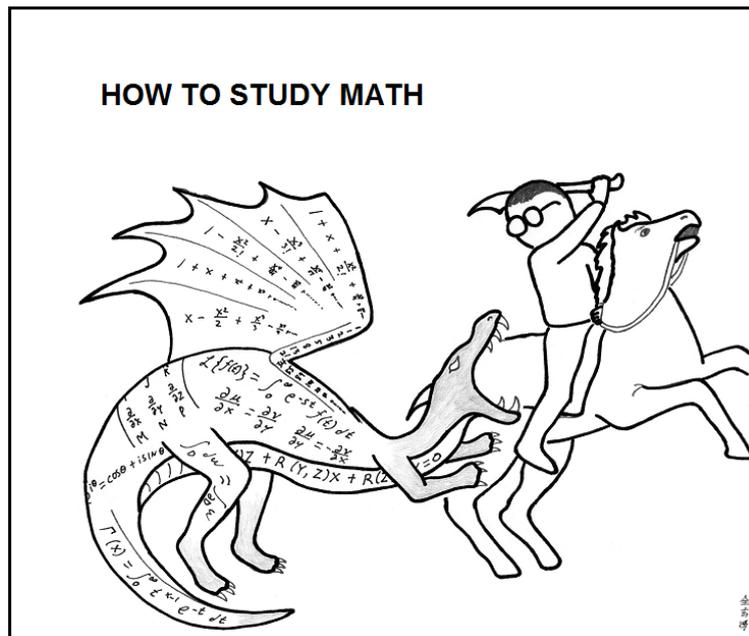


Cours de mathématiques

TSI 2^{ème} année

E. Besson



Don't just read it; fight it!

--- Paul R. Halmos

« Don't just read it; fight it! Ask your own questions, look for your own examples, discover your own proofs. Is the hypothesis necessary? Is the converse true? What happens in the classical special case? What about the degenerate cases? Where does the proof use the hypothesis? »

— Paul R. Halmos

Table des matières

Chapitre 1. Rappels et compléments d'intégration	5
1. Sommes de Riemann	5
2. Positivité et nullité	6
3. Théorèmes d'intégration	7
4. Techniques de calcul	7
Chapitre 2. Compléments d'algèbre linéaire	9
1. Familles de vecteurs et sous-espaces	9
2. Compléments sur les matrices et les applications linéaires	13
3. Déterminant	16
Chapitre 3. Variables aléatoires finies	19
1. Rappels	19
2. Couples de variables aléatoires réelles (finies)	21
3. Variables aléatoires réelles (finies) indépendantes	22
4. Covariance et coefficient de corrélation	23
Chapitre 4. Séries numériques	26
1. Séries, convergence et divergence	26
2. Séries à termes positifs	28
3. Séries absolument convergentes	29
4. Développement décimal des nombres réels	30
Chapitre 5. Fonctions à valeurs vectorielles et courbes paramétrées	32
1. Fonctions à valeurs vectorielles	32
2. Courbes paramétrées	33
Chapitre 6. Réduction des endomorphismes	39
1. Valeurs propres, vecteurs propres et polynôme caractéristique	39
2. Diagonalisation	41
3. Applications	44
Chapitre 7. Intégrale généralisée	46
1. Généralités	46
2. Intégrale généralisée des fonctions positives	47
3. Fonctions intégrables	49
Chapitre 8. Probabilités sur un univers dénombrable	50
1. Généralités	50
2. Conditionnement et indépendance	51
3. Variables aléatoires discrètes	53
4. Espérance et variance	54

5. Lois usuelles	56
Chapitre 9. Séries entières	58
1. Convergence d'une série entière	58
2. Somme d'une série entière	59
3. Fonctions développables en séries entières	60
Chapitre 10. Espaces préhilbertiens	62
1. Définitions	62
2. Norme associée à un produit scalaire	63
3. Orthogonalité	64
4. Familles orthonormales	66
5. Sous-espaces de dimension finie et orthogonalité	67
Chapitre 11. Séries de Fourier	69
1. Un espace préhilbertien	69
2. Coefficients de Fourier	71
3. Série de Fourier	72
Chapitre 12. Isométries des espaces euclidiens	74
1. Rappels	74
2. Isométries vectorielles de E	74
3. Matrices orthogonales	76
4. Diagonalisation des matrices symétriques réelles	78
5. Classification des isométries du plan et de l'espace	79
Chapitre 13. Équations différentielles	83
1. Équations différentielles linéaires du premier ordre	83
2. Équations différentielles linéaires du second ordre	84
3. Systèmes différentiels linéaires	85
Chapitre 14. Fonctions de plusieurs variables réelles	89
1. Éléments de topologie de \mathbf{R}^n	89
2. Limites et continuité	89
3. Dérivées partielles	90
4. Extrema des fonctions de plusieurs variables	91
5. Équations aux dérivées partielles	93

Rappels et compléments d'intégration

« Ça fait trente ans qu'ils font que parler d'intégration. »

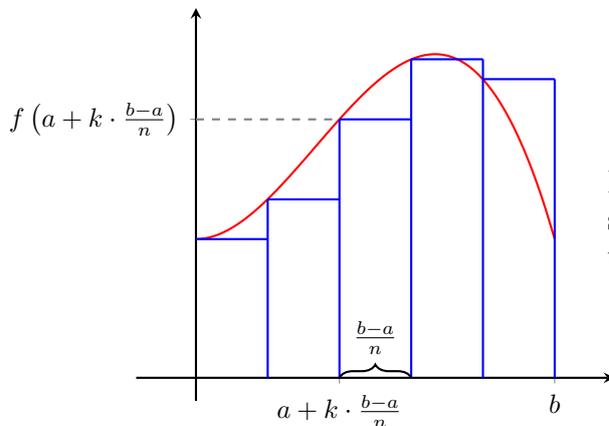
SNIPER, *Brûle* (2006)

1. Sommes de Riemann

THÉORÈME 1.1. Soit f une fonction continue sur un intervalle $[a; b]$. Alors

$$\lim_{n \rightarrow +\infty} \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k \cdot \frac{b-a}{n}\right) = \int_a^b f(t) dt$$

DÉMONSTRATION. C'est simplement la définition de ce qu'est une intégrale. \square



La n -ème somme de Riemann correspond sur le dessin à la somme des aires des rectangles bleus ($n = 5$, $k = 2$).

EXEMPLE(S). Déterminer $\lim_{n \rightarrow +\infty} \sum_{k=0}^{n-1} \sqrt{\frac{n+k}{n^3}}$.

Applications au calcul numérique d'intégrales. Pour calculer une valeur approchée d'une intégrale, on peut simplement calculer la somme de Riemann correspondante pour une valeur élevée de n (« méthode des rectangles »).

def rectangle(a, b, n):

 S = 0

 h = (b - a) / n

```

for k in range(n):
    S = S + f(a + k * h)
return S / n

```

Une version un peu plus efficace de l'algorithme (« méthode des trapèzes ») considère des rectangles dont la hauteur est la moyenne des valeurs de la fonction aux extrémités gauche et droite de la subdivision. On peut montrer que cela revient à faire le calcul par la méthode des rectangles et ajouter $\frac{f(b)-f(a)}{2}$ au résultat.

2. Positivité et nullité

PROPOSITION 2.1 (Positivité). *Soit f une fonction continue et positive sur un intervalle I . Soient $a, b \in I$ et tels que $a \leq b$.*

$$\int_a^b f(t) dt \geq 0.$$

DÉMONSTRATION. L'intégrale est une limite de sommes de Riemann qui sont toutes positives. \square

REMARQUE. L'hypothèse $a \leq b$ est fondamentale! Par exemple $\int_1^0 1 dt = -1$.

PROPOSITION 2.2 (Croissance). *Soient f et g deux fonctions continues sur un intervalle I . On suppose que $f \leq g$ (autrement dit, pour tout $x \in I$, $f(x) \leq g(x)$). Alors, si $a, b \in I$ vérifient $a \leq b$,*

$$\int_a^b f(t) dt \leq \int_a^b g(t) dt.$$

DÉMONSTRATION. On dit que $g = f + (g - f)$ et on applique la proposition précédente à la fonction positive $g - f$. \square

EXEMPLE(S). Montrer que, pour tout $n \in \mathbf{N}^*$,

$$\frac{1}{n+1} \leq \int_n^{n+1} \frac{dt}{t} \leq \frac{1}{n}.$$

Il est évident que l'intégrale d'une fonction constamment nulle est nulle. La réciproque n'est pas toujours vraie, mais on peut énoncer la proposition suivante :

PROPOSITION 2.3 (« aux quatre hypothèses »). *Soit f une fonction définie sur un intervalle I . Soient $a, b \in I$. On suppose que :*

- (1) $a < b$;
- (2) f est continue sur $[a; b]$;
- (3) f est positive sur $[a; b]$;
- (4) $\int_a^b f(t) dt = 0$.

Alors f est constamment nulle sur $[a; b]$ (autrement dit : pour tout $x \in [a; b]$, $f(x) = 0$).

Attention : le résultat cesse d'être vrai dès qu'au moins une des hypothèses n'est pas vérifiée (exercice : trouver un contre-exemple pour chaque hypothèse).

EXEMPLE(S). Soit P un polynôme réel. On suppose que $\int_0^1 P^2(t) dt = 0$. Montrer que P est le polynôme nul.

3. Théorèmes d'intégration

THÉORÈME 3.1 (Théorème fondamental de l'analyse). Soit f une fonction continue sur un intervalle I . Alors :

- (1) Pour tout $a \in I$, la fonction $F_a: x \mapsto \int_a^x f(t) dt$ est une primitive de f sur I (plus précisément, c'est l'unique primitive de f s'annulant en a).
- (2) Si F est une primitive de f sur I , alors pour tous $a, b \in I$,

$$\int_a^b f(t) dt = [F(t)]_a^b = F(b) - F(a).$$

- REMARQUE. • En particulier, si F est une primitive de f sur I , les primitives de f sur I sont exactement les fonctions $F + c$ pour tout $c \in \mathbf{R}$.
- En conséquence de ce théorème, on utilise la notation $\int f(t) dt$ (sans bornes) pour désigner une primitive « générique » de la fonction f .

THÉORÈME 3.2 (Intégration par parties). Soient f une fonction continue et g une fonction de classe \mathcal{C}^1 sur un intervalle $[a; b]$. Alors

$$\int_a^b \overbrace{f(t)} \underbrace{g(t)} dt = [F(t)g(t)]_a^b - \int_a^b F(t)g'(t) dt.$$

- EXEMPLE(S). • Déterminer une primitive de la fonction \ln sur $]0; +\infty[$.
- Calculer $\int_0^\pi t^2 \sin t dt$.

THÉORÈME 3.3 (Changement de variables). Soit f une fonction continue sur un intervalle I . Soit φ une fonction de classe \mathcal{C}^1 sur un intervalle J et à valeurs dans I . Soient $a, b \in J$. Alors

$$\int_a^b f(\varphi(t))\varphi'(t) dt = \int_{\varphi(a)}^{\varphi(b)} f(u) du$$

REMARQUE. En pratique, l'abus de notation « $u = \varphi(t)$, $du = \varphi'(t) dt$ » est courant.

- EXEMPLE(S). • Calculer $\int_0^1 \sqrt{1-t^2} dt$.
- Déterminer une primitive de $f: t \mapsto \frac{e^t}{1+e^{2t}}$ sur \mathbf{R} .

4. Techniques de calcul

4.1. Fractions rationnelles. Pour calculer les intégrales de fonctions de la forme $\frac{P}{Q}$, où P et Q sont deux polynômes, on peut utiliser une technique appelée *décomposition en éléments simples*. En substance, il s'agit de décomposer le dénominateur Q en produit

de polynôme irréductibles sur \mathbf{R} . On peut alors montrer que La fraction $\frac{P}{Q}$ s'écrit comme une somme des « éléments simples » suivants :

- $\frac{a_i}{Q_i(t)}$ avec a_i un coefficient réel à déterminer et Q_i un facteur irréductible de degré 1 de Q ;
- $\frac{a_i t + b_i}{Q_i(t)}$ avec a_i et b_i deux coefficients réels à déterminer et Q_i un facteur irréductible de degré 2 de Q .

EXEMPLE(S). Décomposer en éléments simples les expressions suivantes, puis en déduire leurs primitives (en faisant attention aux intervalles de définition) :

$$F_1(t) = \frac{1}{t^2 - 1}; \quad F_2(t) = \frac{t + 1}{t^2 + 1}; \quad F_3(t) = \frac{t + 1}{t^2 - 3t + 2}; \quad F_4(t) = \frac{1}{t^3 - 1}$$

4.2. Expressions faisant intervenir les fonctions trigonométriques. Pour un produit ou une puissance de fonctions trigonométriques, on peut procéder par *linéarisation* en appliquant les formules d'Euler, puis en développant (à l'aide potentiellement du binôme de Newton) et enfin en simplifiant à l'aide des formules d'Euler « à l'envers ».

EXEMPLE(S). Déterminer une primitive de la fonction \sin^3 par linéarisation.

Pour une fraction de fonctions trigonométriques, on peut utiliser les règles « de Bioche » suivantes :

- si $f(t) dt$ est invariante par la transformation $t \mapsto -t$, poser $u = \cos t$;
- si $f(t) dt$ est invariante par la transformation $t \mapsto \pi - t$, poser $u = \sin t$;
- si $f(t) dt$ est invariante par la transformation $t \mapsto \pi + t$, poser $u = \tan t$.

Si on n'est dans aucun de ces cas, on pose $u = \tan \frac{t}{2}$. En effet, dans ce cas, on peut écrire

$$\cos t = \frac{1 - u^2}{1 + u^2}, \quad \sin t = \frac{2u}{1 + u^2} \quad \text{et} \quad \tan t = \frac{1 + u^2}{1 - u^2}.$$

EXEMPLE(S). Déterminer une primitive de la fonction \sin^3 à l'aide des règles de Bioche.

EXEMPLE(S). Calculer :

- $\int_0^{\pi/2} \frac{\cos^3 t}{1 + \sin t} dt$;
- $\int_{\pi/3}^{\pi/2} \frac{dt}{\sin t}$;
- $\int_0^{\pi/2} \frac{dt}{2 + \sin t}$.

Compléments d'algèbre linéaire

« J'suis déter, moi. »
« Alors je trace. »

JUL, *Davai Davai* (2020)

Dans tout le chapitre, \mathbf{K} désigne \mathbf{R} ou \mathbf{C} ; E est un \mathbf{K} -espace vectoriel.

1. Familles de vecteurs et sous-espaces

1.1. Familles de vecteurs.

DÉFINITION 1.1. Soit $\mathcal{V} = (v_i)_{i \in I}$ une famille (pas nécessairement finie) de vecteurs de E . On appelle *combinaison linéaire* de vecteurs de \mathcal{V} tout vecteur de la forme

$$v = \lambda_1 v_{i_1} + \cdots + \lambda_n v_{i_n},$$

où i_1, \dots, i_n sont un nombre fini d'éléments de I deux à deux distincts et $\lambda_1, \dots, \lambda_n \in \mathbf{K}$.

Si $v = 0$, on parle de combinaison linéaire *nulle*. Si tous les scalaires λ_i sont nuls, on parle de combinaison linéaire *triviale*.

EXEMPLE(S). Dans $E = \mathbf{K}[X]$, si on considère la famille $\mathcal{V} = (1, X, X^2, \dots)$, alors une combinaison linéaire de vecteurs de \mathcal{V} est un polynôme.

REMARQUE. Toute combinaison triviale est nulle, mais la réciproque n'est pas vraie. Par exemple, dans $\mathbf{K}[X]$, si $v_1 = X$ et $v_2 = -X$, la combinaison linéaire $v_1 + v_2$ est nulle mais non triviale.

DÉFINITION 1.2. Soit $\mathcal{V} = (v_i)_{i \in I}$ une famille de vecteurs de E . On appelle *sous-espace vectoriel engendré* par \mathcal{V} l'ensemble des combinaisons linéaires de vecteurs de \mathcal{V} . On le note $\text{Vect}(\mathcal{V})$.

PROPOSITION 1.3. *Si \mathcal{V} est une famille de vecteurs de E , alors $\text{Vect}(\mathcal{V})$ est le plus petit sous-espace vectoriel de E contenant \mathcal{V} . En d'autres termes :*

- $\text{Vect}(\mathcal{V})$ est un sous-espace vectoriel de E ;
- si F est un sous-espace vectoriel de E contenant tous les vecteurs de \mathcal{V} , alors $\text{Vect}(\mathcal{V}) \subset F$.

DÉMONSTRATION. • Commençons par montrer que $\text{Vect}(\mathcal{V})$ est un sev de E .

Le vecteur nul est une combinaison linéaire de \mathcal{V} (prendre par exemple une combinaison linéaire triviale). De plus, si u et v sont deux combinaisons linéaires de \mathcal{V} et $\lambda, \mu \in \mathbf{K}$, alors $\lambda u + \mu v$ est une combinaison linéaire de \mathcal{V} .

- Soit F un sev de E contenant tous les éléments de \mathcal{V} . Étant un sev, F est stable par combinaisons linéaires : toute combinaison linéaire d'éléments de F appartient à F . En particulier (puisque $\mathcal{V} \subset F$), toute combinaison linéaire d'éléments de \mathcal{V} appartient à F ; autrement dit, $\text{Vect}(\mathcal{V}) \subset F$. \square

DÉFINITION 1.4. Soit $\mathcal{V} = (v_i)_{i \in I}$ une famille de vecteurs de E . La famille \mathcal{V} est dite :

- *libre* si toute combinaison linéaire nulle de vecteurs de \mathcal{V} est triviale.
- *génératrice* si tout élément de E peut s'écrire comme combinaison linéaire de vecteurs de \mathcal{V} , autrement dit si $\text{Vect}(\mathcal{V}) = E$.

