{K}$; de plus, g est continue sur I et h est la primitive de g sur I telle que $h(a) = 0$.

Preuve. D'après la proposition 5.1.3.c, il résulte des hypothèses faites que g est continue sur I . On peut donc considérer la primitive h de g sur I telle que $h(a) = 0$, c'est-à-dire l'application $h : I \rightarrow \mathbb{K}$ définie par $h(x) = \int_a^x g(t) dt$ pour tout $x \in I$.

Soit K un segment inclus dans I . Il existe un segment $J = [\alpha, \beta]$ tel que $K \subset J \subset I$ et $a \in J$. Alors pour tout $n \in \mathbb{N}$ et tout $x \in J$, quitte à changer $[a, x]$ en $[x, a]$, on a :

$$\begin{aligned} |h_n(x) - h(x)| &= \left| \int_a^x g_n(t) dt - \int_a^x g(t) dt \right| = \left| \int_a^x (g_n - g)(t) dt \right| \\ &\leq \int_a^x |g_n(t) - g(t)| dt \leq |x - a| \sup_{t \in [a, x]} |g_n(t) - g(t)|. \end{aligned}$$

Mais $[a, x] \subset J = [\alpha, \beta]$, d'où : $|x - a| \leq (\beta - \alpha)$ et $\sup_{[a, x]} |g_n - g| \leq \sup_J |g_n - g|$. On a ainsi montré que :

$$|h_n(x) - h(x)| \leq (\beta - \alpha) \sup_J |g_n - g| \quad \text{pour tous } x \in J, n \in \mathbb{N}.$$

On déduit que $\sup_J |h_n - h| \leq (\beta - \alpha) \sup_J |g_n - g|$. Mais puisque par hypothèse (g_n) converge vers g uniformément sur J , on a $\lim_{n \rightarrow +\infty} \sup_J |g_n - g| = 0$, et donc $\lim_{n \rightarrow +\infty} \sup_J |h_n - h| = 0$.

On conclut que (h_n) converge vers h uniformément sur J , et donc sur K . \square

Ce corollaire est à la base du prochain théorème. Il existe plusieurs résultats concernant la dérivabilité de la fonction limite d'une suite de fonctions dérivables, sous diverses hypothèses. Conformément au programme, on se limite ici à la situation (la plus simple) de fonctions de classe C^1 sur un intervalle.

d) THÉORÈME (convergence uniforme et dérivation). Soit I un intervalle quelconque de \mathbb{R} (non vide, non réduit à un point). Soit (f_n) une suite d'applications $I \rightarrow \mathbb{K}$ convergeant simplement sur I vers une application $f : I \rightarrow \mathbb{K}$. On suppose que chaque f_n est de classe C^1 sur I . On fait l'hypothèse supplémentaire qu'il existe $g : I \rightarrow \mathbb{K}$ telle que la suite (f'_n) converge vers g uniformément sur tout segment $[a, b]$ inclus dans I .

Alors, la suite (f_n) converge vers f uniformément sur tout segment $[a, b]$ inclus dans I ; de plus f est de classe C^1 sur I et $f' = g$. En d'autres termes, on a :

$$\left(\lim_{n \rightarrow +\infty} f_n \right)' = \lim_{n \rightarrow +\infty} f'_n.$$

Preuve. Soit $a \in I$ quelconque. Pour tout $n \in \mathbb{N}$, notons : $g_n = f'_n$ et $h_n = f_n - f_n(a)$. On peut appliquer exactement le corollaire précédent : g est continue sur I et, si h désigne la primitive de g sur I telle que $h(a) = 0$, la suite (h_n) converge vers h uniformément sur tout segment inclus dans I . Comme $f_n = h_n + f_n(a)$ et comme $\lim_n f_n(a) = f(a)$, (puisque (f_n) converge simplement sur I vers f), on en déduit que (f_n) converge uniformément sur tout segment inclus dans I vers $h + f(a)$. En résumé, $f = h + f(a)$. Donc f est, comme h , de classe C^1 sur I , et l'on a : $f' = h' = g$. \square

e) REMARQUES.

- Une limite uniforme de fonctions dérivables sur I n'est pas forcément dérivable sur I .

Contre-exemple : prenons $I = \mathbb{R}$ et $f_n(x) = \sqrt{x^2 + \frac{1}{n}}$ pour tout $n \in \mathbb{N}^*$ et tout $x \in \mathbb{R}$. Chaque f_n est dérivable sur I , la suite (f_n) converge uniformément sur I vers la fonction f définie par $f(x) = |x|$ pour tout $x \in \mathbb{R}$, qui n'est pas dérivable sur \mathbb{R} .

- L'hypothèse de convergence uniforme porte sur la suite des dérivées.

Contre-exemple : prenons $I = [0, \frac{\pi}{2}]$ et $f_n(x) = \frac{\sin nx}{\sqrt{n}}$ pour tout $n \in \mathbb{N}^*$ et tout $x \in I$. La suite (f_n) converge uniformément sur I vers la fonction nulle (car $\sup_{x \in I} |f_n(x)| \leq \frac{1}{\sqrt{n}}$), qui est dérivable sur I de dérivée la fonction nulle elle-même. Par ailleurs, chaque f_n est dérivable sur I avec $f'_n(x) = \sqrt{n} \cos nx$, mais la suite (f'_n) ne converge pas vers la fonction nulle sur I , même simplement, puisque $\lim_n f'_n(0) = \lim_n \sqrt{n} = +\infty$.

- Les énoncés a) et d) restent *a fortiori* vrais lorsqu'on y remplace "converge uniformément sur tout segment inclus dans I " par la condition plus forte "converge uniformément sur I ".

La dernière partie de cette section est consacrée à un analogue du théorème a) ci-dessus mais pour l'intégration sur des intervalles qui ne sont pas nécessairement des segments. Conformément à ce que prévoit le programme, l'énoncé est admis sans détailler de preuve.

f) COMPLÉMENT SUR L'INTERVERSION DE LA LIMITE ET DE L'INTÉGRALE. Le théorème a) ci-dessus consiste à montrer que, sous des hypothèses convenables, on peut intervertir le passage à la limite et le signe d'intégration, c'est-à-dire que :

$$\lim_{n \rightarrow +\infty} \int_I f_n(t) dt = \int_I \lim_{n \rightarrow +\infty} f_n(t) dt.$$

Lorsque I n'est plus supposé être un segment mais un intervalle quelconque, et/ou lorsque la convergence de la suite de fonctions considérée n'est pas uniforme, on a divers résultats dont le suivant (admis) connu comme théorème de convergence dominée de Lebesgue.

THÉORÈME. Soit I un intervalle quelconque (non trivial) de \mathbb{R} . Soit (f_n) une suite d'applications $I \rightarrow \mathbb{K}$ continues sur I convergeant simplement sur I vers une application $f : I \rightarrow \mathbb{K}$ continue sur I . On suppose qu'il existe une application $g : I \rightarrow \mathbb{K}$ que l'on suppose intégrable sur I telle que $|f_n(x)| \leq g(x)$ pour tout $x \in I$ et tout $n \in \mathbb{N}$. Alors, chaque application f_n est intégrable sur I , l'application f est intégrable sur I , et on a $\int_I f(t) dt = \lim_{n \rightarrow +\infty} \int_I f_n(t) dt$.

EXEMPLE. Soit $I = [0, \frac{\pi}{2}]$. Pour tout entier $n \geq 0$, on définit $f_n : I \rightarrow \mathbb{R}$ par $f_n(x) = \sin^n x$ pour tout $x \in I$. Il est clair que (f_n) converge simplement vers la fonction $f : I \rightarrow \mathbb{R}$ définie par $f(x) = 0$ si $x \in [0, \frac{\pi}{2}[$ et $f(\frac{\pi}{2}) = 1$. Comme chaque f_n est continue sur I et que f ne l'est pas, la convergence de la suite (f_n) n'est pas uniforme sur I . On a par ailleurs $|f_n(x)| \leq 1$ pour tous $x \in I, n \in \mathbb{N}$, on peut appliquer le théorème en prenant pour g la fonction constante sur I égale à 1 (qui est bien intégrable sur I). Comme $\int_I f(t) dt = 0$, on conclut que $\lim_{n \rightarrow +\infty} \int_I f_n(t) dt = 0$.

5.1.5 Exercices

EXERCICE 1. Dans chacun des cas suivants, montrer que la suite $(f_n)_{n \geq 1}$ d'applications de I dans \mathbb{R} converge simplement sur I vers une fonction f à déterminer, et préciser si la convergence est ou non uniforme sur I .

a. $I = \mathbb{R}, f_n(x) = \frac{1}{n} \sin nx,$ b. $I = [0, a], f_n(x) = \frac{x}{n+x},$ c. $I = \mathbb{R}, f_n(x) = \frac{x\sqrt{n}}{1+nx^2}.$

EXERCICE 2. Soit a un réel tel que $a > 0$. On considère pour toute entier $n \geq 0$ l'application $f_n : I \rightarrow \mathbb{R}$, avec $I =]0, +\infty[$ définie par $f_n(x) = n^a x e^{-nx}$. Montrer que la suite (f_n) converge simplement vers 0 sur I . Montrer que f_n admet en $\frac{1}{n}$ un maximum que l'on calculera. Montrer que la suite (f_n) converge uniformément sur I si et seulement si $a < 1$.

EXERCICE 3. Montrer que la suite (f_n) d'applications continues sur $I = \mathbb{R}$ définies par $f_n(x) = \frac{1}{1+nx^2}$ pour tout $x \in I$ converge simplement mais non uniformément sur I vers une application f non continue sur I .

Montrer que la suite (g_n) d'applications non continues sur $I =]0, 1]$ définies par $g_n(x) = 0$ si $1 \geq x \geq \frac{1}{n}$ et $g_n(x) = 1$ si $0 < x < \frac{1}{n}$ converge simplement mais non uniformément sur I vers une application g continue sur I .

Montrer que la suite (h_n) d'applications non continues sur $I =]0, 1]$ définies par $h_n(x) = 0$ si $1 \geq x \geq \frac{1}{n}$ et $h_n(x) = \frac{1}{n}$ si $0 < x < \frac{1}{n}$ converge uniformément sur I vers une application g continue sur I .

EXERCICE 4. Pour tout x dans $I = [0, \frac{\pi}{2}]$, on pose $f_n(x) = n \sin x \cos^n x$. Montrer que la suite (f_n) converge simplement sur I vers une fonction f que l'on déterminera. Montre que la convergence n'est pas uniforme sur I ; [indication : considérer $I_n = \int_0^{\pi/2} f_n(x) dx$]. Montrer que la convergence est uniforme sur tout segment inclus dans $]0, \frac{\pi}{2}[$.

EXERCICE 5. Pour tout entier $n \geq 1$, soit $f_n : \mathbb{R} \rightarrow \mathbb{R}$ définie par $f_n(x) = (x^2 + \frac{1}{n^2})^{1/2}$. Montrer que f_n est dérivable sur \mathbb{R} . Montrer que la suite (f_n) converge uniformément sur \mathbb{R}

vers une application f continue mais non dérivable sur \mathbb{R} . De quels théorèmes du cours cet exemple est-il à rapprocher ?

EXERCICE 6. Montrer que la suite de terme général $u_n = \int_0^1 x^n(1-x)^n dx$ converge dans \mathbb{R} vers 0. Montrer que la suite de fonctions (f_n) définie sur $[0, 1]$ par $f_n(x) = n^2 x e^{-nx}$ converge simplement vers la fonction nulle, mais que $\lim_{n \rightarrow +\infty} \int_0^1 f_n(x) dx \neq 0$. Que peut-on en déduire ?

EXERCICE 7. Calculer dans chaque cas la limite de la suite (u_n) considérée :

$$u_n = \int_0^{\pi/4} \tan^n x dx, \quad u_n = \int_0^{+\infty} \frac{dx}{x^n + e^x}, \quad u_n = \int_0^{+\infty} \frac{x^n}{x^{n+2} + 1} dx, \quad u_n = \int_0^{+\infty} \frac{x^n}{x^{2n} + 1} dx.$$

5.2 Séries de fonctions et convergence normale

5.2.1 Convergence simple et uniforme d'une série de fonctions

a) DÉFINITIONS. Soit X une partie de \mathbb{R} . Soit $(f_n)_{n \geq 0}$ une suite d'applications $X \rightarrow \mathbb{K}$. On peut former la suite $(S_n)_{n \geq 0}$ des sommes partielles, avec $S_n = \sum_{k=0}^n f_k$. Chaque S_n est donc une application $X \rightarrow \mathbb{K}$, et tout ce que l'on a développé dans la première partie de ce chapitre s'applique en particulier à la suite (S_n) des sommes partielles.

- On dit que la série de fonctions $\sum f_n$ converge *simplement* sur X lorsque la suite des sommes partielles (S_n) converge simplement sur X .

Dans ce cas, l'application limite $S : X \rightarrow \mathbb{K}$ définie par $S(x) = \lim_{n \rightarrow +\infty} S_n(x)$ est appelée la *somme* de la série de fonctions $\sum_{n \geq 0} f_n$, et est notée $\sum_{k=0}^{+\infty} f_k$. En d'autres termes, si la série de fonctions $\sum f_n$ converge simplement sur X , alors :

$$\sum_{k=0}^{+\infty} f_k \text{ est l'application } X \rightarrow \mathbb{K} \text{ définie par } \sum_{k=0}^{+\infty} f_k(x) = \lim_{n \rightarrow +\infty} \sum_{k=0}^n f_k(x) \text{ pour tout } x \in X.$$

- On dit que la série de fonctions $\sum f_n$ converge *uniformément* sur X lorsque la suite des sommes partielles (S_n) converge uniformément sur X . Il résulte en particulier de 5.1.1.d que : *si la série de fonctions $\sum f_n$ converge uniformément sur X , alors elle converge simplement sur X .*

- On peut de même appliquer à la suite des sommes partielles les notions de condition de Cauchy uniforme (voir 5.1.1.e) ou de convergence uniforme sur tout segment (voir 5.1.3).

b) FORMULATION POUR LES SÉRIES DE FONCTIONS DES PRINCIPAUX THÉORÈMES. Les théorèmes de continuité, d'interversion de limites, d'intégration ou de dérivation vus dans la première partie du chapitre pour les suites de fonctions s'appliquent directement pour les séries de fonctions.

[1] THÉORÈME D'INTERVERSION DE LIMITES. Soient X une partie non vide de \mathbb{R} et $a \in \overline{X}$ un point adhérent à X . Soit (f_n) une suite d'applications $X \rightarrow \mathbb{K}$. On suppose que d'une part, pour tout $n \in \mathbb{N}$, f_n admet une limite ℓ_n en a , et d'autre part que la série $\sum_{n \geq 0} f_n$ converge uniformément sur X . Alors : la série $\sum_{n \geq 0} \ell_n$ converge dans \mathbb{K} , l'application $\sum_{n=0}^{+\infty} f_n$ admet une limite en a , et l'on a :

$$\lim_{x \rightarrow a} \sum_{n=0}^{+\infty} f_n(x) = \sum_{n=0}^{+\infty} \ell_n.$$

Ceci traduit bien l'interversion de limites : $\lim_{x \rightarrow a} \lim_{p \rightarrow +\infty} \sum_{n=0}^p f_n(x) = \lim_{p \rightarrow +\infty} \lim_{x \rightarrow a} \sum_{n=0}^p f_n(x)$.

[2] THÉORÈME DE CONTINUITÉ. Soient X une partie non vide de \mathbb{R} , et $\sum_{n \geq 0} f_n$ une série d'applications $X \rightarrow \mathbb{K}$. On suppose que la série $\sum_{n \geq 0} f_n$ converge uniformément sur X .

(i) Soit $a \in X$; si chaque f_n est continue en a , alors l'application $\sum_{n=0}^{+\infty} f_n$ est continue en a .

(ii) Si chaque f_n est continue sur X , alors l'application $\sum_{n=0}^{+\infty} f_n$ est continue sur X .

[3] THÉORÈME DE CONTINUITÉ. Soit $\sum_{n \geq 0} f_n$ une série d'applications $I \rightarrow \mathbb{K}$, où I est un intervalle non vide de \mathbb{R} . On suppose que la série $\sum_{n \geq 0} f_n$ converge uniformément sur tout segment $[a, b]$ inclus dans I .

Si chaque f_n est continue sur I , alors l'application $\sum_{n=0}^{+\infty} f_n$ est continue sur I .

[4] THÉORÈME D'INTÉGRATION SUR UN SEGMENT. Soient $I = [a, b]$ un intervalle fermé borné de \mathbb{R} , pour des réels fixés $a \leq b$, et $\sum_{n \geq 0} f_n$ une série d'applications de $I \rightarrow \mathbb{K}$. On suppose que chaque f_n est continue sur I et que la série $\sum_{n \geq 0} f_n$ converge uniformément sur I .

Alors l'application $\sum_{n=0}^{+\infty} f_n$ est continue sur I , la série de réels $\sum_{n \geq 0} \int_a^b f_n(t) dt$ converge, et l'on a :

$$\int_a^b \left(\sum_{n=0}^{+\infty} f_n(t) \right) dt = \sum_{n=0}^{+\infty} \left(\int_a^b f_n(t) dt \right).$$

[5] THÉORÈME DE DÉRIVATION. Soit I un intervalle quelconque de \mathbb{R} (non vide, non réduit à un point). Soit $\sum_{n \geq 0} f_n$ une série d'applications $I \rightarrow \mathbb{K}$. On suppose que chaque f_n est de classe C^1 sur I , que la série $\sum_{n \geq 0} f_n$ converge simplement sur I , et que la série $\sum_{n \geq 0} f'_n$ converge uniformément sur tout segment $[a, b]$ inclus dans I .

Alors la série $\sum_{n \geq 0} f_n$ converge uniformément sur tout segment $[a, b]$ inclus dans I , l'application $\sum_{n=0}^{+\infty} f_n$ est de classe C^1 sur I , et l'on a :

$$\left(\sum_{n=0}^{+\infty} f_n(t) \right)' = \sum_{n=0}^{+\infty} f'_n.$$

Les preuves de ces résultats consistent simplement à appliquer directement à la suite des sommes partielles les énoncés 5.1.2.d, 5.1.2.c, 5.1.3.c, 5.1.4.b et 5.1.4.d.

On verra en exercice des exemples d'application de ces résultats. Auparavant, on présente une condition suffisante de convergence uniforme d'une série de fonctions, qui constituera concrètement la principale façon de vérifier qu'une série de fonctions est uniformément convergente.

5.2.2 Convergence normale

a) DEFINITION. Soit X une partie de \mathbb{R} . Soit $(f_n)_{n \geq 0}$ une suite d'applications $X \rightarrow \mathbb{K}$. On dit que la série de fonctions $\sum f_n$ converge *normalement* sur X lorsque les fonctions f_n sont bornées sur X (à partir d'un certain rang) et que la série de réels positifs $\sum \|f_n\|_\infty$ est convergente.

Cette condition équivaut à l'existence d'une suite $(a_n)_{n \geq 0}$ de réels positifs telle que :

$$\left[\forall n \in \mathbb{N}, \forall x \in X, |f_n(x)| \leq a_n \right] \quad \text{et} \quad \left[\text{la série numérique } \sum_{n \geq 0} a_n \text{ est convergente} \right].$$

b) THÉORÈME FONDAMENTAL. Soit $\sum f_n$ une série de fonctions $X \rightarrow \mathbb{R}$. Si elle converge normalement sur X , alors elle converge uniformément sur X .

Preuve. On suppose que $\sum f_n$ est normalement convergente sur X . Il résulte de la dernière remarque ci-dessus que, pour tout $x \in X$, la série numérique $\sum |f_n(x)|$ est convergente (car son terme général est majorée par le terme général a_n d'une série convergente), et donc que cette série numérique $\sum f_n(x)$ est convergente. Cela signifie que la série de fonctions $\sum f_n$ converge simplement sur X . Notons $f = \sum_{n=0}^{+\infty} f_n$, qui est une application $X \rightarrow \mathbb{K}$.

Ceci étant, pour tout entier $N \geq 0$, posons $S_N = \sum_{n=0}^N f_n$ et $R_N = f - S_N = \sum_{n=N+1}^{+\infty} f_n$. Les applications S_N et R_N vérifient donc pour tout $x \in X$:

$$S_N(x) = \sum_{n=0}^N f_n(x) \quad \text{et} \quad R_N(x) = f(x) - S_N(x) = \sum_{n=N+1}^{+\infty} f_n(x).$$

d'où l'on déduit :

$$\|f - S_N\|_\infty = \|R_N\|_\infty = \left\| \sum_{n=N+1}^{+\infty} f_n \right\|_\infty \leq \sum_{n=N+1}^{+\infty} \|f_n\|_\infty = \sum_{n=0}^{+\infty} \|f_n\|_\infty - \sum_{n=0}^N \|f_n\|_\infty$$

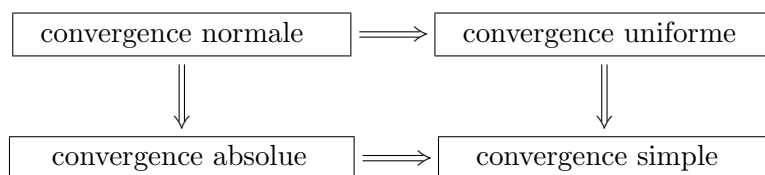
en utilisant l'hypothèse, que la série numérique $\sum \|f_n\|_\infty$ converge; et donc en particulier $\lim_{N \rightarrow +\infty} \|f - S_N\|_\infty = 0$, ce qui démontre le résultat voulu. \square

A noter que l'on a sous l'hypothèse du théorème : $\left\| \sum_{n=0}^{+\infty} f_n \right\|_\infty \leq \sum_{n=0}^{+\infty} \|f_n\|_\infty$.

A noter aussi que l'implication réciproque du théorème est fautive en général (voir contre-exemples ci-dessous).

c) REMARQUE. Soit X une partie de \mathbb{R} . Soit $(f_n)_{n \geq 0}$ une suite d'applications $X \rightarrow \mathbb{K}$. On dit que la série de fonctions $\sum f_n$ converge *absolument* sur X lorsque la série de fonctions $\sum |f_n|$ converge simplement sur X . Cela signifie que, pour tout $x \in X$, la série de réels positifs $\sum |f_n(x)|$ est convergente. En d'autres termes, la série de réels $\sum f_n(x)$ est absolument convergente donc convergente (voir 4.3.1.b) pour tout $x \in X$, ce qui signifie que la série de fonctions $\sum f_n$ est simplement convergente sur X . On a aussi établi au début de la preuve du théorème ci-dessus que toute série de fonctions normalement convergente sur X est absolument convergente sur X .

d) SYNTHÈSE. On peut synthétiser de façon schématique le troisième point de 5.2.1.a, le théorème 5.2.1.b et la remarque 5.2.1.c ci-dessus sous la forme :



e) EXEMPLES ET CONTRE-EXEMPLES.

► **(i)** On considère la série de fonctions $\sum_{n \geq 1} \frac{\cos nx}{n^2+x^2}$ de la variable $x \in \mathbb{R}$. Posons $f_n(x) = \frac{\cos nx}{n^2+x^2}$. Pour tout $x \in \mathbb{R}$ et pour tout entier $n \geq 1$, on a $|f_n(x)| \leq \frac{1}{n^2+x^2} \leq \frac{1}{n^2}$. Donc $|f_n(x)|$ est majoré pour tout $x \in \mathbb{R}$ par le terme général $\frac{1}{n^2}$ d'une série numérique convergente (série de Riemann

avec $\alpha = 2 > 1$) indépendante de x . On conclut que la série de fonctions $\sum f_n$ est normalement convergente sur \mathbb{R} (donc a fortiori uniformément, absolument et simplement convergente sur X).

► (ii) On considère la série de fonctions $\sum_{n \geq 1} \left(\frac{1}{n} - \frac{1}{n+x}\right)$ de la variable $x \in [0, +\infty[$. Posons $f_n(x) = \frac{1}{n} - \frac{1}{n+x}$. Pour tout $x \geq 0$ et pour tout entier $n \geq 1$, $f_n(x) \geq 0$ et $f_n(x) = \frac{x}{n(n+x)}$. Ainsi, pour un $x \geq 0$ fixé quelconque, $f_n(x) \sim \frac{x}{n^2}$ pour n au voisinage de l'infini. Donc, pour tout $x \geq 0$ fixé, la série numérique $\sum f_n(x)$ est convergente. Ceci traduit le fait que la série de fonctions $\sum f_n$ converge simplement sur $[0, +\infty[$ (et aussi absolument puisque $f_n(x) \geq 0$).

En vue d'étudier la convergence normale, on étudie les variations de f_n . C'est une fonction croissante sur $[0, +\infty[$, avec $f_n(0) = 0$ et $\lim_{x \rightarrow +\infty} f_n(x) = \frac{1}{n}$. Donc $\|f_n\|_\infty = \frac{1}{n}$, ce qui prouve que les f_n sont bornées, mais que la série $\sum \|f_n\|_\infty$ est divergente (série harmonique). On conclut donc que la série de fonctions $\sum f_n$ n'est pas normalement convergente sur $[0, +\infty[$.

Considérons maintenant, pour tout réel $a \geq 0$, l'intervalle $[0, a]$. Si l'on restreint les f_n à $[0, a]$, on a $\sup_{x \in [0, a]} |f_n(x)| = \frac{a}{n^2 + na}$, qui est le terme général d'une série numérique convergente (car $\frac{a}{n^2 + na} \sim \frac{a}{n^2}$ pour n au voisinage de l'infini). On conclut donc que la série de fonctions $\sum f_n$ est normalement convergente (donc uniformément convergente) sur tout intervalle $[0, a]$ avec $a \geq 0$.

Comme il est clair que chaque f_n est continue sur $[0, +\infty[$, ce résultat suffit par exemple pour conclure (avec le [3] de 5.2.1.b) que l'application $x \mapsto \sum_{n=1}^{+\infty} \frac{x}{n(n+x)}$ est continue sur $[0, +\infty[$.

► (iii) On considère la série de fonctions $\sum_{n \geq 1} \frac{(-1)^n}{n+x}$ pour $x \in [0, +\infty[$. Posons $f_n(x) = \frac{(-1)^n}{n+x}$.

Pour tout $x \geq 0$ fixé, on a $|f_n(x)| = \frac{1}{n+x}$ pour tout $n \geq 0$, qui est le terme général d'une série numérique divergente (car équivalent pour n au voisinage de l'infini au terme général $\frac{1}{n}$ la série harmonique). Ceci traduit le fait que la série de fonctions $\sum f_n$ ne converge pas absolument sur $[0, +\infty[$ (et donc ne converge pas normalement sur $[0, +\infty[$).

Pour tout $x \geq 0$ fixé, on a $f_n(x) = (-1)^n \alpha_n$ pour tout $n \geq 0$ avec la notation $\alpha_n = \frac{1}{n+x}$. Comme la suite de réels positifs (α_n) décroît en convergeant vers 0, on peut appliquer la règle des séries alternées 4.3.2.b et en déduire que la série numérique $\sum f_n(x)$ est convergente. Ceci traduit le fait que la série de fonctions $\sum f_n$ converge simplement sur $[0, +\infty[$.

La convergence simple que l'on vient de vérifier permet de considérer la fonction somme de la série $f = \sum_{n=1}^{+\infty} f_n$, définie par $f(x) = \sum_{n=1}^{+\infty} f_n(x)$ pour tout $x \in [0, +\infty[$. Notons $S_N = \sum_{n=1}^N f_n$ les sommes partielles de la série de fonctions.

On vérifie que, pour tout $x \in [0; +\infty[$ et tout $N \geq 1$, on a :

$$|f(x) - S_N(x)| = \left| \sum_{n=N+1}^{+\infty} \frac{(-1)^n}{n+x} \right| \leq \frac{1}{N+1+x} \leq \frac{1}{N+1}.$$

Il en résulte que $\lim_{N \rightarrow +\infty} \|f - S_N\|_\infty = 0$, ce qui prouve que la convergence de la suite (S_N) des sommes partielles vers la fonction f est uniforme sur $[0, +\infty[$. Ceci traduit le fait que la série de fonctions $\sum f_n$ converge uniformément sur $[0, +\infty[$.

► On peut retenir de ce dernier exemple la méthode suivante : *une série de fonctions $\sum_{n \geq 0} f_n$ converge uniformément sur une partie X de \mathbb{R} si et seulement si elle converge simplement et vérifie de plus $\lim_{N \rightarrow +\infty} \|f - S_N\|_\infty = 0$ [où l'on a noté $S_N = \sum_{n=0}^N f_n$ les sommes partielles, $f = \lim_{N \rightarrow +\infty} S_N = \sum_{n=0}^{+\infty} f_n$ la fonction somme, et la norme $\|\cdot\|_\infty$ est définie par le sup sur X].*

5.2.3 Exercices

EXERCICE 1. Pour tout entier $n \geq 1$, on considère l'application $f_n : \mathbb{R} \rightarrow \mathbb{R}$ définie par $f_n(x) = \frac{\sin nx}{n!}$. Etudier la convergence sur \mathbb{R} de la série de fonctions $\sum_{n \geq 1} f_n$.

EXERCICE 2. Pour tout entier $n \geq 0$, on considère l'application $f_n : \mathbb{R}_+ \rightarrow \mathbb{R}$ définie par $f_n(x) = nx^2 e^{-x\sqrt{n}}$. Montrer que la série de fonctions $\sum_{n \geq 0} f_n$ converge simplement sur \mathbb{R}_+ . Montrer qu'elle ne converge pas uniformément sur \mathbb{R}_+ . Montrer qu'elle converge normalement sur tout intervalle $[a, +\infty[$ pour tout réel $a > 0$.

EXERCICE 3. Pour tout entier $n \geq 1$, on considère l'application $f_n : \mathbb{R}_+ \rightarrow \mathbb{R}$ définie par $f_n(x) = (-1)^n \ln(1 + \frac{x}{n(1+x)})$. Montrer que la série de fonctions $\sum_{n \geq 1} f_n$ converge simplement sur \mathbb{R}_+ . Montrer qu'elle ne converge pas absolument sur \mathbb{R}_+ . Montrer qu'elle converge uniformément sur \mathbb{R}_+ .

EXERCICE 4. Pour tout entier $n \geq 0$, on considère l'application $f_n : \mathbb{R}_+ \rightarrow \mathbb{R}$ définie par $f_n(x) = \frac{e^{-nx}}{1+n^2}$. Montrer que la série de fonctions $\sum_{n \geq 0} f_n$ converge uniformément sur \mathbb{R}_+ . Montrer que la série de fonctions $\sum_{n \geq 0} f'_n$ converge simplement sur $]0, +\infty[$ et uniformément sur $[a, +\infty[$ pour tout réel $a > 0$. Montrer que la fonction $f = \sum_{n=0}^{+\infty} f_n$ est continue sur $[0, +\infty[$ et dérivable sur $]0, +\infty[$.

EXERCICE 5. Pour tout entier $n \geq 0$, on considère l'application $f_n : \mathbb{R} \rightarrow \mathbb{R}$ définie par $f_n(x) = e^{-x\sqrt{n}}$. Déterminer la partie X de \mathbb{R} maximale telle que la série de fonctions $\sum_{n \geq 0} f_n$ converge simplement sur X . Montrer qu'elle converge normalement sur tout intervalle $[a, +\infty[$ pour tout réel $a > 0$. Montrer que la fonction $f = \sum_{n=0}^{+\infty} f_n$ est dérivable et décroissante sur $]0, +\infty[$. Montrer que la convergence de la série de fonctions n'est pas uniforme sur $]0, +\infty[$.

EXERCICE 6. Pour tout entier $n \geq 1$, on considère l'application $f_n : \mathbb{R} \rightarrow \mathbb{R}$ définie par $f_n(x) = \frac{1}{n^2} \arctan nx$. Montrer que la série de fonctions $\sum_{n \geq 1} f_n$ converge uniformément sur \mathbb{R} . Montrer que la fonction $f = \sum_{n=1}^{+\infty} f_n$ est continue sur \mathbb{R} , et dérivable sur $]0, +\infty[$ et sur $] -\infty, 0[$. Etudier la limite en 0 de $f'(x)$.

EXERCICE 7. Soit $\sum_{n \geq 0} a_n$ une série convergente d'éléments de \mathbb{K} . On considère la série de fonctions $\sum_{n \geq 0} a_n (\frac{x^n}{1+x^n})$ de la variable réelle $x \in [0, +\infty[$. Montrer qu'elle converge normalement sur $]0, \alpha]$ pour tout réel $0 < \alpha < 1$. Montrer qu'elle converge normalement sur $[\beta, +\infty[$ pour tout réel $\beta > 1$. Etudier la continuité sur \mathbb{R}_+ la fonction $\sum_{n=0}^{+\infty} a_n (\frac{x^n}{1+x^n})$.

EXERCICE 8 (*lemme d'Abel uniforme*). Soit D une partie de \mathbb{R} . Soit (α_n) une suite d'applications $D \rightarrow \mathbb{R}$ convergeant uniformément sur D vers l'application nulle. On suppose de plus que, pour tout $x \in D$, la suite de réels $(\alpha_n(x))$ est décroissante. Soit par ailleurs (f_n) une suite d'applications $D \rightarrow \mathbb{K}$ telle qu'il existe un réel $K \geq 0$ majorant $|\sum_{p=0}^n f_p(x)|$ pour tous $n \in \mathbb{N}, x \in D$. Montrer que la série de fonctions $\sum \alpha_n f_n$ converge uniformément sur D .

EXERCICE 9. Pour tout entier $n \geq 0$, on considère l'application $f_n : I \rightarrow \mathbb{R}$ définie par $f_n(x) = \frac{\cos nx}{\sqrt{n+x}}$, où $I =]0, 2\pi[$. En appliquant l'exercice précédent, montrer que la série de fonctions $\sum_{n \geq 0} f_n$ converge uniformément sur tout intervalle de la forme $I_a = [a, 2\pi - a]$ pour tout $0 < a < \frac{\pi}{2}$. Montrer que la fonction $f = \sum_{n=0}^{+\infty} f_n$ est continue sur I .

EXERCICE 10 (*fonction de Weierstraß*). Soient a un entier naturel impair et $0 < b < 1$ un réel tels que $ab > 1 + \frac{3\pi}{2}$. Pour tout $n \geq 0$, on considère l'application $f_n : \mathbb{R} \rightarrow \mathbb{R}$ définie par $f_n(x) = b^n \cos(a^n \pi x)$. Montrer que la série de fonctions $\sum_{n \geq 0} f_n$ converge uniformément sur \mathbb{R} . Montrer que la fonction $f = \sum_{n=0}^{+\infty} f_n$ est continue et bornée sur \mathbb{R} , mais n'est dérivable en aucun point de \mathbb{R} (pour cette dernière question point, des indications complémentaires seront détaillées en cours).

Chapitre 6

Séries entières

Conformément à ce que prévoit le programme, ce chapitre est plus précisément consacré à l'étude des séries entières d'une variable *complexe*, et s'applique donc en particulier aux séries entières d'une variable *réelle*.

Fixons d'abord le cadre de ce chapitre.

On a déjà expliqué à la fin du paragraphe 5.1.2 pourquoi tout ce que l'on a développé en 5.1.1 et jusqu'au théorème 5.1.2.b reste vrai sans changement si l'on considère non plus seulement des suites de fonctions $X \rightarrow \mathbb{K}$ pour X une partie de \mathbb{R} , mais des suites de fonctions $X \rightarrow F$ pour F un evn et X une partie d'un evn E .

En l'appliquant au cas où $E = F = \mathbb{C}$, muni de la norme définie par le module, et en s'intéressant à des suites de fonctions qui sont plus spécifiquement des séries de fonctions, on peut donc pour synthétiser les points suivants :

Soit X une partie de \mathbb{C} . Soit $\sum_{n \geq 0} f_n$ une série de fonctions avec $f_n : X \rightarrow \mathbb{C}$.

• La série $\sum_{n \geq 0} f_n$ converge normalement sur X lorsqu'il existe une suite $(u_n)_{n \geq 0}$ de réels positifs telle que :

$$[\forall n \in \mathbb{N}, \forall z \in X, |f_n(z)| \leq u_n] \quad \text{et} \quad \left[\text{la série numérique } \sum_{n \geq 0} u_n \text{ est convergente} \right].$$

• Si la série $\sum_{n \geq 0} f_n$ converge normalement sur X , alors d'une part elle converge absolument sur X (ce qui signifie que la série de fonctions positives $\sum |f_n|$ converge simplement sur X), d'autre part elle converge uniformément sur X .

• Si la série $\sum_{n \geq 0} f_n$ converge absolument ou uniformément sur X , alors elle converge simplement sur X , et l'on peut donc considérer la fonction f définie comme somme de la série de fonctions :

$$f = \sum_{n=0}^{+\infty} f_n : X \rightarrow \mathbb{C} \quad \text{définie par} \quad f(z) = \sum_{n=0}^{+\infty} f_n(z) \quad \text{pour tout } z \in X.$$

• Si la série $\sum_{n \geq 0} f_n$ converge *uniformément sur* X , ou plus généralement *uniformément sur toute boule fermée incluse dans* X , et si chaque application f_n est continue sur X , alors la fonction $f = \sum_{n=0}^{+\infty} f_n$ est continue sur X (cette dernière observation est un analogue pour \mathbb{C} de la proposition 5.1.3.c)

6.1 Convergence des séries entières

6.1.1 Rayon de convergence d'une série entière

a) DÉFINITION. On appelle *série entière* une série de fonctions $\sum_{n \geq 0} f_n$ où les applications f_n sont de la forme $f_n(z) = a_n z^n$ avec $a_n \in \mathbb{C}$ fixé.

En d'autres termes, une série entière est définie par la donnée d'une suite de nombres complexes $(a_n)_{n \geq 0}$, et l'on note alors la série entière :

$$\sum_{n \geq 0} a_n z^n.$$

Deux problèmes principaux se posent alors naturellement, que l'on va étudier ci-dessous :

- Une série entière $\sum a_n z^n$ étant donnée,
 - déterminer son *domaine de convergence*, c'est-à-dire l'ensemble X sur lequel elle converge simplement (X est donc l'ensemble des $z \in \mathbb{C}$ tels que la série de nombres complexes $\sum a_n z^n$ est convergente),
 - puis étudier la fonction somme $f : X \rightarrow \mathbb{C}$ définie par $z \mapsto f(z) = \sum_{n=0}^{+\infty} a_n z^n$.
- Une application $f : X \rightarrow \mathbb{C}$ avec $X \subseteq \mathbb{C}$ étant donnée, déterminer s'il existe une suite (a_n) telle que $f(z) = \sum_{n=0}^{+\infty} a_n z^n$ pour tout $z \in X$ (ou tout z dans une partie de X).

b) THÉORÈME FONDAMENTAL. Soit $\sum a_n z^n$ une série entière. Il existe un unique élément R qui est, soit un réel positif, soit $+\infty$, tel que l'on ait :

- Pour tout $z \in \mathbb{C}$ tel que $|z| < R$, la série $\sum a_n z^n$ est absolument convergente, et donc convergente.
- Pour tout $z \in \mathbb{C}$ tel que $|z| > R$, la suite $(a_n z^n)$ n'est pas bornée, et donc la série $\sum a_n z^n$ est divergente.

Preuve. Soit E l'ensemble des réels $\rho \geq 0$ tels que la suite $(a_n \rho^n)$ soit bornée. Il est clair que $0 \in E$ et que, si $\rho \in E$, alors $[0, \rho] \subseteq E$; donc E est un intervalle de \mathbb{R}_+ contenant 0. Notons R la borne supérieure de E dans \mathbb{R}_+ lorsque E est majorée, et $R = +\infty$ lorsque $E = [0, +\infty[$. On a par définition $[0, R[\subseteq E \subseteq [0, R]$.

Soit $z \in \mathbb{C}$ tel que $|z| < R$. Il existe $\rho \in E$ tel que $0 \leq |z| < \rho < R$. La suite $(a_n \rho^n)$ est bornée. Il existe $M \in \mathbb{R}_+$ tel que $|a_n \rho^n| \leq M$ pour tout $n \in \mathbb{N}$. On déduit que :

$$|a_n z^n| = \left| a_n \rho^n \left(\frac{z}{\rho} \right)^n \right| \leq M \left(\frac{|z|}{\rho} \right)^n.$$

Or comme $0 \leq \frac{|z|}{\rho} < 1$, la série géométrique $\sum \left(\frac{|z|}{\rho} \right)^n$ est convergente, donc la relation de domination ci-dessus implique que la série $\sum |a_n z^n|$ est convergente. Ceci prouve que R vérifie (i).

Soit $z \in \mathbb{C}$ tel que $|z| > R$. Alors $|z| \notin E$, donc la suite $(a_n z^n)$ n'est pas bornée, ce qui implique que la série $\sum a_n z^n$ est divergente, et prouve que R vérifie (ii).

Il reste à établir l'unicité d'un tel R . Supposons qu'il existe deux réels R_1 et R_2 satisfaisant les conditions (i) et (ii). Si $R_1 < R_2$, le réel $\rho = \frac{1}{2}(R_1 + R_2)$ vérifie $0 \leq \rho < R_2$, de sorte que la série $\sum a_n \rho^n$ est absolument convergente, et vérifie aussi $\rho > R_1$, de sorte que la suite $(a_n \rho^n)$ n'est pas bornée. Les deux conditions sont incompatibles, d'où une contradiction. On conclut de même si $R_1 > R_2$, ce qui achève la preuve. \square

c) DÉFINITIONS. L'élément R déterminé par ce théorème est appelé le *rayon de convergence* de la série entière.

- Dire que $R = 0$ signifie que la série entière ne converge pour aucune valeur non-nulle de z , c'est-à-dire que son domaine de convergence X est réduit à $\{0\}$.

- Au contraire, dire que $R = +\infty$ signifie qu'elle converge (et même absolument) pour tout $z \in \mathbb{C}$, c'est-à-dire que son domaine de convergence X est \mathbb{C} tout entier.

- Supposons enfin que $0 < R < +\infty$. Considérons dans le plan complexe le disque ouvert de centre 0 et de rayon R :

$$D(0, R) = \{z \in \mathbb{C}; |z| < R\}.$$

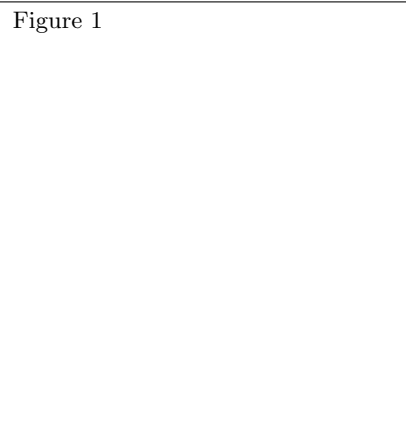
Par définition de R , la série $\sum a_n z^n$ converge absolument en tout point de $D(0, R)$, et diverge en tout point extérieur au disque fermé

$$\overline{D}(0, R) = \{z \in \mathbb{C}; |z| \leq R\}.$$

Mais le théorème ne dit rien sur ce qui se passe sur le cercle

$$C(0, R) = \{z \in \mathbb{C}; |z| = R\}.$$

Ce cercle peut contenir à la fois des points $z \in \mathbb{C}$ pour lesquels la série $\sum a_n z^n$ converge et des points $z \in \mathbb{C}$ pour lesquels elle diverge. Pour cette raison, ce cercle est parfois appelé cercle d'incertitude de la série entière.



d) EXEMPLES (*calcul du rayon de convergence en utilisant seulement la définition*).

► Considérons la série entière $\sum e^{-\sqrt{n}} z^n$. Cherchons son rayon de convergence R .

On a $|e^{-\sqrt{n}} z^n| = \exp(-\sqrt{n} + n \ln |z|) = \exp[\sqrt{n}(\sqrt{n} \ln |z| - 1)]$.

Supposons $|z| > 1$; donc $\ln |z| > 0$, ce qui implique que $|e^{-\sqrt{n}} z^n|$ tend vers $+\infty$ quand $n \rightarrow +\infty$. On a donc nécessairement $\sum e^{-\sqrt{n}} z^n$ divergente. Ainsi la série $\sum e^{-\sqrt{n}} z^n$ diverge pour tout z tel que $|z| > 1$, ce qui prouve d'après le point (i) du théorème b) que $|z| \geq R$. Donc $R \leq 1$.

Supposons $|z| < 1$; donc $\ln |z| < 0$, ce qui implique que $|e^{-\sqrt{n}} z^n|$ tend vers 0 quand $n \rightarrow +\infty$. En particulier, la suite $(e^{-\sqrt{n}} z^n)$ est bornée. Ce qui implique d'après le point (ii) du théorème b) que $|z| \leq R$. Donc $1 \leq R$. On conclut finalement que $R = 1$.

► Considérons la série entière $\sum (\sin n) z^n$. Cherchons son rayon de convergence R .

On a $|(\sin n) z^n| \leq |z|^n$. Si $|z| < 1$, la série géométrique de raison $|z|$ est convergente. Le critère de majoration sur les séries à termes positifs assure donc que la série de terme général $|(\sin n) z^n|$ est convergente. Ainsi la série de terme général $(\sin n) z^n$ est absolument convergente pour tout z tel que $|z| < 1$. Ceci prouve que $1 \leq R$. Prenons maintenant $z = 1$. La série $\sum \sin n$ diverge car son terme général ne tend pas vers 0. Ainsi, on a trouvé le point $z = 1$ en lequel la série entière $\sum \sin n z^n$ est divergente, ce qui suffit à prouver que $R \leq 1$. On conclut finalement que $R = 1$.

e) PROPOSITION (*calcul du rayon de convergence par comparaison*). Soit $\sum a_n z^n$ et $\sum b_n z^n$ deux séries entières, de rayon de convergence respectifs R_a et R_b .

- Majoration : si $|a_n| \leq |b_n|$ à partir d'un certain rang, alors $R_b \leq R_a$.
- Domination : s'il existe $M \in \mathbb{R}_+$ tel que $|a_n| \leq M|b_n|$ à partir d'un certain rang, alors $R_b \leq R_a$.
- Equivalence : si $|a_n| \sim |b_n|$ pour n au voisinage de l'infini, alors $R_a = R_b$.

Preuve. Soit $z \in \mathbb{C}$ tel que $|z| < R_b$. D'après le théorème b), la série $\sum b_n z^n$ est absolument convergente. Or par hypothèse, on a pour n assez grand $|a_n| \leq |b_n|$, donc $|a_n z^n| \leq |b_n z^n|$. On déduit avec le théorème 4.2.1.b que la série $\sum |a_n z^n|$ est convergente, ce qui avec le théorème b), implique $|z| \leq R_a$. On a ainsi prouvé que $[0, R_b[\subseteq [0, R_a[$, c'est-à-dire $R_b \leq R_a$.

Le second point se déduit du premier en considérant la série $\sum M b_n z^n = M \sum b_n z^n$ au lieu de $\sum b_n z^n$, en observant que ces deux séries ont le même rayon de convergence pour $M \in \mathbb{R}_+^*$.

Enfin, si $|a_n| \sim |b_n|$ pour n au voisinage de l'infini, on a à la fois $a_n = O(b_n)$ et $b_n = O(a_n)$ pour n au voisinage de l'infini, d'où en appliquant le second point $R_b \leq R_a$ et $R_a \leq R_b$, et finalement $R_a = R_b$. \square

f) PROPOSITION (calcul du rayon de convergence en utilisant la règle de d'Alembert). Soit $\sum a_n z^n$ une série entière. On suppose que $a_n \neq 0$ à partir d'un certain rang et que :

$$\lim_{n \rightarrow +\infty} \left| \frac{a_{n+1}}{a_n} \right| = \ell, \text{ avec } \ell \in \mathbb{R}_+ \text{ ou } \ell = +\infty.$$

Alors le rayon de convergence de la série $\sum a_n z^n$ est $R = \frac{1}{\ell}$, (avec les conventions $\frac{1}{0} = +\infty$ et $\frac{1}{+\infty} = 0$).

Preuve. Pour tout $z \in \mathbb{C}^*$, on a $\lim_{n \rightarrow +\infty} \left| \frac{a_{n+1} z^{n+1}}{a_n z^n} \right| = \lim_{n \rightarrow +\infty} \left| \frac{a_{n+1}}{a_n} \right| |z| = \ell |z| \in \mathbb{R}_+ \cup \{+\infty\}$.

Si $|z| < \frac{1}{\ell}$, alors $\ell |z| < 1$, donc en appliquant la règle de d'Alembert 4.2.3.a, la série à termes réels positifs $\sum |a_n z^n|$ est convergente. Si $|z| > \frac{1}{\ell}$, on déduit de même que $\sum |a_n z^n|$ est divergente. Ainsi la série entière $\sum a_n z^n$ est absolument convergente pour $|z| < \frac{1}{\ell}$ et non absolument convergente pour $|z| > \frac{1}{\ell}$; cela prouve avec le théorème 6.1.1.b que le rayon de convergence de la série $\sum a_n z^n$ est $R = \frac{1}{\ell}$. \square

► **REMARQUE.** Lorsque $\left| \frac{a_{n+1}}{a_n} \right|$ n'admet pas de limite finie ou infinie pour n tendant vers $+\infty$, cette proposition est inapplicable. Par exemple pour la série entière $\sum (\sin n) z^n$ étudiée par d'autres méthodes précédemment.

► **REMARQUE.** Cette proposition est inapplicable aussi pour les séries entières "lacunaires" du type $\sum a_{2n} z^{2n}$, ou $\sum a_{2n+1} z^{2n+1}$, ou $\sum a_{\sigma(n)} z^{\sigma(n)}$, dans la mesure où il est faux pour de telles séries que tous les a_n sont non-nuls à partir d'un certain rang.

► **EXEMPLE D'APPLICATION.** Toute série entière $\sum a_n z^n$ où a_n est de la forme $F(n)$ pour une certaine fraction rationnelle non-nulle $F \in \mathbb{C}(X)$ est de rayon de convergence égale à 1.

En effet. Notons $F(X) = \frac{\alpha_p X^p + \dots + \alpha_1 X + \alpha_0}{\beta_q X^q + \dots + \beta_1 X + \beta_0}$ avec p le degré du polynôme du numérateur et q le degré du polynôme du dénominateur (et donc $\alpha_p \in \mathbb{C}^*$ et $\beta_q \in \mathbb{C}^*$). Il est clair alors que la suite de terme général $a_n = F(n)$ vérifie $|a_n| \sim \left| \frac{\alpha_p}{\beta_q} \right| n^{p-q}$. D'après le troisième point de la proposition e) ci-dessus, le rayon de convergence R de la série entière $\sum a_n z^n$ est égal au rayon de convergence de la série entière $\sum n^{p-q} z^n$.

Pour cette dernière série entière, on applique la règle de d'Alembert ci-dessus : on a

$$\lim_{n \rightarrow +\infty} \left| \frac{(n+1)^{p-q}}{n^{p-q}} \right| = \lim_{n \rightarrow +\infty} \left(1 + \frac{1}{n} \right)^{p-q} = 1, \text{ donc } R = 1.$$

Ainsi par exemple les séries entières $\sum z^n$, $\sum n z^n$, $\sum (3n^2 + 5n - 2) z^n$, $\sum \frac{1}{n} z^n$, $\sum \frac{n^2 + 3n - 1}{n - 3} z^n$, $\sum \frac{2n}{n^3 - 1} z^n, \dots$ ont toutes pour rayon de convergence 1.

On termine par un dernier exemple de calcul de rayon de convergence, appliqué à une notion qui sera utile par la suite, celle de série dérivée.

g) DÉFINITION ET PROPOSITION (*calcul du rayon de convergence de la série dérivée*). On appelle série dérivée d'une série entière $\sum_{n \geq 0} a_n z^n$ la série $\sum_{n \geq 0} n a_n z^{n-1}$. Ces deux séries entières ont le même rayon de convergence.

Preuve. Soit R le rayon de convergence de $\sum a_n z^n$. Soit R' le rayon de convergence de $\sum n a_n z^{n-1}$. Il est clair que la série $\sum n a_n z^{n-1}$ converge dans \mathbb{C} si et seulement si la série $\sum n a_n z^n = z \times \sum n a_n z^{n-1}$ converge. Donc R' est aussi le rayon de convergence de la série entière $\sum n a_n z^n$.

Pour tout $n \in \mathbb{N}$ et tout $z \in \mathbb{C}$, $|a_n z^n| \leq |n a_n z^n|$. Donc si la série $\sum n a_n z^n$ converge absolument en un point z , il en est de même de la série $\sum a_n z^n$. Ceci prouve que $R' \leq R$.

Soit $z \in \mathbb{C}$ tel que $0 \leq |z| < R$. Il existe un réel $r \geq 0$ tel que $|z| < r < R$. Dès lors, on a pour tout $n \in \mathbb{N}$:

$$|n a_n z^n| = |a_n r^n| \times n \left(\frac{|z|}{r}\right)^n \text{ avec } \begin{cases} (a_n r^n) \text{ suite bornée,} & \text{car } 0 \leq r < R \\ \lim_{n \rightarrow +\infty} n \left(\frac{|z|}{r}\right)^n = 0, & \text{car } 0 \leq \frac{|z|}{r} < 1. \end{cases}$$

Il en résulte que $\lim_{n \rightarrow +\infty} n a_n z^n = 0$, ce qui prouve que $|z| \leq R'$. On a ainsi montré que tout $z \in \mathbb{C}$ tel que $|z| < R$ vérifie aussi $|z| \leq R'$. Ceci implique $R \leq R'$. Et finalement $R = R'$. \square

6.1.2 Convergence normale d'une série entière

Sur le disque de convergence, une série entière converge absolument. L'objet de ce qui suit est de donner des précisions en termes de convergence normale et/ou uniforme. Le résultat suivant, bien que de nature a priori théorique, est d'une grande importance dans les applications.

a) LEMME FONDAMENTAL. Soient $\sum_{n \geq 0} a_n z^n$ une série entière et R son rayon de convergence. La série de fonctions associée converge normalement sur toute boule fermée incluse dans le disque ouvert $D(0, R)$.

Preuve. Comme il est d'usage, on note encore $\sum a_n z^n$ la série de fonctions $\sum f_n$ où f_n est la fonction $D(0, R) \rightarrow \mathbb{C}$ définie par $f_n(z) = a_n z^n$.

Soit K une boule fermée incluse dans $D(0, R)$.

Il existe une boule fermée $\overline{D}(0, r)$ centrée en 0 telle que $K \subseteq \overline{D}(0, r) \subseteq D(0, R)$. Puisque $0 \leq r < R$, la série de réels positifs $\sum |a_n| r^n$ est convergente.

De plus, quel que soit $z \in K$, on a $z \in \overline{D}(0, r)$, donc $|z^n| = |z|^n \leq r^n$, et finalement $|a_n z^n| \leq |a_n| r^n$.

Par définition même de la convergence normale (voir p.107), ceci prouve que la série de fonctions $\sum a_n z^n$ converge normalement sur K . \square

Figure 2

b) REMARQUES.

- Dès lors qu'il existe un $z_0 \in \mathbb{C}$ tel que la série numérique $\sum a_n z_0^n$ est absolument convergente, alors la série de fonctions $\sum a_n z^n$ converge normalement sur la boule fermée $\overline{D}(0, |z_0|)$.
- Une série entière de rayon de convergence R peut ne pas converger normalement, ni même uniformément, sur le disque ouvert $D(0, R)$. C'est le cas par exemple de $\sum z^n$, avec $R = 1$.
- On peut dans le lemme remplacer "converge normalement sur toute boule fermée incluse dans le disque ouvert $D(0, R)$ " par "converge normalement sur toute partie fermée et bornée incluse dans le disque ouvert $D(0, R)$ "

6.1.3 Fonctions définies par la somme d'une série entière

a) PREMIER EXEMPLE PRÉLIMINAIRE (*série exponentielle*). Considérons la série entière $\sum \frac{1}{n!} z^n$. Le rapport $|\frac{a_{n+1}}{a_n}| = \frac{n!}{(n+1)!} = \frac{1}{n+1}$ tend vers la limite $\ell = 0$ quand $n \rightarrow +\infty$. Donc, d'après la proposition 6.1.1.f, le rayon de convergence R est $+\infty$. Il résulte donc du théorème 6.1.1.b que la série $\sum \frac{1}{n!} z^n$ converge absolument pour tout $z \in \mathbb{C}$, et donc que sa somme définit une fonction $\mathbb{C} \rightarrow \mathbb{C}$ appelée la fonction exponentielle complexe, notée e^z ou $\exp z$. On retiendra que :

$$\text{pour tout } z \in \mathbb{C}, \quad e^z = \exp z = \sum_{n=0}^{+\infty} \frac{1}{n!} z^n = 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24} + \dots$$

Rappelons que l'on a déjà démontré en 4.4.1.d, en utilisant la notion de produit de deux séries absolument convergentes, que

$$(\exp z)(\exp z') = \exp(z + z') \text{ pour tous } z, z' \in \mathbb{C}.$$

b) SECOND EXEMPLE PRÉLIMINAIRE (*série entière géométrique*). Soit $a \in \mathbb{C}^*$. Considérons la série entière $\sum a^n z^n$. C'est la série géométrique $\sum (az)^n$ de raison az , dont on sait qu'elle est convergente si et seulement si $|az| < 1$, et que sa somme est alors $\frac{1}{1-az}$. On en déduit que : le rayon de convergence de la série entière $\sum a^n z^n$ est $\frac{1}{|a|}$. Il résulte donc du théorème 6.1.1.b que la série $\sum a^n z^n$ converge absolument pour tout z tel que $|z| < \frac{1}{|a|}$, et donc que sa somme définit une fonction $D(0, \frac{1}{|a|}) \rightarrow \mathbb{C}$.

Les cas $a = 1$ et $a = -1$ sont particulièrement utiles dans la pratique. On retiendra que :

$$\begin{aligned} \text{pour tout } z \in D(0, 1), \quad \frac{1}{1-z} &= \sum_{n=0}^{+\infty} z^n = 1 + z + z^2 + z^3 + z^4 + \dots \\ \text{pour tout } z \in D(0, 1), \quad \frac{1}{1+z} &= \sum_{n=0}^{+\infty} (-1)^n z^n = 1 - z + z^2 - z^3 + z^4 + \dots \end{aligned}$$

c) THÉORÈME (*continuité de la fonction définie par la somme d'une série entière*). Soient $\sum_{n \geq 0} a_n z^n$ une série entière et R son rayon de convergence. L'application $f : D(0, R) \rightarrow \mathbb{C}$ définie par $f(z) = \sum_{n=0}^{+\infty} a_n z^n$ est continue sur le disque ouvert $D(0, R)$.

Preuve. Le résultat étant trivial pour $R = 0$, on peut supposer $R > 0$. Soit $z_0 \in D(0, R)$. Il existe un réel positif r tel que $\overline{D}(z_0, r) \subset D(0, R)$. D'après le lemme 6.1.2.a, la série de fonctions $\sum a_n z^n$ converge normalement sur $\overline{D}(z_0, r)$. Or chaque application $z \mapsto a_n z^n$ est évidemment continues sur \mathbb{C} donc sur $\overline{D}(z_0, r)$. Il suffit d'appliquer le dernier point rappelé p.107 pour conclure que l'application f définie par la somme de la série entière est continue sur $\overline{D}(z_0, r)$. En particulier f est continue en z_0 , et ceci étant fait pour tout $z_0 \in D(0, R)$, on conclut que f est continue sur $D(0, R)$. \square

De façon naturelle, on cherche à étudier de même la dérivabilité de la fonction f définie par la somme d'une série entière. Un obstacle majeur est que l'on n'a pas dans le cadre de ce cours de théorie de la dérivation pour les fonctions d'une variable complexe. On va donc dans ce qui suit se limiter au cas d'une variable réelle. Plus précisément, on considère, pour une série entière $\sum a_n z^n$ avec (a_n) suite de nombre complexes, la restriction à \mathbb{R} de la fonction $f : D(0, R) \rightarrow \mathbb{C}$ définie par la somme de cette série. On obtient donc une fonction encore notée f définie par :

$$f :]-R, R[\rightarrow \mathbb{C}, \quad x \mapsto f(x) = \sum_{n=0}^{+\infty} a_n x^n$$

d) THÉORÈME (dérivabilité de la fonction définie par la somme d'une série entière dans le cas réel). Soient $\sum_{n \geq 0} a_n z^n$ une série entière avec (a_n) est une suite de nombres complexes et R son rayon de convergence. On suppose $R > 0$.

(i) L'application $f :]-R, R[\rightarrow \mathbb{C}$ définie par $f(x) = \sum_{n=0}^{+\infty} a_n x^n$ est dérivable sur l'intervalle ouvert $]-R, R[$, et sa dérivée est :

$$f'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1} \quad \text{pour tout } x \in]-R, R[.$$

(ii) Plus généralement, f est de classe C^∞ sur $]-R, R[$ et l'on a :

$$f^{(k)}(x) = \sum_{n=k}^{+\infty} \frac{n!}{(n-k)!} a_n x^{n-k} \quad \text{pour tout } z \in]-R, R[\text{ et tout } k \in \mathbb{N}.$$

(iii) En particulier, pour tout entier $k \geq 0$, on a : $a_k = \frac{1}{k!} f^{(k)}(0)$.

Preuve. D'après 6.1.1.g, le rayon de convergence de la série entière $\sum n a_n z^{n-1}$ est R . Donc d'après le lemme 6.1.2.a, la série de fonctions $\sum n a_n z^{n-1}$ converge normalement donc uniformément sur toute boule fermée K incluse dans $D(0, R)$. En prenant les restrictions à \mathbb{R} , on en déduit que la série de fonctions $\sum n a_n x^{n-1}$ converge normalement donc uniformément sur tout intervalle $[a, b]$ inclus dans $]-R, R[$. Il suffit alors d'appliquer le théorème de dérivation pour les séries de fonctions formulé en [5] de 5.2.1.b pour obtenir le point (i). Le point (ii) s'obtient en réitérant par une récurrence évidente. Le point (iii) est une conséquence immédiate du (ii). \square

e) REMARQUES ET EXEMPLES. Avec les données et notations ci-dessus, la fonction f est de classe C^∞ sur l'intervalle $]-R, R[$ et l'on a pour tout $x \in]-R, R[$:

$$\begin{aligned} f(x) &= \sum_{n=0}^{+\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + \dots + a_n x^n + \dots \\ f'(x) &= \sum_{n=1}^{+\infty} n a_n x^{n-1} = a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3 + \dots + n a_n x^{n-1} + \dots \\ f''(x) &= \sum_{n=2}^{+\infty} n(n-1) a_n x^{n-2} = 2a_2 + 6a_3 x + 12a_4 x^2 + \dots + n(n-1) a_n x^{n-2} + \dots \\ &\text{etc...} \end{aligned}$$

► Reprenons l'exemple de la série exponentielle traité au a). Appliquons-lui le théorème d), avec $R = +\infty$. L'application exponentielle, définie par : $e^x = \sum_{n=0}^{+\infty} \frac{1}{n!} x^n = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \dots$, est de classe C^∞ sur $]-\infty, +\infty[$, et sa dérivée sur \mathbb{R} est :

$$(e^x)' = 0 + 1 + 2\frac{x}{2} + 3\frac{x^2}{6} + 4\frac{x^3}{24} + \dots = \sum_{n=1}^{+\infty} n \frac{1}{n!} x^{n-1} = \sum_{n=1}^{+\infty} \frac{1}{(n-1)!} x^{n-1} = \sum_{n=0}^{+\infty} \frac{1}{n!} x^n = e^x.$$

On retrouve le fait que : $e^x = (e^x)' = (e^x)'' = \dots = (e^x)^{(n)} = \dots$ pour tout $x \in \mathbb{R}$.

► Reprenons l'exemple de la série géométrique $\sum x^n$ traité au b). Appliquons-lui le théorème d), avec $R = 1$. L'application $\frac{1}{1-x} = \sum_{n=0}^{+\infty} x^n = 1 + x + x^2 + x^3 + x^4 + \dots$ est de classe C^∞ sur $]-1, 1[$, et sa dérivée est : $\frac{1}{(1-x)^2} = \left(\frac{1}{1-x}\right)' = 0 + 1 + 2x + 3x^2 + 4x^3 + \dots = \sum_{n=1}^{+\infty} n x^{n-1}$. On conclut que, pour tout $x \in]-1, 1[$:

$$\frac{1}{(1-x)^2} = \sum_{n=0}^{+\infty} (n+1)x^n, \quad \text{et plus généralement } \frac{1}{(1-x)^{k+1}} = \sum_{n=0}^{+\infty} \binom{n+k}{k} x^n.$$

f) COROLLAIRE (primitive de la fonction définie par la somme d'une série entière). Soient $\sum_{n \geq 0} a_n z^n$ une série entière avec (a_n) est une suite de nombres complexes et R son rayon de convergence. On suppose $R > 0$. Soit $f :]-R, R[\rightarrow \mathbb{C}$ définie par $f(x) = \sum_{n=0}^{+\infty} a_n x^n$. Alors la série entière $\sum_{n \geq 0} \frac{1}{n+1} a_n z^{n+1}$ a pour rayon de convergence R et l'application $F :]-R, R[\rightarrow \mathbb{C}$ définie par $f(x) = \sum_{n=0}^{+\infty} \frac{1}{n+1} a_n x^{n+1}$ est la primitive de f s'annulant en 0.

Preuve. On applique le théorème précédent à F . □

► Reprenons l'exemple de la série géométrique $\sum x^n$ traité au b). Appliquons-lui le corollaire f), avec $R = 1$. On déduit que, pour tout $x \in]-1, 1[$:

$$-\ln(1-x) = \sum_{n=1}^{+\infty} \frac{x^n}{n}.$$

g) REMARQUE. Considérons deux séries entières $\sum a_n x^n$ et $\sum b_n x^n$, de rayon de convergence respectifs R et R' . Soit $f :]-R, R[\rightarrow \mathbb{C}$ et $g :]-R', R'[\rightarrow \mathbb{C}$ les fonctions d'une variable réelle respectivement définies par ces séries.

• S'il existe $\varepsilon > 0$ tel que $f(x) = g(x)$ pour tout $x \in]-\varepsilon, \varepsilon[$, alors $a_n = b_n$ pour tout $n \geq 0$.

En effet. C'est une conséquence évidente du point (iii) du théorème d), puisqu'on a alors $f^{(k)}(0) = g^{(k)}(0)$ pour tout entier $k \geq 0$. □

• f est paire si et seulement si $a_{2p+1} = 0$ pour tout $p \geq 0$. f est impaire si et seulement si $a_{2p} = 0$ pour tout $p \geq 0$.

En effet. La condition est clairement suffisante. Si l'on suppose réciproquement que f est paire, alors $f(x) = f(-x)$ pour tout $x \in]-R, R[$. Comme $x \mapsto f(-x)$ est la fonction définie par la somme de la série entière $\sum (-1)^n a_n x^n$, il résulte de la remarque précédente que $a_n = (-1)^n a_n$ pour tout $n \geq 0$, d'où $a_n = 0$ si n est impair. La preuve est identique si l'on suppose f impaire. □

6.1.4 Cas réel : application aux équations différentielles

On cherche à déterminer, pour une équation différentielle donnée, s'il existe des solutions qui sont la somme d'une série entière sur un intervalle $]-R, R[$ de \mathbb{R} .

EXEMPLE. Considérons l'équation différentielle $4xy''(x) + 2y'(x) - y(x) = 0$.

On cherche une solution $y(x)$ qui soit la somme d'une série entière $\sum a_n x^n$. On a :

$$y(x) = \sum_{n=0}^{+\infty} a_n x^n, \quad y'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1}, \quad y''(x) = \sum_{n=2}^{+\infty} n(n-1) a_n x^{n-2}.$$

$$\begin{aligned} \text{Donc } 4xy''(x) + 2y'(x) - y(x) &= \sum_{n=2}^{+\infty} 4n(n-1)a_n x^{n-1} + \sum_{n=1}^{+\infty} 2n a_n x^{n-1} - \sum_{n=0}^{+\infty} a_n x^n \\ &= \sum_{m=1}^{+\infty} 4(m+1)m a_{m+1} x^m + \sum_{m=0}^{+\infty} 2(m+1)a_{m+1} x^m - \sum_{m=0}^{+\infty} a_m x^m \\ &= (2a_1 - a_0) + \sum_{m=1}^{+\infty} [4(m+1)m a_{m+1} + 2(m+1)a_{m+1} - a_m] x^m. \end{aligned}$$

Il en résulte que $y(x)$ est solution de l'équation différentielle de départ si et seulement si :

$$a_1 = \frac{1}{2}a_0 \quad \text{et} \quad a_{n+1} = \frac{1}{(2n+1)(2n+2)} a_n \quad \text{pour tout } n \geq 1.$$

Imposons de plus la condition initiale $y(0) = 1$. Cela signifie que $a_0 = 1$. D'où l'on tire :

$$a_1 = \frac{1}{2}, \quad a_2 = \frac{1}{3 \times 4} \times \frac{1}{2} = \frac{1}{4!}, \quad a_3 = \frac{1}{5 \times 6} \times \frac{1}{4!} = \frac{1}{6!}, \quad \text{et par récurrence } a_n = \frac{1}{(2n)!}.$$

On conclut que $y(x) = \sum_{n=0}^{+\infty} \frac{1}{(2n)!} x^n$, dont le rayon de convergence est clairement $+\infty$.

REMARQUE. – On définit naturellement à partir de la série exponentielle les deux séries entières :

$$\operatorname{ch} z = \frac{1}{2}(e^z + e^{-z}) = \sum_{n=0}^{+\infty} \frac{1}{(2n)!} z^{2n} \quad \text{et} \quad \cos z = \frac{1}{2}(e^{iz} + e^{-iz}) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n)!} z^{2n},$$

dont le rayon de convergence est $+\infty$. Donc, pour tout $x \in \mathbb{R}$, la solution $y(x)$ trouvée s'exprime comme :

$$y(x) = \sum_{n=0}^{+\infty} \frac{1}{(2n)!} x^n = \begin{cases} \sum_{n=0}^{+\infty} \frac{1}{(2n)!} (\sqrt{x})^{2n} = \operatorname{ch} \sqrt{x} & \text{si } x \geq 0, \\ \sum_{n=0}^{+\infty} \frac{1}{(2n)!} (-\sqrt{-x})^{2n} = \cos \sqrt{-x} & \text{si } x \leq 0. \end{cases}$$

Cette dernière remarque nous introduit à la question qui fait l'objet de la partie 2 de ce chapitre : savoir si, réciproquement à ce qu'on vient de faire dans la partie 1, une fonction donnée est ou non la somme d'une série entière sur un certain intervalle.

6.1.5 Exercices

EXERCICE 1. Dans chacun des cas suivants, déterminer le rayon de convergence de la série entière $\sum_{n \geq 1} a_n z^n$.

$$a_n = \frac{n^n}{n!}, \quad a_n = \frac{\operatorname{sh} n}{\operatorname{ch}^2 n}, \quad a_n = (\cos \frac{1}{n})^{n^\alpha}, \quad a_n = (1 + \frac{(-1)^n}{n})^{n^2}.$$

EXERCICE 2. Déterminer le rayon de convergence des séries entières :

$$\sum_{n \geq 0} \frac{n^2+1}{3^n} z^n, \quad \sum_{n \geq 0} e^{-n^2} z^n, \quad \sum_{n \geq 1} \frac{\ln n}{n^2} z^{2n}, \quad \sum_{n \geq 0} \frac{n^n}{n!} z^{3n}.$$

EXERCICE 3. Déterminer le rayon de convergence des séries entières :

$$\sum_{n \geq 0} n! z^n, \quad \sum_{n \geq 0} \binom{2n}{n} z^n, \quad \sum_{n \geq 1} \frac{(3n)!}{(n!)^3} z^n, \quad \sum_{n \geq 0} ((n+1)^{\frac{1}{n+1}} - n^{\frac{1}{n}}) z^n.$$

EXERCICE 4. Déterminer le rayon de convergence de : $\sum_{n \geq 0} z^{n^2}$, $\sum_{n \geq 0} \sin(n) z^n$, $\sum_{n \geq 1} \frac{\sin n}{n^2} z^n$.

EXERCICE 5. Soit $\sum_{n \geq 0} a_n z^n$ une série entière (avec $a_n \in \mathbb{C}$), de rayon de convergence R .

– Calculer R en fonction des rayons de convergence respectifs R_1 et R_2 des séries entières $\sum_{p \geq 0} a_{2p} z^{2p}$ et $\sum_{p \geq 0} a_{2p+1} z^{2p+1}$.

– On suppose $R \neq 0$. Montrer que la série $\sum_{n \geq 0} \frac{a_n}{n!} z^n$ converge pour tout $z \in \mathbb{C}$.

– On suppose $R \neq 0$. On note $s_n = a_0 + a_1 + \dots + a_n \in \mathbb{C}$ pour tout entier $n \geq 0$. Déterminer le rayon de convergence R' de la série entière $\sum_{n \geq 0} s_n z^n$, en discutant suivant que $R > 1$ ou $R \leq 1$.

EXERCICE 6. Soit $\sum_{n \geq 0} a_n z^n$ et $\sum_{n \geq 0} b_n z^n$ deux séries entières complexes, de rayons de convergence respectifs R_1 et R_2 . Montrer que le rayon de convergence de la série entière $\sum_{n \geq 0} a_n b_n z^n$ (produit de Hadamard) vérifie $R \geq R_1 R_2$. Donner un exemple où l'inégalité est stricte.

EXERCICE 7. Soit $\sum_{n \geq 0} a_n z^n$ une série entière (avec $a_n \in \mathbb{C}$), de rayon de convergence R . On introduit la série entière $\sum_{n \geq 0} b_n$, avec $b_n = \frac{a_n}{1+|a_n|}$, de rayon de convergence R' . Montrer que $R' \geq \max(1, R)$. Vérifier que si $R' > 1$, alors $R' = R$. Exprimer R' en fonction de R .

EXERCICE 8. Soit $\alpha \in \mathbb{R}$. Quel est le rayon de convergence de la série entière réelle $\sum \frac{\cos(n\alpha)}{n} x^n$? (Indication : on pourra considérer la série dérivée).

EXERCICE 9. Déterminer le rayon de convergence de la série entière complexe $\sum_{n \geq 0} (-1)^n z^{2n}$ et calculer sa somme. En déduire que $\arctan x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} x^{2n+1}$ pour tout $x \in]-1, +1[$.

EXERCICE 10. Montrer que $\sum_{n=p}^{+\infty} \binom{n}{p} x^{n-p} = \frac{1}{(1-x)^{p+1}}$ pour tous $p \in \mathbb{N}$ et $x \in]-1, 1[$.

Application : soient $a, b, c \in \mathbb{R}$ et f la fonction polynomiale $x \mapsto a + bx + cx^2$; déterminer le rayon de convergence R de la série entière $\sum_{n \geq 0} f(n)x^n$ et calculer sa somme sur $]-R, R[$.

EXERCICE 11. Déterminer le rayon de convergence des séries entières réelles :

$$y_0(x) = \sum_{p \geq 0} \frac{(-1)^p}{2 \cdot 4 \cdot 6 \cdots 2p} x^{2p} \quad \text{et} \quad y_1(x) = \sum_{p \geq 0} \frac{(-1)^p}{1 \cdot 3 \cdot 5 \cdots (2p+1)} x^{2p+1}.$$

Soit (E) l'équation différentielle : $y''(x) + xy'(x) + y(x) = 0$. Soit $y(x)$ la somme d'une série entière $\sum_{n \geq 0} a_n x^n$ de rayon de convergence infini. Montrer que $y(x)$ est solution de (E) sur \mathbb{R} si et seulement si : $y(x) = a_0 y_0(x) + a_1 y_1(x)$.

EXERCICE 12. Pour tout entier $n \geq 1$, on note d_n le nombre de diviseurs positifs de n et s_n la somme de ces diviseurs. Calculer le rayon de convergence des séries entières $\sum_{n \geq 1} d_n z^n$ et $\sum_{n \geq 1} s_n z^n$. (Indication : on pourra introduire la série entière $\sum nz$).