Si \mathcal{V} est à la fois libre et génératrice, on dit que c'est une *base* de E .

REMARQUE. \mathcal{V} est une base de E si et seulement si tout vecteur de E peut s'écrire de manière unique comme combinaison linéaire de vecteurs de \mathcal{V} .

EXEMPLE(S). Dans $E = \mathbf{R}_2[X]$, la famille $X^2, X + 1$ est libre, mais elle n'est pas génératrice. La famille $1, X, X^2, 1 + X + X^2$ est génératrice, mais pas libre.

REMARQUE. On rappelle qu'un espace vectoriel est de dimension finie ssi il admet une base finie ; dans ce cas, sa dimension est le cardinal d'une de ses bases. Si E est de dimension finie égale à n , toute famille libre de E est de cardinal au plus n ; toute famille génératrice de E est de cardinal au moins n . On peut obtenir une base de E en complétant une famille libre de E , ou en extrayant des vecteurs d'une famille génératrice de E .

PROPOSITION 1.5. Dans $E = \mathbf{K}[X]$ on considère une famille $\mathcal{V} = (P_n)_{n \in I}$ de polynômes. On suppose la famille échelonnée en degré, c'est-à-dire que, pour tous $i, j \in I$, si $i \neq j$, alors $\deg(P_i) \neq \deg(P_j)$. Alors \mathcal{V} est libre.

DÉMONSTRATION. Soit $0 = \lambda_1 P_{i_1} + \dots + \lambda_n P_{i_n}$ une combinaison linéaire nulle de \mathcal{V} ($\lambda_1, \dots, \lambda_n \in \mathbf{K}$).

Par l'absurde : si cette combinaison linéaire n'est pas triviale, considérons l'ensemble des coefficients λ_i non nuls, et choisissons parmi ceux-ci celui tel que le polynôme P_i soit de degré maximal (i est unique en vertu de la propriété d'échelonnement). Notons d_i ce degré. Alors le terme en X^{d_i} dans la combinaison linéaire provient uniquement de $\lambda_i P_i$, puisque tous les autres polynômes ayant des coefficients λ non nuls sont de degré strictement inférieur. Mais, la combinaison linéaire devant être nulle, cela implique qu'on ait $\lambda_i = 0$, ce qui est contradictoire.

La combinaison linéaire est donc triviale. \square

PROPOSITION 1.6. La famille $(1, X, X^2, \dots)$ est une base de $\mathbf{K}[X]$.

DÉMONSTRATION. Par définition de ce qu'est un polynôme, $\text{Vect}(1, X, X^2, \dots) = \mathbf{K}[X]$. La famille $(1, X, \dots)$ est donc génératrice. De plus, c'est une famille échelonnée en degré, elle est donc libre. \square

1.2. Sous-espaces vectoriels.

DÉFINITION 1.7. Soient F_1, \dots, F_p des sous-espaces vectoriels de E . On appelle *somme des F_i* l'ensemble

$$\sum_{i=1}^p F_i = \{v_1 + \dots + v_p \mid v_1 \in F_1, \dots, v_p \in F_p\}.$$

REMARQUE. Autrement dit, un vecteur $v \in E$ appartient à $\sum F_i$ si et seulement s'il existe $v_1 \in F_1, \dots, v_p \in F_p$ tels que $v = v_1 + \dots + v_p$.

PROPOSITION 1.8. Si F_1, \dots, F_p sont des sev de E , alors $\sum_{i=1}^p F_i$ est le plus petit sous-espace vectoriel de E contenant tous les F_i .

ADMIS. □

DÉFINITION 1.9. Soient F_1, \dots, F_p des sous-espaces vectoriels de E . On dit que les F_i sont *en somme directe* si, pour tout $v \in \sum_{i=1}^p F_i$, l'écriture $v = v_1 + \dots + v_p$ avec $v_1 \in F_1, \dots, v_p \in F_p$ est unique. Dans ce cas, on note

$$\bigoplus_{i=1}^p F_i = \sum_{i=1}^p F_i.$$

REMARQUE. Si $p = 2$, on la caractérisation utile : F_1 et F_2 sont en somme directe ssi $F_1 \cap F_2 = \{0\}$. Cette propriété cesse d'être vraie pour $p \geq 3$.

EXEMPLE(S). Dans $E = \mathbf{R}^2$, on considère $F_1 = \mathbf{R}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $F_2 = \mathbf{R}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $F_3 = \mathbf{R}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Vérifier que F_1, F_2 et F_3 sont deux à deux en somme directe, mais que la somme $F_1 + F_2 + F_3$ n'est pas directe.

PROPOSITION 1.10. Soient F_1, \dots, F_p des sous-espaces vectoriels de E . Alors les espaces F_1, \dots, F_p sont en somme directe si et seulement si, pour tous $v_1 \in F_1, \dots, v_p \in F_p$, si $v_1 + \dots + v_p = 0$, alors $v_1 = \dots = v_p = 0$.

REMARQUE. Autrement dit, il suffit de vérifier que la décomposition du vecteur nul sur la somme $F_1 + \dots + F_p$ est unique.

DÉMONSTRATION. Si la somme est directe, la décomposition de tout vecteur de E sur la somme $F_1 + \dots + F_p$ est unique, donc en particulier celle du vecteur nul.

Réciproquement, on suppose que l'unique décomposition du vecteur nul sur $F_1 + \dots + F_p$ est $0 + \dots + 0$. Soit $v \in F_1 + \dots + F_p$. Supposons que v admette deux décompositions : $v = v_1 + \dots + v_p = v'_1 + \dots + v'_p$, avec, pour tout $i \in \llbracket 1; p \rrbracket$, $v_i, v'_i \in F_i$. Alors, en soustrayant les deux décompositions, on obtient

$$0 = \underbrace{v_1 - v'_1}_{\in F_1} + \dots + \underbrace{v_p - v'_p}_{\in F_p}.$$

C'est une décomposition du vecteur nul sur $F_1 + \dots + F_p$, donc par unicité, tous les $v_i - v'_i$ sont nuls. Autrement dit la décomposition de v sur $F_1 + \dots + F_p$ est unique. □

EXEMPLE(S). Dans $E = \mathbf{R}[X]$, montrer que $F_1 = \text{Vect}(1, X)$, $F_2 = \text{Vect}(X^2, X^3)$ et $F_3 = \text{Vect}(X^4, X^5)$ sont en somme directe.

PROPOSITION 1.11. Soient F_1, \dots, F_p des sous-espaces vectoriels de E . On suppose qu'on peut écrire $E = F_1 \oplus \dots \oplus F_p$. Soit, pour tout $i \in \llbracket 1; p \rrbracket$, \mathcal{B}_i une base de F_i . Alors la famille \mathcal{B} obtenue par concaténation des bases \mathcal{B}_i est une base de E , dite adaptée à la décomposition $E = F_1 \oplus \dots \oplus F_p$.

1.3. Hyperplans. Dans tout ce paragraphe, E est de dimension finie $n \in \mathbf{N}^*$.

DÉFINITION 1.12. On appelle *hyperplan* de E tout sous-espace vectoriel de E de dimension $n - 1$.

EXEMPLE(S). Une droite de \mathbf{R}^2 , un plan de \mathbf{R}^3 sont des hyperplans. $\text{Vect}(X, X^2)$ est un hyperplan de $\mathbf{R}_2[X]$. Si E est une droite, l'unique hyperplan de E est $\{0\}$ (mais ce n'est pas très intéressant).

PROPOSITION 1.13. Un sous-espace vectoriel H de E est un hyperplan si et seulement si F admet une droite pour supplémentaire. En d'autres termes, s'il existe G un sous-espace vectoriel de E de dimension 1 tel que $E = H \oplus G$.

DÉMONSTRATION. Si H admet une droite G pour supplémentaire, alors d'après la formule de Grassmann, $\dim E = \dim F + \dim G$, donc $\dim F = \dim E - 1$ et F est donc un hyperplan.

Réciproquement, si H est un hyperplan. L'espace E étant de dimension finie, F admet nécessairement un supplémentaire G . D'après la formule de Grassmann, $\dim E = \dim H + \dim G$ donc $\dim G = \dim E - \dim H = \dim E - (\dim E - 1) = 1$; l'espace G est bien une droite. \square

PROPOSITION 1.14. Soit H un sous-espace vectoriel de E . Alors H est un hyperplan si et seulement s'il existe une forme linéaire $\varphi: E \rightarrow \mathbf{K}$ non nulle telle que $\text{Ker } \varphi = H$. Autrement dit, tel que $H = \{v \in E \mid \varphi(v) = 0\}$. L'équation $\varphi(v) = 0$ est appelée (une) équation de l'hyperplan H .

REMARQUE. On pense évidemment à l'équation cartésienne d'une droite du plan, ou d'un plan de l'espace. Attention : ce ne sont des sous-espaces vectoriels que s'ils passent par l'origine !

DÉMONSTRATION. Si φ est une forme linéaire non nulle, $\text{Im } \varphi$ est un sous-espace vectoriel non nul de \mathbf{K} , donc de dimension 1. D'après le théorème du rang, on a donc $\dim \text{Ker } \varphi = \dim E - 1$, donc $\text{Ker } \varphi$ est un hyperplan de E .

Réciproquement, soit H un hyperplan de E . D'après la proposition 1.14, H admet pour supplémentaire une droite G . Soit e un vecteur non nul de G , la famille (e) est alors une base de G . Autrement dit, pour tout $v \in G$, il existe un unique $\lambda \in \mathbf{K}$ tel que $v = \lambda e$, et l'application coordonnée $\psi: v \mapsto \lambda$ est une application linéaire de G dans \mathbf{K} . Posons alors p la projection sur G parallèlement à H ; l'application $\varphi = \psi \circ p$ répond à la question. \square

EXEMPLE(S). Dans $E = \mathbf{R}^3$, vérifier que $H = \text{Vect} \left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right)$ est un hyperplan et en donner une équation.

2. Compléments sur les matrices et les applications linéaires

2.1. Sous-espaces stables d'un endomorphisme.

DÉFINITION 2.1. Soit f un endomorphisme de E . Soit F un sous-espace vectoriel de E . On dit que F est *stable* par f si $f(F) \subset F$. Autrement dit si, pour tout $v \in F$, $f(v) \in F$.

EXEMPLE(S). (1) les sous-espaces « triviaux » $F = E$ et $F = \{0\}$ sont stables par tout endomorphisme.

(2) Dans $E = \mathbf{R}[X]$, on considère l'endomorphisme $D: P \mapsto P'$. Alors, pour tout $n \in \mathbf{N}$, l'espace $F = \mathbf{R}_n[X]$ est stable par D . En revanche, l'espace $G = \{P \in \mathbf{R}[X] \mid P(0) = 0\}$ n'est pas stable par D .

(3) si E est de dimension finie et que F et G sont deux sev supplémentaires de E , notons p la projection et s la symétrie sur F parallèlement à G . Alors les espaces F et G sont stables par s et p .

PROPOSITION 2.2. On suppose E de dimension finie. Soient f un endomorphisme de E , F un sous-espace vectoriel de E stable par f , et G un supplémentaire de F . Enfin, soit \mathcal{B} une base adaptée à la décomposition $E = F \oplus G$. Alors la matrice de f dans la base \mathcal{B} est de la forme « par blocs »

$$\text{Mat}_{\mathcal{B}} f = \left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array} \right),$$

où A, B, C sont des matrices.

DÉMONSTRATION. Si p est la dimension de F et $\mathcal{B} = (e_1, \dots, e_n)$, alors (e_1, \dots, e_p) est une base de F et (e_{p+1}, \dots, e_n) une base de G . Mais F étant stable par f , alors $f(e_1), \dots, f(e_p)$ sont tous dans F ; autrement dit, $f(e_1), \dots, f(e_p) \in \text{Vect}(e_1, \dots, e_p)$. Ceci explique que, dans les p premières colonnes de la matrice, les lignes $p+1$ à n sont nulles. \square

REMARQUE. Réciproquement, une matrice « par blocs » de la forme ci-dessus peut permettre de détecter un sous-espace stable.

EXEMPLE(S). Dans $E = \mathbf{R}^3$, on considère l'application $f: (x, y, z) \mapsto (y+z, x+z, z)$. Écrire la matrice de f dans la base canonique de \mathbf{R}^3 et en déduire un sous-espace stable par f .

2.2. Transposition.

DÉFINITION 2.3. Soit $A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} \in \mathcal{M}_{m,n}(\mathbf{K})$. On appelle *transposée* de A la matrice

$$A^T = (a_{j,i})_{1 \leq j \leq n, 1 \leq i \leq m} \in \mathcal{M}_{n,m}(\mathbf{K}).$$

Concrètement, c'est la matrice obtenue en transformant chaque ligne de A en colonne, et réciproquement.

EXEMPLE(S). $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$.

Si X, Y sont deux vecteurs colonnes de \mathbf{R}^n , alors $X^T Y$ est le produit scalaire de X et Y . En particulier, $X^T X = \|X\|^2$.

REMARQUE. On utilise parfois la notation ${}^t A$.

PROPOSITION 2.4. *La transposition $A \mapsto A^T$ définit une application linéaire de $\mathcal{M}_{m,n}(\mathbf{K})$ dans $\mathcal{M}_{n,m}(\mathbf{K})$. Autrement dit, si $A, B \in \mathcal{M}_{m,n}(\mathbf{K})$ et $\lambda, \mu \in \mathbf{K}$, alors $(\lambda A + \mu B)^T = \lambda A^T + \mu B^T$.*

PROPOSITION 2.5. *Soient $A \in \mathcal{M}_{m,n}(\mathbf{K})$ et $B \in \mathcal{M}_{n,p}(\mathbf{K})$. Alors $(AB)^T = B^T A^T$.*

REMARQUE. Le changement dans l'ordre des facteurs entre A et B est en quelque sorte « obligatoire » si on veut que les dimensions des matrices soient compatibles avec le produit matriciel.

EXEMPLE(S). Si $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & -6 \\ 3 & -4 \\ 5 & -2 \end{pmatrix}$,

$$AB = \begin{pmatrix} 1+6+15 & -6-8-6 \\ 4+15+30 & -24-20-12 \end{pmatrix} = \begin{pmatrix} 22 & -20 \\ 49 & -56 \end{pmatrix} \quad \text{donc} \quad (AB)^T = \begin{pmatrix} 22 & 49 \\ -20 & -56 \end{pmatrix}.$$

D'autre part, $B^T = \begin{pmatrix} 1 & 3 & 5 \\ -6 & -4 & -2 \end{pmatrix}$ et $A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$ donc

$$B^T A^T = \begin{pmatrix} 1+6+15 & 4+15+30 \\ -6-8-6 & -24-20-12 \end{pmatrix} = (AB)^T.$$

PROPOSITION 2.6. *Soit $A \in \mathcal{M}_n(\mathbf{K})$. Alors A est inversible si et seulement si A^T est inversible ; dans ce cas, on a $(A^T)^{-1} = (A^{-1})^T$.*

DÉMONSTRATION. Si A est inversible, alors $AA^{-1} = I_n$ donc $(AA^{-1})^T = I_n^T = I_n$, autrement dit $(A^{-1})^T A^T = I_n$, ce qui montre que A^T est inversible, d'inverse $(A^{-1})^T$. De même pour la réciproque. \square

DÉFINITION 2.7. Une matrice carrée $A \in \mathcal{M}_n(\mathbf{K})$ est dite :

- *symétrique* si $A^T = A$;
- *antisymétrique* si $A^T = -A$.

EXEMPLE(S). La matrice $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 7 \end{pmatrix}$ est symétrique, la matrice $\begin{pmatrix} 0 & 2 & 3 \\ -2 & 0 & 5 \\ -3 & -5 & 0 \end{pmatrix}$ est antisymétrique.

REMARQUE. La diagonale d'une matrice antisymétrique est nécessairement nulle.

2.3. Trace.

DÉFINITION 2.8. Soit $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbf{K})$. On appelle *trace* de A le nombre

$$\mathrm{Tr}(A) = \sum_{i=1}^n a_{i,i} = a_{1,1} + a_{2,2} + \cdots + a_{n,n} \in \mathbf{K}.$$

Concrètement, la trace de A est la somme de ses coefficients diagonaux.

REMARQUE. Une matrice non carrée n'a pas de diagonale, donc pas de trace!

EXEMPLE(S). $\mathrm{Tr} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = 1 + 5 + 9 = 15.$

PROPOSITION 2.9. *L'application Tr est une forme linéaire sur $\mathcal{M}_n(\mathbf{K})$. Autrement dit, si $A, B \in \mathcal{M}_n(\mathbf{K})$ et $\lambda, \mu \in \mathbf{K}$, alors $\mathrm{Tr}(\lambda A + \mu B) = \lambda \mathrm{Tr}(A) + \mu \mathrm{Tr}(B)$.*

REMARQUE. En particulier, l'ensemble des matrices de trace nulle est un hyperplan de $\mathcal{M}_n(\mathbf{K})$.

DÉMONSTRATION. Les coefficients diagonaux de $\lambda A + \mu B$ sont donné par la combinaison linéaire correspondante des coefficients diagonaux de A et de B . \square

REMARQUE. Pièges à éviter :

- la trace de la matrice identité n'est pas égale à 1, mais à n .
- Il n'est pas vrai en général que $\mathrm{Tr}(AB) = \mathrm{Tr}(A) \mathrm{Tr}(B)$. Par exemple, si on prend $A = B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, on a $\mathrm{Tr}(A) = \mathrm{Tr}(B) = 0$, pourtant $\mathrm{Tr}(AB) = \mathrm{Tr}(A^2) = \mathrm{Tr}(I_2) = 2$.

PROPOSITION 2.10. *Soient $A \in \mathcal{M}_{m,n}(\mathbf{K})$ et $B \in \mathcal{M}_{n,m}(\mathbf{K})$. Alors $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.*

EXEMPLE(S). Si $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & -6 \\ 3 & -4 \\ 5 & -2 \end{pmatrix}$, on peut écrire

$$AB = \begin{pmatrix} 1 + 6 + 15 & -6 - 8 - 6 \\ 4 + 15 + 30 & -24 - 20 - 12 \end{pmatrix}$$

donc $\mathrm{Tr}(AB) = 1 + 6 + 15 - 24 - 20 - 12$.

Par ailleurs, $BA = \begin{pmatrix} 1 - 24 & 2 - 30 & 3 - 36 \\ 3 - 16 & 6 - 20 & 9 - 24 \\ 5 - 8 & 10 - 10 & 15 - 12 \end{pmatrix}$ donc $\mathrm{Tr}(BA) = 1 - 24 + 6 - 20 + 15 - 12 = \mathrm{Tr}(AB)$.

REMARQUE. Pour revenir au deuxième « piège » cité plus haut, on se rend compte dans cet exemple qu'il est même possible de parler de $\mathrm{Tr}(AB)$ sans même que $\mathrm{Tr}(A)$ ou $\mathrm{Tr}(B)$ n'existe (puisque A et B ne sont pas carrées).

Une conséquence de la proposition 2.10 est que si $A \in \mathcal{M}_n(\mathbf{K})$ et $P \in \mathrm{GL}_n(\mathbf{K})$, alors $\mathrm{Tr}(P^{-1}AP) = \mathrm{Tr}(PP^{-1}A) = \mathrm{Tr}(A)$. Autrement dit, la trace est invariante par changement de base.

DÉFINITION 2.11. Soit $f \in \mathcal{L}(E)$. On appelle *trace* de f , notée $\text{Tr}(f)$, la trace de la matrice représentant l'application f dans n'importe quelle base de E .

EXEMPLE(S). On considère l'endomorphisme f de $\mathbf{R}_2[X]$ qui, à un polynôme P , associe sa tangente au point d'abscisse 1. Autrement dit, pour tout $P \in \mathbf{R}_2[X]$, $f(P) = (X - 1)P'(1) + P(1)$. Déterminer la matrice de f dans la base canonique, puis dans la base $(1, X - 1, (X - 1)^2)$. Comparer les traces des deux matrices.

3. Déterminant

3.1. Déterminant d'une matrice carrée. En dimension 2 et 3, on dispose d'outils géométriques pratiques pour déterminer si une famille de vecteurs est libre : le déterminant en dimension 2, le produit mixte en dimension 3. L'objectif de cette partie est de généraliser cet outil à toute dimension.

On souhaite construire une application $\mathcal{M}_n \rightarrow \mathbf{R}$ vérifiant les trois propriétés suivantes :

- (1) linéaire par rapport à chacune des colonnes ;
- (2) nulle si au moins deux colonnes sont égales ;
- (3) égale à 1 sur la matrice identité.

THÉORÈME 3.1. *Il existe une unique application $\mathcal{M}_n \rightarrow \mathbf{K}$ vérifiant les trois propriétés ci-dessus. On l'appelle déterminant, noté \det .*

ADMIS. □

EXEMPLE(S). Soient $a, b, c \in \mathbf{K}$. Calculer $\begin{vmatrix} a & b \\ 0 & c \end{vmatrix} = \det \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$.

REMARQUE. Plus généralement, si A est une matrice triangulaire (notamment, si A est diagonale), alors $\det A$ est le produit des coefficients diagonaux. En particulier, $\det(\lambda I_n) = \lambda^n$ (et non λ).

3.2. Calcul du déterminant.

PROPOSITION 3.2. *Soit $A \in \mathcal{M}_n(\mathbf{K})$. Soit $B \in \mathcal{M}_n(\mathbf{K})$, obtenue en échangeant deux colonnes de A . Alors $\det B = -\det A$.*

DÉMONSTRATION. Notons C_i et C_j ($i < j$) les deux colonnes échangées entre A et B . On note A_1, A_2 et A_3 les « blocs » de A situés respectivement à gauche de C_i , entre C_i et C_j , et à droite de C_j . Ces blocs sont présents aux mêmes positions dans A et dans B . Ainsi, on peut écrire, par blocs, $A = (A_1|C_i|A_2|C_j|A_3)$ et $B = (A_1|C_j|A_2|C_i|A_3)$. Mais alors, en écrivant $C_i = C_i - C_j + C_j$ et $C_j = C_j - C_i + C_i$, on a, par linéarité par

rapport aux colonnes i et j :

$$\begin{aligned}
 \det B &= |A_1|C_j - C_i + C_i|A_2|C_i - C_j + C_j|A_3| \\
 &= |A_1|C_j - C_i|A_2|C_i - C_j + C_j|A_3| + |A_1|C_i|A_2|C_i - C_j + C_j|A_3| \\
 &= \underbrace{|A_1|C_j - C_i|A_2|C_i - C_j|A_3|}_{=0} + \underbrace{|A_1|C_i|A_2|C_i - C_j|A_3|}_{=0} = 0 - \det A \\
 &\quad + \underbrace{|A_1|C_j - C_i|A_2|C_j|A_3|}_{=0-\det A} + \underbrace{|A_1|C_i|A_2|C_j|A_3|}_{=\det A} \\
 &= -\det A. \quad \square
 \end{aligned}$$

COROLLAIRE 3.3. *Les opérations élémentaires sur les colonnes d'une matrice effectuent son déterminant de la façon suivante :*

- $C_i \leftrightarrow C_j$: le déterminant est multiplié par -1 ;
- $C_i \leftarrow C_i + \lambda C_j$ ($i \neq j$) : le déterminant est inchangé ;
- $C_i \leftarrow \lambda C_i$: le déterminant est multiplié par λ .

DÉMONSTRATION. Pour la première opération, c'est exactement la proposition 3.2. Pour la troisième opération, c'est la propriété de linéarité par rapport à la i -ème colonne.

Pour la deuxième opération, par linéarité par rapport à la i -ème colonne, on obtient le déterminant initial, plus λ fois le déterminant où la i -ème colonne est remplacée par la j -ème colonne, qui est nul puisqu'il admet deux colonnes identiques. \square

PROPOSITION 3.4. *Pour tout $A \in \mathcal{M}_n(\mathbf{K})$, $\det A^T = \det A$.*

ADMIS. \square

REMARQUE. En particulier, les opérations élémentaires mentionnées ci-dessus fonctionnent de la même façon sur les lignes.

EXEMPLE(S). Calculer $\det \begin{pmatrix} -2 & 2 & -3 \\ -1 & 1 & 3 \\ 2 & 0 & -1 \end{pmatrix}$.

PROPOSITION 3.5. *Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbf{K})$. Alors $\det A = ad - bc$.*

Soit $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix} \in \mathcal{M}_3(\mathbf{R})$. Alors $\det A = aej + bfg + cdh - gec - ahf - dbj$ (règle de Sarrus).

DÉMONSTRATION. Dans le cas de la dimension 2, par linéarité, on obtient :

$$\det A = ad \underbrace{\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}}_{=1} + ab \underbrace{\begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix}}_{=0} + cd \underbrace{\begin{vmatrix} 0 & 0 \\ 1 & 1 \end{vmatrix}}_{=0} + bc \underbrace{\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}}_{=-1} = ad - bc.$$

De même dans le cas de la dimension 3 (en exercice pour les motivé-es). \square

PROPOSITION 3.6 (Développement par rapport à une colonne). Soit $A = (a_{i,j})_{1 \leq i,j \leq n}$. Pour tous $i, j \in \llbracket 1; n \rrbracket$, on note $A_{i,j}$ la matrice de taille $n - 1$ obtenue en supprimant la i -ème ligne et la j -ème colonne de A . Alors, pour tout $j \in \llbracket 1; n \rrbracket$,

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{i,j} \det A_{i,j}$$

REMARQUE. On peut développer de même suivant une ligne.

ADMIS. □

EXEMPLE(S). Calculer $\det \begin{pmatrix} 1 & 4 & 5 \\ 2 & 9 & 8 \\ 3 & 7 & 6 \end{pmatrix}$. Calculer $\det \begin{pmatrix} 2 & 4 & 0 & 6 \\ -1 & 0 & 7 & 0 \\ 3 & 7 & 0 & 0 \\ 4 & -2 & 0 & 5 \end{pmatrix}$.

3.3. Propriétés du déterminant.

PROPOSITION 3.7. Soient $A, B \in \mathcal{M}_n(\mathbf{K})$. Alors $\det(AB) = \det A \det B$.

ADMIS. □

PROPOSITION 3.8. Soit $A \in \mathcal{M}_n(\mathbf{K})$. Alors A est inversible si et seulement si $\det A \neq 0$. Dans ce cas, $\det(A^{-1}) = \frac{1}{\det A}$.

ADMIS. □

COROLLAIRE 3.9. Si $A \in \mathcal{M}_n(\mathbf{K})$ et $P \in \text{GL}_n(\mathbf{K})$, alors $\det(P^{-1}AP) = \det A$.

DÉFINITION 3.10. Soit $f \in \mathcal{L}(E)$. On appelle *déterminant* de f , noté $\det f$, le déterminant de la matrice de f dans n'importe quelle base de E .

Toutes les propriétés des déterminants des matrices se traduisent immédiatement en termes d'endomorphismes :

PROPOSITION 3.11. Soient $f, g \in \mathcal{L}(E)$. Alors $\det(g \circ f) = \det g \det f$.

PROPOSITION 3.12. L'endomorphisme f est bijectif (autrement dit, est un automorphisme) si et seulement si $\det f \neq 0$. Dans ce cas, $\det(f^{-1}) = \frac{1}{\det f}$.

DÉFINITION 3.13. Si $\mathcal{V} = (v_1, \dots, v_n)$ est une famille de n vecteurs de E (de dimension n), et si \mathcal{B} est une base de E , on appelle *déterminant* de \mathcal{V} dans la base \mathcal{B} le déterminant de la matrice des coordonnées des vecteurs v_i dans la base \mathcal{B} . On le note $\det_{\mathcal{B}}(\mathcal{V})$.

REMARQUE. Attention : contrairement à la matrice d'un endomorphisme, le résultat dépend du choix de la base.

PROPOSITION 3.14. La famille $\mathcal{V} = (v_1, \dots, v_n)$ est une base de E si et seulement si son déterminant dans n'importe quelle base est non nul.

Variables aléatoires finies

« Indépendance ou la mort... Que ces mots sacrés nous rallient, et qu'ils soient le signal des combats et de notre réunion. »

J.-J. DESSALINES,
Proclamation au peuple d'Haïti (1804)

Dans tout le chapitre, (Ω, \mathbf{P}) est un espace probabilisé fini.

1. Rappels

1.1. Généralités.

DÉFINITION 1.1. Étant donné un univers Ω , on appelle *variable aléatoire* (réelle) sur Ω toute fonction $X: \Omega \rightarrow \mathbf{R}$.

EXEMPLE(S). La somme des résultats obtenus lors de deux lancers successifs d'un dé cubique est une variable aléatoire.

DÉFINITION 1.2. Soit X une variable aléatoire sur un univers Ω . Soient $x \in \mathbf{R}$ et $A \subset \mathbf{R}$. On note « $x \in A$ » l'événement $X^{-1}(A)$, « $X = x$ » l'événement $X^{-1}(x)$ et « $X \leq x$ » l'événement $X^{-1}(]-\infty; x])$.

EXEMPLE(S). On note X la somme des résultats obtenus lors de deux lancers successifs d'un dé cubique équilibré. Déterminer $\mathbf{P}(X = 6)$ et $\mathbf{P}(7 \leq X \leq 10)$.

DÉFINITION 1.3. Soit X une variable aléatoire sur un univers Ω . On appelle *loi de probabilité* de X la donnée, pour tout $x \in X(\Omega)$, du nombre $\mathbf{P}(X = x)$.

EXEMPLE(S). Déterminer la loi de probabilité de la somme des résultats obtenus lors de deux lancers successifs d'un dé cubique équilibré.

DÉFINITION 1.4. Soit X une variable aléatoire sur un univers Ω . On appelle *fonction de répartition* de X la fonction $F_X: \mathbf{R} \rightarrow [0; 1]$ définie par $F_X(x) = \mathbf{P}(X \leq x)$.

EXEMPLE(S). Représenter graphiquement la fonction de répartition de la somme des résultats obtenus lors de deux lancers successifs d'un dé cubique équilibré.

PROPOSITION 1.5. Pour tous $x_1, x_2 \in \mathbf{R}$, on a $\mathbf{P}(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1)$.

1.2. Espérance et variance. Dans toute la suite, on se place dans le cas d'un univers Ω fini.

DÉFINITION 1.6. Soit X une variable aléatoire sur Ω . On appelle *espérance* de X le nombre $\mathbf{E}(X) = \sum_{\omega \in \Omega} \mathbf{P}(\{\omega\}) \times X(\omega)$.

Si $\mathbf{E}(X) = 0$, on dit que X est *centrée*.

EXEMPLE(S). Calculer l'espérance de la somme des résultats obtenus lors de deux lancers successifs d'un dé cubique équilibré.

PROPOSITION 1.7. $\mathbf{E}(X) = \sum_{x \in X(\Omega)} x \times \mathbf{P}(X = x)$.

PROPOSITION 1.8 (Linéarité de l'espérance). Soient X et Y deux variables aléatoires sur Ω et $\lambda, \mu \in \mathbf{R}$. Alors $\mathbf{E}(\lambda X + \mu Y) = \lambda \mathbf{E}(X) + \mu \mathbf{E}(Y)$.

EXEMPLE(S). Recalculer l'espérance de la somme des résultats obtenus lors de deux lancers successifs d'un dé cubique équilibré.

PROPOSITION 1.9 (Formule de transfert). Soient X une variable aléatoire sur Ω et φ une fonction $\mathbf{R} \rightarrow \mathbf{R}$. Alors

$$\mathbf{E}(\varphi(X)) = \sum_{x \in X(\Omega)} \varphi(x) \times \mathbf{P}(X = x).$$

EXEMPLE(S). Calculer l'espérance de l'inverse de la somme des numéros obtenus lors du lancer de deux dés cubiques équilibrés.

DÉFINITION 1.10. On appelle *variance* d'une variable aléatoire X sur Ω le nombre $\mathbf{V}(X) = \mathbf{E}((X - \mathbf{E}(X))^2)$. On appelle *écart-type* de X le nombre $\sigma(X) = \sqrt{\mathbf{E}(X)}$.

Si $\mathbf{V}(X) = 1$, on dit que X est *réduite*.

REMARQUE. $\mathbf{V}(X) = 0$ si et seulement si X est constante (elle est alors égale à son espérance).

EXEMPLE(S). Calculer la variance et l'écart-type de la somme des résultats obtenus lors de deux lancers successifs d'un dé cubique équilibré.

PROPOSITION 1.11. Soient X une variable aléatoire sur Ω et $a, b \in \mathbf{R}$. Alors $\mathbf{V}(aX + b) = a^2 \mathbf{V}(X)$.

PROPOSITION 1.12 (Formule de König-Huygens). Soit X une variable aléatoire sur Ω . Alors $\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2$.

DÉMONSTRATION.

$$\begin{aligned} \mathbf{V}(X) &= \mathbf{E}((X - \mathbf{E}(X))^2) = \mathbf{E}(X^2 - 2X\mathbf{E}(X) + \mathbf{E}(X)^2) \\ &= \mathbf{E}(X^2) - 2\mathbf{E}(X)\mathbf{E}(X) + \mathbf{E}(\mathbf{E}(X)^2) \\ &= \mathbf{E}(X^2) - \mathbf{E}(X)^2 \end{aligned} \quad \square$$

PROPOSITION 1.13 (Inégalité de Bienaymé-Tchebychev). Soient X une variable aléatoire sur Ω et $a > 0$. Alors

$$\mathbf{P}(|X - \mathbf{E}(X)| \geq a) \leq \frac{\mathbf{V}(X)}{a^2}.$$

1.3. Loïs de probabilités usuelles.

DÉFINITION 1.14. On dit qu'une variable aléatoire X suit une *loi certaine* s'il existe $a \in \mathbf{R}$ tel que $\mathbf{P}(X = a) = 1$

REMARQUE. Si X suit une loi certaine, alors $\mathbf{E}(X) = a$ et $\mathbf{V}(X) = 0$.

DÉFINITION 1.15. On dit qu'une variable aléatoire X suit une *loi uniforme* s'il existe $n \in \mathbf{N}^*$ et $a_1, \dots, a_n \in \mathbf{R}$ tels que $\mathbf{P}(X = a_k) = \frac{1}{n}$ pour tout $k \in \llbracket 1; n \rrbracket$.

EXEMPLE(S). Le résultat du lancer d'un dé cubique équilibré suit une loi uniforme.

PROPOSITION 1.16. Si une variable aléatoire X suit une loi uniforme sur $\llbracket 1; n \rrbracket$, alors $\mathbf{E}(X) = \frac{n+1}{2}$ et $\mathbf{V}(X) = \frac{n^2-1}{12}$.

DÉFINITION 1.17. On dit qu'une variable aléatoire suit une *loi de Bernoulli de paramètre p* , notée $X \sim \mathcal{B}(p)$ s'il existe $p \in [0; 1]$ tel que $\mathbf{P}(X = 1) = p$ et $\mathbf{P}(X = 0) = 1 - p$.