EXERCICE 13. Pour tout entier $n \geq 1$, on pose $a_n = \int_1^{+\infty} e^{-tn} dt$. Montrer a_n est bien définie, que la suite (a_n) converge dans \mathbb{R}_+ , et que $a_n \sim K \frac{1}{n}$ pour une constante K que l'on déterminera. Déterminer le rayon de convergence de la série $\sum a_n z^n$.

EXERCICE 14. On considère la série entière complexe $S(z) = \sum_{n \geq 2} \left(\frac{1}{n \ln n}\right) z^n$. Etudier la nature des séries numériques $S(1)$ et $S(-1)$; en déduire le rayon de convergence R de $S(z)$. Déterminer la nature de $S(z)$ en tout point z du cercle $C(0, R)$.

Indication. On pourra utiliser le suivant (dit parfois lemme d'Abel) : Soit $\sum u_n$ une série numérique telle que la suite des sommes partielles soit bornée ; soit (α_n) ne suite décroissante de réels positifs qui converge vers 0. Alors la série $\sum \alpha_n u_n$ est convergente.

6.2 Développement en séries entières

6.2.1 Fonction développable en série entière.

a) DÉFINITION. Soit f une fonction d'une variable réelle x définie sur un intervalle I de \mathbb{R} contenant 0. On dit que f est *développable en série entière centrée en 0* lorsqu'il existe un intervalle $]-\alpha, \alpha[$ centré en 0 inclus dans I , et une série entière $\sum a_n z^n$ de rayon de convergence $R \geq \alpha$, tels que :

$$f(x) = \sum_{n=0}^{+\infty} a_n x^n \quad \text{pour tout } x \in]-\alpha, \alpha[.$$

D'après le théorème 6.1.3.d la fonction f est de classe C^∞ sur $]-\alpha, \alpha[$, et l'on a :

$$a_0 = f(0), \quad a_1 = f'(0), \quad a_2 = \frac{1}{2} f''(0), \quad \dots, \quad a_n = \frac{1}{n!} f^{(n)}(0), \quad \dots$$

ce qui prouve en particulier, comme on l'a vu en 6.1.3.g l'unicité du développement en série entière (s'il existe). A noter (avec la même remarque) que le développement en série entière d'une fonction paire (respectivement impaire) ne contient que des termes $a_n x^n$ avec n pair (respectivement impair).

b) CONVENTION TERMINOLOGIQUE. Dans toute la suite, "développable en série entière" signifiera "développable en série entière centrée en 0".

Plus généralement, pour tout $a \in \mathbb{R}$, une fonction f définie sur un intervalle I de \mathbb{R} contenant a est dite développable en série entière centrée en a lorsque la fonction $x \mapsto f(x+a)$ est développable en série entière centrée en 0; c'est pourquoi on se ramène à l'étude du cas $a = 0$.

c) EXEMPLES ET CONTRE-EXEMPLE. Comme on l'a vu en 6.1.3,a et b, les fonction $x \mapsto e^x$ et $x \mapsto \frac{1}{1-ax}$ sont développables en série entière.

Contre-exemple : une fonction peut être de classe C^∞ sur un intervalle I sans être développable en série entière. Prenons $f : \mathbb{R} \rightarrow \mathbb{R}$ définie par $f(x) = e^{-1/x^2}$ si $x \neq 0$ et $f(0) = 0$.

L'application f est de classe C^∞ sur $]-\infty, 0[$ et sur $]0, +\infty[$, et pour tout $x \in \mathbb{R}^*$, $f^{(n)}(x)$ est de la forme $\frac{P_n(x)}{x^3} e^{-1/x^2}$, d'où par application du théorème limite de la dérivée, on déduit que f est de classe C^∞ sur \mathbb{R} , avec $f^{(n)}(0) = 0$. Dès lors, si f était développable en série entière, il existerait $\alpha > 0$ tel que $f(x) = \sum_{n=0}^{+\infty} \frac{1}{n!} f^{(n)}(0) x^n = 0$ pour tout $x \in]-\alpha, \alpha[$, ce qui n'est pas vrai puisque f ne s'annule qu'en 0.

d) CONDITION NÉCESSAIRE ET SUFFISANTE POUR L'EXISTENCE D'UN DÉVELOPPEMENT EN SÉRIE ENTIÈRE. Rappelons d'abord le théorème suivant :

THÉORÈME (formule de Taylor avec reste intégral). *Pour toute application $f :]-\alpha, \alpha[\rightarrow \mathbb{C}$ de classe C^∞ , on a pour tout entier $N \geq 1$ et tout $x \in]-\alpha, \alpha[$:*

$$\begin{aligned} f(x) &= f(0) + f'(0)x + \frac{1}{2!}f''(0)x^2 + \dots + \frac{1}{N!}f^{(N)}(0)x^N + R_N(x) \\ &= \sum_{n=0}^N \frac{1}{n!}f^{(n)}(0)x^n + R_N(x), \end{aligned}$$

où le reste $R_N(x)$ peut s'exprimer sous la forme $R_N(x) = \int_0^x \frac{1}{N!}(x-t)^N f^{(N+1)}(t) dt$.

Sans revenir sur ce résultat vu en première année, on en déduit le résultat suivant :

PROPOSITION. *Avec les données et hypothèses ci-dessus, les conditions suivantes sont équivalentes :*

- (i) f est développable en série entière ;
- (ii) il existe $0 < \beta \leq \alpha$ tel que $\lim_{N \rightarrow +\infty} R_N(x) = 0$ pour tout $x \in]-\beta, \beta[$;
- (iii) il existe $0 < \gamma \leq \alpha$, $A > 0$, $C > 0$ tels que $|f^{(n)}(x)| \leq CA^n n!$ pour tous $x \in]-\gamma, \gamma[$ et $n \in \mathbb{N}$.

Preuve. Supposons que l'on ait (iii). Pour tous $x \in]-\gamma, \gamma[$ et $n \in \mathbb{N}$, on a :

$$\begin{aligned} |R_n(x)| &= \left| \int_0^x \frac{1}{n!}(x-t)^n f^{(n+1)}(t) dt \right| \leq \int_0^x \left| \frac{1}{n!}(x-t)^n \right| |f^{(n+1)}(t)| dt \\ &\leq CA^{n+1}(n+1)! \left| \int_0^x \frac{1}{n!}(x-t)^n \right| = CA^{n+1}|x|^{n+1} \end{aligned}$$

Posons $\beta = \min(\gamma, \frac{1}{A})$. Pour tout $x \in]-\beta, \beta[$, on a $|Ax| < 1$ donc $\lim_{n \rightarrow +\infty} R_n(x) = 0$. Ce qui prouve (ii).

Supposons que l'on ait (ii). Pour tout $x \in]-\beta, \beta[$, on a $\sum_{k=0}^n \frac{1}{k!} f^{(k)}(0)x^k = f(x) - R_n(x)$.

Comme $\lim_{n \rightarrow +\infty} R_n(x) = 0$, on en déduit que $\lim_{n \rightarrow +\infty} \sum_{k=0}^n \frac{1}{k!} f^{(k)}(0)x^k = f(x)$, ce qui prouve que la série entière $\sum \frac{1}{k!} f^{(k)}(0)x^k$ est de rayon de convergence $R \geq \beta > 0$, et donc que f est développable en série entière. La dernière implication est laissée au lecteur en exercice. \square

e) MÉTHODES DE CALCUL. Un raisonnement "théorique" (entre autres celui de la proposition précédente) permettant d'établir l'existence d'un développement en série entière pour certaines fonctions de références (exponentielle, puissances,...), on se ramène à ces fonctions de références en utilisant les propriétés suivantes (dont les énoncés plus explicites et les preuves seront vus en exercices) :

1. la somme de deux fonctions f et g développables en série entière est développable en série entière, et le développement de $f+g$ s'obtient en faisant la somme des deux développements ;
2. idem pour le produit fg ;
3. la dérivée d'une fonction f développable en série entière est développable en série entière, et le développement de f' s'obtient en dérivant terme à terme le développement de f ;
4. idem pour une primitive.

6.2.2 Exemples classiques.

a) PREMIÈRE SÉRIE D'EXEMPLES. Le résultat central (que l'on a déjà démontré en 6.1.3.a) est que la fonction exponentielle est développable en série entière, avec un rayon de convergence infini :

$$e^x = \exp x = \sum_{n=0}^{+\infty} \frac{1}{n!} x^n = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{24}x^4 + \dots \quad \text{pour tout } x \in]-\infty, +\infty[$$

Les fonctions hyperboliques étant définies par $\operatorname{ch}x = \frac{1}{2}(e^x + e^{-x})$ et $\operatorname{sh}x = \frac{1}{2}(e^x - e^{-x})$, les résultats sur les sommes de fonctions développables en séries entières donnent :

$$\begin{aligned} \operatorname{ch}x &= \sum_{n=0}^{+\infty} \frac{1}{(2n)!} x^{2n} = 1 + \frac{1}{2}x^2 + \frac{1}{24}x^4 + \frac{1}{720}x^6 + \dots & \text{pour tout } x \in]-\infty, +\infty[\\ \operatorname{sh}x &= \sum_{n=0}^{+\infty} \frac{1}{(2n+1)!} x^{2n+1} = x + \frac{1}{6}x^3 + \frac{1}{120}x^5 + \frac{1}{5040}x^7 + \dots & \text{pour tout } x \in]-\infty, +\infty[\end{aligned}$$

De même pour les fonctions trigonométriques $\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$ et $\sin x = \frac{1}{2i}(e^{ix} - e^{-ix})$:

$$\begin{aligned} \cos x &= \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 - \frac{1}{720}x^6 + \dots & \text{pour tout } x \in]-\infty, +\infty[\\ \sin x &= \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{1}{6}x^3 + \frac{1}{120}x^5 - \frac{1}{5040}x^7 + \dots & \text{pour tout } x \in]-\infty, +\infty[\end{aligned}$$

b) SECONDE SÉRIE D'EXEMPLES. Un autre résultat fondamental est que la fonction puissance $(1+x)^\alpha$ est développable en série entière pour tout $\alpha \in \mathbb{R}$, avec un rayon de convergence égal à 1 (infini dans le cas particulier où $\alpha \in \mathbb{N}$ comme on l'a vu en 6.1.3.b) :

$$(1+x)^\alpha = 1 + \sum_{n=1}^{+\infty} \frac{\alpha(\alpha-1)(\alpha-2)\dots(\alpha-n+1)}{n!} x^n \quad \text{pour tout } x \in]-1, 1[$$

Preuve. Soit $\alpha \in \mathbb{R}$. Posons $a_0 = 1$ et $a_n = \frac{\alpha(\alpha-1)(\alpha-2)\dots(\alpha-n+1)}{n!}$ pour tout $n \geq 1$. On a $\lim_{n \rightarrow +\infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n \rightarrow +\infty} \left| \frac{\alpha-n}{n+1} \right| = 1$. La série entière $\sum a_n z^n$ a donc 1 pour rayon de convergence, et l'on peut considérer la fonction $f_\alpha :]-1, 1[\rightarrow \mathbb{R}$ définie par $f(x) = \sum_{n=0}^{+\infty} a_n x^n$. On a alors $f'_\alpha(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1}$ pour tout $] -1, 1[\rightarrow \mathbb{R}$. Un calcul simple montre alors que

$$(1+x)f'_\alpha(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1} + \sum_{n=1}^{+\infty} n a_n x^n = \sum_{n=0}^{+\infty} \underbrace{((n+1)a_{n+1} + n a_n)}_{=\alpha a_n} x^n = \alpha f_\alpha(x).$$

Mais on sait par ailleurs que les solutions de l'équation différentielle $(1+x)y' - \alpha y = 0$ linéaire du premier ordre sont les applications de la forme $y(x) = \lambda e^{\alpha \ln(1+x)} = \lambda(1+x)^\alpha$ avec $\lambda \in \mathbb{R}$. Il existe donc $\lambda \in \mathbb{R}$ tel que $f_\alpha(x) = \lambda(1+x)^\alpha$ pour tout $x \in]-1, 1[$. Comme $f_\alpha(0) = a_0 = 1$, on a $\lambda = 1$, ce qui achève la preuve. \square

Pour $\alpha = -1$, on retrouve

$$\frac{1}{1+x} = \sum_{n=0}^{+\infty} (-1)^n x^n \quad \text{pour tout } x \in]-1, 1[$$

et donc en changeant x et $-x$, la formule

$$\frac{1}{1-x} = \sum_{n=0}^{+\infty} x^n \quad \text{pour tout } x \in]-1, 1[$$

En prenant leurs primitives s'annulant en 0, il vient

$$\ln(1+x) = \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} x^n \quad \text{et} \quad -\ln(1-x) = \sum_{n=1}^{+\infty} \frac{1}{n} x^n \quad \text{pour tout } x \in]-1, 1[$$

dont la demie-somme conduit à

$$\frac{1}{2} \ln\left(\frac{1+x}{1-x}\right) = \sum_{n=0}^{+\infty} \frac{1}{2n+1} x^{2n+1} \quad \text{pour tout } x \in]-1, 1[$$

En remplaçant x par x^2 dans $\frac{1}{1+x} = \sum_{n=0}^{+\infty} (-1)^n x^n$ et en prenant la primitive s'annulant en 0, on obtient :

$$\arctan x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} x^{2n+1} \quad \text{pour tout } x \in]-1, 1[$$

Reprenons maintenant la formule de départ avec $\alpha = -\frac{1}{2}$; il vient :

$$\frac{1}{\sqrt{1+x}} = 1 + \sum_{n=1}^{+\infty} (-1)^n \frac{1 \times 3 \times 5 \times \dots \times (2n-1)}{2 \times 4 \times 6 \times \dots \times 2n} x^n \quad \text{pour tout } x \in]-1, 1[$$

En remplaçant x par $-x^2$ et en prenant la primitive s'annulant en 0, on obtient :

$$\arcsin x = x + \sum_{n=1}^{+\infty} \frac{1 \times 3 \times 5 \times \dots \times (2n-1)}{2 \times 4 \times 6 \times \dots \times 2n} \frac{1}{2n+1} x^{2n+1} \quad \text{pour tout } x \in]-1, 1[$$

6.2.3 Exercices

EXERCICE 1. En utilisant des développements en séries entières de fonctions classiques, déterminer le développement en série entière des fonctions réelles suivantes, en précisant le rayon de convergence :

$$f(x) = \sqrt{\frac{1+x}{1-x}}, \quad g(x) = \sin^2 x \cos x, \quad h(x) = \frac{e^x}{1+x}.$$

EXERCICE 2. Par dérivation ou intégration de développements en séries entières connus, déterminer le développement en série entière des fonctions réelles suivantes, en précisant le rayon de convergence :

$$f(x) = \operatorname{argth} \frac{x + \operatorname{sh} \alpha}{\operatorname{ch} \alpha} \quad \text{où } \alpha \text{ réel } > 0 \text{ fixé}, \quad g(x) = (\arctan x)^2.$$

EXERCICE 3. Soit f l'application de $] -1, 1[$ dans \mathbb{R} définie par $f(x) = \frac{\arcsin x}{\sqrt{1-x^2}}$. en intégrant une équation différentielle du premier ordre dont f est solution, déterminer le développement en série entière de f , en précisant le rayon de convergence.

EXERCICE 4. Déterminer le développement en série entière des fonctions rationnelles suivantes, en précisant le rayon de convergence :

$$f(x) = \frac{1}{1+x+x^2}, \quad g(x) = \left(\frac{1+x}{1-x}\right)^2, \quad h(x) = \frac{1}{(1+x^2)(1-x^3)}.$$

EXERCICE 5. Montrer que l'application $f : x \mapsto \int_0^1 t^{xt} dt$ est définie sur \mathbb{R} , et en donner un développement en série entière; *indication* : montrer que, pour tout $x \in \mathbb{R}$, la série $\sum_{n \geq 0} \frac{x^n}{n!} (t \ln t)^n$ de fonctions de la variable t converge uniformément sur $[0, 1]$.

EXERCICE 6. Soit $x \in \mathbb{C}^*$. Soit $\varphi_x : \mathbb{C} \rightarrow \mathbb{C}$ définie par $\varphi_x(u) = \frac{e^{ux}-1}{e^u-1}$ pour tout $u \neq 0$, et $\varphi_x(0) = x$. Montrer que φ_x est développable en série entière.

En notant alors $\varphi_x(u) = \sum_{n=0}^{+\infty} Q_n(x) \frac{1}{n!} u^n$, montrer que chaque Q_n est une fonction polynomiale de x vérifiant : $Q_n(x+1) - Q_n(x) = x^n$ (polynômes de Bernouilli).

Chapitre 7

Géométrie affine en dimension 2 et 3

L'objet de ce chapitre est de réviser, mettre au point, mieux formaliser et compléter les notions de géométrie élémentaire vues lors des années précédentes, en utilisant pour cela l'outil important et efficace que constitue l'algèbre linéaire. Dans ce cadre, la notion de départ est celle d'espace affine, introduite dans la première section du chapitre : les éléments sont des points, définis en lien avec les vecteurs d'un espace vectoriel canoniquement associé. Elle est complétée dans la seconde partie par l'étude des transformations entre points : ces transformations sont les applications affines, définies en lien avec les applications linéaires de l'espace vectoriel associé. Conformément au programme, on se limite dans la mise en œuvre et les exercices à la dimension 2 ou 3 (même s'il est le plus souvent plus commode de donner les définitions et résultats généraux en dimension quelconque).

7.1 Espaces affines

7.1.1 Notion d'espace affine

a) DÉFINITION. Un *espace affine* sur \mathbb{R} est la donnée d'un triplet formé par :

(i) un \mathbb{R} -espace vectoriel E , (ses éléments sont des *vecteurs*; on les notera par des lettres minuscules surmontés d'une flèche $\vec{u}, \vec{v}, \vec{w}, \dots$; en particulier le *vecteur nul* sera noté $\vec{0}$);

(ii) un ensemble non-vide \mathcal{E} , (ses éléments sont des *points*; on les notera par des lettres majuscules A, B, C, M, N, P, \dots);

(iii) une application $\mathcal{E} \times \mathcal{E} \rightarrow E, (A, B) \mapsto \overrightarrow{AB}$ satisfaisant les deux axiomes suivants :

$$(A1) \quad \boxed{\forall A \in \mathcal{E}, \forall \vec{u} \in E, \exists! M \in \mathcal{E}, \overrightarrow{AM} = \vec{u}},$$

$$(A2) \quad \boxed{\forall A, B, C \in \mathcal{E}, \overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}} \quad (\text{relation de Chasles}).$$

TERMINOLOGIE. Dans la définition précédente, on dit par abus de langage que \mathcal{E} est un espace affine sur \mathbb{R} , et que E est le \mathbb{R} -espace vectoriel associé à \mathcal{E} . On dit que \mathcal{E} est un espace affine de dimension finie n lorsque le \mathbb{R} -e.v. E est de dimension finie n . Dans ce cours, on travaillera presque toujours dans un espace affine de dimension 2 ou 3.

b) REMARQUES. Les observations pratiques suivantes découlent directement de la définition.

- Nullité d'un vecteur. $\boxed{\forall A, B \in \mathcal{E}, [\overrightarrow{AB} = \vec{0}] \Leftrightarrow [A = B]}$.

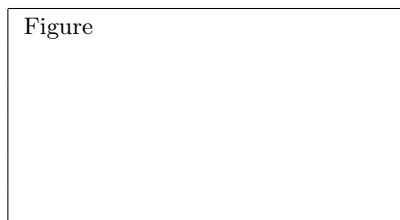
En effet : d'après (A2), $\overrightarrow{AA} = \overrightarrow{AA} + \overrightarrow{AA}$, donc $\overrightarrow{AA} = \vec{0}$. Réciproquement, si $\overrightarrow{AB} = \vec{0}$, alors $\overrightarrow{AB} = \overrightarrow{AA}$, ce qui implique $B = A$ par unicité dans l'axiome (A1).

- Opposé d'un vecteur. $\boxed{\forall A, B \in \mathcal{E}, \overrightarrow{BA} = -\overrightarrow{AB}}$. En effet : $\overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = \vec{0}$.

- Règle du parallélogramme.

$$\boxed{\forall A, B, C, D \in \mathcal{E}, [\overrightarrow{AB} = \overrightarrow{CD}] \Leftrightarrow [\overrightarrow{AC} = \overrightarrow{BD}]}$$

Cela résulte immédiatement de (A2) ; écrire les détails à titre d'exercice.



- L'axiome (A1) équivaut au fait que, pour tout point A fixé dans \mathcal{E} , l'application $\varphi_A : \mathcal{E} \rightarrow E$ définie par $M \mapsto \overrightarrow{AM}$ est bijective ; sa bijection réciproque est alors l'application $\psi_A : E \rightarrow \mathcal{E}$ définie par $\vec{u} \mapsto \psi_A(\vec{u}) :=$ l'unique point $M \in \mathcal{E}$ tel que $\overrightarrow{AM} = \vec{u}$.
- L'axiome (A1) équivaut aussi au fait que, pour tout vecteur \vec{u} fixé dans E , l'application $\tau_{\vec{u}} : \mathcal{E} \rightarrow \mathcal{E}$ définie par $A \mapsto \tau_{\vec{u}}(A) :=$ l'unique point $M \in \mathcal{E}$ tel que $\overrightarrow{AM} = \vec{u}$, est bijective. Cette application est appelée la translation de vecteur \vec{u} (voir plus loin).

COMMENTAIRE. Lorsque $\overrightarrow{AB} = \vec{u}$, on dit que le couple de points (A, B) est un *représentant* du vecteur \vec{u} , dans lequel A est l'*origine* et B l'*extrémité*. Tout autre couple de points (C, D) tel que $\overrightarrow{CD} = \vec{u}$ est un autre représentant de \vec{u} . La translation de vecteur \vec{u} est l'application qui, à tout point A , associe le point B qui est l'extrémité de \vec{u} lorsque son origine est en A .

Le point de vue retenu dans ce cours définit un espace affine à partir de la notion supposée connue d'espace vectoriel. On peut envisager une autre présentation consistant à partir intuitivement d'un ensemble de points \mathcal{E} , et à considérer dans $\mathcal{E} \times \mathcal{E}$ la relation dite d'*équipollence*, définie par : $(A, B) \sim (C, D)$ lorsque $ABDC$ est un parallélogramme. On vérifie que c'est une relation d'équivalence, et un vecteur est alors défini comme une classe d'équivalence pour cette relation [ie. \overrightarrow{AB} est l'ensemble des couples de points (C, D) équipollents à (A, B) , de sorte que l'on retrouve bien la règle du parallélogramme]. Il s'agit ensuite de récupérer géométriquement les diverses opérations sur les vecteurs correspondant à la structure d'espace vectoriel.

7.1.2 Sous-espaces affines

Soit \mathcal{E} un espace affine sur \mathbb{R} , d'espace vectoriel associé E .

- a) DÉFINITION. Une partie \mathcal{F} de \mathcal{E} est un *sous-espace affine* de \mathcal{E} lorsqu'il existe un point A dans \mathcal{F} tel que l'ensemble des vecteurs \overrightarrow{AM} pour M décrivant \mathcal{F} soit un sous-espace vectoriel de E .

Il en résulte en particulier qu'un sous-espace affine n'est jamais vide, et que \mathcal{E} lui-même est un sous-espace affine de \mathcal{E} . Le point fondamental est que le sous-espace vectoriel dans la définition ci-dessus ne dépend en fait pas du point A .

b) PROPOSITION ET DÉFINITION. Soit \mathcal{F} un sous-espace affine de \mathcal{E} ; il existe un sous-espace vectoriel F de E tel que, pour tout $A \in \mathcal{F}$, on ait $F = \{ \overrightarrow{AM} ; M \in \mathcal{F} \}$. On dit que F est le sous-espace vectoriel directeur du sous-espace affine \mathcal{F} , ou encore que \mathcal{F} est dirigé par F .

Preuve. Par hypothèse, il existe $A \in \mathcal{F}$ tel que $\varphi_A(\mathcal{F})$ est un ss-e.v. de E . Fixons $B \in \mathcal{F}$ quelconque et montrons que $\varphi_A(\mathcal{F}) = \varphi_B(\mathcal{F})$. On a $\overrightarrow{AB} = \varphi_A(B) \in \varphi_A(\mathcal{F})$. Quel que soit $M \in \mathcal{F}$, les vecteurs \overrightarrow{AM} et \overrightarrow{AB} appartiennent à $\varphi_A(\mathcal{F})$ qui est un sous-espace vectoriel, donc $\overrightarrow{BM} = \overrightarrow{AM} - \overrightarrow{AB} \in \varphi_A(\mathcal{F})$. Ceci prouve que $\varphi_B(\mathcal{F}) \subseteq \varphi_A(\mathcal{F})$. Pour la réciproque, notons que l'on a aussi $\overrightarrow{AM} + \overrightarrow{AB} \in \varphi_A(\mathcal{F})$; il existe donc $N \in \mathcal{F}$ tel que $\overrightarrow{AM} + \overrightarrow{AB} = \overrightarrow{AN}$ donc $\overrightarrow{AM} = \overrightarrow{AN} - \overrightarrow{AB} = \overrightarrow{BN} \in \varphi_B(\mathcal{F})$. On conclut $\varphi_A(\mathcal{F}) \subseteq \varphi_B(\mathcal{F})$, d'où l'égalité. \square

Réciproquement, la donnée d'un sous-espace vectoriel de E et d'un point de \mathcal{E} détermine un sous-espace affine de \mathcal{E} :

c) THÉORÈME ET DÉFINITION. Soient F un sous-espace vectoriel de E et A un point de \mathcal{E} . Il existe un unique sous-espace affine \mathcal{F} de \mathcal{E} tel que A appartienne à \mathcal{F} et tel que F soit le sous-espace vectoriel directeur de \mathcal{F} . On dit que \mathcal{F} est le sous-espace affine de \mathcal{E} passant par A et dirigé par F .

Preuve. Posons $\mathcal{F} = \varphi_A^{-1}(F) = \{ M \in \mathcal{E} ; \overrightarrow{AM} \in F \}$. On a $A \in \mathcal{F}$ puisque $\overrightarrow{AA} = \vec{0} \in F$; de plus, par construction, $\varphi_A(\mathcal{F}) = F$, donc \mathcal{F} est un ss-e.a. dirigé par F . Pour l'unicité, soit \mathcal{F}' un ss-e.a. de \mathcal{E} passant par A et dirigé par F . D'après la proposition b), cela signifie que $\varphi_A(\mathcal{F}') = F$. Mais alors $\varphi_A(\mathcal{F}) = \varphi_A(\mathcal{F}')$ implique $\mathcal{F} = \mathcal{F}'$ par bijectivité de φ_A . \square

EN RÉSUMÉ, si \mathcal{F} est un sous-espace affine de \mathcal{E} et F le sous-espace vectoriel directeur de \mathcal{F} :

- $\forall A \in \mathcal{F}, \mathcal{F} = \{ M \in \mathcal{E} ; \overrightarrow{AM} \in F \}$ et $F = \{ \overrightarrow{AM} ; M \in \mathcal{F} \}$,
- $\forall A \in \mathcal{F}, \forall \vec{u} \in F, \exists ! M \in \mathcal{F}, \vec{u} = \overrightarrow{AM}$,
- $\forall A \in \mathcal{F}, \forall B \in \mathcal{F}, \overrightarrow{AB} \in F$.

d) DIMENSION D'UN ESPACE AFFINE. On a une notion naturelle et évidente de dimension d'un sous-espace affine, la dimension de \mathcal{F} étant définie comme la dimension du sous-espace vectoriel F de E directeur de \mathcal{F} . En particulier, si E est de dimension finie n , on a $\dim \mathcal{E} = \dim E = n$ et $\dim \mathcal{F} = \dim F \leq n$ pour tout sous-espace affine \mathcal{F} de \mathcal{E} . Il résulte aisément des propositions précédentes que, si \mathcal{F} et \mathcal{F}' sont deux sous-espaces affines, on a :

- $[\mathcal{F}' \subseteq \mathcal{F}] \Rightarrow [\dim \mathcal{F}' \leq \dim \mathcal{F}],$
- $[\mathcal{F}' \subseteq \mathcal{F} \text{ et } \dim \mathcal{F}' = \dim \mathcal{F}] \Rightarrow [\mathcal{F}' = \mathcal{F}].$

Un sous-espace affine de dimension 0 est un singleton $\{A\}$ formé d'un seul point de \mathcal{E} . Un sous-espace affine de dimension 1 s'appelle une *droite affine*. Un sous-espace affine de dimension 2 s'appelle un *plan affine*. Dans le cadre de ce cours, on retiendra que :

- si l'on travaille dans un espace affine \mathcal{E} de dimension 2, les sous-espaces affines sont : les singletons (formés d'un seul point de \mathcal{E}), les droites affines contenues dans \mathcal{E} , et le plan \mathcal{E} lui-même ;
- si l'on travaille dans un espace affine \mathcal{E} de dimension 3, les sous-espaces affines sont : les singletons (formés d'un seul point de \mathcal{E}), les droites affines contenues dans \mathcal{E} , les plans affines contenus dans \mathcal{E} , et l'espace \mathcal{E} lui-même.

7.1.3 Parallélisme

Soit \mathcal{E} un espace affine sur \mathbb{R} , d'espace vectoriel associé E .

a) DÉFINITION. Soient \mathcal{F} et \mathcal{F}' deux sous-espaces affines de \mathcal{E} , de sous-espaces vectoriels directeurs respectifs F et F' . On dit que \mathcal{F} et \mathcal{F}' sont *parallèles* lorsque $F = F'$. On note alors $\mathcal{F} \parallel \mathcal{F}'$.

En particulier, deux sous-espaces affines parallèles sont nécessairement de même dimension. Il est facile de vérifier que le parallélisme est une relation d'équivalence dans l'ensemble des sous-espaces affines de \mathcal{E} .

b) PROPOSITION. *Deux sous-espaces affines parallèles sont nécessairement égaux ou disjoints.*

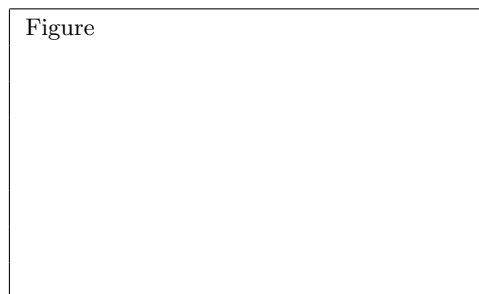
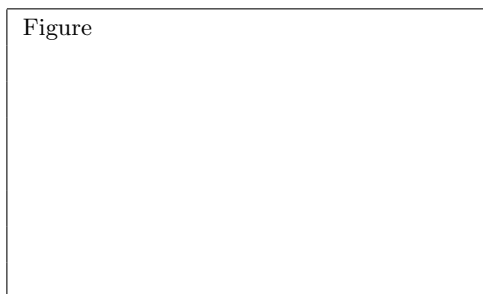
Preuve. Considérons comme ci-dessus $\mathcal{F} \parallel \mathcal{F}'$, avec $F = F'$. Si \mathcal{F} et \mathcal{F}' ne sont pas disjoints, considérons $A \in \mathcal{F} \cap \mathcal{F}'$. Alors \mathcal{F} et \mathcal{F}' passent tous les deux par A en étant dirigés par le même sous-espace vectoriel $F = F'$. On déduit de l'unicité dans le théorème de la section 2 que $\mathcal{F} = \mathcal{F}'$. \square

Attention, la réciproque est trivialement fautive en général ! Dans un espace affine de dimension 3, deux droites qui ne sont pas incluses dans un même plan sont forcément disjointes sans être parallèles. Il y a cependant des arguments de dimensions qui permettent des résultats partiels :

c) OBSERVATION PRATIQUE.

- Si \mathcal{E} est un plan affine, deux droites affines \mathcal{D} et \mathcal{D}' de \mathcal{E} sont parallèles si et seulement si elles sont égales ou disjointes.
- Si \mathcal{E} est un espace affine de dimension 3, deux plans affines \mathcal{P} et \mathcal{P}' de \mathcal{E} sont parallèles si et seulement s'ils sont égaux ou disjoints.

Preuve. Soient \mathcal{D} et \mathcal{D}' deux droites dans un plan affine \mathcal{E} . La proposition précédente montre l'un des sens de l'équivalence, et il est trivial que $\mathcal{D} = \mathcal{D}'$ implique $\mathcal{D} \parallel \mathcal{D}'$. Il s'agit donc de montrer que $\mathcal{D} \cap \mathcal{D}' = \emptyset$ implique $\mathcal{D} \parallel \mathcal{D}'$. Pour cela, supposons que \mathcal{D} et \mathcal{D}' ne sont pas parallèles. Les droites vectorielles D et D' dirigeant \mathcal{D} et \mathcal{D}' respectivement sont donc distinctes. Il en résulte que si l'on choisit $\vec{u} \in D$ et $\vec{v} \in D'$ non-nuls, ils ne sont pas colinéaires, donc forment une famille libre du plan vectoriel E , et donc une base de E . Prenons $A \in \mathcal{D}$ et $A' \in \mathcal{D}'$. Il existe $\lambda, \mu \in \mathbb{R}$ tels que $\overrightarrow{AA'} = \lambda \vec{u} + \mu \vec{v}$. Comme $\lambda \vec{u} \in D$ et $A \in \mathcal{D}$, il existe $M \in \mathcal{D}$ tel que $\overrightarrow{AM} = \lambda \vec{u}$. Donc $\overrightarrow{A'M} = \overrightarrow{AM} - \overrightarrow{AA'} = \lambda \vec{u} - \lambda \vec{u} - \mu \vec{v} = -\mu \vec{v}$. On a ainsi $\overrightarrow{A'M} \in D'$, ce qui, puisque $A' \in \mathcal{D}'$, implique $M \in \mathcal{D}'$. On conclut que $M \in \mathcal{D} \cap \mathcal{D}'$, ce qui achève la preuve du premier point. La preuve du second point est laissée au lecteur. \square



COMMENTAIRE. Une propriété fondamentale dans la formulation axiomatique de la géométrie classique est l'axiome (ou postulat) d'Euclide :

si \mathcal{D} est une droite et A un point n'appartenant pas à \mathcal{D} , alors il passe par A une droite et une seule parallèle à \mathcal{D} .

Dans la mesure où être parallèle signifie avoir le même sous-espace vectoriel directeur (ici la même droite vectorielle directrice), cet énoncé "traditionnel" est une formulation du théorème fondamental 7.1.2.c. Et elle est de ce fait valable pour tout sous-espace affine. Par exemple :

si \mathcal{P} est un plan et A un point n'appartenant pas à \mathcal{P} , alors il passe par A un plan et un seul parallèle à \mathcal{P} .

7.1.4 Bases affines

Soit \mathcal{E} un espace affine sur \mathbb{R} , d'espace vectoriel associé E .

a) PROPOSITION. *Une intersection de sous-espaces affines, à condition qu'elle soit non vide, est un sous-espace affine, dirigé par l'intersection des sous-espaces vectoriels directeurs.*

Preuve. Soit $(\mathcal{F}_i)_{i \in I}$ une famille de sous-espaces affines de \mathcal{E} . Pour tout $i \in I$, notons F_i le sous-espace vectoriel directeur de \mathcal{F}_i . On sait que $F = \bigcap_{i \in I} F_i$ est un sous-espace vectoriel de E . Posons $\mathcal{F} = \bigcap_{i \in I} \mathcal{F}_i$ et supposons que $\mathcal{F} \neq \emptyset$. Prenons $A \in \mathcal{F}$ quelconque. Parce que φ_A est injective (car bijective), on a $\varphi_A(\bigcap_{i \in I} \mathcal{F}_i) = \bigcap_{i \in I} \varphi_A(\mathcal{F}_i)$, c'est-à-dire $\varphi_A(\mathcal{F}) = F$. Comme F est un sous-espace vectoriel de E , ceci prouve que \mathcal{F} est un ss-e.a. de \mathcal{E} dirigé par F . \square

Il en résulte (suivant un argument classique en mathématiques !) que, pour toute partie non-vide \mathcal{X} de \mathcal{E} , on peut considérer le sous-espace affine $\langle \mathcal{X} \rangle$ défini comme l'intersection de tous les sous-espaces affines de \mathcal{E} contenant \mathcal{X} . On vérifie de façon immédiate que c'est le plus petit sous-espace affine de \mathcal{E} contenant \mathcal{X} . On l'appelle le *sous-espace affine de \mathcal{E} engendré par \mathcal{X}* .

Le théorème ci-dessous donne une description explicite du sous-espace affine engendré par un nombre fini de points.

b) THÉORÈME. *Soient A_0, A_1, \dots, A_p des points distincts de \mathcal{E} , avec $p \geq 0$. Le sous-espace affine engendré par $\{A_0, A_1, \dots, A_p\}$ est égal au sous-espace affine de \mathcal{E} passant par A_0 et dirigé par le sous-espace vectoriel de E engendré par $\{\overrightarrow{A_0A_1}, \overrightarrow{A_0A_2}, \dots, \overrightarrow{A_0A_p}\}$. Il est de dimension $\leq p$.*

Preuve. Posons $\mathcal{X} = \{A_0, A_1, \dots, A_p\}$. Soit F le sous-espace vectoriel de E engendré par les p vecteurs $\{\overrightarrow{A_0A_1}, \overrightarrow{A_0A_2}, \dots, \overrightarrow{A_0A_p}\}$. Notons \mathcal{F} le sous-espace affine de \mathcal{E} passant par A_0 et dirigé par F . On a $\mathcal{F} = \varphi_{A_0}^{-1}(F)$. En particulier $\mathcal{X} \subseteq \mathcal{F}$. Soit maintenant \mathcal{H} un sous-espace affine de \mathcal{E} contenant \mathcal{X} . Son sous-espace vectoriel directeur $H = \varphi_{A_0}(\mathcal{H})$ contient les vecteurs $\overrightarrow{A_0A_1}, \overrightarrow{A_0A_2}, \dots, \overrightarrow{A_0A_p}$, donc le sous-espace vectoriel qu'ils engendrent. Ainsi F est un sous-espace vectoriel de H . On en déduit que $\varphi_{A_0}^{-1}(F) \subseteq \varphi_{A_0}^{-1}(H)$, c'est-à-dire $\mathcal{F} \subseteq \mathcal{H}$. On a ainsi montré que le sous-espace affine \mathcal{F} contient \mathcal{X} et est inclus dans tout sous-espace affine de \mathcal{E} contenant \mathcal{X} . On conclut $\mathcal{F} = \langle \mathcal{X} \rangle$. \square

En résumé et en pratique :

Un point M de \mathcal{E} appartient au sous-espace affine engendré par A_0, A_1, \dots, A_p	\Leftrightarrow	$\exists (\alpha_1, \dots, \alpha_p) \in \mathbb{R}^p, \overrightarrow{A_0M} = \sum_{i=1}^p \alpha_i \overrightarrow{A_0A_i}$
--	-------------------	---

Il est clair que, dans ce théorème, A_0 peut être remplacé par n'importe lequel des A_i .