EXEMPLE(S). Le nombre de « pile » obtenus lors du lancer d'une pièce équilibrée suit une loi de Bernoulli de paramètre $\frac{1}{2}$.

PROPOSITION 1.18. Si X suit une loi de Bernoulli de paramètre p , alors $\mathbf{E}(X) = p$ et $\mathbf{V}(X) = p(1 - p)$.

EXEMPLE(S). On lance n fois une pièce équilibrée. Calculer, pour tout $k \in \llbracket 0; n \rrbracket$, la probabilité d'obtenir k fois pile.

DÉFINITION 1.19. On dit qu'une variable aléatoire X suit une *loi binômiale de paramètres n et p* , notée $X \sim \mathcal{B}(n, p)$ s'il existe $n \in \mathbf{N}^*$ et $p \in [0; 1]$ tels que, pour tout $k \in \llbracket 0; n \rrbracket$, $\mathbf{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$.

EXEMPLE(S). Le nombre de « pile » obtenus lors de n lancers successifs d'une pièce équilibrée suit une loi binômiale de paramètres n et $\frac{1}{2}$.

PROPOSITION 1.20. Si X suit une loi binômiale de paramètres n et p , alors $\mathbf{E}(X) = np$ et $\mathbf{V}(X) = np(1 - p)$.

2. Couples de variables aléatoires réelles (finies)

DÉFINITION 2.1. Soient X et Y deux variables aléatoires réelles sur Ω . On appelle :

- *couple* (X, Y) des variables aléatoires réelles X et Y l'application définie pour tout $\omega \in \Omega$ par $(X, Y)(\omega) = (X(\omega), Y(\omega))$;
- *loi conjointe* du couple (X, Y) la donnée, pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, de $\mathbf{P}((X = x) \cap (Y = y))$;
- *lois marginales* du couple (X, Y) les lois de probabilité de X et de Y ;
- pour tout $x \in X(\Omega)$ tel que $\mathbf{P}(X = x) \neq 0$, *loi conditionnelle de Y sachant $(X = x)$* la donnée des valeurs de $\mathbf{P}_{X=x}(Y = y)$ pour tout $y \in Y(\Omega)$.

REMARQUE. L'univers Ω est fini donc les ensembles $X(\Omega)$, $Y(\Omega)$ et $X(\Omega) \times Y(\Omega)$ sont finis.

PROPOSITION 2.2. Soient X et Y deux variables aléatoires réelles sur Ω . Alors la loi conjointe du couple (X, Y) détermine ses lois marginales. Plus précisément, pour tout $x \in X(\Omega)$,

$$\mathbf{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbf{P}((X = x) \cap (Y = y)) = \sum_{y \in Y(\omega)} \mathbf{P}(X = x, Y = y).$$

REMARQUE. On a la formule analogue pour $\mathbf{P}(Y = y)$.

DÉMONSTRATION. Les événements $(Y = y)$ pour tout $y \in Y(\Omega)$ constituent un système complet d'événements. Ainsi, pour tout $x \in X(\Omega)$,

$$\mathbf{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbf{P}((X = x) \cap (Y = y)),$$

ce qui est exactement le résultat cherché. \square

REMARQUE. La réciproque n'est pas vraie : il ne suffit pas de connaître les lois marginales de (X, Y) pour en déduire la loi conjointe du couple.

EXEMPLE(S). On considère les variables aléatoires $X, Y, X', Y' : \Omega \rightarrow \{0, 1\}$ telles que :

- $\mathbf{P}(X = 0, Y = 0) = \mathbf{P}(X = 0, Y = 1) = \mathbf{P}(X = 1, Y = 0) = \mathbf{P}(X = 1, Y = 1) = \frac{1}{4}$;
- $\mathbf{P}(X' = 0, Y' = 0) = \mathbf{P}(X' = 1, Y' = 1) = \frac{1}{8}$, $\mathbf{P}(X' = 1, Y' = 0) = \mathbf{P}(X' = 0, Y' = 1) = \frac{3}{8}$.

Montrer que les lois marginales des couples (X, Y) et (X', Y') sont identiques.

On généralise les définitions à un n -uplet de variables aléatoires :

DÉFINITION 2.3. Soient X_1, \dots, X_n des variables aléatoires sur Ω . On appelle

- n -uplet (X_1, \dots, X_n) l'application définie pour tout $\omega \in \Omega$ par $(X_1, \dots, X_n)(\omega) = (X_1(\omega), \dots, X_n(\omega))$
- loi conjointe du n -uplet (X_1, \dots, X_n) la donnée, pour tout $(x_1, \dots, x_n) \in \prod_{k=1}^n X_k(\Omega)$, du nombre $\mathbf{P}((X_1 = x_1) \cap \dots \cap (X_n = x_n))$;
- lois marginales du n -uplet (X_1, \dots, X_n) les lois de probabilité des X_k , $k \in \llbracket 1; n \rrbracket$.

3. Variables aléatoires réelles (finies) indépendantes

DÉFINITION 3.1. Soient X et Y des variables aléatoires réelles sur Ω . On dit que les variables X et Y sont *indépendantes* si et seulement si, pour tous $x \in X(\Omega)$ et $y \in Y(\Omega)$,

$$\mathbf{P}(X = x, Y = y) = \mathbf{P}(X = x) \times \mathbf{P}(Y = y).$$

PROPOSITION 3.2. Soient X et Y deux variables aléatoires réelles sur Ω , f et g deux fonctions réelles définies respectivement sur $X(\Omega)$ et $Y(\Omega)$. Si X et Y sont indépendantes, alors :

- (1) pour tous $A \subset X(\Omega)$ et $B \subset Y(\Omega)$, $\mathbf{P}(X \in A, Y \in B) = \mathbf{P}(X \in A) \times \mathbf{P}(Y \in B)$;
- (2) $f(X)$ et $f(Y)$ sont des variables aléatoires indépendantes.

ADMIS. □

DÉFINITION 3.3. Soient $n \in \mathbf{N} \setminus \{0; 1\}$ et X_1, \dots, X_n des variables aléatoires réelles sur Ω . On dit que X_1, \dots, X_n sont *mutuellement indépendantes* si, pour tous $x_1 \in X_1, \dots, x_n \in X_n$, on a :

$$\mathbf{P}(X_1 = x_1, \dots, X_n = x_n) = \prod_{k=1}^n \mathbf{P}(X_k = x_k).$$

PROPOSITION 3.4. Soient $n \in \mathbf{N} \setminus \{0; 1\}$ et $p \in \llbracket 1; n-1 \rrbracket$, soient X_1, \dots, X_n des variables aléatoires réelles sur Ω et f, g deux fonctions réelles définies respectivement sur $\prod_{k=1}^p X_k(\Omega)$ et $\prod_{k=p+1}^n X_k(\Omega)$. Si les X_1, \dots, X_n sont mutuellement indépendantes, alors :

(1) pour tous $A_1 \subset X_1(\Omega), \dots, A_n \subset X_n(\Omega)$, les événements $(X_k \in A_k)_{1 \leq k \leq n}$ sont mutuellement indépendants, et en particulier :

$$\mathbf{P}(X_1 \in A_1, \dots, X_n \in A_n) = \prod_{k=1}^n \mathbf{P}(X_k \in A_k);$$

(2) $f(X_1, \dots, X_p)$ et $g(X_{p+1}, \dots, X_n)$ sont des variables aléatoires réelles indépendantes.

ADMIS. □

THÉORÈME 3.5 (Somme de variables aléatoires de Bernoulli indépendantes). Soient $n \in \mathbf{N} \setminus \{0, 1\}$ et X_1, \dots, X_n des variables aléatoires réelles sur Ω . On suppose que X_1, \dots, X_n sont mutuellement indépendantes et suivent chacune la loi $\mathcal{B}(p)$ ($p \in [0; 1]$). Alors la variable aléatoire $X_1 + \dots + X_n$ suit la loi $\mathcal{B}(n, p)$.

4. Covariance et coefficient de corrélation

Dans tout ce paragraphe, X et Y sont deux variables aléatoires réelles sur Ω .

DÉFINITION 4.1. On appelle *covariance* de X et Y le réel

$$\text{Cov}(X, Y) = \mathbf{E}[(X - \mathbf{E}(X))(Y - \mathbf{E}(Y))].$$

Si X et Y sont de variance non nulle, on appelle *coefficient de corrélation* du couple (X, Y) le nombre

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)}.$$

PROPOSITION 4.2. (1) $\text{Cov}(X, Y) = \text{Cov}(Y, X)$;

(2) $\text{Cov}(X, X) = \mathbf{V}(X)$;

(3) $\text{Cov}(X, Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)$;

(4) pour tous $a, b \in \mathbf{R}$, $\mathbf{V}(aX + bY) = a^2\mathbf{V}(X) + 2ab\text{Cov}(X, Y) + b^2\mathbf{V}(Y)$.

DÉMONSTRATION. (1) et (2) sont évidentes.

(3) On part de la définition et on développe :

$$\begin{aligned}\text{Cov}(X, Y) &= \mathbf{E}[(X - \mathbf{E}(X))(Y - \mathbf{E}(Y))] = \mathbf{E}[XY - X\mathbf{E}(Y) - Y\mathbf{E}(X) + \mathbf{E}(X)\mathbf{E}(Y)] \\ &= \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y) - \mathbf{E}(X)\mathbf{E}(Y) + \mathbf{E}(X)\mathbf{E}(Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y).\end{aligned}$$

(4)

$$\begin{aligned}\mathbf{V}(aX + bY) &= \mathbf{E}((aX + bY)^2) - (\mathbf{E}(aX + bY))^2 \\ &= \mathbf{E}(a^2X^2 + 2abXY + b^2Y^2) - (a\mathbf{E}(X) + b\mathbf{E}(Y))^2 \\ &= a^2\mathbf{E}(X^2) + 2ab\mathbf{E}(XY) + b^2\mathbf{E}(Y^2) - a^2\mathbf{E}(X)^2 - 2ab\mathbf{E}(X)\mathbf{E}(Y) - b^2\mathbf{E}(Y)^2 \\ &= a^2(\mathbf{E}(X^2) - \mathbf{E}(X)^2) + 2ab(\mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)) + b^2(\mathbf{E}(Y^2) - \mathbf{E}(Y)^2) \\ &= a^2\mathbf{V}(X) + 2ab\text{Cov}(X, Y) + b^2\mathbf{V}(Y).\end{aligned}\quad \square$$

LEMME 4.3. Si $X \geq 0$ et $\mathbf{E}(X) \geq 0$, alors $X = 0$.

DÉMONSTRATION. On raisonne par l'absurde : supposons $X \neq 0$: il existe donc $\omega \in \Omega$ tel que $X(\omega) \neq 0$, en fait $X(\omega) > 0$. Mais alors $\mathbf{E}(X) \geq \mathbf{P}(\omega)X(\omega) > 0$, absurde. \square

REMARQUE. La démonstration repose sur le fait que l'univers Ω ne contient que des issues *possibles*, c'est-à-dire de probabilité non nulle.

LEMME 4.4 (Inégalité de Cauchy-Schwarz).

$$\mathbf{E}(XY) \leq \sqrt{\mathbf{E}(X^2)\mathbf{E}(Y^2)},$$

avec égalité si et seulement si X et Y sont proportionnelles.

DÉMONSTRATION. On considère la fonction $f: \mathbf{R} \rightarrow \mathbf{R}$ définie par $f(\lambda) = \mathbf{E}((X + \lambda Y)^2)$. Pour tout $\lambda \in \mathbf{R}$, on a $f(\lambda) \geq 0$. Or, si on développe $(X + \lambda Y)^2$, on peut écrire $f(\lambda) = \mathbf{E}(X^2) + 2\lambda\mathbf{E}(XY) + \lambda^2\mathbf{E}(Y^2)$: c'est un polynôme du second degré en λ . On sait qu'un polynôme du second degré de signe constant possède un discriminant négatif ou nul : $4\mathbf{E}(XY)^2 - 4\mathbf{E}(X^2)\mathbf{E}(Y^2) \leq 0$, d'où le résultat.

Le cas d'égalité correspond à un discriminant nul, c'est-à-dire qu'il existe $\lambda \in \mathbf{R}$ tel que $f(\lambda) = 0$, autrement dit que $\mathbf{E}((X + \lambda Y)^2) = 0$. D'après le lemme précédent, la variable aléatoire $(X + \lambda Y)^2$ étant évidemment positive, cela implique que $X + \lambda Y = 0$, autrement dit que $X = -\lambda Y$. \square

REMARQUE. Est-on certain que la fonction f est vraiment polynomiale du *second* degré ?

PROPOSITION 4.5. Si $\mathbf{V}(X)\mathbf{V}(Y) \neq 0$, alors :

- (1) $|\rho(X, Y)| \leq 1$;
- (2) si $\rho(X, Y) = 1$, alors il existe $a > 0$ et $b \in \mathbf{R}$ tels que $Y = aX + b$;
- (3) si $\rho(X, Y) = -1$, alors il existe $a < 0$ et $b \in \mathbf{R}$ tels que $Y = aX + b$.

DÉMONSTRATION. (1) On pose $X' = X - \mathbf{E}(X)$ et $Y' = Y - \mathbf{E}(Y)$. Alors :

$$|\text{Cov}(X, Y)| = |\mathbf{E}((X - \mathbf{E}(X))(Y - \mathbf{E}(Y)))| = |\mathbf{E}(X'Y')| \leq \sqrt{\mathbf{E}(X'^2)\mathbf{E}(Y'^2)},$$

d'après l'inégalité de Cauchy-Schwarz. Mais alors

$$\sqrt{\mathbf{E}(X'^2)\mathbf{E}(Y'^2)} = \sqrt{\mathbf{E}[(X - \mathbf{E}(X))^2]\mathbf{E}[(Y - \mathbf{E}(Y))^2]} = \sqrt{\mathbf{V}(X)\mathbf{V}(Y)} = \sigma(X)\sigma(Y).$$

$$\text{Ainsi } |\rho(X, Y)| = \left| \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)} \right| \leq 1.$$

(2) et (3) On a $\mathbf{V}(X) \neq 0$ donc $\mathbf{E}[(X - \mathbf{E}(X))^2] \neq 0$, c'est-à-dire $\mathbf{E}(X'^2) \neq 0$, d'où $X' \neq 0$. De même, $Y' \neq 0$. Alors :

$$|\rho(X, Y)| = 1 \iff |\text{Cov}(X, Y)| = \sigma(X)\sigma(Y) \iff |\mathbf{E}(X'Y')| = \sqrt{\mathbf{E}(X'^2)\mathbf{E}(Y'^2)}.$$

Il s'agit du cas d'égalité dans l'inégalité de Cauchy-Schwarz : il existe ainsi $a \in \mathbf{R}$ tel que $Y' = aX'$:

$$Y' = aX' \iff Y - \mathbf{E}(Y) = aX - a\mathbf{E}(X) \iff Y = aX + b,$$

où $b = a\mathbf{E}(Y) - a\mathbf{E}(X) \in \mathbf{R}$.

De plus, si $Y = aX + b$ avec $a \neq 0$, alors

$$\begin{aligned} \rho(X, Y) &= \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)} = \frac{\mathbf{E}[X(aX + b)] - \mathbf{E}(X)\mathbf{E}(aX + b)}{|a|\sigma(X)^2} \\ &= \frac{a\mathbf{E}(X^2) + b\mathbf{E}(X) - a\mathbf{E}(X)^2 - b\mathbf{E}(X)}{|a|\mathbf{V}(X)} = \frac{a\mathbf{V}(X)}{|a|\mathbf{V}(X)} = \frac{a}{|a|} \begin{cases} 1 & \text{si } a > 0 \\ -1 & \text{si } a < 0. \end{cases} \end{aligned}$$

□

PROPOSITION 4.6. Si X et Y sont indépendantes, alors :

- (1) $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$;
- (2) $\text{Cov}(X, Y) = 0$;
- (3) $\mathbf{V}(X + Y) = \mathbf{V}(X) + \mathbf{V}(Y)$ (égalité de Bienaymé).

DÉMONSTRATION.

□

REMARQUE. Attention, la réciproque est fautive : deux variables aléatoires peuvent vérifier ces trois propriétés sans pour autant être indépendantes.

CONTRE-EXEMPLE. Soit X une variable aléatoire prenant les valeurs $-1, 0$ et 1 de manière équiprobable ; on note $Y = X^2$. Il est clair que X et Y ne sont pas indépendantes. Pourtant, $\mathbf{E}(XY) = 0 = \mathbf{E}(X)\mathbf{E}(Y)$.

DÉFINITION 4.7. Si $\text{Cov}(X, Y) = 0$, on dit que X et Y sont *décorrélées*.

Séries numériques

« Les séries divergentes sont en général quelque chose de bien fatal et c'est une honte qu'on ose fonder y fonder aucune démonstration. »

N. H. ABEL, lettre à B. H. HOLMBOE (1826)

1. Séries, convergence et divergence

DÉFINITION 1.1. Soit $(u_n)_{n \in \mathbf{N}}$ une suite de nombres réels ou complexes. On appelle *série de terme général* $(u_n)_{n \in \mathbf{N}}$ l'objet formel, noté $\sum u_n$, traduisant le fait d'*essayer* d'additionner tous les termes de la suite $(u_n)_{n \in \mathbf{N}}$.

REMARQUE. Dans tout les énoncés du cours, on supposera que la suite est définie pour tout $n \in \mathbf{N}$. On verra que dans certains cas il faut se restreindre à $n \in \mathbf{N}^*$, mais c'est toujours évident en fonction du contexte, par exemple si on regarde $\sum \frac{1}{n}$.

DÉFINITION 1.2. Soit $\sum u_n$ une série réelle ou complexe. Pour tout $N \in \mathbf{N}$, on appelle *N-ème somme partielle* de la série le nombre

$$S_N = \sum_{n=0}^N u_n.$$

EXEMPLE(S). Calculer les sommes partielles de $\sum n$, $\sum n^2$, $\sum q^n$ pour $q \in \mathbf{C}$.

DÉFINITION 1.3. Soit $\sum u_n$ une série réelle ou complexe. Si la suite $(S_N)_{N \in \mathbf{N}}$ des sommes partielles admet une limite finie, on dit que $\sum u_n$ est *convergente*. Dans ce cas, on note

$$\sum_{n=0}^{+\infty} u_n = \lim_{N \rightarrow +\infty} S_N$$

sa *somme*, et, pour tout $N \in \mathbf{N}$,

$$\sum_{n=N}^{+\infty} u_n = \sum_{n=0}^{+\infty} u_n - \sum_{n=0}^{N-1} u_n$$

son *reste* d'indice N .

Dans le cas contraire, on dit que $\sum u_n$ est *divergente*. Le fait d'être convergente ou divergente est appelé la *nature* de la série $\sum u_n$.

EXEMPLE(S). Étudier la convergence des séries $\sum n$, $\sum n^2$, $\sum q^n$. Le cas échéant, donner l'expression des restes.

PROPOSITION 1.4. Soit $q \in \mathbf{C}$. La série $\sum q^n$ est appelée série géométrique de raison q ; elle est convergente si et seulement si $|q| < 1$. Dans ce cas,

$$\sum_{n=0}^{+\infty} q^n = \frac{1}{1-q}; \quad \forall N \in \mathbf{N} \quad \sum_{n=N}^{+\infty} q^n = \frac{q^N}{1-q}.$$

DÉMONSTRATION. Fait à l'exemple précédent. \square

PROPOSITION 1.5. Si $\mathbf{K} = \mathbf{R}$ ou \mathbf{C} , alors l'ensemble des séries à coefficients dans \mathbf{K} est un \mathbf{K} -espace vectoriel E . L'ensemble des séries convergentes est un sous-espace vectoriel de E , sur lequel l'application somme définit une forme linéaire.

En d'autres termes, si $\sum u_n$ et $\sum v_n$ sont deux séries convergentes, et $\lambda, \mu \in \mathbf{K}$, alors $\sum(\lambda u_n + \mu v_n)$ est convergente et

$$\sum_{n=0}^{+\infty} (\lambda u_n + \mu v_n) = \lambda \sum_{n=0}^{+\infty} u_n + \mu \sum_{n=0}^{+\infty} v_n.$$

REMARQUE. Attention, l'ensemble des séries divergentes n'est pas un sous-espace vectoriel! Autrement dit, on peut faire une combinaison linéaire de séries divergentes et obtenir un résultat convergent. Par exemple, il est évident que $\sum n - \sum n = \sum 0$.

EXEMPLE(S). Étudier la convergence de $\sum(n + \frac{1}{2^n})$.

PROPOSITION 1.6. Soit $\sum u_n$ une série. Si $\sum u_n$ est convergente, alors $\lim_{n \rightarrow +\infty} u_n = 0$.

Par contraposée, une série $\sum u_n$ pour laquelle le terme général ne tend pas vers 0 ne peut être que divergente; on dira qu'elle est grossièrement divergente.

DÉMONSTRATION. On suppose $\sum u_n$ convergente. Alors, pour tout $N \in \mathbf{N}$,

$$u_N = \sum_{n=0}^N u_n - \sum_{n=0}^{N-1} u_n.$$

Lorsque N tend vers $+\infty$, chacune des deux sommes partielles converge vers la somme de la série, u_N tend donc vers 0. \square

REMARQUE. Attention! Une série peut être divergente sans être grossièrement divergente!

EXEMPLE(S). Montrer que, pour tout $x \in \mathbf{R}$, $\ln(1+x) \leq x$. En déduire la convergence de la série $\sum \frac{1}{n}$.

REMARQUE. Si $\sum u_n$ est une série, alors, pour tout $N \in \mathbf{N}$, $u_N = \sum_{n=0}^{N-1} (u_{n+1} - u_n) + u_0$. En particulier, la suite $(u_n)_{n \in \mathbf{N}}$ admet une limite finie ssi la série $\sum (u_{n+1} - u_n)$ est convergente.

2. Séries à termes positifs

Dans toute cette partie, on s'intéresse à des séries dont le terme général est réel positif.

REMARQUE. Si $\sum u_n$ est une série à termes positifs, alors la suite des sommes partielles de $\sum u_n$ est croissante.

PROPOSITION 2.1. *Soit $\sum u_n$ une série à termes positifs. Alors $\sum u_n$ est convergente si et seulement si toutes les sommes partielles sont majorées. En d'autres termes, s'il existe $M \in \mathbf{R}$ tel que, pour tout $N \in \mathbf{N}$,*

$$\sum_{n=0}^N u_n \leq M.$$

DÉMONSTRATION. La suite des sommes partielles est croissante, donc d'après le théorème de la limite monotone, elle est convergente si et seulement si elle est majorée. \square

THÉORÈME 2.2 (Théorème de comparaison des séries à termes positifs, version 1). *Soient $\sum u_n$ et $\sum v_n$ deux séries à termes positifs. On suppose que, pour tout $n \in \mathbf{N}$, $u_n \leq v_n$.*

(1) *Si $\sum v_n$ est convergente, alors $\sum u_n$ est convergente. De plus, dans ce cas,*

$$\sum_{n=0}^{+\infty} u_n \leq \sum_{n=0}^{+\infty} v_n.$$

(2) *Si $\sum u_n$ est divergente, alors $\sum v_n$ est divergente.*

DÉMONSTRATION. On commence par remarquer que, pour tout $N \in \mathbf{N}$,

$$\sum_{n=0}^N u_n \leq \sum_{n=0}^N v_n.$$

(1) Dans ce cas, les sommes partielles de $\sum v_n$ sont majorées, donc celles de $\sum u_n$ le sont également par transitivité. Et on conclut par passage à la limite dans les inégalités.

(2) Dans ce cas, les sommes partielles de $\sum u_n$ ne sont pas majorées, donc celles de v_n ne le sont pas non plus. \square

EXEMPLE(S). Étudier la convergence des séries $\sum \frac{1}{n!}$, puis $\sum \frac{\ln(n^n)}{n!}$.

THÉORÈME 2.3 (Théorème de comparaison des séries à termes positifs, version 2). *Soient $\sum u_n$ et $\sum v_n$ deux séries à termes positifs.*

(1) *On suppose que $u_n = o(v_n)$. Alors, si $\sum u_n$ est divergente, $\sum v_n$ l'est ; si $\sum v_n$ est convergente, $\sum u_n$ l'est.*

(2) *On suppose que $u_n \sim v_n$. Alors $\sum u_n$ est convergente si et seulement si $\sum v_n$ l'est.*

DÉMONSTRATION. (1) il existe $(\alpha_n)_{n \in \mathbf{N}}$ une suite réelle positive tendant vers 0 telle que $u_n = \alpha_n v_n$ pour tout $n \in \mathbf{N}$. En particulier, la suite $(\alpha_n)_{n \in \mathbf{N}}$ est bornée. Notons $\alpha \in \mathbf{R}$ un majorant de $(\alpha_n)_{n \in \mathbf{N}}$, on a donc $u_n \leq \alpha v_n$ pour tout $n \in \mathbf{N}$.

Si $\sum v_n$ est convergente, alors $\sum \alpha v_n$ l'est. Par théorème de comparaison des séries à termes positifs V1, on en déduit que $\sum u_n$ est convergente.

Si $\sum u_n$ est divergente, alors en particulier la suite $(u_n)_{n \in \mathbf{N}}$ n'est pas constamment nulle à partir d'un certain rang, donc $\alpha > 0$. On peut alors écrire $\frac{u_n}{\alpha} \leq v_n$ pour tout $n \in \mathbf{N}$, et la série $\sum \frac{u_n}{\alpha}$ est de même nature que $\sum u_n$; on conclut d'après le théorème de comparaison V1.

(2) $\frac{u_n}{v_n} \rightarrow 1$ donc ne s'annule pas à partir d'un certain rang n_0 . Notons m et M respectivement un minorant et un majorant non nuls de $\frac{u_n}{v_n}$ pour tout $n \geq n_0$. On peut alors écrire, pour tout $n \geq n_0$, $mv_n \leq u_n \leq Mv_n$. Les séries $\sum mv_n$ et $\sum Mv_n$ sont de même nature que $\sum v_n$, on peut donc conclure en fonction du cas d'après la V1 du théorème de comparaison.

□

PROPOSITION 2.4. Soit $\alpha > 0$. La série de Riemann $\sum \frac{1}{n^\alpha}$ est convergente si et seulement si $\alpha > 1$. Dans ce cas, on note sa somme $\zeta(\alpha)$.

DÉMONSTRATION. • Si $\alpha \leq 1$, alors $\frac{1}{n^\alpha} \leq \frac{1}{n}$ qui est le terme général d'une série divergente (vu en exemple), donc d'après le théorème de comparaison des séries à termes positifs, $\sum \frac{1}{n^\alpha}$ est divergente.

• si $\alpha > 1$, on majore les sommes partielles à l'aide d'une intégrale : pour tout $N \geq 2$,

$$\sum_{n=1}^N \frac{1}{n^\alpha} \leq 1 + \int_1^N \frac{dt}{t^\alpha} = 1 + \frac{1}{\alpha-1} \left[\frac{1}{t^{\alpha-1}} \right]_1^N \leq 1 + \frac{1}{\alpha-1},$$

la série $\sum \frac{1}{n^\alpha}$ est donc convergente.

□

PROPOSITION 2.5 (Critère de d'Alembert). Soit $\sum u_n$ une série à termes positifs. On suppose que $u_n \neq 0$ pour tout $n \in \mathbf{N}$ et que le quotient $\frac{u_{n+1}}{u_n}$ admet une limite finie $\ell \in \mathbf{R}$. Alors :

- si $\ell < 1$, alors $\sum u_n$ est convergente ;
- si $\ell > 1$, alors $\sum u_n$ est grossièrement divergente.

DÉMONSTRATION. Par comparaison avec une série géométrique.

□

REMARQUE. Il faut garder en tête que le critère reste muet dans le cas où $\ell = 1$ ou si le quotient $\frac{u_{n+1}}{u_n}$ n'a pas de limite, ce qui restreint fortement le nombre de situations où il peut se montrer utile. Par exemple, on a $\ell = 1$ pour toutes les séries de Riemann, pourtant certaines sont convergentes et d'autres sont divergentes.

3. Séries absolument convergentes

Dans cette partie, on s'intéresse à des séries à termes dans \mathbf{R} ou \mathbf{C} . Dans ce cas général, le comportement de la suite des sommes partielles peut être beaucoup plus

compliqué que pour les séries à termes positifs : les sommes partielles peuvent avoir une limite finie, infinie, et même pas de limite du tout, et une simple majoration ne permet pas d'avoir suffisamment d'informations pour conclure.

DÉFINITION 3.1. Soit $\sum u_n$ une série à termes réels ou complexes. On dit que $\sum u_n$ est *absolument convergente* si la série $\sum |u_n|$ est convergente.

REMARQUE. La série $\sum |u_n|$ est à termes positifs, donc on peut lui appliquer tous les résultats de la partie précédente.

EXEMPLE(S). $\sum \frac{e^{in}}{n^2}$ est absolument convergente, $\sum \frac{(-1)^n}{n}$ ne l'est pas.

PROPOSITION 3.2. *L'ensemble des séries absolument convergentes est un sous-espace vectoriel de l'ensemble des séries. Autrement dit, une combinaison linéaire de séries absolument convergentes est absolument convergente.*

DÉMONSTRATION. Soient $\sum u_n$ et $\sum v_n$ deux séries absolument convergentes et $\lambda, \mu \in \mathbf{C}$. Il s'agit de montrer que la série $\sum |\lambda u_n + \mu v_n|$ est convergente. Or, d'après l'inégalité triangulaire, pour tout $n \in \mathbf{N}$, $|\lambda u_n + \mu v_n| \leq |\lambda| |u_n| + |\mu| |v_n|$. Mais $\sum |\lambda| |u_n|$ et $\sum |\mu| |v_n|$ sont convergentes donc, par théorème de comparaison des séries à termes positifs, $\sum |\lambda u_n + \mu v_n|$ l'est également. \square

THÉORÈME 3.3. *Soit $\sum u_n$ une série à termes réels ou complexes. On suppose $\sum u_n$ absolument convergente. Alors $\sum u_n$ est convergente. De plus,*

$$\left| \sum_{n=0}^{+\infty} u_n \right| \leq \sum_{n=0}^{+\infty} |u_n|.$$

ADMIS. \square

EXEMPLE(S). $\sum \frac{e^{in}}{n^2}$ est convergente.

REMARQUE. L'implication réciproque n'est pas vraie en général : une suite peut être convergente sans être absolument convergente (on parle alors de série *semi-convergente*).

EXEMPLE(S). Montrer que $\sum \frac{(-1)^n}{n}$ est convergente.

4. Développement décimal des nombres réels

LEMME 4.1. *Soit $(u_n)_{n \in \mathbf{N}^*}$ une suite réelle. On suppose que, pour tout $n \in \mathbf{N}^*$, $0 \leq u_n \leq 9$. Alors $\sum \frac{u_n}{10^n}$ est convergente.*

DÉMONSTRATION. Pour tout $n \in \mathbf{N}^*$, $\frac{u_n}{10^n} \leq \frac{10}{10^n} \leq \frac{1}{10^{n-1}}$ qui est le terme général d'une série géométrique convergente ($q = \frac{1}{10}$); par théorème de comparaison des séries à termes positifs, $\sum \frac{u_n}{10^n}$ est convergente. \square

THÉORÈME 4.2. *Soit $\alpha \in [0; 1[$. Alors il existe une unique suite $(u_n)_{n \in \mathbf{N}^*}$ vérifiant les propriétés suivantes :*

- $\forall n \in \mathbf{N}^* \quad u_n \in \llbracket 0; 9 \rrbracket$;

- $\sum_{n=1}^{+\infty} \frac{u_n}{10^n} = \alpha ;$

- la suite $(u_n)_{n \in \mathbf{N}}$ n'est pas constante égale à 9 à partir d'un certain rang.

La suite $(u_n)_{n \in \mathbf{N}}$ est appelée développement décimal du nombre réel α . On le note

$$\alpha = 0, u_0 u_1 \dots u_n \dots$$

ADMIS. □

REMARQUE. Le théorème se généralise immédiatement à tout nombre réel x , puisqu'on peut toujours écrire $x = \lfloor x \rfloor + \alpha$ avec $\alpha \in [0; 1[$.

La condition « $(u_n)_{n \in \mathbf{N}^*}$ non constante égale à 9 à partir d'un certain rang » est nécessaire pour avoir unicité du développement décimal.

EXEMPLE(S). $0,9999\dots = 1$.

PROPOSITION 4.3. *Un nombre réel est rationnel si et seulement si son développement décimal est périodique à partir d'un certain rang.*

EXEMPLE(S). Déterminer le développement décimal de $\frac{1}{44}$. Trouver deux nombres entiers p et q tels que $\frac{p}{q} = 0,31612612612\dots$

Fonctions à valeurs vectorielles et courbes paramétrées

« Durant des années, quand je présentais mes travaux à la Régie Renault ou ailleurs, j'invoquais les recherches d'un professeur mythique que j'avais appelé Durand. Je lui avais attribué les résultats de mes propres réflexions, ce qui donnait confiance aux gens. Parce que si j'avais dit qu'il s'agissait de polynômes définis par moi-même, je crois que je serais devenu une abomination pour la maison ! Alors je parlais des fonctions de Durand et les gens regardaient les courbes, très satisfaits. »

P. BÉZIER, interview pour *Science & Vie* (1990)

1. Fonctions à valeurs vectorielles

DÉFINITION 1.1. On appelle *fonction à valeurs vectorielles* toute fonction $f: I \rightarrow \mathbf{R}^n$, où I est un intervalle de \mathbf{R} et n un entier supérieur ou égal à 2.

Si $f = (f_1, \dots, f_n)$, les fonctions $f_1, \dots, f_n: I \rightarrow \mathbf{R}$ sont appelées *fonctions coordonnées* de f .

EXEMPLE(S). $f: \mathbf{R} \rightarrow \mathbf{R}^2, t \mapsto \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$, $g:]0; +\infty[\rightarrow \mathbf{R}^3, t \mapsto \begin{pmatrix} \ln t \\ \frac{1}{t} \\ \sqrt{t} \end{pmatrix}$.

DÉFINITION 1.2. Soient $f = (f_1, \dots, f_n): I \rightarrow \mathbf{R}^n$ et $a \in I$. On dit que f est *continue* (resp. *dérivable*) en a si et seulement si les fonctions coordonnées f_1, \dots, f_n le sont.

EXEMPLE(S). Étudier la continuité et la dérivabilité de $f: \mathbf{R}_+ \rightarrow \mathbf{R}, t \mapsto \begin{pmatrix} t^2 \\ \sqrt{t} \end{pmatrix}$.

DÉFINITION 1.3. Soit $f = (f_1, \dots, f_n): I \rightarrow \mathbf{R}^n$. On dit que f est *continue* (resp. *dérivable*, *de classe \mathcal{C}^k*) sur I si et seulement si les fonctions coordonnées f_1, \dots, f_n le sont.

On note $\mathcal{C}^k(I, \mathbf{R}^n)$ l'ensemble des fonctions de classe \mathcal{C}^k sur I à valeurs dans \mathbf{R}^n . C'est un \mathbf{R} -espace vectoriel.