La question qui se pose alors est celle de savoir si la famille de vecteurs $X_0 = \{\overrightarrow{A_0A_1}, \overrightarrow{A_0A_2}, \dots, \overrightarrow{A_0A_p}\}$ est libre ou non, dans la mesure où il est clair que :

$$[X_0 \text{ libre}] \Leftrightarrow [X_0 \text{ base de } F] \Leftrightarrow [\dim F = p] \Leftrightarrow [\dim \mathcal{F} = p].$$

Là encore, c'est une propriété qui ne dépend pas du choix de A_0 , comme le montre le lemme suivant.

c) LEMME. Soient A_0, A_1, \dots, A_p des points de \mathcal{E} deux à deux distincts. Les conditions suivantes sont équivalentes :

- (i) la famille $X_0 = \{\overrightarrow{A_0A_1}, \overrightarrow{A_0A_2}, \dots, \overrightarrow{A_0A_p}\}$ est libre dans E ;
- (ii) pour tout $0 \leq j \leq p$, la famille $X_j = \{\overrightarrow{A_jA_0}, \dots, \overrightarrow{A_jA_{j-1}}, \overrightarrow{A_jA_{j+1}}, \dots, \overrightarrow{A_jA_p}\}$ est libre dans \mathcal{E} ;
- (iii) aucun des points A_i n'appartient au sous-espace affine engendré par les p autres points.

Preuve. Supposons X_0 libre, et fixons $0 < j \leq p$. Soient $\lambda_0, \dots, \lambda_{j-1}, \lambda_{j+1}, \dots, \lambda_p \in \mathbb{R}$ tels que $\sum_{i \neq j} \lambda_i \overrightarrow{A_jA_i} = \vec{0}$. En décomposant $\overrightarrow{A_jA_i} = \overrightarrow{A_jA_0} + \overrightarrow{A_0A_i}$, il vient : $(\sum_{i \neq j} \lambda_i) \overrightarrow{A_jA_0} + \sum_{i \neq j} \lambda_i \overrightarrow{A_0A_i} = \vec{0}$. Parce que X_0 est libre, on déduit $\sum_{i \neq j} \lambda_i = 0$ et $\lambda_i = 0$ pour tout $1 \leq i \leq p$ distinct de j . D'où finalement $\lambda_i = 0$ pour tout $0 \leq i \leq p$ distinct de j . Ceci prouve que (i) \Rightarrow (ii), et donc (i) \Leftrightarrow (ii).

Supposons maintenant qu'il existe $0 \leq i \leq p$ tel que A_i appartienne au sous-espace affine engendré par $A_0, A_1, \dots, A_{i-1}, A_{i+1}, \dots, A_p$. Alors, pour tout $0 \leq j \neq i \leq p$, il existe des coefficients $\alpha_k \in \mathbb{R}$ pour $0 \leq k \neq i, k \neq j \leq p$ tel que $\overrightarrow{A_jA_i} = \sum_k \alpha_k \overrightarrow{A_jA_k}$, ou encore $\sum_k \alpha_k \overrightarrow{A_jA_k} - \overrightarrow{A_jA_i} = \vec{0}$, ce qui prouve que la famille X_j est liée. Par contraposée, ceci montre que (ii) \Rightarrow (iii). La réciproque s'obtient par des calculs analogues. \square

d) DÉFINITIONS. Une famille \mathcal{X} de $p+1$ points deux à deux distincts de \mathcal{E} est dite *affinement libre* si elle satisfait les conditions équivalentes de la proposition précédente.

Lorsque \mathcal{X} est affinement libre, on dit que \mathcal{X} est une *base affine* du sous-espace affine $\mathcal{F} = \langle \mathcal{X} \rangle$ engendré par \mathcal{X} .

$$\boxed{\{A_0, A_1, \dots, A_p\} \text{ est une base affine de } \mathcal{F}} \Leftrightarrow \boxed{\forall M \in \mathcal{F}, \exists ! (\alpha_1, \dots, \alpha_p) \in \mathbb{R}^p, \overrightarrow{A_0M} = \sum_{i=1}^p \alpha_i \overrightarrow{A_0A_i}}$$

Il est clair dans ce cas que \mathcal{F} est de dimension p (puisque X_0 est alors une base du sous-espace vectoriel F directeur de \mathcal{F}), et que l'on peut remplacer A_0 par n'importe lequel des A_i .

- *Un premier cas particulier.* Prenons deux points A et B distincts dans \mathcal{E} . Alors $X = \{\overrightarrow{AB}\}$ est libre, donc le sous-espace affine engendré par $\mathcal{X} = \{A, B\}$ est de dimension 1 : on l'appelle la droite affine passant par A et B , noté (AB) . En particulier, deux points sont toujours alignés. La droite affine (AB) est dirigée par la droite vectorielle Δ de base $\{\overrightarrow{AB}\}$. On a pour tout $M \in \mathcal{E}$:

$$[A, B, M \text{ alignés}] \Leftrightarrow [M \in (AB)] \Leftrightarrow [\overrightarrow{AM} \in \Delta] \Leftrightarrow [\{\overrightarrow{AB}, \overrightarrow{AM}\} \text{ liée}] \Leftrightarrow [\exists \lambda \in \mathbb{R}, \overrightarrow{AM} = \lambda \overrightarrow{AB}].$$

$$[A, B, M \text{ non alignés}] \Leftrightarrow [\{A, B, M\} \text{ affinement libre}].$$

- *Un second cas particulier.* Prenons trois points A, B, C non alignés dans \mathcal{E} . Alors $X = \{\overrightarrow{AB}, \overrightarrow{AC}\}$ est libre, donc le sous-espace affine engendré par $\mathcal{X} = \{A, B, C\}$ est de dimension 2 : on l'appelle le plan affine passant par A, B et C , noté (ABC) . En particulier trois points sont toujours coplanaires.

Le plan affine (ABC) est dirigé par le plan vectoriel Π de base $\{\overrightarrow{AB}, \overrightarrow{AC}\}$. Pour tout $M \in \mathcal{E}$:

$$[A, B, C, M \text{ coplanaires}] \Leftrightarrow [M \in (ABC)] \Leftrightarrow [\overrightarrow{AM} \in \Pi] \\ \Leftrightarrow [\{\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AM}\} \text{ liée}] \Leftrightarrow [\exists \lambda, \mu \in \mathbb{R}, \overrightarrow{AM} = \lambda \overrightarrow{AB} + \mu \overrightarrow{AC}].$$

$$[A, B, C, M \text{ non coplanaires}] \Leftrightarrow [\{A, B, C, M\} \text{ affinement libre}].$$



7.1.5 Coordonnées et représentations paramétriques des sous-espaces affines

Soit \mathcal{E} un espace affine sur \mathbb{R} de dimension finie n , d'espace vectoriel associé E .

a) DÉFINITIONS. Un repère cartésien de \mathcal{E} est un couple $\mathcal{R} = (O, \mathcal{B})$ formé par un point fixé quelconque $O \in \mathcal{E}$, appelé l'origine du repère, et une base $\mathcal{B} = (\overrightarrow{e}_1, \dots, \overrightarrow{e}_n)$ de E .

Pour tout point $M \in \mathcal{E}$, les composantes du vecteurs \overrightarrow{OM} dans la base \mathcal{B} sont appelées les *coordonnées cartésiennes* du point M dans le repère \mathcal{R} . On note : $M(x_1, \dots, x_n)$. Ainsi :

$$[M(x_1, \dots, x_n) \text{ dans le repère } (O, \overrightarrow{e}_1, \dots, \overrightarrow{e}_n)] \iff [\overrightarrow{OM} = \sum_{i=1}^n x_i \overrightarrow{e}_i].$$

Pour tout $1 \leq i \leq n$, soit A_i le point de \mathcal{E} tel que $\overrightarrow{OA_i} = \overrightarrow{e}_i$. Comme $\mathcal{B} = \{\overrightarrow{OA_1}, \overrightarrow{OA_2}, \dots, \overrightarrow{OA_n}\}$ est libre, la famille de $n+1$ points $\mathcal{X} = \{O, A_1, A_2, \dots, A_n\}$ est affinement libre. Comme de plus \mathcal{B} engendre l'espace vectoriel E , le sous-espace affine de \mathcal{E} engendré par \mathcal{X} n'est autre que \mathcal{E} lui-même. En d'autres termes, \mathcal{X} est une base affine de \mathcal{E} .

Une conséquence triviale mais très utile pratiquement est que :

si $M(x_1, \dots, x_n)$ et $N(y_1, \dots, y_n)$ dans le repère $\mathcal{R} = (O, \mathcal{B})$, alors les composantes du vecteur $\overrightarrow{MN} = \overrightarrow{ON} - \overrightarrow{OM}$ dans la base \mathcal{B} sont $(y_1 - x_1, \dots, y_n - x_n)$.

b) REMARQUE ET DÉFINITION. Le repère $\mathcal{R} = (O, \mathcal{B})$ avec $\mathcal{B} = (\overrightarrow{e}_1, \dots, \overrightarrow{e}_n)$ étant fixé, considérons un sous-espace affine \mathcal{F} de \mathcal{E} , de dimension p , passant par un point donné A , et dont le sous-espace vectoriel directeur F est donné par une base $\mathcal{C} = \{\overrightarrow{v}_1, \dots, \overrightarrow{v}_p\}$. Chaque \overrightarrow{v}_i se décompose dans \mathcal{B} en $\overrightarrow{v}_i = \alpha_{i,1} \overrightarrow{e}_1 + \alpha_{i,2} \overrightarrow{e}_2 + \dots + \alpha_{i,n} \overrightarrow{e}_n$, avec $\alpha_{i,j} \in \mathbb{R}$ pour tout $1 \leq j \leq n$ et tout $1 \leq i \leq p$. Donc, pour $\lambda_1, \dots, \lambda_p \in \mathbb{R}$:

$$\sum_{i=1}^p \lambda_i \overrightarrow{v}_i = \sum_{i=1}^p \lambda_i \left(\sum_{j=1}^n \alpha_{i,j} \overrightarrow{e}_j \right) = \sum_{j=1}^n \left(\sum_{i=1}^p \lambda_i \alpha_{i,j} \right) \overrightarrow{e}_j.$$

Notons $A(a_1, \dots, a_n)$, de sorte que pour tout $M(x_1, \dots, x_n)$, on a $\overrightarrow{AM} = \sum_{j=1}^n (x_j - a_j) \overrightarrow{e}_j$.

L'équivalence $(M \in \mathcal{F} \Leftrightarrow \overrightarrow{AM} \in F)$ devient donc :

$$M \in \mathcal{F} \iff \exists \lambda_1, \dots, \lambda_p \in \mathbb{R}, \quad x_j = a_j + \sum_{i=1}^p \lambda_i \alpha_{i,j} \text{ pour tout } 1 \leq j \leq n \quad (*)$$

On dit que les relations (*) constituent une *représentation paramétrique* du sous-espace affine \mathcal{F} dans le repère \mathcal{R} . Les relations (*) traduisent que $(\lambda_1, \dots, \lambda_p)$ sont les coordonnées du point M de \mathcal{F} dans ce repère (A, \mathcal{C}) de \mathcal{F} .

• EXEMPLE. Plaçons-nous dans un espace affine \mathcal{E} de dimension 2 rapporté à un repère ; une représentation paramétrique de la droite passant par $A(a_1, a_2)$ et de vecteur directeur

$$\vec{v}(\alpha_1, \alpha_2) \text{ est : } \begin{cases} x_1 = a_1 + \lambda \alpha_1 \\ x_2 = a_2 + \lambda \alpha_2 \end{cases}, \text{ avec } \lambda \in \mathbb{R}.$$

• EXEMPLE. Plaçons-nous dans un espace affine \mathcal{E} de dimension 3 rapporté à un repère ; – une représentation paramétrique de la droite passant par $A(a_1, a_2, a_3)$ et de vecteur directeur

$$\vec{v}(\alpha_1, \alpha_2, \alpha_3) \text{ est : } \begin{cases} x_1 = a_1 + \lambda \alpha_1 \\ x_2 = a_2 + \lambda \alpha_2 \\ x_3 = a_3 + \lambda \alpha_3 \end{cases}, \text{ avec } \lambda \in \mathbb{R}.$$

– une représentation paramétrique du plan passant par $A(a_1, a_2, a_3)$ et dirigé par le plan vectoriel

$$\text{de base } \{\vec{v}, \vec{w}\} \text{ avec } \vec{v}(\alpha_1, \alpha_2, \alpha_3) \text{ et } \vec{w}(\beta_1, \beta_2, \beta_3) \text{ est : } \begin{cases} x_1 = a_1 + \lambda \alpha_1 + \mu \beta_1 \\ x_2 = a_2 + \lambda \alpha_2 + \mu \beta_2 \\ x_3 = a_3 + \lambda \alpha_3 + \mu \beta_3 \end{cases}, \text{ avec } \lambda, \mu \in \mathbb{R}.$$

7.1.6 Equations cartésiennes des sous-espaces affines

Soit \mathcal{E} un espace affine sur \mathbb{R} de dimension finie n , d'espace vectoriel associé E . On fixe un repère cartésien $\mathcal{R} = (O, \mathcal{B})$ d'origine $O \in \mathcal{E}$ avec $\mathcal{B} = (\vec{e}_1, \dots, \vec{e}_n)$ une base de E .

a) DÉFINITION. On appelle *hyperplan affine* de \mathcal{E} tous sous-espace affine de \mathcal{E} de dimension $n - 1$.

Dans un plan affine, les hyperplans sont les droites affines. Dans un espace affine de dimension 3, les hyperplans sont les plans affines. Le sous-espace directeur H d'un hyperplan affine \mathcal{H} est un sous-espace vectoriel de dimension $n - 1$ de E , ie. un hyperplan vectoriel. Le théorème ci-dessous est fondé sur le fait bien connu en algèbre linéaire (conséquence évidente de la formule du rang) que les hyperplans vectoriels sont les noyaux des formes linéaires non-nulles.

b) THÉORÈME.

(i) Pour tout hyperplan affine \mathcal{H} de \mathcal{E} , il existe $(a_1, \dots, a_n, a_{n+1}) \in \mathbb{R}^{n+1}$ avec $(a_1, \dots, a_n) \neq (0, \dots, 0)$, tel que \mathcal{H} soit l'ensemble des points $M(x_1, \dots, x_n)$ de \mathcal{E} vérifiant :

$$a_1 x_1 + a_2 x_2 + \dots + a_n x_n + a_{n+1} = 0 \quad (**)$$

(ii) Réciproquement, si $(a_1, \dots, a_n, a_{n+1}) \in \mathbb{R}^{n+1}$ tel que $(a_1, \dots, a_n) \neq (0, \dots, 0)$, l'ensemble des points $M(x_1, \dots, x_n)$ de \mathcal{E} vérifiant (**) est un hyperplan de \mathcal{E} .

Preuve Pour montrer (i), soit \mathcal{H} un hyperplan affine. Son ss-e.v. directeur H est un hyperplan vectoriel de E ; donc il existe une forme linéaire non-nulle $f : E \rightarrow \mathbb{R}$ telle que $H = \text{Ker } f$. Si l'on note (a_1, \dots, a_n) les composantes de f dans la base duale $\{e_1^*, \dots, e_n^*\}$ de E^* , on a $(a_1, \dots, a_n) \neq (0, \dots, 0)$ et, pour tout $\vec{u}(y_1, \dots, y_n)$, on peut calculer :

$$f(\vec{u}) = \sum_{i=1}^n a_i e_i^*(\vec{u}) = \sum_{i=1}^n a_i e_i^* \left(\sum_{j=1}^n y_j \vec{e}_j \right) = \sum_{i=1}^n \sum_{j=1}^n a_i y_j e_i^*(\vec{e}_j) = \sum_{i=1}^n a_i y_i.$$

Donc H est l'hyperplan vectoriel de E d'équation $a_1 y_1 + \dots + a_n y_n = 0$ dans \mathcal{B} . Soit $B(b_1, \dots, b_n) \in \mathcal{H}$. On a : $[M(x_1, \dots, x_n) \in \mathcal{H}] \Leftrightarrow [\overrightarrow{BM} \in H] \Leftrightarrow [a_1(x_1 - b_1) + \dots + a_n(x_n - b_n) = 0]$, d'où le résultat en posant $a_{n+1} = -(a_1 b_1 + \dots + a_n b_n)$.

Pour (ii), supposons donné $(a_1, \dots, a_n, a_{n+1}) \in \mathbb{R}^{n+1}$ avec $(a_1, \dots, a_n) \neq (0, \dots, 0)$ et notons \mathcal{H} l'ensemble des points $M(x_1, \dots, x_n)$ vérifiant (**). Soit $B(b_1, \dots, b_n) \in \mathcal{H}$; (il en existe toujours car l'un au moins des a_i est non-nul). Comme $a_1 b_1 + \dots + a_n b_n + a_{n+1} = 0$, on a pour tout $M(x_1, \dots, x_n) \in \mathcal{H}$:

$$[M \in \mathcal{H}] \Leftrightarrow [a_1 x_1 + \dots + a_n x_n + a_{n+1} = a_1 b_1 + \dots + a_n b_n + a_{n+1}] \Leftrightarrow [a_1(x_1 - b_1) + \dots + a_n(x_n - b_n) = 0].$$

Ceci signifie que $M \in \mathcal{H}$ équivaut à $\overrightarrow{BM} \in \text{Ker } f$, où l'on note f la forme linéaire $f = a_1 e_1^* + \dots + a_n e_n^*$, qui est non-nulle d'après l'hypothèse sur les a_i . En notant H l'hyperplan vectoriel $\text{Ker } f$ dans E , on a finalement : $M \in \mathcal{H}$ ssi $\overrightarrow{BM} \in H$, ce qui prouve que \mathcal{H} est le sous-espace affine passant par B et dirigé par H , donc que \mathcal{H} est un hyperplan affine. \square

La relation (**) est appelée une *équation cartésienne* de l'hyperplan \mathcal{H} dans le repère. Comme on l'a vu dans la preuve, \mathcal{H} est dirigé par l'hyperplan vectoriel H dans E d'équation $a_1 x_1 + a_2 x_2 + \dots + a_n x_n = 0$ dans la base \mathcal{B} . On en déduit immédiatement :

COROLLAIRE. Soient \mathcal{H} et \mathcal{H}' deux hyperplans de \mathcal{E} d'équations $a_1 x_1 + a_2 x_2 + \dots + a_n x_n + a_{n+1} = 0$ et $a'_1 x_1 + a'_2 x_2 + \dots + a'_n x_n + a'_{n+1} = 0$ respectivement. Alors :

$\mathcal{H} \parallel \mathcal{H}'$ si et seulement s'il existe $\lambda \in \mathbb{R}, \lambda \neq 0$, tel que $a'_i = \lambda a_i$ pour tout $1 \leq i \leq n$.

$\mathcal{H} = \mathcal{H}'$ si et seulement s'il existe $\lambda \in \mathbb{R}, \lambda \neq 0$, tel que $a'_i = \lambda a_i$ pour tout $1 \leq i \leq n+1$.

► **Cas des droites en dimension 2.** On suppose \mathcal{E} de dimension 2, rapporté à un repère cartésien.

Une équation cartésienne d'une droite affine \mathcal{D} de \mathcal{E} est de la forme

$$ax + by + c = 0 \quad \text{avec } (a, b) \neq (0, 0).$$

Une base de la droite vectorielle Δ dirigeant \mathcal{D} est alors $\{\vec{u}\}$ avec $\vec{u}(-b, a)$.

EXEMPLES D'APPLICATIONS. (Ecrire à titre d'exercice le détail des calculs correspondants).

• *Condition d'alignement.* Soient $A(\alpha, \beta)$ et $B(\alpha', \beta')$. Pour tout $M(x, y)$, on a :

$$(M, A, B) \text{ alignés} \Leftrightarrow \{\overrightarrow{AM}, \overrightarrow{AB}\} \text{ lié} \Leftrightarrow \begin{vmatrix} x-\alpha & \alpha'-\alpha \\ y-\beta & \beta'-\beta \end{vmatrix} = 0 \Leftrightarrow \boxed{(\beta' - \beta)x + (\alpha - \alpha')y + (\alpha'\beta - \alpha\beta') = 0},$$

ce qui, pour $A \neq B$, donne une équation de la droite (AB) . Ou encore, en remarquant que les points sont alignés si et seulement si leurs coordonnées vérifient une équation de droite :

$$(M, A, B) \text{ alignés} \Leftrightarrow \exists a, b, c \in \mathbb{R}, (a, b) \neq (0, 0), \begin{cases} ax + by + c = 0 \\ a\alpha + b\beta + c = 0 \\ a\alpha' + b\beta' + c = 0 \end{cases} \Leftrightarrow \begin{vmatrix} x & y & 1 \\ \alpha & \beta & 1 \\ \alpha' & \beta' & 1 \end{vmatrix} = 0.$$

• *Position relative de deux droites.* Soient deux droites \mathcal{D} et \mathcal{D}' d'équations respectives $ax + by + c = 0$ et $a'x + b'y + c' = 0$. Notons (S) le système $\begin{cases} ax + by + c = 0 \\ a'x + b'y + c' = 0 \end{cases}$, et $\delta = \begin{vmatrix} a & b \\ a' & b' \end{vmatrix}$ son déterminant.

(i) $(\mathcal{D} \text{ parallèle à } \mathcal{D}') \Leftrightarrow (\exists \lambda \in \mathbb{R}^*, a' = \lambda a, b' = \lambda b) \Leftrightarrow (\delta = 0)$.

– Si l'on a aussi $c' = \lambda c$, les deux équations sont équivalentes, donc $\mathcal{D} = \mathcal{D}'$.

– Sinon, (S) n'est pas compatible, donc $\mathcal{D} \cap \mathcal{D}' = \emptyset$.

(ii) (\mathcal{D} non parallèle à \mathcal{D}') \Leftrightarrow ($\delta \neq 0$) \Leftrightarrow ($\mathcal{D} \cap \mathcal{D}' = \{\Omega\}$) avec $\Omega \left(\frac{bc' - b'c}{ab' - a'b}, \frac{a'c - ac'}{ab' - a'b} \right)$.

• *Condition de concours de trois droites.* Soient trois droites \mathcal{D} , \mathcal{D}' et \mathcal{D}'' d'équations respectives $ax + by + c = 0$, $a'x + b'y + c' = 0$ et $a''x + b''y + c'' = 0$. On suppose que les trois droites sont deux à deux non parallèles, c'est-à-dire que $ab' - a'b$, $a'b'' - a''b'$ et $a''b' - a'b''$ sont tous les trois non-nuls. On dit qu'elles sont concourantes lorsqu'elles se coupent en un même point, c'est-à-dire lorsque leurs trois points d'intersection deux à deux sont confondus ; ceci équivaut à dire que les coordonnées du point d'intersection Ω de \mathcal{D} et \mathcal{D}' (voir ci-dessus) sont solutions de l'équation $a''x + b''y + c'' = 0$. Il vient après calcul :

$$(\mathcal{D}, \mathcal{D}', \mathcal{D}'' \text{ concourantes}) \Leftrightarrow \left(\begin{vmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{vmatrix} = 0 \right).$$

► **Cas des plans en dimension 3.** On suppose \mathcal{E} de dimension 3, rapporté à un repère cartésien.

Une équation cartésienne d'un plan affine \mathcal{P} de \mathcal{E} est de la forme

$$ax + by + cz + d = 0 \quad \text{avec } (a, b, c) \neq (0, 0, 0).$$

Une base du plan vectoriel Π dirigeant \mathcal{P} est $\{\vec{u}, \vec{v}\}$ avec $\vec{u}(-b, a, 0)$ et $\vec{v}(-c, 0, a)$.

EXEMPLES D'APPLICATIONS. (Ecrire le détail des calculs correspondants et faire des figures).

• *Condition de coplanarité.* Soient $A(\alpha, \beta, \gamma)$, $B(\alpha', \beta', \gamma')$, $C(\alpha'', \beta'', \gamma'')$. Pour tout $M(x, y, z)$, on a :

$$(M, A, B, C) \text{ coplanaires} \Leftrightarrow \begin{vmatrix} x-\alpha & \alpha'-\alpha & \alpha''-\alpha \\ y-\beta & \beta'-\beta & \beta''-\beta \\ z-\gamma & \gamma'-\gamma & \gamma''-\gamma \end{vmatrix} = 0 \Leftrightarrow \begin{vmatrix} x & y & z & 1 \\ \alpha & \beta & \gamma & 1 \\ \alpha' & \beta' & \gamma' & 1 \\ \alpha'' & \beta'' & \gamma'' & 1 \end{vmatrix} = 0$$

ce qui, quand $\{\vec{AB}, \vec{AC}\}$ est libre, donne en développant une équation du plan passant par A, B, C .

• *Position relative de deux plans.* Soient deux plans \mathcal{P} et \mathcal{P}' d'équations respectives $ax + by + cz + d = 0$ et $a'x + b'y + c'z + d' = 0$. Notons (S) le système $\begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases}$, et $\mu = \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}$ sa matrice.

(i) (\mathcal{P} parallèle à \mathcal{P}') \Leftrightarrow ($\exists \lambda \in \mathbb{R}^*$, $a' = \lambda a$, $b' = \lambda b$, $c' = \lambda c$) \Leftrightarrow ($\text{rg } \mu = 1$).

– Si l'on a aussi $d' = \lambda d$, les deux équations sont équivalentes, donc $\mathcal{P} = \mathcal{P}'$.

– Sinon, (S) n'est pas compatible, donc $\mathcal{P} \cap \mathcal{P}' = \emptyset$.

(ii) Supposons que $\text{rg } \mu \neq 1$. Les deux plans \mathcal{P} et \mathcal{P}' ne sont donc pas parallèles. Comme $(a, b, c) \neq (0, 0, 0) \neq (a', b', c')$, on a $\text{rg } \mu \neq 0$, donc $\text{rg } \mu = 2$. Donc l'un au moins des 3 mineurs $ab' - a'b$, $bc' - c'b$, $ca' - c'a$ est non-nul. Il en résulte en particulier que $\mathcal{P} \cap \mathcal{P}' \neq \emptyset$.

En effet, si par exemple $ab' - a'b \neq 0$, l'ensemble des solutions de (S) est $\{(x(z), y(z), z) ; z \in \mathbb{R}\}$, avec $x(z), y(z)$ donnés par les formules de Cramer dans le système $\begin{cases} ax + by = -cz - d \\ a'x + b'y = -c'z - d' \end{cases}$.

Dès lors, $\mathcal{P} \cap \mathcal{P}'$ est un sous-espace affine dirigé par le ss-e.v. $\Pi \cap \Pi'$. Or $\vec{u}(x, y, z) \in \Pi \cap \Pi'$ si et seulement si (x, y, z) est solution du système homogène $(S_0) \begin{cases} ax + by + cz = 0 \\ a'x + b'y + c'z = 0 \end{cases}$. On a $\text{rg}(S_0) = \text{rg } \mu = 2$ donc l'espace vectoriel des solutions de (S_0) est de dimension $3 - 2 = 1$. On conclut que $\Pi \cap \Pi'$ est une droite vectorielle, donc que $\mathcal{P} \cap \mathcal{P}'$ est une droite affine.

On retiendra que : *l'intersection de deux plans affines non parallèles est une droite affine.*

► **Cas des droites en dimension 3.** On suppose \mathcal{E} de dimension 3, rapporté à un repère cartésien.

PROPOSITION. \mathcal{D} est une droite affine de \mathcal{E} si et seulement s'il existe (a, b, c) et (a', b', c') deux triplets linéairement indépendants dans \mathbb{R}^3 , et deux scalaires $d, d' \in \mathbb{R}$ tels que \mathcal{D} soit exactement l'ensemble des points $M(x, y, z)$ dont les coordonnées sont solutions du système (S) suivant :

$$\boxed{\begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases}}, \quad \text{avec } (a', b', c') \neq \lambda(a, b, c) \text{ pour tout } \lambda \in \mathbb{R}, \quad (a, b, c) \neq (0, 0, 0).$$

Preuve Un sens résulte de ce que l'on vient de voir à la fin du paragraphe précédent. Réciproquement, soient \mathcal{D} une droite affine, A un point de \mathcal{D} et \vec{u} un vecteur non-nul de la droite vectorielle Δ dirigeant \mathcal{D} . On peut compléter \vec{u} en une base $\{\vec{u}, \vec{v}, \vec{w}\}$ de E . On introduit le plan \mathcal{P} passant par A et de sous-espace vectoriel directeur le plan vectoriel Π de base $\{\vec{u}, \vec{v}\}$. De même, soit \mathcal{P}' passant par A et de sous-espace vectoriel directeur le plan vectoriel Π' de base $\{\vec{u}, \vec{w}\}$. On a $\{\vec{u}, \vec{v}, \vec{w}\}$ libre, donc $\Pi \neq \Pi'$, donc \mathcal{P} non parallèle à \mathcal{P}' . D'après ce que l'on a vu précédemment, $\mathcal{P} \cap \mathcal{P}'$ est alors une droite affine, dirigée par la droite vectorielle $\Pi \cap \Pi'$. Comme $A \in \mathcal{P} \cap \mathcal{P}'$ et $\vec{u} \in \Pi \cap \Pi'$, on conclut $\mathcal{P} \cap \mathcal{P}' = \mathcal{D}$. On introduit des équations cartésiennes de \mathcal{P} et de \mathcal{P}' pour achever la preuve. \square

On dit que (S) est un système d'équations cartésiennes de \mathcal{D} dans le repère. Il exprime que : toute droite affine de \mathcal{E} est (d'une infinité de façons) l'intersection de deux plans de \mathcal{E} .

EXERCICES D'APPLICATION. (Ecrire le détail des calculs et faire des figures).

• *Position relative d'une droite et d'un plan.* Soient \mathcal{D} une droite et \mathcal{P} un plan dans \mathcal{E} . En utilisant des équations cartésiennes, montrer que les seuls cas possibles sont :

– la droite vectorielle Δ dirigeant \mathcal{D} est un sous-espace vectoriel du plan vectoriel Π dirigeant \mathcal{P} ;

 dans ce cas : ou bien $\mathcal{D} \subset \mathcal{P}$, alors $\mathcal{D} \cap \mathcal{P} = \mathcal{D}$;

 ou bien $\mathcal{D} \not\subset \mathcal{P}$, alors $\mathcal{D} \cap \mathcal{P} = \emptyset$;

– la droite vectorielle Δ n'est pas un sous-espace vectoriel de Π ; dans ce cas $\mathcal{D} \cap \mathcal{P}$ est un singleton.

• *Position relative de deux droites.* Soient \mathcal{D} et \mathcal{D}' deux droites dans \mathcal{E} . En utilisant des équations cartésiennes, montrer que les seuls cas possibles sont :

– les droites \mathcal{D} et \mathcal{D}' sont confondues, alors $\mathcal{D} \cap \mathcal{D}' = \mathcal{D}$;

– les droites \mathcal{D} et \mathcal{D}' sont parallèles mais non confondues, alors $\mathcal{D} \cap \mathcal{D}' = \emptyset$;

– les droites \mathcal{D} et \mathcal{D}' ne sont pas parallèles, alors : ($\mathcal{D} \cap \mathcal{D}' = \emptyset$) ou ($\mathcal{D} \cap \mathcal{D}'$) est un singleton.

7.1.7 Barycentre

Soit \mathcal{E} un espace affine sur \mathbb{R} , d'espace vectoriel associé E . La notion de barycentre d'une famille de points est un outil essentiel du travail dans les espaces affines, dont le rôle est comparable à celui de combinaison linéaire d'une famille de vecteurs dans le cadre des espaces vectoriels.

a) THÉORÈME. Soit $\mathcal{A} = (A_i, \alpha_i)_{1 \leq i \leq n}$ une famille finie de points pondérés, (ceci signifie que, pour tout $1 \leq i \leq n$, A_i est un point de \mathcal{E} , et α_i est un réel appelé le poids ou la masse affecté à A_i). On pose : $\sigma = \sum_{i=1}^n \alpha_i$, appelée la masse totale de la famille. On suppose que $\sigma \neq 0$; alors il existe un unique point G de \mathcal{E} vérifiant l'une des conditions suivantes équivalentes :

$$(1) \quad \boxed{\sum_{i=1}^n \alpha_i \overrightarrow{GA_i} = \vec{0}}, \quad (2) \quad \boxed{\exists M_0 \in \mathcal{E}, \overrightarrow{M_0G} = \frac{1}{\sigma} \sum_{i=1}^n \alpha_i \overrightarrow{M_0A_i}}, \quad (3) \quad \boxed{\forall M \in \mathcal{E}, \overrightarrow{MG} = \frac{1}{\sigma} \sum_{i=1}^n \alpha_i \overrightarrow{MA_i}}.$$

Preuve. Soit f l'application $\mathcal{E} \rightarrow E$ définie par $f(M) = \sum_{i=1}^n \alpha_i \overrightarrow{MA_i}$ pour tout $M \in \mathcal{E}$. Pour $M, N \in \mathcal{E}$, on a $f(M) - f(N) = \sum_{i=1}^n \alpha_i \overrightarrow{MA_i} - \sum_{i=1}^n \alpha_i \overrightarrow{NA_i} = \sum_{i=1}^n \alpha_i (\overrightarrow{MA_i} + \overrightarrow{A_iN}) = \sum_{i=1}^n \alpha_i \overrightarrow{MN}$. On retient :

$$\text{pour tous } M, N \in \mathcal{E}, \quad f(M) - f(N) = \sigma \overrightarrow{MN}. \quad (*)$$

Il en résulte que f est injective (en effet $f(M) = f(N) \Rightarrow \overrightarrow{MN} = \frac{1}{\sigma}(f(M) - f(N)) = \vec{0} \Rightarrow M = N$).

– On montre que les 3 assertions sont équivalentes. Si G vérifie (i), on a $f(G) = \vec{0}$; d'après (*) on a alors $f(M) = f(G) + \sigma \overrightarrow{MG} = \sigma \overrightarrow{MG}$ pour tout $M \in \mathcal{E}$, donc (iii) est vérifié. Il est clair que (iii) \Rightarrow (ii). Enfin, si G vérifie (ii), on a d'après (*) : $f(G) = f(M_0) + \sigma \overrightarrow{GM_0} = \sigma \overrightarrow{M_0G} + \sigma \overrightarrow{GM_0} = \vec{0}$.

– On montre l'existence de G . Soit $A \in \mathcal{E}$ quelconque ; alors pour le vecteur $\frac{1}{\sigma} f(A) \in E$, il existe $G \in \mathcal{E}$ tel que $\overrightarrow{AG} = \frac{1}{\sigma} f(A)$. En utilisant (*), il vient : $f(G) = f(A) + \sigma \overrightarrow{GA} = \sigma \overrightarrow{AG} + \sigma \overrightarrow{GA} = \vec{0}$.

– On montre l'unicité de G . Si G' est un autre point de \mathcal{E} vérifiant (i), on a $f(G) = \vec{0} = f(G')$, d'où $G = G'$ par injectivité de f . \square

DÉFINITION. Sous les hypothèses du théorème précédent, le point G est appelé le *barycentre* du système de points pondérés $(A_i, \alpha_i)_{1 \leq i \leq n}$. On le note :

$$G = \text{Bar}(A_i, \alpha_i)_{1 \leq i \leq n} \quad \text{ou} \quad G = \text{Bar} \left(\begin{matrix} A_1, \dots, A_i, \dots, A_n \\ \alpha_1, \dots, \alpha_i, \dots, \alpha_n \end{matrix} \right),$$

Homogénéité du barycentre. Il est clair que, pour tout $\lambda \in \mathbb{R}^*$, $\text{Bar}(A_i, \lambda \alpha_i)_{1 \leq i \leq n} = \text{Bar}(A_i, \alpha_i)_{1 \leq i \leq n}$. Donc, quitte à multiplier chaque poids par $\frac{1}{\sigma}$, on peut toujours supposer que $\sigma = 1$.

Un cas particulier important est celui où les masses sont toutes égales. Par homogénéité, on peut les prendre égales à 1. La somme des masses est $n \neq 0$, ce qui autorise la définition suivante.

b) DÉFINITION. Soient A_1, \dots, A_n des points de \mathcal{E} . Le barycentre $G = \text{Bar}(A_i, 1)_{1 \leq i \leq n}$ est appelé l'*isobarycentre* du n -uplet (A_1, \dots, A_n) . Il est défini par :

$$\boxed{\forall M \in \mathcal{E}, \quad \overrightarrow{MG} = \frac{1}{n} \sum_{i=1}^n \overrightarrow{MA_i}} \quad , \quad \text{ou encore} \quad \boxed{\sum_{i=1}^n \overrightarrow{GA_i} = \vec{0}}$$

Comme $\text{Bar}(A_i, 1)_{1 \leq i \leq n} = \text{Bar}(A_{s(i)}, 1)_{1 \leq i \leq n}$ pour toute permutation s de $\{1, 2, \dots, n\}$, on peut parler de l'isobarycentre des points A_1, A_2, \dots, A_n , indépendamment de l'ordre des points.