PROPOSITION 1.4 (Formule de Taylor-Young). Soient $F \in \mathcal{C}^k(I, \mathbf{R}^n)$ et $a \in I$. Alors, pour $t \in I$ tendant vers a , on a

$$F(t) = \sum_{i=0}^k \frac{(t-a)^i}{i!} F^{(i)}(a) + o((t-a)^k).$$

REMARQUE. Dans cette formule, bien garder en tête que le $o((t-a)^k)$ est lui-même une fonction à valeurs dans \mathbf{R}^n .

En pratique, on utilise le plus souvent les formules des développements limités usuels coordonnée par coordonnée, mais en identifiant avec la formule de Taylor-Young, cela peut permettre d'en déduire les valeurs des dérivées successives d'une fonction en un point (on verra dans la prochaine partie en quoi cela peut être utile).

EXEMPLE(S). Considérons $F:]-1; +\infty[\rightarrow \mathbf{R}^2$ définie par $f(t) = \begin{pmatrix} \sqrt{1+t} \\ \ln(1+t) \end{pmatrix}$. On sait qu'au voisinage de $t = 0$, on a $\sqrt{1+t} = 1 + \frac{t}{2} - \frac{t^2}{8} + o(t^2)$ et $\ln(1+t) = t - \frac{t^2}{2} + o(t^2)$, donc

$$F(t) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1/2 \\ 1 \end{pmatrix} + t^2 \begin{pmatrix} -1/8 \\ -1/2 \end{pmatrix} + o(t^2).$$

En identifiant avec la formule de Taylor, on en déduit

$$F(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad tF'(0) = t \begin{pmatrix} 1/2 \\ 1 \end{pmatrix} \quad \text{et} \quad \frac{t^2}{2}F''(0) = t^2 \begin{pmatrix} -1/8 \\ -1/2 \end{pmatrix},$$

donc

$$F(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad F'(0) = \begin{pmatrix} 1/2 \\ 1 \end{pmatrix} \quad \text{et} \quad F''(0) = \begin{pmatrix} -1/4 \\ -1 \end{pmatrix}.$$

2. Courbes paramétrées

Dans toute la suite, on supposera que les fonctions étudiées sont dérivables suffisamment de fois pour que les formules présentées aient un sens.

DÉFINITION 2.1. Soit $F = (f_1, f_2): I \rightarrow \mathbf{R}^2$. On appelle *arc paramétré* par F la partie du plan $\Gamma \subset \mathbf{R}^2$ définie par $\Gamma = \{F(t) \mid t \in I\}$. On le note le plus souvent

$$\Gamma: \begin{cases} x = f_1(t) \\ y = f_2(t) \end{cases} \quad (t \in I).$$

Le plus souvent, pour $t \in I$, on note $M(t)$ le point (parfois dit « courant ») de coordonnées $F(t)$, et on note $x(t)$ et $y(t)$ les deux fonctions coordonnées de F .

DÉFINITION 2.2. Soit $t \in I$. Le point $M(t)$ est dit :

- *régulier* si $F'(t) \neq 0$;
- *stationnaire* sinon.

L'étude des points stationnaires est un ingrédient fondamental du tracé d'arcs paramétrés. En effet, au voisinage d'un point régulier, le point courant possède un vecteur vitesse non nul et suit donc une trajectoire « lisse ». En revanche, en un point stationnaire, le vecteur vitesse s'annule et la trajectoire peut donc posséder des variations brutales. On peut faire une analogie en regardant la trajectoire d'une voiture sur l'autoroute (la vitesse ne s'annule pas, la trajectoire est lisse) et en la comparant avec la même voiture faisant un créneau (la voiture s'arrête plusieurs fois en cours de manœuvre et repart chaque fois avec une direction différente).

2.1. Étude au voisinage d'un point. On considère un instant $t_0 \in I$ et on s'intéresse à l'aspect de l'arc paramétré au voisinage du point $M(t_0)$. Si on fait un développement limité de la fonction F pour t tendant vers t_0 , on peut écrire

$$F(t) = F(t_0) + \frac{(t - t_0)^p}{p!} F^{(p)}(t_0) + o((t - t_0)^p),$$

où p désigne l'ordre du premier vecteur dérivé de F non nul en t_0 (pour un point régulier, $p = 1$). Autrement dit,

$$F(t) - F(t_0) \sim \frac{(t - t_0)^p}{p!} F^{(p)}(t_0),$$

c'est-à-dire que le vecteur $\overrightarrow{M(t_0)M(t)}$ est, localement, quasiment colinéaire à $F^{(p)}(t_0)$. On en déduit que

PROPOSITION 2.3. *La tangente à la courbe à l'instant t_0 admet comme vecteur directeur $F^{(p)}(t_0)$, où $p = \min\{k \in \mathbf{N}^* \mid F^{(k)}(t_0) \neq 0\}$.*

La composante $\frac{(t-t_0)^p}{p!} F^{(p)}(t_0)$ du développement limité décrit donc le déplacement du point $M(t)$ « le long du vecteur tangent ». Pour connaître l'allure du déplacement du point dans la direction transversale, il faut donc pousser le développement limité jusqu'à trouver un vecteur dérivé, disons $F^{(q)}(t_0)$, *non colinéaire* à $F^{(p)}(t_0)$. L'allure du déplacement de $M(t)$ transversalement au vecteur tangent de la courbe est alors déterminée par le terme $\frac{(t-t_0)^q}{q!} F^{(q)}(t_0)$.

On rencontre alors quatre cas possibles, en fonction des parités des entiers p et q . En effet, la parité de p détermine si le facteur $(t - t_0)^p$ va, ou non, changer de signe au voisinage de t_0 . Si p est impair, le point $M(t)$ va passer « d'un côté à l'autre » du point $M(t_0)$, tandis que si p est pair, il restera toujours « du même côté ». De même, si q est impair, le point $M(t)$ va « traverser » la tangente de la courbe en t_0 tandis que, si q est pair, il va rester localement « du même côté » de sa tangente.

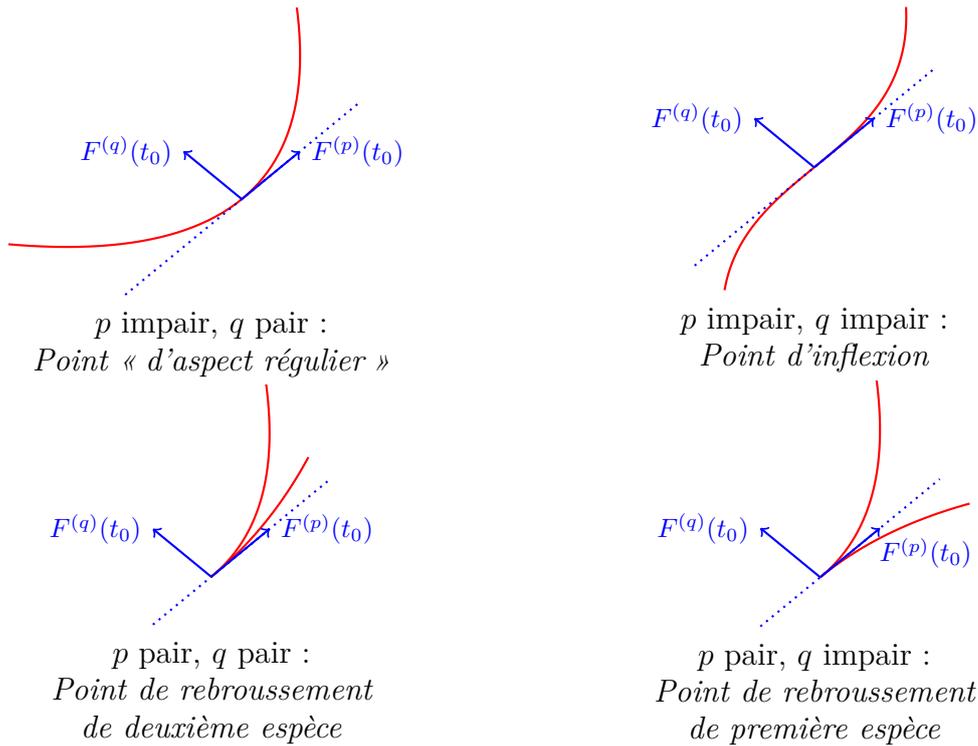
REMARQUE. Notons que, dans tout ce paragraphe, on a fait comme s'il était toujours possible de trouver deux vecteurs dérivés non colinéaires. Les cas d'un point où tous les vecteurs dérivés successifs sont nuls, ou alors tous colinéaires entre eux, requiert d'autres méthodes qu'on ne mentionnera pas ici.

2.2. Plan complet d'étude. Nous allons décrire la stratégie permettant de tracer de A à Z un arc paramétré. On prendra pour exemple la courbe suivante (*astroïde*) :

$$\mathcal{C}: \begin{cases} x = \cos^3 t \\ y = \sin^3 t \end{cases} \quad (t \in \mathbf{R}).$$

2.2.1. Intervalle d'étude. On vérifie la régularité de la fonction F sur l'intervalle I . Ici : les fonctions x et y sont de classe \mathcal{C}^∞ sur \mathbf{R} .

Puis on tente de convertir les propriétés de parité, périodicité, etc. des fonctions x et y en symétries sur la courbe \mathcal{C} , pour réduire autant que possible l'intervalle de travail I . Ici, on repère par exemple que les fonctions x et y sont 2π -périodiques. Autrement dit,



pour tout $t \in \mathbf{R}$, $M(t + 2\pi) = M(t)$. Il suffit donc de se restreindre à n'importe quel intervalle de longueur 2π , par exemple $I = [-\pi; \pi]$, puisque les valeurs de t en-dehors de cet intervalle ne font que repasser sur cette partie de la courbe.

De plus, on remarque que x est paire et y impaire. On a donc, pour tout $t \in \mathbf{R}$, $x(-t) = x(t)$ mais $y(-t) = -y(t)$: on peut donc se restreindre aux valeurs positives de t , et ensuite compléter la courbe par symétrie par rapport à l'axe des abscisses. On travaille donc sur $I = [0; \pi]$.

On pourrait exploiter d'autres propriétés des fonctions x et y pour restreindre davantage (exercice : montrer qu'on peut se restreindre à $I = [0; \frac{\pi}{4}]$), mais contentons-nous de cela pour l'instant.

2.2.2. *Variations.* Il s'agit de calculer les dérivées de x et y et d'étudier leurs signes afin de dresser un tableau de variations. Ici, pour tout $x \in \mathbf{R}$, $x'(t) = -3 \sin t \cos^2 t$ et $y'(t) = 3 \cos t \sin^2 t$, donc :

t	0	$\pi/2$	π
$x'(t)$	0	-	0
$y'(t)$	0	+	0
$x(t)$	1	\searrow	0
$y(t)$	0	\nearrow	1
			\searrow
			0

Ce tableau de variation permet de repérer :

- les points stationnaires : ce sont les points où $x'(t) = y'(t) = 0$; ici, il y en a en $t = 0, \frac{\pi}{2}$ et π .

- certaines tangentes horizontales ou verticales : ce sont les points où l'une des deux dérivées s'annule, mais pas l'autre.

2.2.3. *Points stationnaires.* Pour chaque point stationnaire repéré à l'étape précédente, on effectue un développement limité des fonctions x et y afin de déterminer la tangente et l'apparence de la courbe au voisinage du point. Les développements limités doivent être poussés à un ordre suffisant jusqu'à trouver deux vecteurs dérivés non colinéaires.

- au voisinage de $t = 0$: $\cos t = 1 - \frac{t^2}{2} + o(t^3)$ donc $\cos^3 t = 1 - \frac{3}{2}t^2 + o(t^3)$, $\sin t = t - \frac{t^3}{6} + o(t^3)$ donc $\sin^3 t = t^3 + o(t^3)$. On en déduit le développement limité :

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t^2 \begin{pmatrix} -3/2 \\ 0 \end{pmatrix} + t^3 \begin{pmatrix} 0 \\ -1/6 \end{pmatrix} + o(t^3).$$

Le premier vecteur dérivé non nul est d'ordre $p = 2$. Le vecteur dérivé d'ordre 3 ne lui est pas colinéaire donc $q = 3$. On est donc dans le cas d'un point de rebroussement de première espèce, avec une tangente dirigée par le premier vecteur dérivé non nul, donc par $\begin{pmatrix} -3/2 \\ 0 \end{pmatrix}$, ou encore par $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ (on ne change pas la direction du vecteur en le multipliant par un scalaire non nul).

- de même au voisinage de $t = \pi/2$: à l'aide des formules $\sin(t + \pi/2) = \cos t$ et $\cos(t + \pi/2) = -\sin t$, on trouve

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + (t - \frac{\pi}{2})^2 \begin{pmatrix} 0 \\ -3/2 \end{pmatrix} + (t - \frac{\pi}{2})^3 \begin{pmatrix} 1/6 \\ 0 \end{pmatrix} + o((t - \frac{\pi}{2})^3),$$

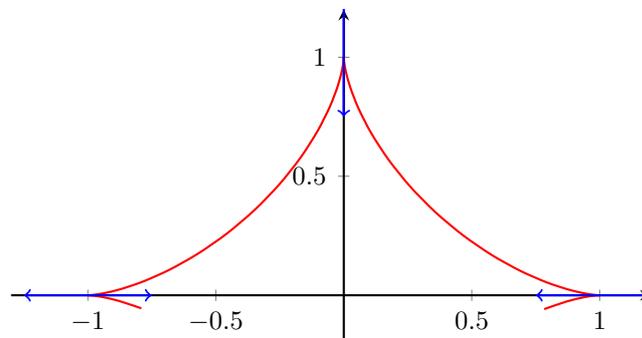
donc un point de rebroussement de première espèce, de vecteur tangent $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

- de même au voisinage de $t = \pi$: $\cos(t + \pi) = -\cos t$ et $\sin(t + \pi) = -\sin t$ donc

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} + (t - \pi)^2 \begin{pmatrix} 3/2 \\ 0 \end{pmatrix} + (t - \pi)^3 \begin{pmatrix} 0 \\ 1/6 \end{pmatrix} + o((t - \pi)^3),$$

donc un point de rebroussement de première espèce, de vecteur tangent $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

2.2.4. *Première ébauche de tracé.* On place tous les points et vecteurs tangents calculés aux étapes précédentes, puis on trace la courbe en suivant les tangentes et en respectant les variations des fonctions x et y :



L'objectif de ce tracé est de repérer toutes les informations qui pourraient manquer pour un tracé définitif. On s'intéresse en particulier aux éventuels points doubles (la

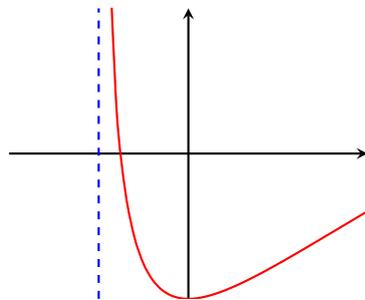
courbe « se croise elle-même »), points d'inflexion et branches infinies, mais on peut parfois souhaiter pour une raison ou une autre connaître les coordonnées d'un point spécifique, la direction d'une tangente, ou d'autres choses.

2.2.5. *Points doubles, points d'inflexion, branches infinies.* Pour trouver un point double (si le premier tracé en laisse soupçonner un), on cherche deux instants $t_1 \neq t_2 \in I$ tels que $M(t_1) = M(t_2)$. On peut alors calculer les vecteurs tangents en t_1 et t_2 pour déterminer l'aspect de la courbe au voisinage du point double.

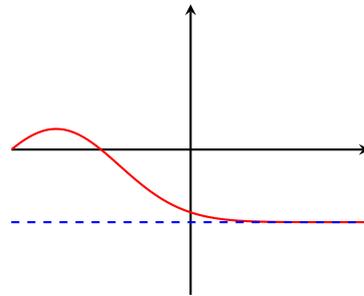
Pour rechercher un point d'inflexion, il faut trouver un instant t tel que $F'(t)$ et $F''(t)$ soient colinéaires.

On a une *branche infinie* en t_0 si $x(t)$ ou $y(t)$ tend vers $\pm\infty$ au voisinage de $t = t_0$. Plusieurs cas sont alors possibles :

- Si $x(t) \rightarrow \pm\infty$ mais $y(t) \rightarrow \ell \in \mathbf{R}$, la courbe admet une *asymptote horizontale* d'équation $y = \ell$. De même avec une *asymptote verticale* si $x(t) \rightarrow \ell$ et $y(t) \rightarrow \pm\infty$.

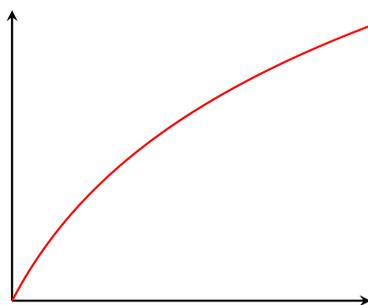


Asymptote verticale

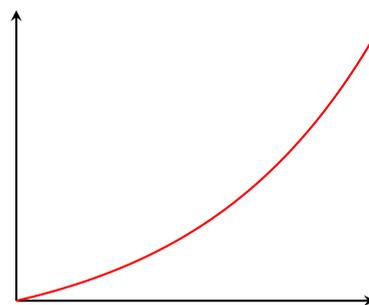


Asymptote horizontale

- Si $x(t)$ et $y(t)$ tendent tous les deux vers $\pm\infty$, on étudie le quotient $\frac{y(t)}{x(t)}$. S'il tend vers 0, on a une *branche parabolique* de direction (Ox) . S'il tend vers $\pm\infty$, on a une *branche parabolique* de direction (Oy) .