L'isobarycentre de 2 points de \mathcal{E} s'appelle leur milieu. Si A, B sont deux points de \mathcal{E} ,

$$[I \text{ milieu de } (A, B)] \iff [\overrightarrow{IA} + \overrightarrow{IB} = \vec{0}] \iff [\overrightarrow{AI} = \overrightarrow{IB}]$$

• La propriété suivante, évidente mais précieuse sur le plan pratique, indique que l'on peut regrouper les points par paquets, et que le barycentre global est alors le barycentre des barycentres partiels, affectés des sommes partielles des masses correspondantes.

c) PROPOSITION (associativité du barycentre). Soit $\mathcal{A} = (A_i, \alpha_i)_{1 \leq i \leq n}$ une famille finie de points pondérés de masse totale $\sum_{i=1}^n \alpha_i$ non-nulle. Notons G son barycentre. On suppose que, pour un entier $1 \leq p < n$, on ait $\sum_{i=1}^p \alpha_i \neq 0$, et on considère le barycentre partiel $G' = \text{Bar} \left(\begin{matrix} A_1, A_2, \dots, A_p \\ \alpha_1, \alpha_2, \dots, \alpha_p \end{matrix} \right)$. Alors :

$$G = \text{Bar} \left(\begin{matrix} A_1, \dots, A_p, A_{p+1}, \dots, A_n \\ \alpha_1, \dots, \alpha_p, \alpha_{p+1}, \dots, \alpha_n \end{matrix} \right) = \text{Bar} \left(\begin{matrix} G', & A_{p+1}, \dots, A_n \\ \alpha_1 + \dots + \alpha_p, & \alpha_{p+1}, \dots, \alpha_n \end{matrix} \right).$$

$$\text{Preuve. } \vec{0} = \sum_{i=1}^n \alpha_i \overrightarrow{GA_i} = \sum_{i=1}^p \alpha_i \overrightarrow{GA_i} + \sum_{i=p+1}^n \alpha_i \overrightarrow{GA_i} = \left(\sum_{i=1}^p \alpha_i \right) \overrightarrow{GG'} + \sum_{i=p+1}^n \alpha_i \overrightarrow{GA_i} \quad \square$$

★ *Exemple d'application.* Dans \mathcal{E} supposé de dimension ≥ 2 , soient A, B, C trois points non alignés, et A', B', C' les milieux respectifs de $(B, C), (C, A), (A, B)$. Par associativité, l'isobarycentre du triangle (A, B, C) est $G = \text{Bar} \left(\begin{smallmatrix} A, B, C \\ 1, 1, 1 \end{smallmatrix} \right) = \text{Bar} \left(\begin{smallmatrix} C', C \\ 2, 1 \end{smallmatrix} \right)$, de sorte que $\overrightarrow{CG} = 2\overrightarrow{GC'}$, d'où $G \in (CC')$. De même $G \in (AA')$ et $G \in (BB')$. On a prouvé : *les trois médianes d'un triangle se coupent en l'isobarycentre (ou centre de gravité) du triangle, situé au tiers de chacune d'elles à partir du côté.*

★ *Exercice d'application.* De la même façon, montrer que, dans \mathcal{E} supposé de dimension ≥ 3 , pour A, B, C, D quatre points non coplanaires, les 3 droites passant par le milieu d'une des 6 arêtes du tétraèdre et le milieu de l'arête opposée se coupent en l'isobarycentre G du tétraèdre. Montrer que G est aussi le point de concurrence des quatre droites joignant chaque sommet au centre de gravité du triangle opposé.

- On donne maintenant une caractérisation en termes de barycentre de la notion de sous-espace affine, exprimant qu'un sous-espace affine est une partie stable par barycentre.

d) PROPOSITION. Soit \mathcal{F} une partie non-vide de \mathcal{E} . Les conditions suivantes sont équivalentes :

- \mathcal{F} est un sous-espace affine de \mathcal{E} ;
- le barycentre de toute famille finie de points pondérés de \mathcal{F} appartient encore à \mathcal{F} .

Preuve. Supposons (i). Notons F le sous-espace vectoriel directeur de \mathcal{F} . Soit $(A_i, \alpha_i)_{1 \leq i \leq n}$ une famille finie de points pondérés dans \mathcal{F} telle que $\sigma = \alpha_1 + \dots + \alpha_n \neq 0$. Le barycentre G des (A_i, α_i) vérifie $\overrightarrow{MG} = \frac{1}{\sigma} \sum_{i=1}^n \alpha_i \overrightarrow{MA_i}$ pour tout $M \in \mathcal{E}$. Si $M \in \mathcal{F}$, on a $\overrightarrow{MA_i} \in F$ pour tout $1 \leq i \leq n$. Donc \overrightarrow{MG} étant combinaison linéaire des $\overrightarrow{MA_i}$, on a $\overrightarrow{MG} \in F$. Comme $M \in \mathcal{F}$, ceci implique que $G \in \mathcal{F}$.

Supposons (ii). Choisissons $A \in \mathcal{F}$. Il s'agit de montrer que $F := \{\overrightarrow{AM}; M \in \mathcal{F}\}$ est un sous-espace vectoriel de E . Pour cela, considérons $\overrightarrow{AM}, \overrightarrow{AN} \in F$ quelconques, avec $M, N \in \mathcal{F}$, et $\alpha, \beta \in \mathbb{R}$. Soit G le barycentre de (A, M, N) affectés des coefficients $(1 - \alpha - \beta, \alpha, \beta)$. D'après l'hypothèse (ii), $G \in \mathcal{F}$. Par définition du barycentre, $\overrightarrow{AG} = \alpha \overrightarrow{AM} + \beta \overrightarrow{AN}$. Mais $\overrightarrow{AG} \in F$ puisque $G \in \mathcal{F}$. On a ainsi vérifié que : $\alpha \overrightarrow{AM} + \beta \overrightarrow{AN} \in F$, ce qui prouve que F est un sous-espace vectoriel de E , et donc que \mathcal{F} est un sous-espace affine de \mathcal{E} . \square

- On termine par une formulation en termes de barycentre de la notion de base affine.

PROPOSITION. Soit $\mathcal{X} = (A_0, A_1, \dots, A_p)$ une famille affinement libre de points de \mathcal{E} . Soit \mathcal{F} le sous-espace affine de \mathcal{E} engendré par \mathcal{X} . Alors :

$$\forall M \in \mathcal{F}, \exists! (\alpha_0, \alpha_1, \dots, \alpha_p) \in \mathbb{R}^{p+1}, \sum_{i=0}^p \alpha_i = 1 \text{ et } M = \text{Bar}(A_i, \alpha_i)_{0 \leq i \leq p}.$$

Preuve. Comme on l'a vu en 7.1.4, \mathcal{X} est une base affine de \mathcal{F} , et $X = \{\overrightarrow{A_0A_1}, \overrightarrow{A_0A_2}, \dots, \overrightarrow{A_0A_p}\}$ est une base du sous-espace vectoriel F de E directeur de \mathcal{F} . Soit $M \in \mathcal{F}$. Donc $\overrightarrow{A_0M} \in F$. Soient $(\alpha_1, \dots, \alpha_p)$ les composantes de $\overrightarrow{A_0M}$ dans la base X . Soit $\alpha_0 = 1 - \sum_{i=1}^p \alpha_i$. On a $\overrightarrow{A_0M} = \sum_{i=0}^p \alpha_i \overrightarrow{A_0A_i}$ et $\sum_{i=0}^p \alpha_i = 1$. L'unicité des α_i résulte de la liberté de X . \square

La base affine \mathcal{X} de \mathcal{F} est parfois appelé un *repère barycentrique* de \mathcal{F} , et les réels $(\alpha_0, \alpha_1, \dots, \alpha_p)$ sont appelés les *coordonnées barycentriques* du point M de \mathcal{F} relativement à \mathcal{X} .

Par exemple, l'isobarycentre G d'un triangle (A, B, C) a pour coordonnées barycentriques $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ dans la famille affinement libre (A, B, C) . De nombreux exercices de géométrie affine se résolvent d'autant plus simplement que l'on choisit un repère barycentrique bien adapté au problème.

7.1.8 Cas particulier des espaces affines euclidiens.

On étudie dans cette dernière section des espaces affines qui sont munis d'une structure supplémentaire, à savoir une structure euclidienne. Ce sont donc des espaces affines dans lesquels tout ce que l'on vient d'étudier a du sens et s'applique, mais dans lesquels, en plus, on a des notions de distance, d'orthogonalité, d'angles,... qui n'auraient pas de sens dans le cas général car se déduisent de la structure euclidienne que l'on rajoute en plus.

a) NOTION D'ESPACE AFFINE EUCLIDIEN. On appelle *espace affine euclidien* tout espace affine \mathcal{E} sur \mathbb{R} dont l'espace vectoriel directeur E est euclidien, c'est-à-dire muni d'un produit scalaire (voir 2.1.1.c) et de dimension finie. On note $\|\cdot\|$ la norme et \cdot le produit scalaire.

En particulier \mathcal{E} est, comme E , de dimension finie. On notera $n = \dim \mathcal{E} = \dim E$. On supposera dans la suite $n \geq 2$.

Pour tous points A, B dans \mathcal{E} , on appelle *distance* de A à B le réel positif :

$$d(A, B) = \|\overrightarrow{AB}\| \in \mathbb{R}_+; \quad \text{on note aussi : } AB = d(A, B).$$

On déduit des propriétés de la norme (voir (voir 2.1.1.d et f) que, pour tous points $A, B, C \in \mathcal{E}$:

$$d(A, B) = d(B, A), \quad d(A, B) = 0 \Leftrightarrow A = B, \quad d(A, C) \leq d(A, B) + d(B, C).$$

L'existence d'une distance dans \mathcal{E} , associée au produit scalaire dans E , permet de considérer de façon naturelle, entre autres :

- *Propriété de Pythagore* (voir exercice 2 de 2.1.5) si $\overrightarrow{AB} \perp \overrightarrow{AC}$, alors $BC^2 = AB^2 + AC^2$

Le triangle (ABC) est alors dit rectangle en A . Le côté $[BC]$ est l'hypothénuse.

- *Repère orthogonal ou orthonormé.* Soit $\mathcal{R} = (0, \mathcal{B})$ un repère cartésien de \mathcal{E} . On dit que \mathcal{R} est un *repère orthogonal* (resp. *orthonormal*) lorsque la base \mathcal{B} de E est *orthogonale* (resp. *orthonormale*) au sens de (voir 2.1.2.c).

- *Sous-espaces affines orthogonaux.* Soient \mathcal{F} et \mathcal{H} deux sous-espaces affines de \mathcal{E} , de sous-espaces vectoriels directeurs respectifs F et H dans E . On dit que \mathcal{F} et \mathcal{H} sont orthogonaux lorsque F et H le sont (tout vecteur de F est orthogonal à tout vecteur de H .)

- *Vecteur normal à un hyperplan affine.* Soit $\mathcal{R} = (0, \mathcal{B})$ un repère orthonormé de \mathcal{E} . Soit \mathcal{H} un hyperplan de \mathcal{E} . On considère une équation de \mathcal{H} p/r à \mathcal{R} :

$$a_1x_1 + \cdots + a_nx_n + a_{n+1} = 0, \quad \text{avec } (a_1, \dots, a_n) \text{ non-nul dans } \mathbb{R}^n \text{ et } a_{n+1} \in \mathbb{R}.$$

On appelle *vecteur normal* à \mathcal{H} tout vecteur normal à l'hyperplan vectoriel H directeur de \mathcal{H} , c'est-à-dire tout vecteur dirigeant la droite H^\perp , c'est-à-dire encore (voir exercice 3 de 2.1.5) tout vecteur non-nul colinéaire au vecteur \vec{n} de composantes (a_1, \dots, a_n) dans la base \mathcal{B} de E .

b) DISTANCE ENTRE UN POINT ET UN SOUS-ESPACE AFFINE

- *Définition.* Soit \mathcal{F} un sous-espace affine de \mathcal{E} et soit A un point de \mathcal{E} . On appelle distance de A à \mathcal{F} le réel positif : $d(A, \mathcal{F}) = \inf_{M \in \mathcal{F}} d(A, M)$. En particulier : $A \in \mathcal{F} \Leftrightarrow d(A, \mathcal{F}) = 0$.

- *Proposition.* Soit \mathcal{F} un sous-espace affine de \mathcal{E} , de sous-espace vectoriel directeur F . Soit A un point de \mathcal{E} . Il existe un unique point A' de \mathcal{F} tel que $d(A, \mathcal{F}) = d(A, A')$. Ce point A' est le point d'intersection de \mathcal{F} avec le sous-espace affine passant par A et dirigé par F^\perp .

A' est appelé le *projeté orthogonal* de A sur \mathcal{F} .

Preuve Comme F et F^\perp sont supplémentaires orthogonaux dans E (voir 2.1.2.b), l'intersection de \mathcal{F} avec le sous-espace affine de \mathcal{E} passant par A et dirigé par F^\perp est un singleton. Notons-le $\{A'\}$. Ainsi $A' \in \mathcal{F}$ et $\overrightarrow{AA'} \in F^\perp$. Pour tout $M \in \mathcal{F}$, $\overrightarrow{A'M} \in F$ d'où $\overrightarrow{AA'} \cdot \overrightarrow{A'M} = 0$. Dès lors, avec la propriété de Pythagore : $(AM)^2 = (AA')^2 + (A'M)^2 \geq (AA')^2$ et $AM = AA'$ si et seulement si $A'M = 0$, ie. $M = A'$. Ceci prouve que $AA' = \inf\{AM; M \in \mathcal{F}\}$. \square

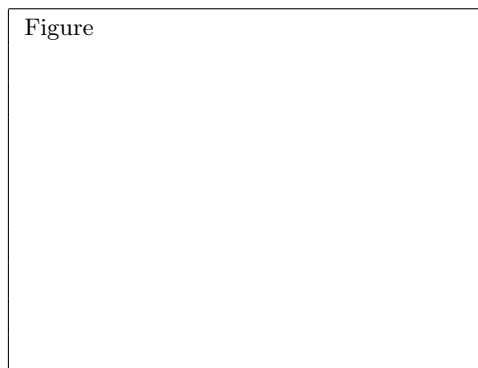
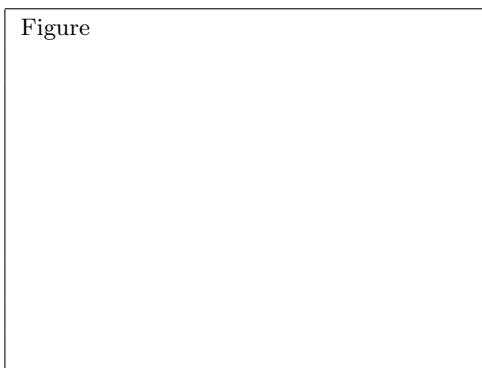
• *Corollaire.* (Calcul explicite dans le cas d'un hyperplan)

Soit \mathcal{R} un repère orthonormé de \mathcal{E} . Soit $A \in \mathcal{E}$ quelconque, de coordonnées $(\alpha_1, \dots, \alpha_n)$ par rapport à \mathcal{R} . Soit \mathcal{H} un hyperplan de \mathcal{E} d'équation cartésienne $a_1x_1 + \dots + a_nx_n + a_{n+1} = 0$ par rapport à \mathcal{R} , avec (a_1, \dots, a_n) non-nul dans \mathbb{R}^n . Alors :

$$d(A, \mathcal{H}) = \frac{|a_1\alpha_1 + \dots + a_n\alpha_n + a_{n+1}|}{\sqrt{a_1^2 + \dots + a_n^2}}$$

Preuve Notons A' le projeté orthogonal de A sur \mathcal{F} et $(\beta_1, \dots, \beta_n)$ ses coordonnées par rapport à \mathcal{R} . On sait que $\vec{n}(a_1, \dots, a_n)$ est normal à \mathcal{H} , et A' vérifie par définition $A' \in \mathcal{H}$ et $\overrightarrow{AA'}$ est colinéaire à \vec{n} . Il existe donc $\lambda \in \mathbb{R}$ tel que $\overrightarrow{AA'} = \lambda \vec{n}$. D'après la proposition précédente, $d(A, \mathcal{H}) = AA' = |\lambda| \times \|\vec{n}\|$, avec $\|\vec{n}\| = \sqrt{a_1^2 + \dots + a_n^2}$.

D'une part $A' \in \mathcal{H}$, donc : $a_1\beta_1 + \dots + a_n\beta_n + a_{n+1} = 0$. D'autre part $\overrightarrow{AA'} = \lambda \vec{n}$, donc : $(\beta_1 - \alpha_1, \dots, \beta_n - \alpha_n) = \lambda(a_1, \dots, a_n)$. D'où : $a_1(\alpha_1 + \lambda a_1) + \dots + a_n(\alpha_n + \lambda a_n) + a_{n+1} = 0$, ou encore : $\lambda = -\frac{a_1\alpha_1 + \dots + a_n\alpha_n + a_{n+1}}{a_1^2 + \dots + a_n^2}$. Ce qui prouve le résultat voulu. \square



► Si $\dim \mathcal{E} = 2$, $\mathcal{H} = \mathcal{D}$ d'équation $ax + by + c = 0$; $\vec{n}(a, b)$ normal ;

si $A(\alpha, \beta) \in \mathcal{E}$, alors : $d(A, \mathcal{D}) = \frac{|a\alpha + b\beta + c|}{\sqrt{a^2 + b^2}}$.

► Si $\dim \mathcal{E} = 3$, $\mathcal{H} = \mathcal{P}$ d'équation $ax + by + cz + d = 0$; $\vec{n}(a, b, c)$ normal ;

si $A(\alpha, \beta, \gamma) \in \mathcal{E}$, alors : $d(A, \mathcal{P}) = \frac{|a\alpha + b\beta + c\gamma + d|}{\sqrt{a^2 + b^2 + c^2}}$.

c) HYPERPLAN MÉDIATEUR

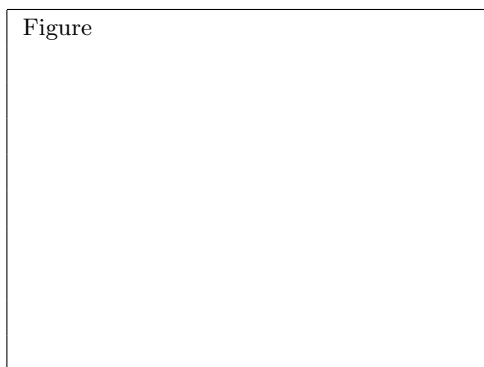
• *Lemme.* Soient $\vec{v} \in E$ non-nul, $O \in \mathcal{E}$, $\lambda \in \mathbb{R}$. Alors l'ensemble $\mathcal{H} = \{M \in \mathcal{E}; \vec{v} \cdot \overrightarrow{OM} = \lambda\}$ est un hyperplan affine de \mathcal{E} , et \vec{v} est un vecteur normal à \mathcal{H} .

Preuve On choisit une base orthonormale \mathcal{B} de E . On considère le repère orthonormal $\mathcal{R} = (O, \mathcal{B})$. Soit (a_1, \dots, a_n) les composantes de \vec{v} par rapport à \mathcal{B} . Pour tout point $M \in \mathcal{E}$ de coordonnées (x_1, \dots, x_n) par rapport à \mathcal{R} , on a : $\vec{v} \cdot \overrightarrow{OM} = \lambda$ si et seulement si $a_1x_1 + \dots + a_nx_n = \lambda$, d'où le résultat. \square

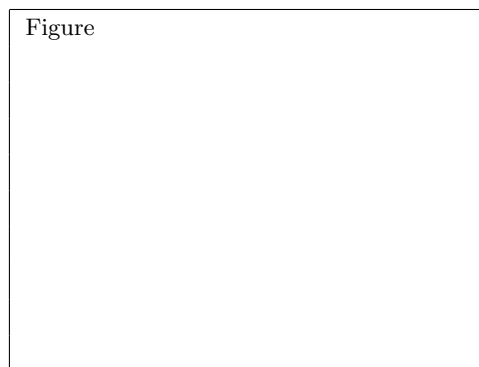
• *Théorème et définition.* Soient A, B deux points distincts de \mathcal{E} . Alors :

- (i) L'ensemble $\mathcal{H} = \{M \in \mathcal{E} ; d(A, M) = d(B, M)\}$ des points équidistants de A et B est un hyperplan affine de \mathcal{E} , appelé l'hyperplan médiateur de (A, B)
- (ii) \mathcal{H} est l'hyperplan affine passant par le milieu I de (A, B) est dirigé par l'hyperplan vectoriel $H = \Delta^\perp$ de E , où Δ est la droite vectorielle de E dirigée par \overrightarrow{AB} .

Preuve $M \in \mathcal{H}$ si et seulement si $AM^2 = BM^2$; or $AM^2 - BM^2 = \overrightarrow{AM} \cdot \overrightarrow{AM} - \overrightarrow{BM} \cdot \overrightarrow{BM} = \overrightarrow{AM} \cdot \overrightarrow{AM} - (\overrightarrow{AM} - \overrightarrow{AB}) \cdot (\overrightarrow{AM} - \overrightarrow{AB}) = -\overrightarrow{AB} \cdot \overrightarrow{AB} + 2\overrightarrow{AB} \cdot \overrightarrow{AM}$. Ainsi $M \in \mathcal{H}$ si et seulement si $\overrightarrow{AB} \cdot \overrightarrow{AM} = \frac{1}{2}AB^2$, d'où le résultat en appliquant le lemme précédent avec $\vec{v} = \overrightarrow{AB} \neq \vec{0}_E$ et $\lambda = \frac{1}{2}AB^2 \in \mathbb{R}_+$. \square



dim $\mathcal{E} = 2$: médiatrice



$\mathcal{E} = 3$: plan médiateur

d) SPHÈRES

• DÉFINITION. Soient $O \in \mathcal{E}$ et $r \in \mathbb{R}_+$. La sphère de centre O et de rayon r est :

$$S(O, r) = \{M \in \mathcal{E} ; d(O, M) = r\}.$$

Dans le cas où $\dim \mathcal{E} = 2$, $S(O, r)$ est appelé cercle de centre O et de rayon r , noté $C(O, r)$.

• THÉORÈME FONDAMENTAL. Soit (A_0, A_1, \dots, A_n) une famille de $(n + 1)$ points affinement libre dans \mathcal{E} . Il existe une unique sphère de \mathcal{E} passant par A_0, A_1, \dots, A_n , dont le centre est le point d'intersection des hyperplans médiateurs des bipoints (A_i, A_j) , $1 \leq i \neq j \leq n$.

Preuve Rappelons que n est ici la dimension de \mathcal{E} . Pour tout $1 \leq i \leq n$, notons :

- $\mathcal{H}_i = \{M \in \mathcal{E} ; d(A_0, M) = d(A_i, M)\}$ l'hyperplan médiateur de (A_0, A_i) ,
- H_i l'hyperplan vectoriel de E directeur de \mathcal{H}_i ,
- $\vec{n}_i = \overrightarrow{A_0A_i}$ qui est normal à \mathcal{H}_i d'après 1.4.b.

Par hypothèse, la famille $\mathcal{N} = \{\vec{n}_1, \dots, \vec{n}_n\}$ est libre dans E (donc est base de E). Fixons un repère orthonormal $\mathcal{R} = (O, \mathcal{B})$ de \mathcal{E} , et notons $(a_{i,1}, \dots, a_{i,n}) \in \mathbb{R}^n$ les composantes de \vec{n}_i dans \mathcal{B} . On sait qu'alors il existe $a_{i,n+1} \in \mathbb{R}$ telle qu'une équation de \mathcal{H}_i dans \mathcal{R} soit $(e_i) : a_{i,1}x_1 + \dots + a_{i,n}x_n + a_{i,n+1} = 0$, et ceci pour tout $1 \leq i \leq n$. Posons :

$$\mathcal{F} = \{M \in \mathcal{E} ; d(A_0, M) = d(A_1, M) = \dots = d(A_n, M)\} = \bigcap_{1 \leq i \leq n} \mathcal{H}_i.$$

Il est clair que \mathcal{F} est l'intersection des hyperplans médiateurs de tous les bipoints (A_i, A_j) , $1 \leq i \neq j \leq n$. Soit $M \in \mathcal{E}$, soit (x_1, \dots, x_n) ses coordonnées dans le repère \mathcal{R} . Alors $M \in \mathcal{F}$ ssi (x_1, \dots, x_n) est solution du système linéaire (Σ) formé par les n équations à n inconnues

$(e_1), \dots, (e_n)$. D'après ce qui précède, la matrice de (Σ) n'est autre que la transposée de la matrice de la base \mathcal{N} dans la base \mathcal{B} . Elle est donc inversible, d'où système de Cramer et \mathcal{F} est un singleton $\{O\}$. Ainsi $d(A_0, O) = d(A_1, O) = \dots = d(A_n, O)$, et en notant r cette valeur commune et \mathcal{S} la sphère de centre O et de rayon r , on a bien $A_i \in \mathcal{S}$ pour tout $0 \leq i \leq n$.

• Montrons maintenant l'unicité. Soit \mathcal{S}' une sphère de centre O' et de rayon r' passant par A_0, A_1, \dots, A_n . En particulier, $d(A_0, O') = d(A_i, O') = r'$ pour tout $1 \leq i \leq n$, ce qui prouve que $O' \in \mathcal{H}_i$ pour tout $1 \leq i \leq n$, c'est-à-dire $O' \in \mathcal{F}$. D'après l'étape précédente, on conclut $O' = O$. Dès lors, $r' = d(A_0, O') = d(A_0, O) = r$, d'où finalement $\mathcal{S}' = \mathcal{S}$. \square

• *Corollaire.* (Cas de la dimension 2 : cocyclicité) On suppose $\dim \mathcal{E} = 2$. Soient A, B, C trois points non alignés de \mathcal{E} .

(i) les médiatrices de (A, B) , (B, C) et (C, A) sont concourantes en un point O .

(ii) Il existe un unique cercle passant par A, B, C , dit cercle circonscrit au triangle (ABC) ; son centre est O .

Des points du plan \mathcal{E} étant dits cocycliques s'il appartiennent à un même cercle, le corollaire ci-dessus exprime que trois points non-alignés du plan sont toujours cocycliques.

Figure



• *Corollaire.* (Cas de dimension 3). Si $\dim \mathcal{E} = 3$, par quatre points non coplanaires il passe une unique sphère, dont le centre est l'intersection des plans médiateurs des côtés du tétraèdre.

7.1.9 Exercices

EXERCICE 1. Soit \mathcal{E} un espace affine euclidien de dimension $n \geq 2$. On fixe un réel $a > 0$. On considère dans \mathcal{E} trois points distincts non alignés A, B, C tels que $AB = 3a$, $AC = 4a$ et $BC = 5a$. On note ψ l'application $\mathcal{E} \rightarrow \mathbb{R}$ définie par $\psi(M) = -5MA^2 + 4MB^2 + 3MC^2$ pour tout $M \in \mathcal{E}$.

a) Soit G le barycentre de $\{(A, -5), (B, 4), (C, 3)\}$. Exprimer les coordonnées de G dans le repère orthogonal $(A; \overrightarrow{AB}, \overrightarrow{AC})$ du plan (ABC) . En déduire la valeur de $\psi(G)$ en fonction de a , puis pour tout point $M \in \mathcal{E}$ l'expression de $\psi(M)$ en fonction de a et de MG .

b) Déterminer pour tout $k \in \mathbb{R}$ l'ensemble S_k des points $M \in \mathcal{E}$ vérifiant $\psi(M) = ka^2$.

EXERCICE 2. Soit \mathcal{E} un espace affine de dimension 3 sur \mathbb{R} . On fixe quatre points non coplanaires A, B, C, D dans \mathcal{E} . On note \mathcal{R} le repère cartésien $(A; \overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD})$ de \mathcal{E} . On introduit les points : E le barycentre de $\{(A, 1), (B, a)\}$, F le barycentre de $\{(B, 1), (C, b)\}$, G le barycentre de $\{(C, 1), (D, c)\}$ et H le barycentre de $\{(D, 1), (A, d)\}$, où $a, b, c, d \in \mathbb{R}$ différents de -1 .

a) Donner une condition nécessaire et suffisante portant sur a, b, c, d pour que les quatre points E, F, G, H soient coplanaires.

b) Expliciter une équation cartésienne par rapport à \mathcal{R} de chacun des plans (ECD) , (FDA) , (GAB) et (HBC) . En déduire une condition nécessaire et suffisante portant sur a, b, c, d pour que ces quatre plans aient un point commun.

EXERCICE 3. Soit \mathcal{E} un espace affine euclidien de dimension 2. On fixe un repère orthonormé \mathcal{R} de \mathcal{E} (toutes les coordonnées considérées sont par rapport à ce repère). On considère trois réels $b > 0$, $c > 0$ et $k \neq -1$, ainsi que les points $B(b, 0)$ et $C(0, c)$. On note M le barycentre de $\{(B, 1), (C, k)\}$.

- Calculer les coordonnées de M et la distance OM en fonction de b, c, k .
- Déterminer la valeur de k et de OM : (i) lorsque (OM) est la médiane issue de O du triangle (OBC) , (ii) lorsque (OM) est la hauteur issue de O du triangle (OBC) , (iii) lorsque (OM) est la bissectrice intérieure issue de O du triangle (OBC) .
- Calculer les coordonnées des points E et D tels que $(BCDE)$ soit un carré (attention, il y a plusieurs solutions).
- Soient N et P les projetés orthogonaux de M sur les (OB) et (OC) respectivement. Exprimer l'aire \mathcal{A} du rectangle $(ONMP)$ en fonction de b, c, k ; pour quelle valeur de k est-elle maximale?

EXERCICE 4. Soient A, B, C trois points non alignés d'un plan affine \mathcal{E} , et a, b, c trois réels tels que $a + b \neq 0$, $b + c \neq 0$ et $a + c \neq 0$. On considère : A' le barycentre de $\{(B, b), (C, c)\}$, B' le barycentre de $\{(A, a), (C, c)\}$ et C' le barycentre de $\{(A, a), (B, b)\}$.

- Soient α, β, γ trois réels de somme non-nulle, et M le barycentre de $\{(A, \alpha), (B, \beta), (C, \gamma)\}$. Montrer que $M \in (AA')$ si et seulement si $c\beta = b\gamma$.
- Montrer que : (i) si $a + b + c \neq 0$, alors les trois droites (AA') , (BB') et (CC') sont concourantes; (ii) si $a + b + c = 0$, alors les trois droites (AA') , (BB') et (CC') sont parallèles.

EXERCICE 5. Soit \mathcal{E} un espace affine euclidien de dimension 3. On fixe un repère orthonormé \mathcal{R} de \mathcal{E} (toutes les coordonnées de points et équations de sous-espaces considérées sont par rapport à ce repère). On fixe $a \in \mathbb{R}$, $a \neq 0$. On considère les quatre plans \mathcal{F}_i , pour $1 \leq i \leq 4$, d'équations respectives :

$$\mathcal{F}_1 : x + y + z = 3a \quad \mathcal{F}_2 : x - y + z = 7a \quad \mathcal{F}_3 : x + y - z = a \quad \mathcal{F}_4 : -x + y + z = a.$$

- Montrer que $\mathcal{F}_2 \cap \mathcal{F}_3 \cap \mathcal{F}_4$ est un singleton $\{A\}$ et déterminer les coordonnées de A . Déterminer de même $\mathcal{F}_1 \cap \mathcal{F}_3 \cap \mathcal{F}_4 = \{B\}$, $\mathcal{F}_1 \cap \mathcal{F}_2 \cap \mathcal{F}_4 = \{C\}$ et $\mathcal{F}_1 \cap \mathcal{F}_2 \cap \mathcal{F}_3 = \{D\}$. Montrer que A, B, C, D ne sont pas coplanaires.
- Déterminer une équation de chacun des trois plans $\mathcal{H}, \mathcal{H}', \mathcal{H}''$, plans médiateurs respectifs des bipoints (A, B) , (A, C) et (A, D) . Déterminer le centre Ω et le rayon de la sphère circonscrite au tétraèdre $ABCD$.
- On note A' le projeté orthogonal de A sur \mathcal{F}_1 . Donner une représentation paramétrique de la droite (AA') et montrer qu'elle passe par Ω ; déterminer les coordonnées de A' et calculer la distance $d(\Omega, \mathcal{F}_1)$. Calculer de même $d(\Omega, \mathcal{F}_i)$ pour $i = 2, 3, 4$.
- Déterminer l'isobarycentre des points A, B, C, D .

EXERCICE 6. Soit ABC un triangle d'un plan affine \mathcal{E} . On fixe un réel $p \in [0, \frac{1}{2}]$ et on considère les points A', B', C' de \mathcal{E} définis par $\overrightarrow{AC'} = p\overrightarrow{AB}$, $\overrightarrow{BA'} = p\overrightarrow{BC}$ et $\overrightarrow{CB'} = p\overrightarrow{CA}$. Les droites (AA') et (BB') se coupent en un point K , les droites (BB') et (CC') en un point I , les droites (AA') et (CC') en un point J . Le but de l'exercice est de comparer les aires des triangles IJK et ABC .

- Ecrire les points A', B', C' comme barycentres des points A, B, C . Montrer que le point I est barycentre de A, B, C affectés des coefficients respectifs $p, \frac{p^2}{1-p}$ et $1 - p$. Exprimer de même J et K comme des barycentres de A, B, C .

b) Montrer que I est le barycentre des deux points C et J affectés des coefficients respectifs $1 - 2p$ et p . Exprimer de même J comme barycentre de A et K , et K comme barycentre de B et I .

c) On note $\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{A}_4, \mathcal{A}_5, \mathcal{A}_6, \mathcal{A}_7$ les aires respectives des triangles $ABC, ABJ, AJI, IJK, BJK, BCK, CIK$ et ACI . Montrer que $\mathcal{A}_2/\mathcal{A}_7 = IJ/CI = (\mathcal{A}_3 + \mathcal{A}_4)/(\mathcal{A}_5 + \mathcal{A}_6)$. En posant $\lambda = IJ/CI$, écrire quatre autres égalités du même type pour en déduire $\mathcal{A}_3/\mathcal{A}_0$ en fonction de λ . Vérifier que $\lambda = \frac{1-2p}{p}$ et conclure que : Aire(IJK)/Aire(ABC) = $\frac{(1-2p)^2}{p^2-p+1}$.

Que devient ce rapport lorsque $p = 0$? $p = \frac{1}{2}$? $p = \frac{1}{3}$?

7.2 Applications affines

7.2.1 Notion d'application affine

Bien que le contenu de cette section puisse être dans sa presque totalité rédigé pour des applications affines d'un espace affine \mathcal{E} vers un espace affine \mathcal{F} , on se limite, en vue des applications concrètes qui sont celles du programme de ce cours, au cas où l'espace d'arrivée \mathcal{F} est le même que l'espace de départ \mathcal{E} .

On se fixe pour toute la suite un espace affine \mathcal{E} sur \mathbb{R} , d'espace vectoriel associé E .

a) DÉFINITION. Une application $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ est une *application affine*, ou un *endomorphisme affine*, lorsqu'il existe une application linéaire $f : E \rightarrow E$, dite *associée* à φ , telle que :

$$\overrightarrow{\varphi(A)\varphi(B)} = f(\overrightarrow{AB}) \quad \text{pour tous } A, B \in \mathcal{E}$$

Le théorème fondamental suivant, très utile dans la pratique, exprime qu'une application affine est entièrement déterminée par son application linéaire associée et par l'image d'un point.

b) THÉORÈME. Soient A et B deux points de \mathcal{E} , et f une application linéaire $E \rightarrow E$. Alors il existe une unique application affine $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ telle que $\varphi(A) = B$ et telle que f soit l'application linéaire associée à φ .

Preuve. Montrons d'abord l'unicité; Pour cela, soit $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ affine d'application linéaire associée f et telle que $\varphi(A) = B$. Pour tout $M \in \mathcal{E}$, on a : $\overrightarrow{B\varphi(M)} = \overrightarrow{\varphi(A)\varphi(M)} = f(\overrightarrow{AM})$, ce qui définit de façon unique $\varphi(M)$ pour tout $M \in \mathcal{E}$. D'où l'unicité de φ .

Réciproquement, définissons $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ en définissant, pour tout $M \in \mathcal{E}$, $\varphi(M)$ comme le point de \mathcal{E} tel que $\overrightarrow{B\varphi(M)} = f(\overrightarrow{AM})$. Il en résulte en particulier que $\overrightarrow{B\varphi(A)} = f(\overrightarrow{AA}) = f(\overrightarrow{0}) = \overrightarrow{0}$, ce qui implique $\varphi(A) = B$. D'autre part, pour tous $M, N \in \mathcal{E}$, on a :

$$\overrightarrow{\varphi(M)\varphi(N)} = \overrightarrow{B\varphi(N)} - \overrightarrow{B\varphi(M)} = f(\overrightarrow{AN}) - f(\overrightarrow{AM}) = f(\overrightarrow{AN} - \overrightarrow{AM}) = f(\overrightarrow{MN}),$$

Ceci prouve que φ est affine, d'application linéaire associée f . □

c) EXEMPLES. Les translations, les homothéties, les projections, les symétries, sont des endomorphismes affines, que l'on détaillera dans les paragraphes suivants. Si de plus l'espace affine \mathcal{E} est supposé euclidien, on voit apparaître parmi les endomorphismes affines tous les types d'isométries (rotations,...) ou de similitudes.

Avant de développer géométriquement certains de ces exemples, on donne une série de propriétés générales des applications affines, qui ne découlent en fait que de la définition.

d) PROPOSITION (Conservation des sous-espaces affines). Soit $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ une application affine d'application linéaire associée $f : E \rightarrow E$.

- (i) Soit \mathcal{H} un sous-espace affine de \mathcal{E} dirigé par le sous-espace vectoriel H de E .
Alors $\varphi(\mathcal{H})$ est un sous-espace affine de \mathcal{E} , dirigé par le sous-espace vectoriel $f(H)$ de E .
- (ii) Soit \mathcal{H}' un sous-espace affine de \mathcal{E} dirigé par le sous-espace vectoriel H' de E .
Si $\varphi^{-1}(\mathcal{H}') \neq \emptyset$, alors $\varphi^{-1}(\mathcal{H}')$ est un sous-espace affine de \mathcal{E} , dirigé par le sous-espace vectoriel $f^{-1}(H')$ de E .