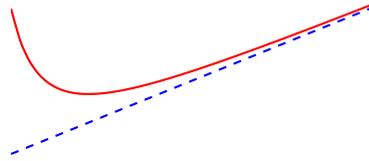


*Branche parabolique
de direction (Ox)*

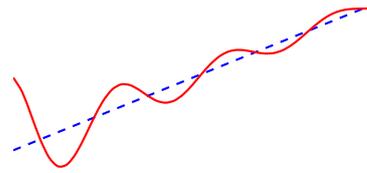


*Branche parabolique
de direction (Oy)*

- Enfin, si $y(t)$ et $x(t)$ tendent vers $\pm\infty$ mais que le quotient $\frac{y(t)}{x(t)}$ admet une limite finie non nulle a , on étudie la limite de $y(t) - ax(t)$. Si celle-ci est également finie égale à b , alors la droite d'équation $y = ax + b$ est une asymptote de la courbe. On peut alors étudier le signe de $y(t) - ax(t) - b$ pour déterminer la *position relative* de la courbe par rapport à l'asymptote (au-dessus, en-dessous, croisements...).

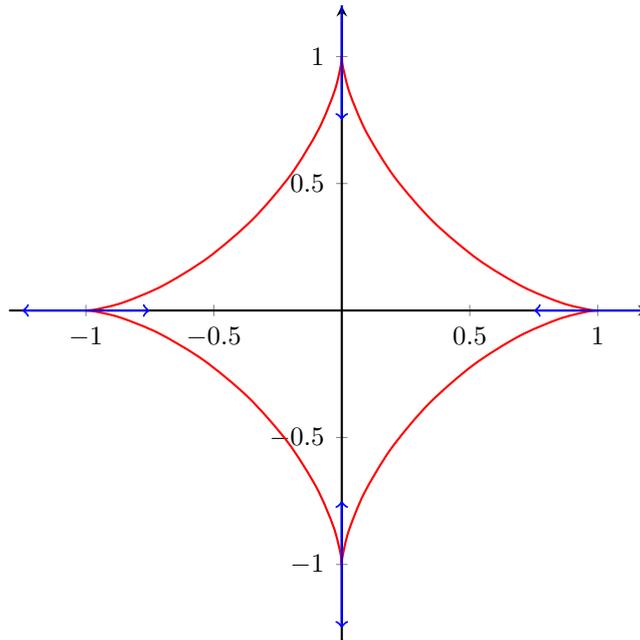


Asymptote oblique,
 $y(t) \geq ax(t) + b$



Asymptote oblique
signe de $y(t) - ax(t) - b$
non constant

2.2.6. *Tracé définitif.* On recommence le tracé, en tenant compte le cas échéant de toutes les informations supplémentaires récoltées à l'étape précédente. Enfin, on complète le tracé à l'aide des symétries établies à la première étape.



2.3. Longueur d'une courbe.

PROPOSITION 2.4. *La longueur de l'arc paramétré par F entre les instants t_1 et t_2 est donnée par*

$$L = \int_{t_1}^{t_2} \|F'(t)\| dt.$$

Réduction des endomorphismes

« Quelque chose m'échappe concernant l'équation de Schrödinger : pourquoi lorsque l'opérateur hamiltonien est indépendant du temps (le système est donc conservatif) on se limite à l'étude de l'équation aux valeurs propres $H\Psi = E\Psi$? »

JORISJOJO, post sur le forum [jeuxvideo.com](https://www.jeuxvideo.com) (2020)

Dans tout le chapitre, $\mathbf{K} = \mathbf{R}$ ou \mathbf{C} .

1. Valeurs propres, vecteurs propres et polynôme caractéristique

1.1. Valeurs propres et vecteurs propres.

DÉFINITION 1.1. Soit $A \in \mathcal{M}_n(\mathbf{K})$. Soit $X \in \mathbf{K}^n \setminus \{0\}$. On dit que X est un *vecteur propre* de la matrice A s'il existe $\lambda \in \mathbf{K}$ tel que $AX = \lambda X$. Dans ce cas, le scalaire λ est appelé une *valeur propre* de la matrice A associée au vecteur propre X .

EXEMPLE(S). Si $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$, montrer que $X_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ et $X_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ sont des vecteurs propres de A et déterminer les valeurs propres correspondantes. Montrer que $Y = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ n'est pas un vecteur propre.

REMARQUE. Dire que X est un vecteur propre de A revient à dire que la droite $\mathbf{K}X$ est stable par A .

DÉFINITION 1.2. L'ensemble des valeurs propres de A est appelé *spectre* de A , noté $\text{Sp}(A)$.

DÉFINITION 1.3. Montrer que $\text{Sp} \left(\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \right) = \{1, 2\}$.

DÉFINITION 1.4. Soit $\lambda \in \text{Sp}(A)$. On appelle *sous-espace propre* de A associé à la valeur propre λ l'ensemble $E_\lambda = \{X \in \mathbf{R}^n \mid AX = \lambda X\}$. C'est un sous-espace vectoriel de \mathbf{K}^n .

EXEMPLE(S). Soit $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$. Déterminer les sous-espaces propres de A .

REMARQUE. Les éléments de E_λ sont tous les vecteurs propres de A associés à la valeur propre λ , ainsi que le vecteur nul (qui n'est pas un vecteur propre par définition).

DÉMONSTRATION. $E_\lambda = \text{Ker } \varphi$, où φ est l'endomorphisme de \mathbf{K}^n défini par $\varphi(X) = AX - \lambda X$. □

PROPOSITION 1.5. Soit $A \in \mathcal{M}_n(\mathbf{K})$. Soient $\lambda_1, \dots, \lambda_p$ des valeurs propres deux à deux distinctes de A . Alors les sous-espaces propres $E_{\lambda_1}, \dots, E_{\lambda_p}$ sont en somme directe.

DÉMONSTRATION. On procède par récurrence sur p , le nombre de sous-espaces propres considérés.

Le résultat est trivial pour $p = 0$ ou 1 . Pour $p = 2$, il s'agit d'une somme de deux espaces donc il suffit de montrer que $E_{\lambda_1} \cap E_{\lambda_2} = \{0\}$. Soit donc $X \in E_{\lambda_1} \cap E_{\lambda_2}$. Alors $AX = \lambda_1 X = \lambda_2 X$, donc $(\lambda_2 - \lambda_1)X = 0$, mais $\lambda_2 \neq \lambda_1$ donc $X = 0$. Ainsi E_{λ_1} et E_{λ_2} sont en somme directe.

Soit $p \in \mathbf{N}$. On suppose que toute somme de p sous-espaces propres associés à des valeurs propres distinctes est directe. On considère $p + 1$ valeurs propres deux à deux distinctes $\lambda_1, \dots, \lambda_{p+1}$. Soient $X_1 \in E_{\lambda_1}, \dots, X_{p+1} \in E_{\lambda_{p+1}}$. On suppose que $X_1 + \dots + X_{p+1} = 0$. Alors $A(X_1 + \dots + X_{p+1}) = 0$, c'est-à-dire $AX_1 + \dots + AX_{p+1} = 0$, ou encore $\lambda_1 X_1 + \dots + \lambda_{p+1} X_{p+1} = 0$.

En multipliant la première ligne par λ_{p+1} , on obtient $\lambda_{p+1} X_1 + \dots + \lambda_{p+1} X_{p+1} = 0$, qu'on peut soustraire à la deuxième ligne : on a donc $(\lambda_{p+1} - \lambda_1)X_1 + \dots + (\lambda_{p+1} - \lambda_p)X_p = 0$. Par hypothèse de récurrence, tous les $(\lambda_{p+1} - \lambda_i)X_i$ sont nuls, mais comme le nombre $\lambda_{p+1} - \lambda_i$ n'est pas nul car les valeurs propres sont deux à deux distinctes, on en déduit que $X_1 = \dots = X_p = 0$. La première ligne devient alors $X_{p+1} = 0$, ce qui achève la démonstration. \square

1.2. Polynôme caractéristique.

REMARQUE. Soient $A \in \mathcal{M}_n(\mathbf{K})$, $X \in \mathbf{K}^n \setminus \{0\}$ et $\lambda \in \mathbf{K}$. Alors

$$\begin{aligned} X \text{ valeur propre de } A \text{ associé à la valeur propre } \lambda &\iff AX = \lambda X \\ &\iff \lambda X - AX = 0 \\ &\iff (\lambda I_n - A)X = 0 \\ &\iff X \in \text{Ker}(\lambda I_n - A) \end{aligned}$$

Ainsi, λ est valeur propre de A si et seulement si $\text{Ker}(\lambda I_n - A) \neq \{0\}$, autrement dit si $\lambda I_n - A$ n'est pas inversible. En utilisant le déterminant, on obtient la caractérisation suivante :

PROPOSITION 1.6. Soit $A \in \mathcal{M}_n(\mathbf{K})$. Soit $\lambda \in \mathbf{K}$. Alors

$$\lambda \in \text{Sp}(A) \iff \det(\lambda I_n - A) = 0.$$

DÉFINITION 1.7. Soit $A \in \mathcal{M}_n(\mathbf{K})$. On appelle *polynôme caractéristique* de A la fonction

$$\chi_A: \mathbf{K} \rightarrow \mathbf{K}, \quad \lambda \mapsto \det(\lambda I_n - A).$$

REMARQUE. Les valeurs propres de A sont donc exactement les racines de χ_A . La multiplicité d'une racine de χ_A est appelée *multiplicité* de la valeur propre.

EXEMPLE(S). Soit $A = \begin{pmatrix} 1 & \\ & \frac{1}{3} \end{pmatrix}$. Calculer le polynôme caractéristique de A et en déduire son spectre.

Même question avec $B = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 1 & 0 \\ 2 & 0 & 2 \end{pmatrix}$.

PROPOSITION 1.8. *Si $A \in \mathcal{M}_n(\mathbf{K})$, alors χ_A est un polynôme unitaire de degré n à coefficients dans \mathbf{K} .*

DÉMONSTRATION. Par récurrence sur n , en développant par rapport à la première colonne. \square

REMARQUE. En particulier, une matrice de taille n possède au maximum n valeurs propres.

1.3. Version géométrique. Tout ce qui vient d'être dit pour les matrices carrées de taille n se retranscrit immédiatement pour les endomorphismes d'un espace E de dimension n . Plus précisément, si $u \in \mathcal{L}(E)$:

- un *vecteur propre* de u est un vecteur $x \in E$ non nul tel qu'il existe $\lambda \in \mathbf{K}$ vérifiant $u(x) = \lambda x$; le scalaire λ est appelé *valeur propre* de u ; le spectre de u est l'ensemble de ses valeurs propres.
- si λ est une valeur propre, le *sous-espace propre* associé à λ est $E_\lambda = \{x \in E \mid u(x) = \lambda x\}$. Les sous-espaces propres associés à des valeurs propres distinctes sont en somme directe.
- le *polynôme caractéristique* de u est l'application définie sur \mathbf{K} par $\chi_u(\lambda) = \det(\lambda \text{id} - u)$; c'est un polynôme unitaire de degré n dont les racines sont exactement les valeurs propres de u .

EXEMPLE(S). Dans $E = \mathbf{R}_2[X]$, on considère l'endomorphisme $\phi: P \mapsto (X+1)P' + P$. Déterminer les valeurs propres de ϕ et les sous-espaces propres correspondants.

2. Diagonalisation

2.1. Matrices diagonalisables.

DÉFINITION 2.1. Soit $A \in \mathcal{M}_n(\mathbf{K})$. On dit que A est *diagonalisable* s'il existe $P \in \text{GL}_n(\mathbf{K})$ et $D \in \mathcal{M}_n(\mathbf{K})$ diagonale telle que $A = PDP^{-1}$.

Diagonaliser une matrice, c'est donner (le cas échéant) les matrices P et D .

REMARQUE. Soit $A \in \mathcal{M}_n(\mathbf{K})$. On suppose A diagonalisable. Alors les colonnes de P sont des vecteurs propres de A , et les coefficients diagonaux de la matrice D les valeurs propres correspondantes. En effet, si e_i désigne le i -ème vecteur de la base canonique de \mathbf{K}^n , alors Pe_i est la i -ème colonne de P .

THÉORÈME 2.2. *Soit $A \in \mathcal{M}_n(\mathbf{K})$. Alors A est diagonalisable si et seulement s'il existe une base \mathcal{B} de \mathbf{K}^n dont les vecteurs sont tous propres pour A . Dans ce cas, on peut écrire $A = PDP^{-1}$, où les colonnes de P sont les vecteurs de la base \mathcal{B} , et D est la matrice diagonale dont les coefficients sont les valeurs propres correspondantes.*

DÉMONSTRATION.

\square

EXEMPLE(S). Dire si les matrices suivantes sont diagonalisables, les diagonaliser le cas échéant :

$$A = \begin{pmatrix} 0 & 2 & -1 \\ 3 & -2 & 0 \\ -2 & 2 & 1 \end{pmatrix}; \quad B = \begin{pmatrix} 0 & 3 & 2 \\ -2 & 5 & 2 \\ 2 & -3 & 0 \end{pmatrix}; \quad C = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 3 & 3 \\ 0 & 0 & 3 \end{pmatrix}.$$

PROPOSITION 2.3. Soit $A \in \mathcal{M}_n(\mathbf{K})$ et E_1, \dots, E_p ses sous-espaces propres. Alors A est diagonalisable si et seulement si $E_1 \oplus E_2 \oplus \dots \oplus E_p = \mathbf{K}^n$.

DÉMONSTRATION. □

REMARQUE. On sait que E_1, \dots, E_p sont en somme directe. D'après la formule de Grassmann, A est donc diagonalisable ssi $\dim E_1 + \dots + \dim E_p = n$.

PROPOSITION 2.4. Soit $A \in \mathcal{M}_n(\mathbf{K})$. Soit λ une valeur propre de A , on note m_λ sa multiplicité et d_λ la dimension du sous-espace propre associé. Alors $1 \leq d_\lambda \leq m_\lambda \leq n$.

THÉORÈME 2.5. Soit $A \in \mathcal{M}_n(\mathbf{K})$. Alors A est diagonalisable si et seulement si χ_A est scindé et si, pour tout $\lambda \in \text{Sp}(A)$, la multiplicité de λ est égale à la dimension du sous-espace propre associé.

EXEMPLE(S). Diagonaliser $C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 2 \end{pmatrix}$.

DÉMONSTRATION. □

COROLLAIRE 2.6. Soit $A \in \mathcal{M}_n(\mathbf{K})$. Si χ_A est scindé à racines simples, alors A est diagonalisable.

DÉMONSTRATION. □

EXEMPLE(S). Soit $A = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. La matrice A est-elle diagonalisable ? Si oui, la diagonaliser.

2.2. Trigonalisation.

DÉFINITION 2.7. Une matrice $A \in \mathcal{M}_n(\mathbf{K})$ est dite *trigonalisable* s'il existe deux matrices $P \in \text{GL}_n(\mathbf{K})$ et $T \in \mathcal{M}_n(\mathbf{K})$ triangulaire telles que $A = PTP^{-1}$.

REMARQUE. Une matrice diagonale étant un cas particulier de matrice triangulaire, toute matrice diagonalisable est trigonalisable.

THÉORÈME 2.8. Soit $A \in \mathcal{M}_n(\mathbf{K})$. Alors A est trigonalisable si et seulement si son polynôme caractéristique est scindé. Dans ce cas, les coefficients diagonaux de la matrice T sont les valeurs propres de A , comptées avec multiplicité.

ADMIS. □

EXEMPLE(S). Trigonaliser la matrice $A = \begin{pmatrix} 1 & 4 & -2 \\ 0 & 6 & -3 \\ -1 & 4 & 0 \end{pmatrix}$.

COROLLAIRE 2.9. *Toute matrice à coefficients complexes est trigonalisable.*

DÉMONSTRATION. Comme tout polynôme, le polynôme caractéristique est scindé sur \mathbf{C} . \square

REMARQUE. Toute matrice à coefficients réels est trigonalisable, à condition d'autoriser les valeurs propres complexes.

EXEMPLE(S). Montrer que la matrice suivante est trigonalisable sur \mathbf{C} mais pas sur \mathbf{R} :

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

PROPOSITION 2.10. *Soit $A \in \mathcal{M}_n(\mathbf{C})$. Soient $\lambda_1, \dots, \lambda_n$ les valeurs propres (comptées avec multiplicité) de A . Alors*

$$\sum_{i=1}^n \lambda_i = \text{Tr } A; \quad \prod_{i=1}^n \lambda_i = \det A.$$

DÉMONSTRATION. \square

EXEMPLE(S). Soit $A = \frac{1}{3} \begin{pmatrix} 2 & -1 & 2 \\ 2 & 2 & -1 \\ -1 & 2 & 2 \end{pmatrix}$. Par des arguments géométriques, on

montre que A est semblable à une matrice de la forme $\begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Que dire de θ ?

2.3. Endomorphismes diagonalisables et trigonalisables. Tous les énoncés des parties précédentes se retranscrivent naturellement pour les endomorphismes d'un \mathbf{K} -espace vectoriel E de dimension finie n .

- f est *diagonalisable* si et seulement s'il existe une base \mathcal{B} de E telle que la matrice de f dans la base \mathcal{B} soit diagonale. Autrement dit, une base de vecteurs propres. *Diagonaliser f* , c'est exhiber une telle base, et les valeurs propres correspondantes.
- L'endomorphisme f est diagonalisable si et seulement si la somme de ses sous-espaces propres est égale à E . Pour cela, il faut et il suffit que la dimension de chaque sous-espace propre soit égal à la multiplicité de la valeur propre correspondante. En particulier, si toutes les valeurs propres de f sont simples, alors f est diagonalisable.
- L'endomorphisme f est trigonalisable si et seulement s'il existe une base \mathcal{B} de E telle que la matrice de f dans \mathcal{B} soit triangulaire; c'est vrai si et seulement si le polynôme caractéristique de f est scindé.