Preuve. Soient $A \in \mathcal{H} \subset \mathcal{E}$, et $B = \varphi(A) \in \varphi(\mathcal{H})$. Pour tout $M \in \mathcal{E}$, on a :

$$\begin{aligned} [M \in \varphi(\mathcal{H})] &\Leftrightarrow [\exists N \in \mathcal{H}, M = \varphi(N)] \\ &\Leftrightarrow [\exists N \in \mathcal{H}, \overrightarrow{BM} = \overrightarrow{\varphi(A)\varphi(N)}] \Leftrightarrow [\exists N \in \mathcal{H}, \overrightarrow{BM} = f(\overrightarrow{AN})], \end{aligned}$$

Ce qui, comme $A \in \mathcal{H}$ et \mathcal{H} dirigé par H , équivaut à l'existence de $\vec{u} \in H$ tel que $\overrightarrow{BM} = f(\vec{u})$. Ceci prouve que l'ensemble $\{\overrightarrow{BM}; M \in \varphi(\mathcal{H})\}$ est égal à $f(H)$, qui est un sous-espace vectoriel de E (car image directe d'un sous-espace vectoriel par une application linéaire). Le point (i) est démontré.

Pour (ii), supposons qu'il existe $A \in \varphi^{-1}(\mathcal{H}')$. Donc $\varphi(A) \in \mathcal{H}'$. Pour tout $M \in \mathcal{E}$, on a : $M \in \varphi^{-1}(\mathcal{H}')$ si et seulement si $\varphi(M) \in \mathcal{H}'$, ce qui équivaut à $\overrightarrow{\varphi(A)\varphi(M)} \in \overrightarrow{H'}$ car $\varphi(A) \in \mathcal{H}'$ et \mathcal{H}' dirigé par H' . En résumé, $M \in \varphi^{-1}(\mathcal{H}')$ si et seulement si $f(\overrightarrow{AM}) \in H'$. On déduit que l'ensemble $\{\overrightarrow{AM}; M \in \varphi^{-1}(\mathcal{H}')$ est égal à $f^{-1}(H')$, qui est un sous-espace vectoriel de E comme image réciproque d'un sous-espace vectoriel par une application linéaire. \square

e) COROLLAIRE (Conservation du parallélisme). Soit $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ une application affine, et $\mathcal{H}_1, \mathcal{H}_2$ deux sous-espaces affines de \mathcal{E} . Si \mathcal{H}_1 et \mathcal{H}_2 sont parallèles, alors $\varphi(\mathcal{H}_1)$ et $\varphi(\mathcal{H}_2)$ sont parallèles.

Preuve. Si \mathcal{H}_1 et \mathcal{H}_2 sont parallèles, on a $H_1 = H_2$ dans E . Donc $f(H_1) = f(H_2)$. Or $f(H_1)$ est le sous-espace vectoriel directeur de $\varphi(\mathcal{H}_1)$, et $f(H_2)$ celui de $\varphi(\mathcal{H}_2)$. D'où $\varphi(\mathcal{H}_1)$ et $\varphi(\mathcal{H}_2)$ parallèles. \square

f) COROLLAIRE (Conservation de l'alignement). Soit $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ une application affine. Quels que soient A, B, C trois points distincts alignés dans \mathcal{E} , les points $\varphi(A), \varphi(B), \varphi(C)$ sont alignés ou confondus dans \mathcal{E} .

Preuve. Si $\overrightarrow{AC} = \lambda \overrightarrow{AB}$ avec $\lambda \in \mathbb{R}, \lambda \neq 0, \lambda \neq 1$, alors $f(\overrightarrow{AC}) = \lambda f(\overrightarrow{AB})$, et donc $\overrightarrow{\varphi(A)\varphi(C)} = \lambda \overrightarrow{\varphi(A)\varphi(B)}$, ce qui prouve le résultat voulu. \square

En particulier une bijection affine transforme trois points alignés en trois points alignés. On peut montrer par ailleurs (on ne le fera pas ici) que réciproquement, une bijection $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ qui conserve l'alignement est nécessairement affine.

g) PROPOSITION (Conservation des barycentres). Une application $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ est affine si et seulement si elle vérifie la propriété suivante : pour toute famille $(A_i, \alpha_i)_{1 \leq i \leq n}$ de points pondérés de \mathcal{E} admettant un barycentre G , le point $\varphi(G)$ est le barycentre de la famille $(\varphi(A_i), \alpha_i)_{1 \leq i \leq n}$.

Preuve. Supposons φ affine et notons $f : E \rightarrow E$ l'application linéaire associée à φ . Soit $(A_i, \alpha_i)_{1 \leq i \leq n}$ est une famille de points pondérés de \mathcal{E} avec $\sigma = \alpha_1 + \dots + \alpha_n \neq 0$, et soit G son barycentre. Pour tout $M \in \mathcal{E}$, on a : $\overrightarrow{MG} = \frac{1}{\sigma} \sum_{i=1}^n \alpha_i \overrightarrow{MA_i}$ donc : $f(\overrightarrow{MG}) = f(\frac{1}{\sigma} \sum_{i=1}^n \alpha_i \overrightarrow{MA_i}) = \frac{1}{\sigma} \sum_{i=1}^n \alpha_i f(\overrightarrow{MA_i})$, c'est-à-dire : $\overrightarrow{\varphi(M)\varphi(G)} = \frac{1}{\sigma} \sum_{i=1}^n \alpha_i \overrightarrow{\varphi(M)\varphi(A_i)}$. Donc $\varphi(G)$ est bien le barycentre de la famille image.

Réciproquement supposons que φ conserve les barycentres. Fixons A un point de \mathcal{E} . Pour tout $\vec{u} \in E$, il existe $M \in \mathcal{E}$ unique tel que $\overrightarrow{AM} = \vec{u}$. Posons alors $f(\vec{u}) = \overrightarrow{\varphi(A)\varphi(M)}$. On définit ainsi une application $f : E \rightarrow F$; il s'agit de montrer qu'elle est linéaire.

Soient donc $\vec{u}, \vec{v} \in E$, $\lambda, \mu \in \mathbb{R}$. Il existe $M, N \in \mathcal{E}$ tels que $\vec{u} = \overrightarrow{AM}$ et $\vec{v} = \overrightarrow{AN}$. En appliquant l'hypothèse que φ conserve les barycentres, on peut considérer :

$$G = \text{Bar} \left(\begin{array}{c} A, M, N \\ 1-\lambda-\mu, \lambda, \mu \end{array} \right) \quad \text{et} \quad \varphi(G) = \text{Bar} \left(\begin{array}{c} \varphi(A), \varphi(M), \varphi(N) \\ 1-\lambda-\mu, \lambda, \mu \end{array} \right)$$

D'une part $\overrightarrow{AG} = (1-\lambda-\mu)\overrightarrow{AA} + \lambda\overrightarrow{AM} + \mu\overrightarrow{AN} = \lambda\vec{u} + \mu\vec{v}$, ce qui implique par définition de f que : $f(\lambda\vec{u} + \mu\vec{v}) = \overrightarrow{\varphi(A)\varphi(G)}$. D'autre part, $\overrightarrow{\varphi(A)\varphi(G)} = (1-\lambda-\mu)\overrightarrow{\varphi(A)\varphi(A)} + \lambda\overrightarrow{\varphi(A)\varphi(M)} + \mu\overrightarrow{\varphi(A)\varphi(N)} = \lambda f(\vec{u}) + \mu f(\vec{v})$. On a ainsi vérifié que $f(\lambda\vec{u} + \mu\vec{v}) = \lambda f(\vec{u}) + \mu f(\vec{v})$. \square

h) REMARQUE. La détermination et l'étude géométrique concrète des applications affines passe souvent par la détermination de leurs points fixes. Avant de le voir sur des exemples, donnons quelques résultats généraux qui découlent simplement de la définition. Pour toute application $\varphi : \mathcal{E} \rightarrow \mathcal{E}$, on note :

$$\text{Fix } \varphi = \{M \in \mathcal{E} ; \varphi(M) = M\}.$$

Par ailleurs, pour toute application linéaire $f : E \rightarrow E$, on considère le sous-espace propre associé à la valeur propre 1 :

$$\text{Fix } f = \text{Ker}(f - \text{id}_E) = \{\vec{u} \in E ; f(\vec{u}) = \vec{u}\}.$$

i) PROPOSITION (Points fixes d'une application affine). Soit $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ une application affine, d'application linéaire associée f . Alors :

- ou bien φ n'admet aucun point fixe,
- ou bien $\text{Fix } \varphi$ est un sous-espace affine de \mathcal{E} dont le sous-espace vectoriel directeur est $\text{Fix } f$.

Preuve. Supposons $\text{Fix } \varphi \neq \emptyset$. Considérons un point $A \in \mathcal{E}$ tel que $\varphi(A) = A$. Pour tout $M \in \mathcal{E}$, on a :

$$M = \varphi(M) \Leftrightarrow \overrightarrow{A\varphi(M)} = \overrightarrow{AM} \Leftrightarrow \overrightarrow{\varphi(A)\varphi(M)} = \overrightarrow{AM} \Leftrightarrow f(\overrightarrow{AM}) = \overrightarrow{AM}.$$

En d'autres termes, $M \in \text{Fix } \varphi$ si et seulement si $\overrightarrow{AM} \in \text{Fix } f$, ce qui prouve que $\text{Fix } \varphi$ est le sous-espace affine de \mathcal{E} passant par A et dirigé par $\text{Fix } f$. \square

On termine cette section par un résultat exprimant que, lorsque \mathcal{E} est de dimension finie (ce qui sera toujours le cas dans la pratique), il suffit, pour connaître une application affine, de connaître les images d'un nombre fini de points formant une base affine de \mathcal{E} .

j) PROPOSITION (Détermination d'une application affine). L'espace affine \mathcal{E} étant supposé de dimension finie, une application affine $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ est déterminée (entièrement et de façon unique) par l'image d'un repère de \mathcal{E} , c'est-à-dire par l'image d'une base affine de \mathcal{E} .

Preuve. Soit $\mathcal{R} = (O, \mathcal{B})$ un repère de \mathcal{E} . Notons $\mathcal{B} = (\vec{e}_1, \dots, \vec{e}_n)$ qui est une base de E . Reprenons les observations faites à la section 4. Pour tout $1 \leq i \leq n$, notons A_i l'unique point de \mathcal{E} tel que $\overrightarrow{OA_i} = \vec{e}_i$. Notons \mathcal{X} la famille de $n+1$ points $\mathcal{X} = \{O, A_1, A_2, \dots, A_n\}$. Comme $\mathcal{B} = \{\overrightarrow{OA_1}, \overrightarrow{OA_2}, \dots, \overrightarrow{OA_n}\}$ est une base de E , \mathcal{X} est une base affine de \mathcal{E} .

Ceci étant, soit $\mathcal{Y} = \{B_0, B_1, B_2, \dots, B_n\}$ une famille de $n+1$ points de \mathcal{E} . On peut construire la famille de n vecteurs $\mathcal{C} = \{\overrightarrow{B_0B_1}, \overrightarrow{B_0B_2}, \dots, \overrightarrow{B_0B_n}\}$ de E . D'après un résultat bien connu d'algèbre linéaire, il existe une unique application linéaire $f : E \rightarrow E$ telle que, pour tout

$1 \leq i \leq n$, on ait : $f(\overrightarrow{OA_i}) = \overrightarrow{B_0B_i}$. En appliquant le théorème du début de cette section, on peut considérer l'unique application affine $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ dont l'application linéaire associée est f , et telle que $\varphi(O) = B_0$. Pour tout $1 \leq i \leq n$, on a : $\overrightarrow{B_0\varphi(A_i)} = \overrightarrow{\varphi(O)\varphi(A_i)} = f(\overrightarrow{OA_i}) = \overrightarrow{B_0B_i}$; d'où $\varphi(A_i) = B_i$. \square

A titre d'exemples de conséquences pratiques de cet énoncé, citons :

- (1) Deux applications affines qui coïncident en deux points distincts A et B [respectivement en trois points non alignés A, B, C] coïncident en tout point de la droite (AB) [respectivement du plan (ABC)].
- (2) Une application affine qui fixe 3 points non alignés dans l'espace \mathcal{E} de dimension 2 [respectivement 4 points non coplanaires dans l'espace \mathcal{E} de dimension 3] est égale à $\text{id}_{\mathcal{E}}$.

7.2.2 Applications affines bijectives, groupe affine

\mathcal{E} est toujours un espace affine sur \mathbb{R} , d'espace vectoriel associé E . On commence par deux résultats élémentaires, mais fondamentaux.

a) LEMME (Composée de deux applications affines). Soient $\varphi, \psi : \mathcal{E} \rightarrow \mathcal{E}$ deux applications affines d'applications linéaires associées respectives $f, g : E \rightarrow E$. Alors $\psi \circ \varphi$ est affine d'application linéaire associée $g \circ f$.

Preuve. Pour tous $A, B \in \mathcal{E}$, on a : $\overrightarrow{\psi(\varphi(A))\psi(\varphi(B))} = g(\overrightarrow{\varphi(A)\varphi(B)}) = g(f(\overrightarrow{AB}))$. \square

b) LEMME. (Bijektivité d'une application affine). Soit $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ affine d'application linéaire associée $f : E \rightarrow E$. Alors :

φ injective $\Leftrightarrow f$ injective, φ surjective $\Leftrightarrow f$ surjective, φ bijective $\Leftrightarrow f$ bijective.

Preuve. Supposons f injective. Soient $A, B \in \mathcal{E}$ tels que $\varphi(A) = \varphi(B)$. Alors $\overrightarrow{0} = \overrightarrow{\varphi(A)\varphi(A)} = \overrightarrow{\varphi(A)\varphi(B)} = f(\overrightarrow{AB})$. D'où $\overrightarrow{AB} = \overrightarrow{0}$ c'est-à-dire $A = B$. Supposons réciproquement que φ est injective. Soit $\overrightarrow{u} \in \text{Ker } f$. Soit $A \in \mathcal{E}$ et M l'unique point de \mathcal{E} tel que $\overrightarrow{AM} = \overrightarrow{u}$. Donc $\overrightarrow{0} = f(\overrightarrow{u}) = f(\overrightarrow{AM}) = \overrightarrow{\varphi(A)\varphi(M)}$. D'où $\varphi(A) = \varphi(M)$ et donc $A = M$ et $\overrightarrow{u} = \overrightarrow{AM} = \overrightarrow{0}$. Le reste de la preuve est clair et laissé en exercice. \square

c) COROLLAIRE. Soit $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ affine, d'application linéaire associée $f : E \rightarrow E$. Alors :

- (i) φ est bijective si et seulement si elle transforme une base affine de \mathcal{E} en une base affine de \mathcal{E} .
- (ii) Si φ est bijective, alors sa réciproque φ^{-1} est affine d'application linéaire associée f^{-1} .

Preuve. Le point (i) découle du lemme b) ci-dessus et de la proposition 7.2.1.j. Pour (ii), considérons $M, N \in \mathcal{E}$ quelconques. Par bijectivité de φ , il existe $A, B \in \mathcal{E}$ uniques tels que $\varphi(A) = M$ et $\varphi(B) = N$. Donc $\overrightarrow{MN} = \overrightarrow{\varphi(A)\varphi(B)} = f(\overrightarrow{AB})$ puisque φ est affine. Donc $\overrightarrow{AB} = f^{-1}(\overrightarrow{MN})$ puisque f est bijective d'après le lemme b). On déduit $\overrightarrow{\varphi^{-1}(M)\varphi^{-1}(N)} = f^{-1}(\overrightarrow{MN})$. \square

d) DÉFINITION ET PROPOSITION. Une application affine $\mathcal{E} \rightarrow \mathcal{E}$ qui est bijective est appelée un automorphisme affine de \mathcal{E} . L'ensemble des automorphismes affines de \mathcal{E} est un groupe pour la loi de composition, appelé groupe affine de \mathcal{E} , et noté $\text{GA}(\mathcal{E})$.

Preuve. $\text{GA}(\mathcal{E})$ est un sous-ensemble du groupe des bijections de \mathcal{E} sur \mathcal{E} , non vide (il contient $\text{id}_{\mathcal{E}}$), stable pour la loi \circ d'après le lemme 1, et stable par passage à l'inverse d'après le point (ii) du corollaire. C'est donc un sous-groupe. \square

En résumé, $\boxed{\text{GA}(\mathcal{E}) = \{ \varphi : \mathcal{E} \rightarrow \mathcal{E}; \varphi \text{ affine et bijective} \}}$ est un groupe

e) COMPLÉMENT. D'après le lemme b), la bijectivité de φ équivaut à celle de l'endomorphisme vectoriel associé $f : \varphi \in \text{GA}(\mathcal{E}) \Leftrightarrow f \in \text{GL}(E)$.

On peut donc considérer l'application $\ell : \text{GA}(\mathcal{E}) \rightarrow \text{GL}(E)$ qui, à tout automorphisme affine φ de \mathcal{E} , associe son application linéaire associée qui est un automorphisme d'espace vectoriel de E . D'après le lemme a), c'est un morphisme de groupes (si $\ell(\varphi) = f$ et $\ell(\psi) = g$, alors $\ell(\psi \circ \varphi) = g \circ f$). Le théorème 7.2.1.b montre que ℓ est surjective. Il reste à déterminer son noyau, c'est-à-dire l'ensemble des applications affines $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ dont l'application linéaire associée $f = \ell(\varphi)$ est l'application identité id_E . Or on va précisément voir au début de la section suivante que les applications affines dont l'application linéaire associée est l'identité de E sont exactement les translations de \mathcal{E} . Donc, en anticipant sur le début du paragraphe suivant : $\text{Ker } \ell = \{ \varphi \in \text{GA}(\mathcal{E}); f = \text{id}_E \} = \text{T}(\mathcal{E})$, le groupe des translations de \mathcal{E} .

7.2.3 Homothéties, translations

\mathcal{E} est un espace affine sur \mathbb{R} , d'espace vectoriel associé E .

a) DÉFINITION. Soit \vec{u} un vecteur de E . On appelle *translation* de vecteur \vec{u} l'application $\tau_{\vec{u}} : \mathcal{E} \rightarrow \mathcal{E}$ qui, à tout point M de \mathcal{E} , associe le point M' défini par $\overrightarrow{MM'} = \vec{u}$.

$$\boxed{M' = \tau_{\vec{u}}(M) \Leftrightarrow \overrightarrow{MM'} = \vec{u}}$$

Remarquons que $\text{id}_{\mathcal{E}} = \tau_{\vec{0}}$.

b) PROPOSITION.

- (i) Toute translation de \mathcal{E} est une application affine $\mathcal{E} \rightarrow \mathcal{E}$, et son application linéaire associée est id_E .
- (ii) Réciproquement, toute application affine $\mathcal{E} \rightarrow \mathcal{E}$ dont l'application linéaire associée est id_E est une translation de \mathcal{E} .

Preuve. Soit $\tau = \tau_{\vec{u}}$ où $\vec{u} \in E$. Pour tous $A, B \in \mathcal{E}$, on a $\overrightarrow{\tau(A)\tau(B)} = \overrightarrow{\tau(A)\vec{A} + \vec{A}\vec{B} + \vec{B}\tau(B)} = -\vec{u} + \vec{A}\vec{B} + \vec{u} = \vec{A}\vec{B}$, ce qui prouve le point (i). Pour (ii), considérons une application affine $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ dont l'application linéaire associée est $f = \text{id}_E$. Soit $A \in \mathcal{E}$ arbitrairement choisi; posons $\vec{u} = A\varphi(A)$. Pour tout $M \in \mathcal{E}$, on a alors $\overrightarrow{M\varphi(M)} = \overrightarrow{M\vec{A} + \vec{A}\varphi(A) + \varphi(A)\varphi(M)}$. Or par hypothèse, $\overrightarrow{\varphi(A)\varphi(M)} = f(\overrightarrow{AM}) = \overrightarrow{AM}$. On obtient donc $\overrightarrow{M\varphi(M)} = \overrightarrow{M\vec{A} + \vec{u} + \vec{A}\vec{M}} = \vec{u}$. On conclut que $\varphi = \tau_{\vec{u}}$. \square

D'après le lemme 7.2.2.b, on déduit de (i) que toute translation est une bijection de \mathcal{E} sur \mathcal{E} . Donc l'ensemble $\text{T}(\mathcal{E})$ de toutes les translations de \mathcal{E} est inclus dans $\text{GA}(\mathcal{E})$. De plus, il résulte de (ii) que, comme on l'avait annoncé à la fin de 7.2.2.d, $\text{T}(\mathcal{E})$ n'est autre que le noyau du morphisme $\ell : \text{GA}(\mathcal{E}) \rightarrow \text{GL}(E)$ qui associe à toute application affine son application linéaire. En particulier, puisque $\text{T}(\mathcal{E}) = \text{Ker } \ell$, $\text{T}(\mathcal{E})$ est un sous-groupe de $\text{GA}(\mathcal{E})$. Concrètement, la composée de deux translations, ainsi que la réciproque d'une translation, sont des translations. On peut préciser que, quels que soient $\vec{u}, \vec{v} \in E$, on a :

$$\tau_{\vec{u}} \circ \tau_{\vec{v}} = \tau_{\vec{v}} \circ \tau_{\vec{u}} = \tau_{\vec{u} + \vec{v}} \quad \text{et} \quad \tau_{\vec{u}}^{-1} = \tau_{-\vec{u}}$$

En effet. Prenons $M \in \mathcal{E}$ quelconque, et posons :

$$M' = \tau_{\vec{u}}(M) \text{ et } M'' = \tau_{\vec{v}}(M') = \tau_{\vec{v}}(\tau_{\vec{u}}(M)).$$

On a $\overrightarrow{MM''} = \overrightarrow{MM'} + \overrightarrow{M'M''} = \vec{u} + \vec{v}$. Ceci prouve que $\tau_{\vec{v}} \circ \tau_{\vec{u}} = \tau_{\vec{u} + \vec{v}}$. Puisque $\vec{u} + \vec{v} = \vec{v} + \vec{u}$, il en résulte que $\tau_{\vec{v}} \circ \tau_{\vec{u}} = \tau_{\vec{u}} \circ \tau_{\vec{v}}$.

Par ailleurs, pour tous $\vec{u} \in E$ et $M, M' \in \mathcal{E}$, on a :

$$(M' = \tau_{\vec{u}}(M)) \Leftrightarrow (\overrightarrow{MM'} = \vec{u}) \Leftrightarrow (\overrightarrow{M'M} = -\vec{u}) \Leftrightarrow (M = \tau_{-\vec{u}}(M')) \quad \square$$

Figure

On retiendra que :

$$\mathbf{T}(\mathcal{E}) = \{\tau_{\vec{u}}; \vec{u} \in E\} = \text{Ker } \ell \text{ est un sous-groupe abélien de } \text{GA}(\mathcal{E})$$

c) DÉFINITION. Soit A un point de \mathcal{E} . Soit λ un réel non-nul. On appelle *homothétie affine* de centre A et de rapport λ l'application $\vartheta_{A,\lambda} : \mathcal{E} \rightarrow \mathcal{E}$ qui, à tout $M \in \mathcal{E}$, associe l'unique point $\vartheta_{A,\lambda}(M) = M'$ défini par : $\overrightarrow{AM'} = \lambda \overrightarrow{AM}$.

Dans le cas particulier où $\lambda = 1$, on a $\vartheta_{A,1} = \text{id}_{\mathcal{E}}$.

Dans le cas particulier où $\lambda = -1$, l'homothétie $\vartheta_{A,-1}$ est appelée la symétrie centrale de centre A ; elle associe à tout point M le point M' tel que A est le milieu de (M, M') .

$$M' = \vartheta_{A,\lambda}(M) \Leftrightarrow \overrightarrow{AM'} = \lambda \overrightarrow{AM}$$

Figure

$$\lambda > 0$$

Figure

$$\lambda < 0$$

Figure

$$\lambda = -1$$

d) PROPOSITION.

- (i) Toute homothétie $\vartheta_{A,\lambda}$ est une application affine, d'application linéaire associée λid_E .
- (ii) Toute homothétie est une bijection, donc appartient au groupe affine $\text{GA}(\mathcal{E})$.
- (iii) L'ensemble des points fixes d'une homothétie $\vartheta_{A,\lambda}$ distincte de l'identité (ie. de rapport différent de 1) est réduit au singleton $\{A\}$ formé par le centre.

Preuve. Fixons $A \in \mathcal{E}$ et $\lambda \in \mathbb{R}$, et notons $\vartheta = \vartheta_{A,\lambda}$. Pour tous $M, N \in \mathcal{E}$, on a $\overrightarrow{\vartheta(M)\vartheta(N)} = \overrightarrow{A\vartheta(N)} - \overrightarrow{A\vartheta(M)} = \lambda \overrightarrow{AN} - \lambda \overrightarrow{AM} = \lambda \overrightarrow{MN}$, ce qui montre le point (i). Le point (ii) s'en déduit puisque λid_E est une bijection de E sur E . Enfin, $\vartheta(M) = M$ équivaut à $\lambda \overrightarrow{AM} = \overrightarrow{AM}$, donc $(\lambda - 1)\overrightarrow{AM} = \vec{0}$, d'où $A = M$ dès lors que $\lambda \neq 1$. \square

e) DÉFINITION. On appelle *homothétie-translation* toute application affine $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ dont l'application linéaire associée f est de la forme $f = \lambda \text{id}_E$, avec $\lambda \in \mathbb{R}^*$. Le réel non-nul λ s'appelle le *rapport* de l'homothétie-translation.

f) THÉORÈME.

- (i) Une homothétie-translation de rapport 1 est une translation. Une homothétie-translation de rapport $\lambda \neq 1$ est une homothétie de rapport λ , dont le centre est uniquement déterminé.
- (ii) L'ensemble $H(\mathcal{E})$ des homothéties-translations est un sous-groupe de $GA(\mathcal{E})$, égal à la réunion du sous-ensemble des homothéties et du sous-groupe $T(\mathcal{E})$ des translations de \mathcal{E} .

Preuve. Soit $\varphi \in GA(\mathcal{E})$ d'application linéaire associée $f = \lambda \text{id}_E$ avec $\lambda \in \mathbb{R}^*$. On a déjà montré au début de cette section que, si $\lambda = 1$, alors φ est une translation.

Supposons donc maintenant $\lambda \neq 1$. Soit $B \in \mathcal{E}$ fixé, et $B' = \varphi(B)$. Pour tout $M \in \mathcal{E}$, on a $\overrightarrow{B'\varphi(M)} = \overrightarrow{\varphi(B)\varphi(M)} = \overrightarrow{f(\overrightarrow{BM})} = \lambda \overrightarrow{BM}$. En particulier $\varphi(M) = M$ si et seulement si $\overrightarrow{B'M} = \lambda \overrightarrow{BM}$, c'est-à-dire avec la relation de Chasles si et seulement si $(1 - \lambda)\overrightarrow{BM} = \overrightarrow{BB'}$. Ceci montre que φ admet un unique point fixe A , qui est défini par $\overrightarrow{BA} = (1 - \lambda)^{-1}\overrightarrow{BB'}$. Pour tout $M \in \mathcal{E}$, on a alors : $\overrightarrow{A\varphi(M)} = \overrightarrow{\varphi(A)\varphi(M)} = \overrightarrow{f(\overrightarrow{AM})} = \lambda \overrightarrow{AM}$, ce qui prouve que φ est l'homothétie de centre A et de rapport λ . Le point (i) est démontré.

Il est clair que $\{\lambda \text{id}_E ; \lambda \in \mathbb{R}^*\}$ est un sous-groupe de $GL(E)$. Le fait que $H(\mathcal{E})$ soit un sous-groupe de $GA(\mathcal{E})$ résulte alors du lemme 7.2.2.a. Le point (i) se traduit par le fait que $H(\mathcal{E})$ est la réunion du sous-groupe $T(\mathcal{E})$ des translations et du sous-ensemble des homothéties. \square

Le point (i) du théorème se traduit sur le plan pratique par le fait que :

- une application affine φ vérifiant $\overrightarrow{\varphi(M)\varphi(N)} = \overrightarrow{MN}$ pour tous $M, N \in \mathcal{E}$ est une translation ; on trouve son vecteur en prenant un point quelconque A et en considérant le vecteur $\overrightarrow{A\varphi(A)}$,
- une application affine φ pour laquelle il existe $\lambda \in \mathbb{R}^*$, $\lambda \neq 1$, vérifiant $\overrightarrow{\varphi(M)\varphi(N)} = \lambda \overrightarrow{MN}$ pour tous $M, N \in \mathcal{E}$ est une homothétie de rapport λ ; on trouve son centre en déterminant son unique point fixe.

Quant au point (ii) du théorème, il convient de préciser comment se composent entre eux les différents éléments de $H(\mathcal{E})$. C'est ce qu'explicitent les assertions suivantes, dont les preuves (et les dessins qui les accompagnent !) sont laissés au lecteur à titre d'exercice. On prendra garde en particulier au fait que la composée de deux homothéties n'est pas forcément une homothétie (les homothéties ne forment pas un sous-groupe de $H(\mathcal{E})$).

★ composée de deux translations : $\tau_{\vec{u}} \circ \tau_{\vec{v}} = \tau_{\vec{u} + \vec{v}}$

★ composée d'une translation et d'une homothétie :

$$\text{Si } \lambda = 1, \tau_{\vec{u}} \circ \vartheta_{A,\lambda} = \tau_{\vec{u}}$$

$$\text{Si } \lambda \neq 1, \tau_{\vec{u}} \circ \vartheta_{A,\lambda} = \vartheta_{B,\lambda} \quad \text{où } B \text{ est le point défini par } \overrightarrow{AB} = (1 - \lambda)^{-1}\vec{u}$$

★ composée de deux homothéties de même centre : $\vartheta_{A,\lambda} \circ \vartheta_{A,\lambda'} = \vartheta_{A,\lambda\lambda'}$

★ composée de deux homothéties non nécessairement de même centre :

$$\text{Si } \lambda\lambda' = 1, \vartheta_{A',\lambda'} \circ \vartheta_{A,\lambda} = \tau_{(1-\lambda')\overrightarrow{AA'}}$$

$$\text{Si } \lambda\lambda' \neq 1, \vartheta_{A',\lambda'} \circ \vartheta_{A,\lambda} = \vartheta_{B,\lambda\lambda'} \quad \text{où } B \text{ est le point défini par } \overrightarrow{AB} = \frac{(1-\lambda')}{(1-\lambda\lambda')} \overrightarrow{AA'}$$

7.2.4 Projections, symétries

► *Un rappel d'algèbre linéaire.* Soit E un espace vectoriel. Soient F et H deux sous-espaces vectoriels supplémentaires dans E , ie. tels que $E = F \oplus H$. Tout vecteur $\vec{u} \in E$ se décompose de façon unique en une somme $\vec{u} = \vec{v} + \vec{w}$ avec $\vec{v} \in F$ et $\vec{w} \in H$. L'application $p : E \rightarrow E$ qui, à tout $\vec{u} \in E$ ainsi décomposé, associe sa composante \vec{v} sur F s'appelle la *projection* (vectorielle) de E sur F parallèlement à H . L'application $s : E \rightarrow E$ qui, à tout $\vec{u} \in E$ ainsi décomposé, associe le vecteur $\vec{v} - \vec{w}$ s'appelle la *symétrie* (vectorielle) par rapport à F parallèlement à H .

$$\begin{aligned} E &= F \oplus H, \\ \forall \vec{u} \in E, \exists! \vec{v} \in F, \exists! \vec{w} \in H, \vec{u} &= \vec{v} + \vec{w} \\ p(\vec{u}) &= \vec{v}, \quad s(\vec{u}) = \vec{v} - \vec{w} \end{aligned}$$

Figure

Il est clair que p et s sont des applications linéaires.

$$\begin{aligned} \text{Ker } p &= H, & \text{Im } p &= F, & \text{Fix } p &= F, & p \circ p &= p, \\ \text{Ker } s &= \{\vec{0}\}, & \text{Im } s &= E, & \text{Fix } s &= F, & s \circ s &= \text{id}_E. \end{aligned}$$

On considère maintenant \mathcal{E} un espace affine sur \mathbb{R} , d'espace vectoriel associé E .

a) LEMME PRÉLIMINAIRE. Soient \mathcal{F} et \mathcal{H} deux sous-espaces affines de \mathcal{E} dont les sous-espaces vectoriels directeurs F et H vérifient $E = F + H$. Alors $\mathcal{F} \cap \mathcal{H} \neq \emptyset$.

Preuve. Soient $A \in \mathcal{F}$ et $B \in \mathcal{H}$. Comme $E = F + H$, il existe $\vec{v} \in F$ et $\vec{w} \in H$ tels que $\overrightarrow{AB} = \vec{v} + \vec{w}$. On a $A \in \mathcal{F}$ et $\vec{v} \in F$, donc il existe $C \in F$ tel que $\overrightarrow{AC} = \vec{v}$. On réécrit alors $\overrightarrow{AB} = \overrightarrow{AC} + \vec{w}$ sous la forme $\vec{w} = \overrightarrow{AB} - \overrightarrow{AC} = \overrightarrow{CB}$. Ainsi, $\overrightarrow{CB} \in H$ avec $B \in \mathcal{H}$, d'où $C \in \mathcal{H}$. On conclut que $C \in \mathcal{F} \cap \mathcal{H}$. \square

b) THÉORÈME ET DÉFINITION. Soient \mathcal{F} et \mathcal{H} deux sous-espaces affines de \mathcal{E} . On suppose que leurs sous-espaces vectoriels directeurs F et H sont supplémentaires dans E . Alors

- (i) Pour tout point M de \mathcal{E} , il existe un unique point $M' \in \mathcal{F}$ tels que $\overrightarrow{MM'}$ appartienne à H . Ce point M' est le point d'intersection de \mathcal{F} avec le sous-espace affine passant par M et parallèle à \mathcal{H} . Le point M' est appelé le projeté de M sur \mathcal{F} parallèlement à \mathcal{H} .
- (ii) L'application $\pi : \mathcal{E} \rightarrow \mathcal{E}$ qui, à tout point M associe son projeté M' défini ci-dessus s'appelle la *projection affine* sur \mathcal{F} parallèlement à \mathcal{H} ; c'est une application affine, d'application linéaire associée la projection vectorielle $p : E \rightarrow E$ sur F parallèlement à H .
- (iii) On a $\pi \circ \pi = \pi$.
- (iv) L'ensemble des points fixes de π est $\text{Fix } \pi = \mathcal{F}$.

Preuve. Soit M un point de \mathcal{E} . Notons \mathcal{H}' le sous-espace affine de \mathcal{E} passant par M et parallèle à \mathcal{H} , c'est-à-dire dirigé par H . Comme $E = F \oplus H$ par hypothèse, on applique le lemme préliminaire pour déduire que $\mathcal{F} \cap \mathcal{H}'$ n'est pas vide. C'est donc un sous-espace affine de sous-espace vectoriel directeur $F \cap H$. Or $F \cap H = \{\vec{0}\}$, donc $\mathcal{F} \cap \mathcal{H}'$ est un singleton; notons $\mathcal{F} \cap \mathcal{H}' = \{M'\}$. On a $M \in \mathcal{H}'$ et $M' \in \mathcal{H}'$, donc $\overrightarrow{MM'} \in H$.

Réciproquement, si N est un point de \mathcal{F} vérifiant $\overrightarrow{MN} \in H$, on a $N \in \mathcal{H}'$ (puisque \mathcal{H}' passe par M et est dirigé par H), d'où $N \in \mathcal{F} \cap \mathcal{H}'$, et donc $N = M'$. Ceci prouve le point (i).

Les points (iii) et (iv) en découlent immédiatement. Pour (ii), considérons $A \in \mathcal{F}$, vérifiant donc $\pi(A) = A$. Soient $M \in \mathcal{E}$ quelconque, et $M' = \pi(M)$. D'une part $M' \in \mathcal{F}$ donc $\overrightarrow{AM'} \in F$. D'autre part $\overrightarrow{MM'} \in H$. Ainsi $\overrightarrow{AM} = \overrightarrow{AM'} + \overrightarrow{M'M}$ avec $\overrightarrow{AM'} \in F$ et $\overrightarrow{M'M} \in H$, ce qui prouve que $\overrightarrow{AM'} = p(\overrightarrow{AM})$, ie. $A\pi(M) = p(\overrightarrow{AM})$.

Dès lors, pour tous $M, N \in \mathcal{E}$, on a : $\overrightarrow{\pi(M)\pi(N)} = \overrightarrow{A\pi(N)} - \overrightarrow{A\pi(M)} = p(\overrightarrow{AN}) - p(\overrightarrow{AM}) = p(\overrightarrow{AN} - \overrightarrow{AM}) = p(\overrightarrow{MN})$, ce qui montre (ii) et achève la preuve. \square

Figure

c) THÉORÈME ET DÉFINITION. Soient \mathcal{F} et \mathcal{H} deux sous-espaces affines de \mathcal{E} . On suppose que leurs sous-espaces vectoriels directeurs F et H sont supplémentaires dans E . Alors

(i) Pour tout point M de \mathcal{E} , il existe un unique point $M'' \in \mathcal{E}$ tels que le projeté $M' = \pi(M)$ défini précédemment soit le milieu de (M, M'') . Ce point M'' est l'unique point vérifiant $\overrightarrow{M\pi(M)} = \overrightarrow{\pi(M)M''}$. Le point M' est appelé le symétrique de M par rapport à \mathcal{F} parallèlement à \mathcal{H} .

(ii) L'application $\sigma : \mathcal{E} \rightarrow \mathcal{E}$ qui, à tout point M associe son symétrique M'' défini ci-dessus s'appelle la symétrie affine par rapport à \mathcal{F} parallèlement à \mathcal{H} ; c'est une application affine dont l'application linéaire associée est la symétrie vectorielle $s : E \rightarrow E$ par rapport à F parallèlement à H .

(iii) On a : $\sigma \circ \sigma = \text{id}_{\mathcal{E}}$, d'où il résulte que σ est bijective avec $\sigma^{-1} = \sigma$.

(iv) L'ensemble des points fixes de σ est $\text{Fix } \sigma = \mathcal{F}$.

Preuve. Tout à fait analogue à celle du théorème précédent; laissée en exercice. \square

Figure

d) QUELQUES REMARQUES.

(1) Dans le cas particulier où \mathcal{F} est un singleton $\{A\}$, alors $F = \{\vec{0}\}$ et $H = E$, donc π est l'application constante qui envoie tout point M de \mathcal{E} sur A , et σ est la symétrie centrale de centre A qui envoie tout point M de \mathcal{E} sur le point M'' tel que A soit le milieu de (M, M'') .

- (2) Dans le cas particulier où \mathcal{F} est l'espace \mathcal{E} tout entier, alors $F = E$ et $H = \{\vec{0}\}$, donc π et σ sont égales à l'application identité de \mathcal{E} (qui envoie tout point de M sur lui-même).
- (3) Hormis ces cas extrêmes, les "vraies" situations que l'on rencontrera feront intervenir en dimension 2 des symétries par rapport à une droite \mathcal{D} parallèlement à une autre droite \mathcal{D}' non parallèle à \mathcal{D} , et en dimension 3 des symétries par rapport à un plan \mathcal{P} parallèlement à une droite \mathcal{D} ou par rapport à une droite \mathcal{D} parallèlement à un plan \mathcal{P} (avec la droite vectorielle dirigeant \mathcal{D} non incluse dans le plan vectoriel dirigeant \mathcal{P}). De même bien sûr pour les projections.
- (4) On peut vérifier aisément (la preuve est laissée en exercice) que :
- une application affine $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ est une projection si et seulement si $\varphi \circ \varphi = \varphi$;
 - une application affine $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ est une symétrie si et seulement si $\varphi \circ \varphi = \text{id}_{\mathcal{E}}$.

7.2.5 Le cas des isométries affines.

On se place maintenant dans le contexte d'un espace affine \mathcal{E} qui est euclidien c'est-à-dire (comme on l'a vu en 7.1.8) dont l'espace vectoriel directeur E est euclidien. On sait que, dans E , on a une notion d'isométrie vectorielle (voir chapitre 2). On peut alors définir une notion naturelle d'isométrie affine dans \mathcal{E} comme une application affine $\mathcal{E} \rightarrow \mathcal{E}$ dont l'application linéaire associée est une isométrie vectorielle $E \rightarrow E$. La description détaillée que l'on a donnée en 2.4 des isométries vectorielles en dimension 2 et 3 permet de déduire des résultats comparables pour les isométries affines. On en donne ci-dessous un bref résumé, sans détailler les preuves. Ces dernières sont laissées au lecteur à titre d'exercice et se déduisent des résultats généraux sur les applications affines vus aux paragraphes précédents, et de l'étude des isométries vectorielles en dimension 2 et 3 menée en 2.4.

On fixe donc dans tout ce qui suit un espace affine euclidien \mathcal{E} , dont l'espace vectoriel euclidien directeur est noté E . On définit à partir de la norme euclidienne sur E la distance dans \mathcal{E} par :

$$\text{pour tous } A, B \in \mathcal{E}, d(A, B) = \|\overrightarrow{AB}\|; \text{ on note encore } d(A, B) = AB.$$

a) DÉFINITIONS. Une application affine $\mathcal{E} \rightarrow \mathcal{E}$ est appelée une *isométrie affine* de \mathcal{E} lorsqu'elle conserve la distance :

$$\varphi \text{ isométrie} \Leftrightarrow d(\varphi(A), \varphi(B)) = d(A, B) \text{ pour tous } A, B \in \mathcal{E}.$$

Il est facile de vérifier que φ est une isométrie affine de \mathcal{E} si et seulement si son application linéaire associée f est une isométrie vectorielle de E . Toute isométrie affine est donc une bijection de \mathcal{E} sur \mathcal{E} , et les isométries affines forment un sous-groupe de $\text{GA}(\mathcal{E})$. On note $\text{Is}(\mathcal{E})$ ce groupe.

$$\text{Is}(\mathcal{E}) = \{\varphi \in \text{GA}(\mathcal{E}); \varphi \text{ conserve la distance}\} = \{\varphi \in \text{GA}(\mathcal{E}); f \in \text{O}(E)\}.$$

On appelle *isométrie directe* de \mathcal{E} , ou *déplacement* de \mathcal{E} , toute isométrie $\varphi \in \text{Is}(\mathcal{E})$ dont l'application linéaire associée f est une isométrie vectorielle directe de E , ie telle que $f \in \text{O}^+(E)$. On note $\text{Is}^+(\mathcal{E})$ l'ensemble des déplacements de \mathcal{E} ; il est clair que c'est un sous-groupe de $\text{Is}(\mathcal{E})$.

De même, une isométrie $\varphi \in \text{Is}(\mathcal{E})$ telle que $f \in \text{O}^-(E)$ s'appelle une *isométrie indirecte* ou un *antidépacement* de \mathcal{E} . Bien sûr, l'ensemble $\text{Is}^-(\mathcal{E})$ des antidépacements de \mathcal{E} n'est pas un sous-groupe de $\text{Is}(\mathcal{E})$.

b) PREMIERS EXEMPLES.

► *Toute translation de \mathcal{E} est une isométrie de \mathcal{E} .*

C'est évident puisque toute translation est affine, d'application linéaire associée id_E (voir 7.2.3.b) et est donc une isométrie directe de \mathcal{E} . En particulier, le groupe $T(\mathcal{E})$ des translations de \mathcal{E} est un sous-groupe abélien de $\text{Is}^+(\mathcal{E})$.

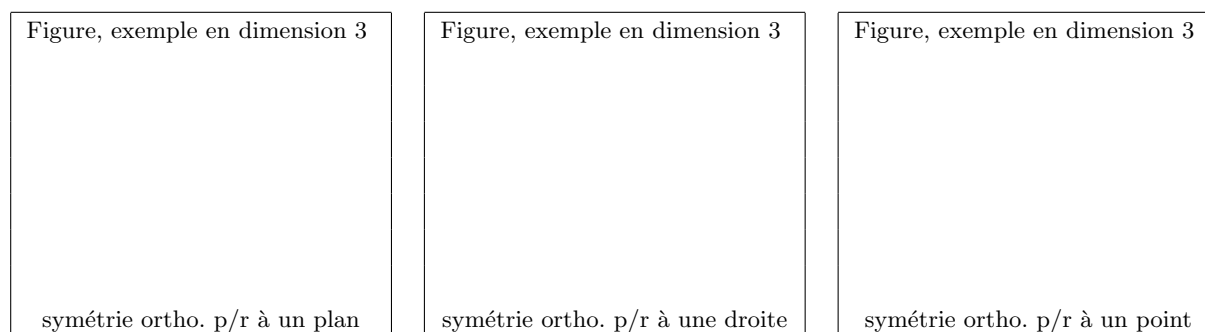
► *Toute symétrie affine orthogonale est une isométrie de \mathcal{E} .*

C'est évident d'après le point (ii) de 7.2.4.c, et la proposition 2.2.3.c

Le fait que la symétrie affine orthogonale $\sigma_{\mathcal{F}}$ par rapport à un sous-espace affine \mathcal{F} soit directe ou indirecte dépend des valeurs de n et $\dim \mathcal{F}$ [voir 2.2.3.c].

Dans le cas où \mathcal{F} est un hyperplan de \mathcal{E} , on dit que $\sigma_{\mathcal{F}}$ est une symétrie hyperplane affine ; les symétries hyperplanes affines sont des antidéplacements.

Dans le cas où \mathcal{F} est un singleton $\{A\}$, on dit que σ_A est la symétrie centrale de centre A ; à noter que σ_A est simplement l'application qui, à tout point $M \in \mathcal{E}$, associe le point $M'' \in \mathcal{E}$ tel que A est le milieu de $[MM'']$.



c) LE GROUPE DES ISOMÉTRIES AFFINES EN DIMENSION 2.

On suppose ici que l'espace affine euclidien \mathcal{E} est de dimension 2 et orienté. En particulier, les symétries orthogonales par rapport aux droites sont des antidéplacements.

(i) - Pour tout $A \in \mathcal{E}$ et $\theta \in \mathbb{R}$, on appelle *rotation affine* de centre A et d'angle θ l'application affine $\rho_{A,\theta} : \mathcal{E} \rightarrow \mathcal{E}$ telle que son application linéaire associée soit la rotation vectorielle r_θ de E (voir 2.4.1 et telle que $\rho_{A,\theta}(A) = A$).

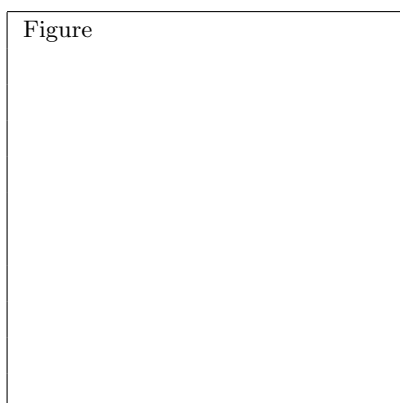
Elle existe et est unique d'après la remarque finale de 7.2.1.a.

$$M' = \rho_{A,\theta}(M) \Leftrightarrow \overrightarrow{AM'} = r_\theta(\overrightarrow{AM}).$$

Il est clair que :

$$\rho_{A,\theta} \in \text{Is}^+(\mathcal{E}) \text{ pour tous } A \in \mathcal{E}, \theta \in \mathbb{R}.$$

Si $\theta \equiv 0$ modulo 2π , alors $\rho_{A,\theta} = \text{id}_{\mathcal{E}}$,
sinon, A est l'unique point fixe de $\rho_{A,\theta}$



THÉORÈME. *Le groupe $\text{Is}^+(\mathcal{E})$ des déplacements du plan affine \mathcal{E} euclidien est formé par les translations et les rotations.*

(ii) - Pour toute droite affine \mathcal{D} de \mathcal{E} et tout vecteur \vec{u} de E appartenant à la droite vectorielle Δ directrice de \mathcal{D} , on appelle *symétrie glissée* de vecteur \vec{u} par rapport à \mathcal{D} la composée φ de la translation $\tau_{\vec{u}}$ de vecteur \vec{u} et de la symétrie orthogonale $\sigma_{\mathcal{D}}$ par rapport à \mathcal{D} . On a :

$$\varphi = \tau_{\vec{u}} \circ \sigma_{\mathcal{D}} = \sigma_{\mathcal{D}} \circ \tau_{\vec{u}}.$$

Il est clair que φ est affine d'application linéaire associée $f = \text{id}_E \circ s_{\Delta}$, donc $f = s_{\Delta} \in \text{O}^-(E)$, d'où $\varphi \in \text{Is}^-(\mathcal{E})$.

Si $\vec{u} = \vec{0}$, alors φ est la symétrie $\sigma_{\mathcal{D}}$, donc l'ensemble des points fixes de φ est \mathcal{D} . Sinon, φ n'a pas de point fixe.

Figure

THÉORÈME. L'ensemble $\text{Is}^-(\mathcal{E})$ des antidéplacements du plan affine euclidien \mathcal{E} est formé par les symétries glissées.

(iii) - On termine en rappelant le résultat de composition suivant.

PROPOSITION. Soient \mathcal{D} et \mathcal{D}' deux droites du plan affine euclidien \mathcal{E} .

- (i) Si \mathcal{D} et \mathcal{D}' sont parallèles, alors $\sigma_{\mathcal{D}'} \circ \sigma_{\mathcal{D}} = \tau_{2\vec{u}}$, où \vec{u} est le vecteur orthogonal à la direction commune de \mathcal{D} et \mathcal{D}' tel que $\tau_{\vec{u}}(\mathcal{D}) = \mathcal{D}'$.
- (ii) Si \mathcal{D} et \mathcal{D}' ne sont pas parallèles, alors $\sigma_{\mathcal{D}'} \circ \sigma_{\mathcal{D}} = \rho_{A,2\theta}$, où A est le point d'intersection de \mathcal{D} et \mathcal{D}' , et $\theta = \widehat{(\mathcal{D}, \mathcal{D}')} \text{ modulo } \pi$.
- (iii) En particulier, si \mathcal{D} et \mathcal{D}' sont perpendiculaires en A , alors $\sigma_{\mathcal{D}'} \circ \sigma_{\mathcal{D}}$ est la symétrie centrale de centre A .

Figure

Figure

Figure

COROLLAIRE. Tout déplacement du plan affine euclidien \mathcal{E} est produit de deux symétries orthogonales par rapport à des droites. Tout antidéplacement du plan affine euclidien \mathcal{E} est produit de une ou trois symétries orthogonales par rapport à des droites.

On exprime cette propriété en disant que : le groupe $\text{Is}^{\mathcal{E}}$ est engendré par les symétries orthogonales par rapport aux droites.

d) LE GROUPE DES ISOMÉTRIES AFFINES EN DIMENSION 3. On suppose ici que l'espace affine euclidien \mathcal{E} est de dimension 3 et orienté. En particulier, les symétries orthogonales par rapport aux plans sont des antidéplacements.

(i) - Pour toute droite affine orientée \mathcal{D} de \mathcal{E} et $\theta \in \mathbb{R}$, on appelle *rotation affine* d'axe \mathcal{D} et d'angle θ l'application affine $\rho_{\mathcal{D},\theta} : \mathcal{E} \rightarrow \mathcal{E}$ telle que son application linéaire associée soit la rotation vectorielle $r_{\Delta,\theta}$ de E d'axe la droite vectorielle orientée Δ de E directrice de \mathcal{D} et d'angle θ (voir 2.4.3), et telle que $\rho_{\mathcal{D},\theta}(M) = M$ pour tout point M de la droite \mathcal{D} .

Elle existe et est unique d'après remarque finale de 7.2.1.a. Il est clair que :

$$\rho_{\mathcal{D},\theta} \in \text{Is}^+(\mathcal{E}) \text{ pour tous } \theta \in \mathbb{R}, \mathcal{D} \text{ droite orientée de } \mathcal{E}.$$

Si $\theta \equiv 0$ modulo 2π , alors $\rho_{\mathcal{D},\theta} = \text{id}_{\mathcal{E}}$,
sinon, les points fixes de $\rho_{\mathcal{D},\theta}$ sont les points de \mathcal{D} .

Figure

(ii) - Pour toute droite affine orientée \mathcal{D} de \mathcal{E} et $\theta \in \mathbb{R}$, on appelle *vissage* de \mathcal{E} tout produit d'une rotation $\rho = \rho_{\mathcal{D},\theta}$ au sens précédent par une translation $\tau = \tau_{\vec{u}}$ telle que le vecteur \vec{u} de translation appartienne à la direction Δ de \mathcal{D} . Parce que $\vec{u} \in \Delta$, on a $\tau \circ \rho = \rho \circ \tau$, et τ et ρ sont uniques. Il est clair qu'un vissage est un déplacement de \mathcal{E} .

Une rotation est un vissage dont le vecteur est nul. Une translation est un vissage dont l'angle est nul (modulo 2π). Un vissage qui n'est pas une rotation n'admet aucun point fixe.

Figure

THÉORÈME. Le groupe $\text{Is}^+(\mathcal{E})$ des déplacements de l'espace affine euclidien \mathcal{E} de dimension 3 est formé par les vissages.

THÉORÈME. L'ensemble $\text{Is}^-(\mathcal{E})$ des antidéplacements de l'espace affine euclidien \mathcal{E} de dimension 3 est formé par :

1. les symétries orthogonales par rapport à un plan,
2. les composées d'une symétrie orthogonale par rapport à un plan avec une rotation dont l'axe est une droite perpendiculaire au plan de la symétrie,
3. les composées d'une symétrie orthogonale par rapport à un plan avec une translation dont le vecteur appartient au plan vectoriel directeur du plan de la symétrie (symétrie glissée).

COROLLAIRE. Toute isométrie de l'espace affine euclidien \mathcal{E} de dimension 3 est un produit d'un nombre fini de symétries orthogonales par rapport à des plans.

On exprime cette propriété en disant que : le groupe $\text{Is } \mathcal{E}$ est engendré par les symétries orthogonales par rapport aux plans.

7.2.6 Exercices

EXERCICE 1 (*translations, homothéties*). Soit \mathcal{E} un espace affine sur \mathbb{R} .

- a) Soient A, B deux points distincts de \mathcal{E} sur \mathbb{R} , et a, b, c des réels fixés tels que $a + b + c \neq 0$. Pour tout $M \in \mathcal{E}$, on note $\varphi(M)$ le barycentre de $\{(A, a), (B, b), (M, c)\}$. Montrer que l'application $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ ainsi définie est soit une homothétie (déterminer son centre et son rapport) soit une translation (déterminer son vecteur). Faire un dessin illustrant le résultat précédent lorsque \mathcal{E} est de dimension 2 pour $(a, b, c) = (1, 2, 3)$ puis pour $(a, b, c) = (1, -1, 2)$.
- b) Soient O, A_1, A_2, \dots, A_n des points distincts de \mathcal{E} sur \mathbb{R} , et $\alpha_1, \alpha_2, \dots, \alpha_n$ des réels fixés tels que $\sigma = \sum_{i=1}^n \alpha_i \neq 0$. Pour tout $M \in \mathcal{E}$, on note $\varphi(M)$ le point M' défini par $\overrightarrow{OM'} = \sum_{i=1}^n \alpha_i \overrightarrow{A_i M}$. Montrer que l'application $\varphi : \mathcal{E} \rightarrow \mathcal{E}$ ainsi définie est soit une homothétie (déterminer son centre et son rapport) soit une translation (déterminer son vecteur).

EXERCICE 2 (*composée d'homothéties et théorème de Ménelaüs*). Soit A, B, C trois points non alignés d'un plan affine. Soient $A' \in (BC)$, $B' \in (AC)$ et $C' \in (AB)$ tels que soient définis les réels :

$$\alpha = \frac{A'B}{A'C} \quad \beta = \frac{B'C}{B'A} \quad \text{et} \quad \gamma = \frac{C'A}{C'B}.$$

On note η_1, η_2, η_3 les homothéties de centres A', B', C' et de rapports α, β, γ respectivement.

- a) On pose $\varphi = \eta_1 \circ \eta_2 \circ \eta_3$. Vérifier que $\varphi(A') \in (A'B)$. Montrer que, si A', B', C' sont alignés sur une même droite \mathcal{D} , alors $\varphi(\mathcal{D}) = \mathcal{D}$. Dédire qu'alors $\varphi(A') = A'$; calculer $\alpha\beta\gamma$.
- b) Supposons que $\alpha\beta\gamma = 1$. Montrer que $\eta_2 \circ \eta_3$ est une homothétie dont le centre appartient à $(B'C')$. En déduire que $A' \in (B'C')$.
- c) Conclure que A', B', C' sont alignés si et seulement si $\alpha\beta\gamma = 1$.

EXERCICE 3 (*projections affines, homothéties, théorème de Thalès, théorème de Pappus*).

a) Soient \mathcal{D} et \mathcal{D}' deux droites d'un plan affine \mathcal{E} , sécantes en un point C . Soient A, B deux points de \mathcal{D} , et A', B' deux points de \mathcal{D}' . Montrer qu'il existe deux homothéties ϑ et ϑ' de même centre C telles que $\vartheta(A) = B$ et $\vartheta'(A') = B'$; quels sont leurs rapports? En introduisant la projection affine π sur \mathcal{D}' parallèlement à la droite (AA') , démontrer que les trois assertions suivantes sont équivalentes :

$$(i) \quad (AA') \parallel (BB') \quad (ii) \quad \frac{AC}{BC} = \frac{A'C}{B'C} \quad (iii) \quad \vartheta = \vartheta'$$

b) Soient \mathcal{D} et \mathcal{D}' deux droites distinctes du plan affine \mathcal{E} . Soient A, B, C trois points de \mathcal{D} et A', B', C' trois points de \mathcal{D}' . Montrer que si $(AB') \parallel (BA')$ et $(BC') \parallel (CB')$, alors $(AC') \parallel (CA')$.

EXERCICE 4 (*Homothéties, translations, théorème de Desargues*). Soient ABC et $A'B'C'$ deux triangles sans sommet commun dont les côtés sont deux à deux strictement parallèles [avec $(AB) \parallel (A'B')$, $(BC) \parallel (B'C')$ et $(CA) \parallel (C'A')$]. Montrer que les trois droites (AA') , (BB') et (CC') sont concourantes ou parallèles.

Indication : Si (AA') et (BB') se coupent en un point O , montrer qu'il existe une homothétie ϑ de centre O telle que $\vartheta(A) = A'$ et $\vartheta(B) = B'$. Vérifier qu'alors $\vartheta(C) = C'$ et conclure que (AA') , (BB') et (CC') sont concourantes. Si $(AA') \parallel (BB')$, reprendre un raisonnement de même nature avec des translations au lieu d'homothéties.

EXERCICE 5 (*applications affines et alignement*). Soit \mathcal{E} un plan affine sur \mathbb{R} (supposé de plus euclidien pour la dernière question).

a) Montrer qu'une application affine conserve l'alignement. Montrer que, \mathcal{E} étant rapporté à un repère cartésien, l'application $M(x, y) \mapsto M'(x^3, 0)$ conserve l'alignement mais n'est pas affine.

b) Soient A et B deux points distincts fixés de \mathcal{E} . Montrer qu'il existe une application affine de \mathcal{E} dans \mathcal{E} qui envoie tout point M non aligné avec A et B sur le centre de gravité du triangle (ABM) .

c) Soient A et B deux points distincts fixés de \mathcal{E} . Montrer qu'il n'existe pas d'application affine de \mathcal{E} dans \mathcal{E} qui envoie tout point M non aligné avec A et B sur l'orthocentre du triangle (ABM) (construire un contre-exemple).

EXERCICE 6 (*projections affines*). Soit \mathcal{E} un espace affine de dimension 3 sur \mathbb{R} . Soit E son espace vectoriel directeur. Soit φ une application affine de \mathcal{E} dans \mathcal{E} telle que φ n'est pas constante mais $\varphi^2 = \varphi \circ \varphi$ est constante sur \mathcal{E} . On note A le point tel que $\varphi^2(M) = A$ pour tout $M \in \mathcal{E}$.

a) Montrer que l'application linéaire f associée à φ vérifie $f \neq O$ et $f^2 = O$ dans $\text{End } E$. Montrer que $\text{Im } f \subset \text{Ker } f$ avec $\dim \text{Im } f = 1$ et $\dim \text{Ker } f = 2$.

b) Déterminer l'ensemble des points de \mathcal{E} invariants de φ . Montrer qu'il existe au moins un point B dont l'image $C = \varphi(B)$ est différente de A . Montrer que les trois points A, B, C sont non-alignés. Montrer que l'image par φ de l'espace \mathcal{E} est égale à la droite (AC) .

c) Soit \mathcal{P} l'ensemble des points $M \in \mathcal{E}$ tels que $\varphi(M) = A$. Montrer que \mathcal{P} est un sous-espace affine de \mathcal{E} et déterminer son sous-espace vectoriel directeur. En déduire que \mathcal{P} est un plan affine contenant la droite (AC) et ne contenant pas le point B .

d) Soit π la projection affine sur la droite (AB) parallèlement au plan \mathcal{P} , et soit π' la projection affine sur le plan \mathcal{P} parallèlement à la droite (BC) . Démontrer que $\varphi = \pi' \circ \pi$.

EXERCICE 7 (*affinités*). Soit \mathcal{E} un espace affine de dimension 3 sur \mathbb{R} , d'espace vectoriel directeur E . On fixe une base \mathcal{B} de E , un point O de \mathcal{E} , et l'on considère le repère (O, \mathcal{B}) de \mathcal{E} . Toutes les coordonnées et équations de sous-espaces sont relatives à ce repère. On note φ l'application de \mathcal{E} dans \mathcal{E} qui, à tout point $M(x, y, z)$ associe $M' = \varphi(M)$ dont les coordonnées (x', y', z') sont :

$$\begin{cases} x' &= 3x + 4y + 2z - 4 \\ y' &= -2x - 3y - 2z + 4. \\ z' &= 4x + 8y + 5z - 8 \end{cases}$$

a) Montrer que l'ensemble des points de \mathcal{E} invariants par φ est un plan affine \mathcal{P} . Montrer qu'il existe une droite vectorielle Δ de E que l'on déterminera telle que $\overrightarrow{M\varphi(M)} \in \Delta$ pour tout $M \in \mathcal{E}$.

b) On note π la projection affine sur \mathcal{P} parallèlement à Δ . Montrer que $\overrightarrow{M\pi(M)} = -\frac{1}{2}\overrightarrow{M\varphi(M)}$.

c) En déduire qu'il existe un unique réel λ , que l'on déterminera, tel que l'on ait pour tout point $M \in \mathcal{E}$ l'égalité : $\overrightarrow{\pi(M)\varphi(M)} = \lambda \overrightarrow{\pi(M)M}$. Qu'en conclure pour φ ?

EXERCICE 8 (*cercle des neuf points*). Soient A, B, C trois points non alignés d'un plan affine euclidien \mathcal{E} . On note A_1, B_1, C_1 les projetés orthogonaux respectifs de A, B, C sur les droites $(BC), (CA), (AB)$. On note A', B', C' les milieux respectifs des $(B, C), (A, C), (A, B)$. On note G l'isobarycentre de (ABC) , O le centre du cercle circonscrit à (ABC) et Ω le centre du cercle circonscrit à $(A'B'C')$.

a) Montrer qu'il existe $A'', B'', C'' \in \mathcal{E}$ tels que A, B, C sont les milieux respectifs de $(B'', C''), (A'', C'')$ et (A'', B'') . Montrer que les hauteurs $(AA_1), (BB_1), (CC_1)$ de (ABC) sont les médiatrices des côtés du triangle $(A''B''C'')$; en déduire qu'elles sont concourantes en un point H (l'orthocentre).

b) Soit η l'homothétie de centre G et de rapport $-1/2$. Montrer que $\eta(A) = A', \eta(B) = B', \eta(C) = C'$, d'où $\eta(O) = \Omega$. Montrer que G est l'isobarycentre de $(A''B''C'')$, d'où $\eta(H) = O$. Conclure que G, H, O, Ω sont alignés et que Ω est le milieu de (O, H) .

c) Soit π la projection orthogonale sur (BC) . Montrer que $\pi(H) = A_1$, $\pi(O) = A'$, d'où $\Omega A_1 = \Omega A'$.

d) Montrer que $\overrightarrow{\Omega A} + \overrightarrow{\Omega H} = \overrightarrow{OA} = 2\overrightarrow{A'\Omega}$; en déduire que, si A_2 désigne le milieu de (A, H) , alors Ω est le milieu de (A', A_2) .

e) Conclure que, pour tout triangle ABC , il existe un cercle (dit *cercle des neuf points du triangle*) passant par les milieux A', B', C' des côtés, les pieds A_1, B_1, C_1 des hauteurs, et les milieux A_2, B_2, C_2 des bipoints $(A, H), (B, H), (C, H)$, où H est l'orthocentre.

EXERCICE 9. (*Détermination d'une isométrie par l'image d'une base affine*) Soit \mathcal{E} un plan affine euclidien. Soient ABC un triangle non aplati de \mathcal{E} . Soient A', B', C' trois points de \mathcal{E} tels que $A'B' = AB$, $B'C' = BC$ et $C'A' = CA$. Montrer que $A'B'C'$ est un triangle non aplati et qu'il existe une unique isométrie de \mathcal{E} qui envoie A sur A' , B sur B' et C sur C' .

Formuler et démontrer un résultat analogue en dimension 3.

EXERCICE 10. (*Forme réduite d'une isométrie*). Soit \mathcal{E} un plan affine euclidien. En utilisant la classification des isométries de \mathcal{E} , vérifier que toute isométrie affine φ s'écrit de façon unique sous la forme $\varphi = \tau \circ \varphi_0 = \varphi_0 \circ \tau$ où φ_0 est une isométrie ayant des points fixes et τ une translation de \mathcal{E} . Vérifier que le vecteur de τ appartient au sous-espace vectoriel directeur du sous-espace affine des points fixes de φ_0 .

Formuler et démontrer un résultat analogue en dimension 3.

EXERCICE 11. (*Conservation de la distance et applications affines*). Soit \mathcal{E} un espace affine euclidien, d'espace vectoriel directeur E . Soit φ une application de \mathcal{E} dans \mathcal{E} conservant la distance. Fixons $A \in \mathcal{E}$ et considérons l'application $f : E \rightarrow E$ qui à tout vecteur \vec{u} associe $f(\vec{u}) := \overrightarrow{\varphi(A)\varphi(M)}$, où M est le point de \mathcal{E} tel que $\overrightarrow{AM} = \vec{u}$.

Montrer que f ainsi construite vérifie $f(\vec{0}) = \vec{0}$ et $\|f(\vec{u}) - f(\vec{v})\| = \|\vec{u} - \vec{v}\|$ pour tous $\vec{u}, \vec{v} \in E$, puis $f(\vec{u}) \cdot f(\vec{v}) = \vec{u} \cdot \vec{v}$; en déduire que f est linéaire.

Conclure que, s'il est d'usage de définir une isométrie affine de \mathcal{E} comme une application affine de \mathcal{E} dans \mathcal{E} conservant la distance, on peut dans cette définition omettre le mot "affine".

EXERCICE 12 (*Déplacement*). Soit \mathcal{E} un espace affine euclidien orienté de dimension 3, d'espace vectoriel associé E . On fixe un repère orthonormé direct \mathcal{R} . Soit φ l'application $\mathcal{E} \rightarrow \mathcal{E}$ qui, à tout point M de coordonnées (x', y', z') associe $\varphi(M) = M'$ dont les coordonnées (x', y', z') sont :

$$x' = -z - 2, \quad y' = -x + 1, \quad z' = y + 1.$$

1) Montrer que φ est un endomorphisme affine sans points fixes, dont l'application linéaire associée $f \in \text{End } E$ est une rotation vectorielle de E dont on déterminera l'axe Δ et l'angle θ (moyennant une orientation de Δ). En déduire que $\varphi \in \text{Is}^+(\mathcal{E})$.

2) Déterminer l'ensemble \mathcal{D} des points $M \in \mathcal{E}$ tels que $\overrightarrow{M\varphi(M)} \in \Delta$. Montrer qu'il existe $\vec{u} \in \Delta$ tel que $\overrightarrow{M\varphi(M)} = \vec{u}$ pour tout $M \in \mathcal{D}$. En notant τ la translation de vecteur \vec{u} dans \mathcal{E} et $\rho = \tau^{-1} \circ \varphi$, montrer que ρ est la rotation affine d'axe \mathcal{D} et d'angle θ ; conclure que φ est un vissage.

EXERCICE 13 (*Déplacement*). Soit \mathcal{E} un espace affine euclidien orienté de dimension 3, d'espace vectoriel associé E . On fixe un repère orthonormé direct \mathcal{R} . Soit φ l'application $\mathcal{E} \rightarrow \mathcal{E}$ qui, à tout point M de coordonnées (x, y, z) associe $\varphi(M) = M'$ dont les coordonnées (x', y', z') sont :

$$x' = -x + 2, \quad y' = z + 1, \quad z' = y + 1.$$

1) Montrer que φ est un endomorphisme affine sans points fixes. Exprimer pour tout point $M \in \mathcal{E}$ la distance $M\varphi(M)$ en fonction des coordonnées de M , déterminer l'ensemble \mathcal{D} des points M pour lesquels cette distance est minimale, et montrer que \mathcal{D} est un sous-espace affine de \mathcal{E} stable par φ .

2) Déterminer $\vec{u} \in E$ tel que $\overrightarrow{M\varphi(M)} = \vec{u}$ pour tout $M \in \mathcal{D}$. Conclure que φ est un vissage.

EXERCICE 14 (*Déplacement*). Soit \mathcal{E} un espace affine euclidien orienté de dimension 3, d'espace vectoriel associé E . On fixe un repère orthonormé direct \mathcal{R} . Soit φ l'application $\mathcal{E} \rightarrow \mathcal{E}$ qui, à tout point M de coordonnées (x', y', z') associe $\varphi(M) = M'$ dont les coordonnées (x', y', z') sont :

$$x' = \frac{1}{3}(-2x - 2y + z + 1), \quad y' = \frac{1}{3}(-2x + y - 2z + 2), \quad z' = \frac{1}{3}(x - 2y - 2z + 1).$$

Montrer que φ est un vissage dont on déterminera l'axe, l'angle, le vecteur.

EXERCICE 15 (*Antidéplacement*). Soit \mathcal{E} un espace affine euclidien de dimension 3, d'espace vectoriel associé E . On fixe un repère orthonormé $\mathcal{R} = (O, \mathcal{B})$ avec $\mathcal{B} = (\vec{i}, \vec{j}, \vec{k})$.

1) Pour tous $a, b \in \mathbb{R}$, soit $f_{a,b} \in \text{End } E$ dont la matrice par rapport à \mathcal{B} est :

$$M_{a,b} = \begin{pmatrix} 3a+b & -4a & 0 \\ -4a & -3a+b & 0 \\ 0 & b & 1 \end{pmatrix}$$

1) Pour quelles valeurs de a et b a-t-on $f_{a,b} \in \text{GL}(E)$? $f_{a,b} \in \text{O}(E)$? Pour $a = \frac{1}{5}$ et $b = 0$, montrer que $f_{1/5,0}$ est une symétrie orthogonale par rapport à un sous-espace vectoriel F de E .

2) Soit φ l'application $\mathcal{E} \rightarrow \mathcal{E}$ qui, à tout point M de coordonnées (x', y', z') par rapport à \mathcal{R} associe $\varphi(M) = M'$ dont les coordonnées (x', y', z') sont définies par :

$$x' = \frac{3}{5}x - \frac{4}{5}y + 4, \quad y' = -\frac{4}{5}x - \frac{3}{5}y - 3, \quad z' = z.$$

(i) Montrer que $\varphi \in \text{Is}^-(\mathcal{E})$. (ii) Soit \mathcal{P} le plan affine passant par O et dirigé par le plan vectoriel de base (\vec{i}, \vec{j}) ; montrer que \mathcal{P} est globalement invariant par φ . On note ψ la restriction de φ à \mathcal{P} . Déterminer l'unique droite affine \mathcal{D} de \mathcal{P} et l'unique vecteur non-nul \vec{v} de la droite vectorielle Δ dirigeant \mathcal{D} tels que $\psi = \psi_0 \circ \tau_{\vec{v}} = \tau_{\vec{v}} \circ \psi_0$ où ψ_0 désigne la symétrie orthogonale par rapport à \mathcal{D} dans \mathcal{P} . (iii) Soit \mathcal{F} le plan affine contenant la droite \mathcal{D} et dirigé par le plan vectoriel F de base (\vec{v}, \vec{k}) . Déterminer l'unique vecteur $\vec{u} \in F$ tel que $\varphi = \sigma_{\mathcal{F}} \circ \tau_{\vec{u}} = \tau_{\vec{u}} \circ \sigma_{\mathcal{F}}$ où $\sigma_{\mathcal{F}}$ désigne la symétrie orthogonale par rapport à \mathcal{F} dans \mathcal{E} .

EXERCICE 16. Soit \mathcal{E} un espace affine euclidien orienté de dimension 3, d'espace vectoriel directeur E . On fixe $\vec{w} \in E$ non-nul et on note τ la translation de \mathcal{E} de vecteur \vec{w} . On considère l'ensemble \mathcal{A} des déplacements $\varphi \in \text{Is}^+(\mathcal{E})$ tels que $\varphi = \tau \circ \varphi \circ \tau$.

1) Montrer que $\varphi \in \mathcal{A}$ si et seulement l'application linéaire f associée à φ vérifie $f(\vec{w}) = -\vec{w}$.

2) Montrer qu'une application $f \in \text{O}^+(\mathcal{E})$ vérifie $f(\vec{w}) = -\vec{w}$ si et seulement si f est une rotation d'angle π et d'axe Δ inclus dans le plan $\text{Vect}\{\vec{w}\}^\perp$. En déduire l'ensemble \mathcal{A} .

EXERCICE 17 (*Isométries échangeant deux droites non coplanaires*). Soient \mathcal{D}_1 et \mathcal{D}_2 deux droites non coplanaires d'un espace affine euclidien \mathcal{E} de dimension 3, d'espace vectoriel associé E . Soit \mathcal{D} la droite de \mathcal{E} perpendiculaire à \mathcal{D}_1 et \mathcal{D}_2 ; on note $\{A_1\} = \mathcal{D}_1 \cap \mathcal{D}$ et $\{A_2\} = \mathcal{D}_2 \cap \mathcal{D}$.

Soit $\varphi \in \text{Is}(\mathcal{E})$ telle que $\varphi(\mathcal{D}_1) = \mathcal{D}_2$ et $\varphi(\mathcal{D}_2) = \mathcal{D}_1$. Soit $f \in \text{O}(\mathbb{E})$ son application linéaire associée.

Montrer que $\varphi(\mathcal{D}) = \mathcal{D}$, $\varphi(A_1) = A_2$ et $\varphi(A_2) = A_1$. En déduire que φ admet un point fixe O . Déterminer φ , en distinguant suivant que \mathcal{D}_1 et \mathcal{D}_2 sont ou non orthogonales.

Appendice : archives de sujets de devoirs en temps limité

LICENCE SCIENCES-LANGUES, DEUXIÈME ANNÉE
Mathématiques premier semestre

Devoir surveillé n° 1

Durée : deux heures. *Sans documents.*

EXERCICE 1

Soient a, b, c, d, e, f des réels. On considère dans $\mathcal{M}_4(\mathbb{R})$ les matrices :

$$A = \begin{pmatrix} 1 & a & b & c \\ 0 & 1 & d & e \\ 0 & 0 & -1 & f \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad B = A - I_4, \quad C = A + I_4.$$

- 1) Déterminer, suivant les valeurs des paramètres a, b, c, d, e, f , le rang de la matrice B .
- 2) Déterminer, suivant les valeurs des paramètres a, b, c, d, e, f , le rang de la matrice C .
- 3) Déterminer les valeurs propres de la matrice A .
- 4) Dédire des trois questions précédentes une condition nécessaire et suffisante portant sur les paramètres a, b, c, d, e, f , pour que A soit diagonalisable.

EXERCICE 2

Soient q un réel. On considère dans $\mathcal{M}_3(\mathbb{R})$ la matrice :

$$A = \begin{pmatrix} 3 - q & q - 5 & q \\ -q & q - 2 & q \\ 5 & -5 & -2 \end{pmatrix}.$$

- 1) Calculer le polynôme caractéristique de A .
- 2) Pour quelle(s) valeur(s) de q la matrice A est-elle inversible ?
- 3) Montrer qu'il existe une unique valeur q_0 de q , que l'on déterminera, pour laquelle A est diagonalisable.
- 4) On prend ici $q = q_0$. Diagonaliser A . Calculer explicitement A^n pour tout entier $n \geq 0$. La formule trouvée ci-dessus est-elle encore valable pour $n = -1$?

suite au verso...

EXERCICE 3

Soient f et g deux endomorphismes d'un \mathbb{K} -espace vectoriel E de dimension finie. Montrer que $g \circ f$ et $f \circ g$ ont les mêmes valeurs propres.

Indications. Soit λ une valeur propre de $f \circ g$.

Supposer d'abord que $\lambda = 0$; montrer que $f \circ g$ n'est pas bijective, en déduire qu'il en est de même pour $g \circ f$, et conclure que 0 est aussi valeur propre de $g \circ f$.

Supposer ensuite que $\lambda \neq 0$; prendre un vecteur propre x de $f \circ g$ associé à la valeur propre λ , et considérer le vecteur $g(x)$.

EXERCICE 4

On rappelle que $\mathcal{M}_3(\mathbb{R})$ est un espace vectoriel de dimension 9 sur \mathbb{R} , dont la base canonique est $\mathcal{B} = (E_{11}, E_{12}, E_{13}, E_{21}, E_{22}, E_{23}, E_{31}, E_{32}, E_{33})$, où E_{ij} désigne la matrice "élémentaire" dont le coefficient de la i -ième ligne et j -ième colonne est 1, et tous les autres sont 0.

1) Pour toute matrice $A \in \mathcal{M}_3(\mathbb{R})$, on note $\mathcal{C}(A)$ l'ensemble des matrices $M \in \mathcal{M}_3(\mathbb{R})$ qui commutent avec A , c'est-à-dire telles que $AM = MA$. Montrer que $\mathcal{C}(A)$ est toujours un sous-espace vectoriel de $\mathcal{M}_3(\mathbb{R})$.

2) On considère dans $\mathcal{M}_3(\mathbb{R})$ les matrices :

$$A = \begin{pmatrix} 2 & -1 & -2 \\ 2 & -1 & -4 \\ -1 & 1 & 3 \end{pmatrix} \quad \text{et} \quad D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

a) Démontrer qu'il existe dans $\mathcal{M}_3(\mathbb{R})$ une matrice inversible P telle que $A = PDP^{-1}$ (on n'aura pas besoin dans la suite du calcul explicite de P).

b) Montrer que l'application $\phi : \mathcal{M}_3(\mathbb{R}) \rightarrow \mathcal{M}_3(\mathbb{R})$ définie par $\phi(M) = P^{-1}MP$ pour tout $M \in \mathcal{M}_3(\mathbb{R})$ est un endomorphisme (c'est-à-dire est linéaire).

c) Démontrer que, pour toute $M \in \mathcal{M}_3(\mathbb{R})$, on a $M \in \mathcal{C}(A)$ si et seulement si $\phi(M) \in \mathcal{C}(D)$.

d) Démontrer que ϕ définit un isomorphisme d'espaces vectoriels de $\mathcal{C}(A)$ sur $\mathcal{C}(D)$.

3) Montrer qu'une matrice quelconque $M' = \begin{pmatrix} a & b & c \\ d & e & f \\ i & j & k \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$ appartient à $\mathcal{C}(D)$ si et seulement si elle est de la forme $M' = \begin{pmatrix} a & 0 & 0 \\ 0 & e & f \\ 0 & j & k \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$. En déduire une base et la dimension de $\mathcal{C}(D)$.

4) Déduire des questions 2.d) et 3) la dimension de $\mathcal{C}(A)$, et une méthode pour en trouver une base.

LICENCE SCIENCES-LANGUES, DEUXIÈME ANNÉE
Mathématiques premier semestre

Devoir surveillé n° 2

Durée : deux heures. *Sans documents.*

EXERCICE 1

On considère l'espace vectoriel euclidien $E = \mathbb{R}^3$ muni du produit scalaire canonique. On note \mathcal{B} la base canonique de E .

Pour tout $a \in \mathbb{R}$, avec $a \neq 1$, on considère l'endomorphisme f_a de E dont la matrice par rapport à la base canonique est :

$$M_a = \frac{1}{3} \begin{pmatrix} a+2 & a-1 & a-1 \\ a-1 & a+2 & a-1 \\ a-1 & a-1 & a+2 \end{pmatrix}$$

- 1) Montrer que l'ensemble des vecteurs u de E qui vérifient $f_a(u) = u$ est un plan P de E .
- 2) Montrer que a est valeur propre simple de f_a . Déterminer le sous-espace propre associé D .
- 3) En déduire que M_a admet deux valeurs propres distinctes et que les sous-espaces propres associés sont orthogonaux.
- 4) Donner une base de E qui soit constituée de vecteurs propres de M_a et qui soit orthogonale. En déduire une base \mathcal{B}' de E qui soit constituée de vecteurs propres de M_a et qui soit orthonormale.
- 5) Soit P la matrice de passage de \mathcal{B} à \mathcal{B}' . Calculer le produit de P par sa transposée. Que peut-on en déduire pour P ? Donner la matrice D_a de f_a dans la base \mathcal{B}' .
- 6) En déduire que, pour $a = 0$, f_a est une projection orthogonale sur un sous-espace vectoriel F de E que l'on déterminera.
- 7) Montrer que, pour $a = -1$, f_a est la symétrie orthogonale par rapport à F .

suite au verso...

EXERCICE 2

On se place dans un espace vectoriel euclidien. On note $\langle \cdot | \cdot \rangle$ le produit scalaire et $n = \dim E$.

I. On fixe un vecteur a dans E tel que $\|a\| = 1$. Pour tout $\alpha \in \mathbb{R}$, on considère l'application $f_\alpha : E \rightarrow E$ définie par :

$$f_\alpha(x) = x + \alpha \langle a | x \rangle a \quad \text{pour tout } x \in E.$$

1) Propriétés des applications f_α .

- a) Montrer que, pour tout $\alpha \in \mathbb{R}$, l'application f_α est linéaire.
- b) Montrer que, pour tous $\alpha, \beta \in \mathbb{R}$, il existe un réel γ que l'on exprimera en fonction de α et β , tel que $f_\alpha \circ f_\beta = f_\beta \circ f_\alpha = f_\gamma$.
- c) Montrer qu'il existe une unique valeur $\alpha_0 \in \mathbb{R}$ telle que $f_{\alpha_0} = \text{id}_E$.
- d) Pour quelles valeurs de α l'endomorphisme f_α est-il bijectif? Montrer qu'alors l'endomorphisme réciproque f_α^{-1} est de la forme $f_{\alpha'}$ pour une valeur de α' que l'on exprimera en fonction de α .
- e) En résumant les questions précédentes, que peut-on dire de $G = \{f_\alpha ; \alpha \in \mathbb{R}, \alpha \neq -1\}$?

2) Valeurs propres des endomorphismes f_α . On fixe $\alpha \in \mathbb{R}$.

- a) Montrer que tout vecteur non-nul de E qui est orthogonal à a est un vecteur propre de f_α associé à une valeur propre que l'on précisera.
- b) Montrer que tout vecteur non-nul de E qui est colinéaire à a est un vecteur propre de f_α associé à une valeur propre que l'on précisera.
- c) En déduire toutes les valeurs propres de f_α et les sous-espaces propres associés. L'endomorphisme f_α est-il diagonalisable?

II. On fixe deux vecteurs a et b dans E tel que $\|a\| = \|b\| = 1$. On considère l'application $g : E \rightarrow E$ définie par :

$$g(x) = x - \langle a | x \rangle b \quad \text{pour tout } x \in E.$$

1) Vérifier que $\langle a | b \rangle = 1$ si et seulement si $a = b$.

► *Indication* : penser au cas d'égalité dans l'inégalité de Cauchy-Schwarz.

2) Donner une condition nécessaire et suffisante portant sur les vecteurs a et b pour que g soit bijective.

► *Indication* : calculer $\langle a | x \rangle$ pour $x \in \text{Ker } g$.

Lorsque cette condition est satisfaite, donner pour tout $y \in E$ une expression explicite de $g^{-1}(y)$ en fonction de y , a et b .

3) Déterminer que les seules valeurs propres possibles de g sont 1 et $1 - \langle a | b \rangle$.

► *Indication* : montrer que, si x est un vecteur propre associé à une valeur propre λ de g , alors

$$\langle a | x \rangle \langle a | b \rangle = (1 - \lambda) \langle a | x \rangle.$$

En déduire que g est diagonalisable si et seulement si les vecteurs a et b ne sont pas orthogonaux.

LICENCE SCIENCES-LANGUES, DEUXIÈME ANNÉE
Mathématiques premier semestre

Devoir surveillé n° 3

Durée : deux heures. *Sans documents.*

EXERCICE 1

On se place dans un espace vectoriel euclidien orienté E de dimension 3. On fixe $\mathcal{B} = (e_1, e_2, e_3)$ une base orthonormale directe de E . On fixe trois réels a, b, c tels que $a^2 + b^2 + c^2 = 1$. On note u le vecteur de E de composantes (a, b, c) dans la base \mathcal{B} , qui est donc unitaire.

1) Montrer que l'endomorphisme f de E dont la matrice par rapport à la base \mathcal{B} est :

$$A = \begin{pmatrix} a^2 & ab - c & ac + b \\ ab + c & b^2 & bc - a \\ ac - b & bc + a & c^2 \end{pmatrix}$$

est une isométrie vectorielle [c'est-à-dire un élément de $O(E)$].

2) Montrer que l'application $g : E \rightarrow E$ définie par :

$$g(x) = \langle x|u \rangle u \quad \text{pour tout } x \in E.$$

est linéaire. Ecrire sa matrice B dans la base \mathcal{B} .

3) Montrer que l'application $h : E \rightarrow E$ définie par :

$$h(x) = u \wedge x \quad \text{pour tout } x \in E.$$

est linéaire. Ecrire sa matrice C dans la base \mathcal{B} .

4) Dédire des questions précédentes que :

$$f(x) = \langle x|u \rangle u + u \wedge x \quad \text{pour tout } x \in E.$$

5) Calculer $f(u)$. Montrer que l'on peut choisir v, w dans E tels que $\mathcal{B}' = (u, v, w)$ soit une base orthonormale directe de E . Calculer $f(v)$ et $f(w)$. Conclure que f est une rotation d'axe dirigé et orienté par u , et dont on déterminera l'angle.

suite au verso...

EXERCICE 2

On se place dans un espace vectoriel euclidien E de dimension $n \geq 3$. On fixe deux vecteurs unitaires a et b , que l'on suppose non colinéaires. On appelle H_a l'hyperplan de E dont a est un vecteur normal, et H_b l'hyperplan de E dont b est un vecteur normal.

1) Montrer que $H_a \neq H_b$ et calculer la dimension du sous-espace vectoriel $H_a \cap H_b$.

2) Montrer que l'application $f : E \rightarrow E$ définie par :

$$f(x) = \langle x|b \rangle a \quad \text{pour tout } x \in E.$$

est linéaire. Montrer que l'endomorphisme transposé ${}^t f$ vérifie :

$${}^t f(y) = \langle y|a \rangle b \quad \text{pour tout } y \in E.$$

3) On introduit l'endomorphisme $g = f + {}^t f$ de E .

a) Calculer $\text{Ker } g$. En déduire que 0 est valeur propre d'ordre au moins $n - 2$ de g .

b) Calculer $g(a + b)$ et $g(a - b)$; en déduire que g admet deux valeurs propres λ et μ , non-nulles et distinctes, que l'on déterminera explicitement.

c) Conclure que g est diagonalisable et que les sous-espaces propres sont deux à deux orthogonaux.

d) Le résultat du c) était-il prévisible sans aucun calcul ?

EXERCICE 3

On se place dans un espace vectoriel euclidien orienté E de dimension 3. On fixe $\mathcal{B} = (e_1, e_2, e_3)$ une base orthonormale directe de E . Pour tous réel p, q, r , on note $f_{p,q,r}$ l'endomorphisme de E dont la matrice par rapport à la base \mathcal{B} est :

$$M_{p,q,r} = \begin{pmatrix} p & q & r \\ r & p & q \\ q & r & p \end{pmatrix}.$$

1) Montrer que $f_{p,q,r} \in \text{O}(E)$ si et seulement si $|p + q + r| = 1$ et $pq + qr + rp = 0$.

2) On dit que $f_{p,q,r}$ est involutif si et seulement si $f_{p,q,r} \circ f_{p,q,r} = \text{id}_E$. Montrer qu'un endomorphisme $f_{p,q,r}$ qui appartient à $\text{O}(E)$ est involutif si et seulement si $q = r$.

3) On note G l'ensemble des endomorphismes $f_{p,q,r}$ qui appartiennent à $\text{O}(E)$ et qui sont involutifs.

a) Montrer que G est formé de quatre éléments, dont une symétrie orthogonale s par rapport à un plan P que l'on déterminera, et une rotation r dont l'axe est la droite $D = P^\perp$.

b) Ecrire sous forme d'une table à quatre lignes et quatre colonnes tous les produits deux à deux (pour la loi \circ) des éléments de G . Que peut-on dire de G ?

LICENCE SCIENCES-LANGUES, DEUXIÈME ANNÉE
Mathématiques premier semestre

Devoir surveillé n° 4

Durée : deux heures. Sans documents.

EXERCICE 1

Déterminer la nature des séries de terme général ci-dessous :

a) $u_n = \frac{1}{n \cos^2 n}$, $v_n = \left(\frac{1}{n}\right)^{1+\frac{1}{n}}$, $w_n = \frac{1}{(\ln n)^{\ln n}}$.

Indication : on pourra considérer les suites nv_n et n^2w_n .

b) $u_n = \sin\left(n\pi + \frac{\pi}{n}\right)$, $v_n = \frac{1}{\sin \frac{1}{n}} \left(1 - \cos\left(\frac{1}{n\sqrt{\ln n}}\right)\right)$.

c) $u_n = \left(\frac{\lambda^n}{1 + \lambda^{2n}}\right)$, $v_n = \left(\frac{\lambda^{2n}}{1 + \lambda^{2n}}\right)$, $w_n = \left(\frac{1}{1 + \lambda^{2n}}\right)$, avec $\lambda \in \mathbb{R}_+$.

Indication : distinguer suivant que $\lambda = 1$, $\lambda > 1$ ou $\lambda < 1$.

EXERCICE 2

1) Pour tout entier $n \geq 1$, on note $H(n) = \sum_{k=1}^n \frac{1}{k}$. Rappeler comment est définie la constante d'Euler γ . En déduire que, pour tout $n \geq 1$:

$$H(n) = \ln n + \gamma + \varepsilon(n) \quad \text{avec} \quad \lim_{n \rightarrow +\infty} \varepsilon(n) = 0.$$

2) On considère la série $\sum_{n \geq 1} u_n$ où $u_n = \frac{5n+6}{n(n+1)(n+2)}$. Montrer que cette série converge.

3) Déterminer explicitement trois entiers $a, b, c \in \mathbb{Z}$ tels que : $u_n = \frac{a}{n} + \frac{b}{n+1} + \frac{c}{n+2}$.

4) Déduire des questions 1) et 3) qu'il existe un entier d que l'on déterminera tel que, pour tout $n \geq 1$:

$$\sum_{k=1}^n u_k = \ln \frac{n^3}{(n+1)(n+2)^2} + d + \varepsilon'(n) \quad \text{avec} \quad \lim_{n \rightarrow +\infty} \varepsilon'(n) = 0.$$

5) Calculer $S = \sum_{k=1}^{+\infty} u_k$.

suite au verso...

EXERCICE 3

On fixe un réel θ non multiple de 2π . Pour tout entier $n \geq 1$, on pose :

$$S_n = \sum_{k=0}^n \cos k\theta \quad \text{et} \quad u_n = \frac{\cos n\theta}{n}.$$

1) Montrer que la suite (S_n) est bornée.

Indication : on pourra introduire les sommes partielles de la série complexe $\sum_{k \geq 0} \exp ik\theta$.

2) Exprimer u_n en fonction de S_n et S_{n-1} . En déduire que, pour tout $N \in \mathbb{N}$:

$$\sum_{n=1}^N u_n = \sum_{n=1}^N \frac{S_n}{n(n+1)} - S_0 + \frac{S_N}{N+1}.$$

En déduire que la série $\sum_{n \geq 1} u_n$ est convergente.

3) Montrer que $|u_n| \geq u_{2n} + \frac{1}{2n}$ pour tout $n \geq 1$.

Indication : on pourra utiliser le fait que $|\cos x| \geq \cos^2 x$ pour tout $x \in \mathbb{R}$.

En déduire que la série $\sum_{n \geq 1} u_n$ n'est pas absolument convergente.

EXERCICE 4

Déterminer la nature des séries de terme général $u_n = \frac{(-1)^n}{\ln[n + (-1)^n]}$ et $v_n = \frac{(-1)^n}{\ln n + (-1)^n}$.

Indication : on pourra

- introduire la série de terme général $a_n = \frac{(-1)^n}{\ln n}$ et déterminer sa nature,
- écrire u_n et v_n sous la forme $u_n = a_n - b_n$ et $v_n = a_n - c_n$, pour des suites (b_n) et (c_n) de réels choisies de façon appropriée,
- déterminer la nature des séries $\sum b_n$ et $\sum c_n$,
- conclure en déterminant la nature des séries $\sum u_n$ et $\sum v_n$.

LICENCE SCIENCES-LANGUES, DEUXIÈME ANNÉE
Mathématiques premier semestre

Complément au devoir surveillé n° 4

Durée : 45 minutes *Sans documents.*

Déterminer la nature (convergente ou divergente) des séries réelles de terme général u_n suivant :

1) $u_n = \frac{n}{\sqrt{n^3 + 1}}$.

2) $u_n = \frac{\operatorname{ch} n}{\operatorname{ch} 2n}$. (On rappelle que $\operatorname{ch} x = \frac{1}{2}(e^x + e^{-x})$ pour tout $x \in \mathbb{R}$).

3) $u_n = e - \left(1 + \frac{1}{n}\right)^n$.

4) $u_n = \left(\frac{2n+1}{3n+4}\right)^n$.

5) $u_n = \left(1 + \frac{x}{n}\right)^{-n^2}$, où $x \in \mathbb{R}$ fixé. (Distinguer suivant les valeurs de x).

6) $u_n = \frac{a^n}{(1+a)(1+a)^2 \cdots (1+a)^n}$ où $a \in \mathbb{R}_+$ fixé.

7) $u_n = (-1)^n (\sqrt{n^2 + 1} - n)$

8) $u_n = \frac{1}{n^2 + 1} [(-1)^n + k]$, où $k \in \mathbb{R}$ fixé. (Ecrire $\sum u_n$ comme somme de deux séries).

9) $u_n = \ln \left(1 + \frac{(-1)^n}{n^a}\right)$, où $a \in \mathbb{R}_+^*$ fixé.

10) $u_n = \left(\frac{\sqrt{n}}{\sqrt{n} + 1}\right)^n$. (Etudier $n^2 u_n$).

LICENCE SCIENCES-LANGUES, DEUXIÈME ANNÉE
Mathématiques premier semestre

Devoir surveillé n° 5

Durée : deux heures. Sans documents.

EXERCICE 1

Soit E un espace vectoriel normé.

- 1) Rappeler (sans démonstration) l'énoncé dans E de la proposition connue sous le nom de caractérisation séquentielle de l'adhérence.
- 2) Soit F une partie de E et \overline{F} son adhérence. Montre que si F est un sous-espace vectoriel de E , alors il en est de même de \overline{F} .
- 3) Pour toute partie A de E , on note $\text{Vect } A$ le sous-espace vectoriel de E engendré par A . Montrer que l'on a $\text{Vect}(\overline{A}) \subseteq \overline{\text{Vect } A}$.

EXERCICE 2

Soit E le \mathbb{R} -espace vectoriel des suites de réels qui sont bornées, muni de la norme $\|\cdot\|_\infty$ définie par $\|x\|_\infty = \sup_{n \geq 0} |x_n|$ pour toute $x = (x_n)_{n \geq 0}$ dans E .

Soit F le sous-espace vectoriel de E formé des suites de réels qui sont convergentes dans E .

Soit $u = (u_n)_{n \geq 0}$ la suite définie par $u_n = (-1)^n$ pour tout entier $n \geq 0$.

- 1) Vérifier que $u \in E$ mais $u \notin F$. On se propose donc de calculer dans la suite la distance entre u et F dans E , c'est-à-dire $d = \inf_{x \in F} \|x - u\|_\infty$.
- 2) Soit $x = (x_n)_{n \geq 0}$ une suite appartenant à F et ℓ sa limite.
 - a) Montrer que $|x_{2p} - u_{2p}| \leq \|x - u\|_\infty$ pour tout $p \geq 0$. En déduire que $|\ell - 1| \leq \|x - u\|_\infty$.
 - b) Montrer de même que $|\ell + 1| \leq \|x - u\|_\infty$. En déduire que $\|x - u\|_\infty \geq 1$.
- 3) Déduire du 2) que $d \geq 1$. Calculer la distance entre u et 0_E , en déduire que $d \leq 1$, et conclure.

EXERCICE 3

Soit E un espace vectoriel normé. Soit $(u_n)_{n \geq 1}$ une suite d'éléments de E . On définit une nouvelle suite $(v_n)_{n \geq 1}$ d'éléments de E en posant :

$$v_n = \frac{1}{n}(u_1 + u_2 + \cdots + u_n) \quad \text{pour tout } n \geq 1.$$

- 1) Montrer que, si la suite (u_n) converge dans E vers une limite ℓ , alors la suite (v_n) converge dans E vers la même limite ℓ .
- 2) Montrer par un contre-exemple que la réciproque peut être fautive (*indication* : utiliser la suite $(-1)^n$ dans \mathbb{R}).

suite au verso...

EXERCICE 4

On considère le \mathbb{R} -espace vectoriel $E = \mathcal{C}^0([0, 1], \mathbb{R})$ des applications $[0, 1] \rightarrow \mathbb{R}$ continues sur l'intervalle réel $[0, 1]$. On considère dans E la norme $\|\cdot\|_1$ définie par $\|f\|_1 = \int_0^1 |f(t)| dt$. On considère la suite $(f_n)_{n \geq 2}$ d'applications $[0, 1] \rightarrow \mathbb{R}$ définie par :

$$f_n(x) = \begin{cases} 1 & \text{si } x \in [0, \frac{1}{2}], \\ -nx + \frac{n+2}{2} & \text{si } x \in [\frac{1}{2}, \frac{1}{2} + \frac{1}{n}], \\ 0 & \text{si } x \in [\frac{1}{2} + \frac{1}{n}, 1]. \end{cases}$$

- 1) Dessiner la représentation graphique de f_n , et montrer que $f_n \in E$ pour tout entier $n \geq 2$.
- 2) On considère deux entiers $q \geq p \geq 2$. On notera si besoin dans les calculs $q = p + r$ où $r = q - p \geq 0$.
 - a) Dessiner sur une même figure les représentations graphiques de f_p et f_q .
 - b) Montrer que : $\|f_p - f_q\|_1 = \frac{r}{2p(p+r)} \leq \frac{1}{2p}$.
 - c) En déduire que la suite (f_n) est de Cauchy dans l'evn E muni de la norme $\|\cdot\|_1$.
- 3) On suppose qu'il existe une fonction $f \in E$ telle que la suite (f_n) converge vers f dans E (toujours au sens de la norme $\|\cdot\|_1$).
 - a) Montrer que $\|f - f_n\|_1 \geq \int_0^{\frac{1}{2}} |f(t) - 1| dt$. En déduire que $f(x) = 1$ pour tout $x \in [0, \frac{1}{2}]$.
 - b) Montrer que $\|f - f_n\|_1 \geq \int_{\frac{1}{2}}^1 |f(t)| dt$ pour tout réel $\alpha \in]\frac{1}{2}, 1]$ et tout entier $n \geq 0$ assez grand pour que $\frac{1}{2} + \frac{1}{n} \leq \alpha$. En déduire que $f(x) = 0$ pour tout $x \in]\frac{1}{2}, 1]$.
 - c) Dégager des questions a) et b) une contradiction, et conclure que la suite (f_n) n'est pas convergente pas dans l'evn E pour la norme $\|\cdot\|_1$.
- 4) Quelle propriété de E peut-on déduire des questions 2.c) et 3.c) ?

EXERCICE 5

Soit E l'ensemble des applications $[0, 1] \rightarrow \mathbb{R}$ qui sont de classe C^1 (dérivables à dérivée continue) sur \mathbb{R} et qui vérifient $f(0) = 0$.

- 1) Montrer que E est un sous-espace vectoriel de l'espace vectoriel des applications $[0, 1] \rightarrow \mathbb{R}$ qui sont bornées sur \mathbb{R} .
- 2) Pour toute $f \in E$, on pose $N(f) = \|3f + f'\|_\infty$. Montrer que N est une norme sur E .
- 3) Montrer qu'il existe un réel $\alpha > 0$ tel que $\|f\|_\infty \leq \alpha N(f)$ pour tout $f \in E$.
(Indication : vérifier que $f(x) = e^{-3x} \int_0^x (3f(t) + f'(t))e^{3t} dt$ pour tout $x \in [0, 1]$).
- 4) Pour tout entier $n \geq 1$, calculer $\|f_n\|_\infty$ et $N(f_n)$ pour l'application $f_n : [0, 1] \rightarrow \mathbb{R}$ définie par $f_n(x) = x^n$. En déduire que N et $\|\cdot\|_\infty$ ne sont pas équivalentes.

LICENCE SCIENCES-LANGUES, DEUXIÈME ANNÉE
Mathématiques premier semestre

Devoir surveillé n° 6

Durée : deux heures. *Sans documents.*

EXERCICE 1

1) Pour tout entier $n \geq 1$, soit f_n l'application $[0, 1] \rightarrow \mathbb{R}$ définie par $f_n(x) = x^n \ln x$ pour $x \in]0, 1]$ et $f_n(0) = 0$.

– Etudier les variations de f_n .

– En déduire que la suite (f_n) converge uniformément sur $[0, 1]$ vers une application f que l'on déterminera.

2) Pour tout entier $n \geq 1$, soit g_n l'application $[0, +\infty[\rightarrow \mathbb{R}$ définie par $g_n(x) = e^{-nx} \sin(nx)$ pour tout $x \in [0, +\infty[$.

– Montrer que la suite (g_n) converge simplement sur $[0, +\infty[$ vers une application g que l'on déterminera.

– Calculer $g_n(\frac{\pi}{2n})$ et en déduire que la suite (g_n) ne converge pas uniformément sur $[0, +\infty[$.

– Montrer que la suite (g_n) converge uniformément sur tout intervalle $[a, +\infty[$ inclus dans $[0, +\infty[$ (avec $a > 0$).

3) Pour tout entier $n \geq 1$, soit h_n l'application $[0, 1] \rightarrow \mathbb{R}$ définie par $h_n(x) = n^2 x(1 - nx)$ si $0 \leq x \leq \frac{1}{n}$ et $h_n(x) = 0$ si $\frac{1}{n} \leq x \leq 1$.

– Montrer que la suite (h_n) converge simplement sur $[0, 1]$ vers une application h que l'on déterminera.

– Soit $I_n = \int_0^1 h_n(t) dt$ pour tout $n \geq 1$. Montrer que la suite (I_n) converge vers une limite $\ell \in \mathbb{R}$ que l'on déterminera. En déduire que la suite (h_n) ne converge pas uniformément sur $[0, 1]$.

– Montrer que la suite (h_n) converge uniformément sur tout intervalle $[a, 1]$ avec $a > 0$.

EXERCICE 2

Soit (f_n) une suite d'applications polynomiales de \mathbb{R} dans \mathbb{R} . On suppose que la suite (f_n) converge uniformément sur \mathbb{R} vers une application f .

1) Montrer qu'il existe un entier $N \geq 0$ tel que, pour tout $n \geq N$, on ait $\|f_n - f\|_\infty \leq 1$.

2) En déduire que, pour $n \geq N$ fixé quelconque, l'application polynomiale $f_n - f_N$ est bornée sur \mathbb{R} , et qu'il existe alors un réel c_n tel que $f_n(x) = f_N(x) + c_n$ pour tout $x \in \mathbb{R}$.

3) Conclure que f est une application polynomiale.

suite au verso...

EXERCICE 3

On considère trois suites de réels $(a_n)_{n \geq 1}$, $(b_n)_{n \geq 1}$ et $(c_n)_{n \geq 1}$ qui vérifient, pour tout $n \geq 1$:

$$0 < a_n - b_n < a_n < a_n + b_n \quad \text{et} \quad c_n > 0.$$

On considère pour tout $n \geq 1$ l'application affine par intervalles $f_n : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ définie par

$$f_n(x) = \begin{cases} 0 & \text{si } 0 \leq x \leq a_n - b_n \text{ ou si } x \geq a_n + b_n \\ \frac{c_n}{b_n}(x - a_n + b_n) & \text{si } a_n - b_n \leq x \leq a_n \\ -\frac{c_n}{b_n}(x - a_n - b_n) & \text{si } a_n \leq x \leq a_n + b_n \end{cases}.$$

1) Dessiner la représentation graphique de f_n . Montrer que f_n est bornée sur \mathbb{R} , intégrable sur \mathbb{R} , et telle que f_n^2 est intégrable sur \mathbb{R} .

2) Pour tout $n \geq 1$, calculer en fonction de a_n , b_n et c_n :

$$N_\infty(f_n) = \sup_{t \in \mathbb{R}} |f_n(t)|, \quad N_1(f_n) = \int_{\mathbb{R}} |f_n(t)| dt \quad \text{et} \quad N_2(f_n) = \sqrt{\int_{\mathbb{R}} |f_n(t)|^2 dt}.$$

Indication : montrer qu'il existe un réel $\lambda > 0$, que l'on déterminera, tel quel $N_2(f) = \lambda c_n \sqrt{b_n}$.

3) Montrer que si l'on choisit $a_n = n^3$, $b_n = n^3$ et $c_n = \frac{1}{n}$, alors la suite (f_n) obtenue vérifie :

$$\lim_{n \rightarrow +\infty} N_\infty(f_n) = 0 \quad \text{mais} \quad \lim_{n \rightarrow +\infty} N_1(f_n) \neq 0 \quad \text{et} \quad \lim_{n \rightarrow +\infty} N_2(f_n) \neq 0.$$

4) Montrer que si l'on choisit $a_n = 1$, $b_n = \frac{1}{n^3}$ et $c_n = n^2$, alors la suite (f_n) obtenue vérifie :

$$\lim_{n \rightarrow +\infty} N_1(f_n) = 0 \quad \text{mais} \quad \lim_{n \rightarrow +\infty} N_2(f_n) \neq 0 \quad \text{et} \quad \lim_{n \rightarrow +\infty} N_\infty(f_n) \neq 0.$$

5) Montrer que si l'on choisit $a_{2p} = 1$ et $a_{2p+1} = (2p+1)^3$, $b_{2p} = \frac{1}{p^3}$ et $b_{2p+1} = (2p+1)^3$, $c_{2p} = p$ et $c_{2p+1} = \frac{1}{(2p+1)^2}$, alors la suite (f_n) obtenue vérifie :

$$\lim_{n \rightarrow +\infty} N_2(f_n) = 0 \quad \text{mais} \quad \lim_{n \rightarrow +\infty} N_1(f_n) \neq 0 \quad \text{et} \quad \lim_{n \rightarrow +\infty} N_\infty(f_n) \neq 0.$$

EXERCICE 4

1) Rappeler sans démonstration l'énoncé du théorème de convergence dominée.

2) Pour tout $n \geq 1$, soit f_n l'application $\mathbb{R}_+ \rightarrow \mathbb{R}$ définie par $f_n(x) = (1 + \frac{x}{n})^n e^{-2x}$ pour $0 \leq x \leq n$, et $f_n(x) = 0$ si $x > n$.

– Montrer que la suite (f_n) converge simplement sur \mathbb{R}_+ vers une application f que l'on déterminera.

– Montrer que, pour tout $n \geq 1$ et tout $x \in \mathbb{R}_+$, on a : $|f_n(x)| \leq f(x)$.

– En déduire que $\lim_{n \rightarrow +\infty} \int_0^n (1 + \frac{x}{n})^n e^{-2x} dx$ existe dans \mathbb{R} et calculer sa valeur.

3) Pour tout $n \geq 1$, soit g_n l'application $[1, +\infty[\rightarrow \mathbb{R}$ définie par $g_n(x) = \frac{|\sin^n x|}{x^2}$ pour tout $x \geq 1$.

– Montrer que la suite (g_n) converge simplement sur $[1, +\infty[$ vers une application g que l'on déterminera.

– En déduire que la suite (g_n) ne converge pas uniformément sur $[1, +\infty[$.

– Montrer que $\lim_{n \rightarrow +\infty} \int_1^{+\infty} \frac{\sin^n x}{x^2} dx$ existe dans \mathbb{R} et calculer sa valeur.

LICENCE SCIENCES-LANGUES, DEUXIÈME ANNÉE
Mathématiques premier semestre

Devoir surveillé n° 7

Durée : deux heures. Sans documents.

EXERCICE 1

On considère la série entière $\sum_{n \geq 2} a_n x^n$ avec $a_n = \frac{(-1)^n}{n(n-1)}$.

- 1) Déterminer le rayon de convergence de cette série.
- 2) Montrer que l'application $f :]-1, 1[\rightarrow \mathbb{R}$ définie par $f(x) = \sum_{n=2}^{+\infty} a_n x^n$ est bien définie, et que la convergence de la série entière est normale sur $]-1, 1[$.
- 3) Exprimer $f(x)$ à l'aide des fonctions usuelles lorsque $x \in]-1, 1[$. [*Indication* : on pourra exprimer d'abord la dérivée de f].
- 4) Montrer que f est continue sur $[-1, 1]$ et en déduire les valeurs de $f(1)$ et $f(-1)$.

EXERCICE 2

On considère la série entière $\sum_{n \geq 1} a_n x^n$ avec $a_n = \frac{1}{n} \cos \frac{2n\pi}{3}$.

- 1) Déterminer le rayon de convergence de sa série entière dérivée et en déduire que l'application $f :]-1, 1[\rightarrow \mathbb{R}$ définie par $f(x) = \sum_{n=1}^{+\infty} a_n x^n$ est bien définie.
- 2) Montrer que, pour tout $x \neq 0$ dans $]-1, 1[\rightarrow \mathbb{R}$, on a : $f'(x) = \frac{1}{x} \operatorname{Re} \left(\sum_{n=1}^{+\infty} (x e^{2i\pi/3})^n \right)$.
[*Indication* : on rappelle que $e^{it} = \cos t + i \sin t$ pour tout $t \in \mathbb{R}$].
- 3) En déduire que, pour tout x dans $]-1, 1[$, on a : $f'(x) = -\frac{2x+1}{2(x^2+x+1)}$.
- 4) En déduire $f(x)$ pour tout x dans $]-1, 1[$.

suite au verso...

EXERCICE 3

On considère l'intervalle $I =]-1, +\infty[$ et l'application $f : I \rightarrow \mathbb{R}$ définie par $f(x) = [\ln(1+x)]^2$ pour tout $x \in I$.

- 1) Montrer que f est la solution sur I de l'équation différentielle $(1+x)y'(x) = 2\ln(1+x)$ vérifiant $y(0) = 0$.
- 2) Rechercher toutes les solutions sur I de cette équation différentielle qui sont développables en série entière (en zéro). [*Indication* : on exprimera le terme général de cette série entière en fonction de $s_n = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n}$; on utilisera aussi le développement en série entière de $\ln(1+x)$ (en zéro)].
- 3) En déduire que f est développable en série entière (en zéro). Expliciter la série entière correspondante et déterminer son rayon de convergence.
- 4) Retrouver les résultats du 3) en utilisant le carré du développement en série entière de $\ln(1+x)$.

EXERCICE 4

Soit $\alpha \in \mathbb{R}$ fixé. Pour tout entier $n \geq 1$ et tout $x \in \mathbb{R}$, on pose $f_n(x) = \frac{n^\alpha x e^{-nx}}{n^2 + 1}$.

- 1) Montrer que la série de fonctions $\sum_{n \geq 1} f_n$ converge simplement sur \mathbb{R}_+ .
- 2) Montrer qu'elle converge normalement sur tout segment inclus dans \mathbb{R}_+^* .
- 3) Pour tout $n \geq 1$, calculer $\|f_n\|_\infty = \sup_{x \in \mathbb{R}_+} |f_n(x)|$. En déduire que la série numérique $\sum_{n \geq 1} \|f_n\|_\infty$ converge si et seulement $\alpha < c$, où c est un entier fixé que l'on déterminera.
- 4) On suppose $\alpha \geq c$. Montrer que la suite $(\sum_{n=N+1}^{+\infty} f_n(\frac{1}{N}))_{N \geq 1}$ ne converge pas vers 0 dans \mathbb{R} .
- 5) On note f l'application $\mathbb{R}_+ \rightarrow \mathbb{R}$ définie par $f(x) = \sum_{n=0}^{+\infty} f_n(x)$ pour tout $x \in \mathbb{R}_+$. On introduit pour tout $N \geq 0$ les applications $\mathbb{R}_+ \rightarrow \mathbb{R}$ définies par $S_N = \sum_{n=0}^N f_n$ et $R_N = f - S_N = \sum_{n=N+1}^{+\infty} f_n$. Déduire des questions 3) et 4) que la série de fonctions $\sum_{n \geq 1} f_n$ est normalement convergente sur \mathbb{R}_+ lorsque $\alpha < c$, et non uniformément convergente sur \mathbb{R}_+ lorsque $\alpha \geq c$.
- 6) Etudier la continuité de f sur \mathbb{R}_+^* , puis sur \mathbb{R}_+ .