- La somme des valeurs propres complexes de f (comptées avec multiplicité) est égale à sa trace ; leur produit est égal au déterminant.

3. Applications

3.1. Calcul des puissances d'une matrice. Soient $A \in \mathcal{M}_n(\mathbf{K})$ et $n \in \mathbf{N}$. On souhaite calculer A^n . Si A est diagonalisable, on la diagonalise sous la forme $A = PDP^{-1}$, et dans ce cas

$$A^n = PD \underbrace{P^{-1}P}_{I_n} DP^{-1} \dots PDP^{-1} = PD^n P^{-1}$$

Une application est l'étude des systèmes de suites vérifiant une relation de récurrence linéaire à coefficients constants. Supposons deux suites $(u_n)_{n \in \mathbf{N}}$ et $(v_n)_{n \in \mathbf{N}}$ telles que, pour tout $n \in \mathbf{N}$,

$$\begin{cases} u_{n+1} = au_n + bv_n \\ v_{n+1} = cu_n + dv_n \end{cases}$$

Pour tout $n \in \mathbf{N}$, on peut poser $X_n = \begin{pmatrix} u_n \\ v_n \end{pmatrix}$, et on aura dans ce cas, pour tout $n \in \mathbf{N}$, $X_{n+1} = AX_n$, où $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. On en déduit, pour tout $n \in \mathbf{N}$, $X_n = A^n X_0$. Si la matrice A est diagonalisable, on peut facilement calculer la matrice A^n et en déduire l'expression des termes généraux de $(u_n)_{n \in \mathbf{N}}$ et $(v_n)_{n \in \mathbf{N}}$.

3.2. Suites vérifiant une relation de récurrence linéaire d'ordre deux à coefficients constants. Soient $a, b \in \mathbf{C}$. On s'intéresse à l'ensemble E des suites $(u_n)_{n \in \mathbf{N}}$ complexes vérifiant la relation de récurrence $(R): u_{n+2} = au_{n+1} + bu_n$ pour tout $n \in \mathbf{N}$.

Une première approche est de considérer la suite $(v_n)_{n \in \mathbf{N}}$ définie pour tout $n \in \mathbf{N}$ par $v_n = u_{n+1}$. Dans ce cas, (R) se réécrit

$$\begin{cases} u_{n+1} = v_n \\ v_{n+1} = av_n + bu_n \end{cases}$$

et, en posant $X_n = \begin{pmatrix} u_n \\ v_n \end{pmatrix}$, on se ramène ainsi à la méthode esquissée au paragraphe précédent. Notons que dans ce cas, la matrice A s'écrit $A = \begin{pmatrix} 0 & 1 \\ b & a \end{pmatrix}$, donc son polynôme caractéristique est $\chi_A(\lambda) = \lambda^2 - a\lambda - b$. La question de la diagonalisabilité de A se ramène ainsi à l'étude des solutions de l'équation $\lambda^2 = a\lambda + b$.

Une autre approche consiste à s'intéresser à la structure de l'ensemble E .

PROPOSITION 3.1. *L'ensemble E est un espace vectoriel de dimension 2.*

DÉMONSTRATION. E est inclus dans l'ensemble des suites complexes, qui est un \mathbf{C} -espace vectoriel. Il suffit donc de montrer que E est un sous-espace vectoriel.

- La suite constamment nulle vérifie la relation (R) .
- On vérifie facilement par le calcul que la combinaison linéaire de deux suites vérifiant la relation (R) vérifie également la relation (R) .

L'ensemble E est donc bien un espace vectoriel. Reste à montrer qu'il est de dimension deux. Pour cela, on considère les deux suites $(a_n)_{n \in \mathbf{N}}$ et $(b_n)_{n \in \mathbf{N}}$, définies par $a_0 = 1$, $a_1 = 0$, $b_0 = 0$, $b_1 = 1$ et la relation de récurrence (R) . On montre sans difficulté que $((a_n)_{n \in \mathbf{N}}, (b_n)_{n \in \mathbf{N}})$ est une famille libre et génératrice de E . C'est donc une base, et E est donc de dimension deux. \square

PROPOSITION 3.2. *Si $\lambda \in \mathbf{C}$ vérifie l'équation $\lambda^2 = a\lambda + b$, alors la suite $(\lambda^n)_{n \in \mathbf{N}}$ appartient à E .*

Si $\lambda \in \mathbf{C}$ est une racine double de l'équation $\lambda^2 = a\lambda + b$, alors la suite $(n\lambda^n)_{n \in \mathbf{N}}$ appartient à (E) .

DÉMONSTRATION. Il suffit de vérifier par un simple calcul que ces suites vérifient la relation (R) . Dans le deuxième cas, noter que λ est une racine double, donc $\lambda = \frac{a}{2}$ et $b = -\frac{a^2}{4}$. \square

THÉORÈME 3.3. *Si l'équation $\lambda^2 = a\lambda + b$ admet deux solutions complexes λ_1 et λ_2 , alors, pour toute suite $(u_n)_{n \in \mathbf{N}}$ de E , il existe $A, B \in \mathbf{C}$ tels que, pour tout $n \in \mathbf{N}$, $u_n = A\lambda_1^n + B\lambda_2^n$.*

Si l'équation $\lambda^2 = a\lambda + b$ admet une unique solution λ , alors, pour toute suite $(u_n)_{n \in \mathbf{N}}$ de E , il existe $A, B \in \mathbf{C}$ tels que, pour tout $n \in \mathbf{N}$, $u_n = (An + B)\lambda^n$.

DÉMONSTRATION. C'est une conséquence des deux propositions précédentes. Dans le premier cas, $(\lambda_1^n)_{n \in \mathbf{N}}$ et $(\lambda_2^n)_{n \in \mathbf{N}}$ forment une famille libre de E , donc une base, donc tout élément de E s'écrit comme combinaison linéaire de ces deux suites.

Dans le deuxième cas, la base est formée par les suites $(\lambda^n)_{n \in \mathbf{N}}$ et $(n\lambda^n)_{n \in \mathbf{N}}$. \square

REMARQUE. Les coefficients A et B se déterminent à l'aide des conditions initiales, c'est-à-dire (en général) des valeurs de u_0 et u_1 .

PROPOSITION 3.4. *Dans le cas particulier où $a, b \in \mathbf{R}$ et $\lambda^2 = a\lambda + b$ admet deux solutions complexes conjuguées $\lambda = \rho e^{\pm i\theta}$, pour toute suite réelle $(u_n)_{n \in \mathbf{N}}$ de E , il existe $A, B \in \mathbf{R}$ tels que, pour tout $n \in \mathbf{N}$, $u_n = \rho^n (A \cos(n\theta) + B \sin(n\theta))$.*

REMARQUE. On peut aussi utiliser la forme alternative $u_n = A\rho^n \cos(n\theta + \varphi)$, où A et φ sont les deux constantes à déterminer.

Intégrale généralisée

« Par cette révolution, il se forme une espèce de fuseau infiniment long, et cependant Torricelli qui lui a donné ce nom, a démontré évidemment qu'il est égal à un solide ou corps fini. »

J. D'ALEMBERT, *Encyclopédie* (1751)

Dans tout le chapitre, on notera $\bar{\mathbf{R}} = \mathbf{R} \cup \{\pm\infty\}$.

1. Généralités

DÉFINITION 1.1. Soient $a < b \in \bar{\mathbf{R}}$. On appelle *intégrale impropre* toute intégrale de la forme $\int_a^b f(t) dt$, où f est une fonction continue sur $]a; b[$ mais pas sur $[a; b]$ tout entier. On dit que l'intégrale admet une *impropreté* en a (reps. en b) si la fonction f n'est pas continue en a (resp. en b).

REMARQUE. C'est en particulier le cas dès que a ou b est une borne infinie.

EXEMPLE(S). Déterminer les impropretés des intégrales suivantes :

$$\int_0^{+\infty} e^{-t} dt; \quad \int_0^1 \ln t dt; \quad \int_0^{+\infty} \frac{dt}{t}.$$

DÉFINITION 1.2. On considère une intégrale impropre $\int_a^b f(t) dt$.

- Si l'unique impropreté est en b , on dira que l'intégrale est *convergente* si la fonction $x \mapsto \int_a^x f(t) dt$ admet une limite finie lorsque x tend vers b par valeurs inférieures. Dans ce cas, on notera

$$\int_a^b f(t) dt = \lim_{x \rightarrow b, x < b} \int_a^x f(t) dt.$$

- De même, si l'unique impropreté est en a , on dira que l'intégrale est *convergente* si la fonction $x \mapsto \int_x^b f(t) dt$ admet une limite finie lorsque x tend vers a par valeurs supérieures. Dans ce cas, on notera

$$\int_a^b f(t) dt = \lim_{x \rightarrow a, x > a} \int_x^b f(t) dt.$$

- Enfin, si l'intégrale est impropre en a et en b , on choisit $c \in]a; b[$, et on dit que l'intégrale est *convergente* si et seulement si les deux intégrales $\int_a^c f(t) dt$ et

$\int_c^b f(t) dt$ (qui possèdent chacune une seule impropreté) sont convergentes. Dans ce cas, on notera

$$\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt.$$

Une intégrale impropre non convergente est dite *divergente*.

REMARQUE. Dans le troisième cas, on peut montrer facilement (d'après la relation de Chasles) que le résultat ne dépend pas du choix de c .

EXEMPLE(S). Donner la nature (et, le cas échéant, la valeur) des intégrales impropres de l'exemple précédent.

Dans toute la suite, on étudiera uniquement le cas des intégrales $\int_a^b f(t) dt$ avec une seule impropreté en b , mais tous les théorèmes se généralisent de manière évidente au cas où l'impropreté est en a (et, par découpage, au cas de double impropreté).

PROPOSITION 1.3. Soit $\int_a^b f(t) dt$ une intégrale admettant une impropreté en b . On suppose que $b \neq +\infty$ et que la fonction f admet une limite finie en b . Alors l'intégrale est convergente (on dira qu'elle est faussement impropre).

DÉMONSTRATION. La fonction f est continue sur $[a; b[$ et possède une limite finie en b , on peut donc la prolonger en une fonction g continue sur $[a; b]$. Alors, pour tout $x \in [a; b[$,

$$\int_a^x f(t) dt = \int_a^x g(t) dt \xrightarrow{x \rightarrow b} \int_a^b g(t) dt \in \mathbf{R},$$

l'intégrale est donc convergente. □

EXEMPLE(S). Étudier la convergence de l'intégrale $\int_0^{\pi/2} \frac{\sin t}{t} dt$.

THÉORÈME 1.4. (Intégrales de référence)

- l'intégrale $\int_0^{+\infty} e^{\alpha t} dt$ est convergente si et seulement si $\alpha < 0$;
- l'intégrale $\int_1^{+\infty} \frac{dt}{t^\alpha}$ est convergente si et seulement si $\alpha > 1$;
- l'intégrale $\int_0^1 \frac{dt}{t^\alpha}$ est convergente si et seulement si $\alpha < 1$.

2. Intégrale généralisée des fonctions positives

Dans cette partie, on se restreint au cas où la fonction f est à valeurs dans \mathbf{R}_+ . Dans ce cas, la fonction $x \mapsto \int_a^x f(t) dt$ est croissante ; d'après le théorème de la limite monotone, elle admet donc une limite finie en b si et seulement si elle est majorée. On en déduit les théorèmes de comparaison :

THÉORÈME 2.1 (Théorème de comparaison, première version). Soient f, g deux fonctions continues et positives sur $[a; b[$. On suppose que, pour tout $t \in [a; b[$, $f(t) \leq g(t)$. Alors :

- si $\int_a^b f(t) dt$ est divergente, alors $\int_a^b g(t) dt$ est divergente.
- si $\int_a^b g(t) dt$ est convergente, alors $\int_a^b f(t) dt$ est convergente. De plus, $\int_a^b f(t) dt \leq \int_a^b g(t) dt$.

REMARQUE. Le théorème reste valable si l'inégalité est vraie uniquement au voisinage de b .

EXEMPLE(S). Étudier la convergence de l'intégrale $\int_0^{+\infty} e^{-t} \sin^2 t dt$.

THÉORÈME 2.2 (Théorème de comparaison, deuxième version). Soient f, g deux fonctions continues et positives sur $[a; b[$. On suppose que, pour t tendant vers b , $f(t) = o(g(t))$. Alors :

- si $\int_a^b f(t) dt$ est divergente, alors $\int_a^b g(t) dt$ est divergente.
- si $\int_a^b g(t) dt$ est convergente, alors $\int_a^b f(t) dt$ est convergente.

EXEMPLE(S). Étudier la convergence de l'intégrale $\int_0^{+\infty} e^{-t^2} dt$.

THÉORÈME 2.3 (Théorème de comparaison, troisième version). Soient f, g deux fonctions continues et positives sur $[a; b[$. On suppose que, pour t tendant vers b , $f(t) \sim g(t)$. Alors les intégrales $\int_a^b f(t) dt$ et $\int_a^b g(t) dt$ sont de même nature.

EXEMPLE(S). Étudier la convergence de l'intégrale $\int_1^{+\infty} \sin\left(\frac{1}{t}\right) dt$

THÉORÈME 2.4 (Théorème de changement de variables). Soient $f:]a; b[\rightarrow \mathbf{R}$ et $\varphi:]\alpha; \beta \rightarrow]a; b[$, de classe \mathcal{C}^1 et strictement croissante. Alors les intégrales

$$\int_a^b f(t) dt \quad \text{et} \quad \int_\alpha^\beta f \circ \varphi(t) \varphi'(t) dt$$

sont de même nature ; si elles convergent, on a alors

$$\int_a^b f(t) dt = \int_\alpha^\beta f \circ \varphi(t) \varphi'(t) dt.$$

REMARQUE. Le théorème reste vrai pour une fonction φ strictement décroissante, à condition de replacer les bornes dans le bon sens. En revanche, il cesse d'être vrai en général si la fonction φ n'est pas strictement monotone.

EXEMPLE(S). Prouver la convergence et calculer la valeur de $\int_0^{+\infty} \frac{dx}{x\sqrt{1+x^2}}$.

PROPOSITION 2.5 (Comparaison série-intégrale). Soit f une fonction continue, positive et décroissante sur \mathbf{R}_+ . Alors la série $\sum f(n)$ et l'intégrale $\int_0^{+\infty} f(t) dt$ sont de même nature.

EXEMPLE(S). Étudier la convergence de $\int_1^{+\infty} \frac{dt}{t \ln t}$. En déduire la nature de la série (dite série de Bertrand) $\sum \frac{1}{n \ln n}$.

3. Fonctions intégrables

DÉFINITION 3.1. Soit f une fonction continue sur un intervalle $I =]a; b[$ ($a, b \in \bar{\mathbf{R}}$). On dit que f est *intégrable* sur I si l'intégrale impropre $\int_a^b |f(t)| dt$ est convergente.

THÉORÈME 3.2. Si f est intégrable sur $]a; b[$, alors l'intégrale impropre $\int_a^b f(t) dt$ est convergente. De plus,

$$\left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt.$$

ADMIS. □

REMARQUE. Si la fonction f est intégrable sur $]a; b[$, on dit parfois que $\int_a^b f(t) dt$ est *absolument convergente*.

EXEMPLE(S). Étudier la convergence des intégrales $\int_0^{+\infty} \frac{\sin t}{(1+t^2)\sqrt{t}} dt$ et $\int_0^1 \sin\left(\frac{1}{t}\right) dt$.

REMARQUE. La réciproque n'est pas vraie : une fonction peut avoir une intégrale impropre qui converge, sans être intégrable pour autant !

EXEMPLE(S). Montrer que $\int_0^{+\infty} \frac{\sin t}{t} dt$ est convergente, mais que la fonction $t \mapsto \frac{\sin t}{t}$ n'est pas intégrable sur \mathbf{R}_+ .

PROPOSITION 3.3. L'ensemble des fonctions continues intégrables sur I est un sous-espace vectoriel de l'ensemble des fonctions continues sur I .

L'application $f \mapsto \int_a^b f(t) dt$ définit une forme linéaire sur cet espace. Autrement dit, si f et g sont deux fonctions intégrables sur I et si $\lambda, \mu \in \mathbf{R}$, alors $\lambda f + \mu g$ est intégrable sur I et

$$\int_a^b (\lambda f(t) + \mu g(t)) dt = \lambda \int_a^b f(t) dt + \mu \int_a^b g(t) dt.$$

Probabilités sur un univers dénombrable

« Un exemple frappant est le nombre de soldats tués chaque année par des coups de sabot de cheval dans les corps d'armée prussiens. »

E. J. GUMBEL, *L. von Bortkiewicz* (1931)

1. Généralités

1.1. Ensembles dénombrables.

DÉFINITION 1.1. Un ensemble E est dit *dénombrable* s'il existe une partie une bijection $\varphi: \mathbf{N} \rightarrow E$. Autrement dit, si on peut numéroter ses éléments par des nombres entiers.

EXEMPLE(S). \mathbf{N} et \mathbf{Z} sont dénombrables, \mathbf{R} et $\mathcal{P}(\mathbf{N})$ ne sont pas dénombrables.

REMARQUE. Une partie d'un ensemble dénombrable est finie ou dénombrable. Une union, une intersection, un produit cartésien (finis) d'ensembles dénombrables est (fini ou) dénombrable.

1.2. Langage des probabilités. Dans ce paragraphe (et dans ce chapitre), on ne supposera pas nécessairement l'univers Ω fini.

Si $(A_n)_{n \in \mathbf{N}}$ est une suite d'événements, alors

- $\cup_{n \in \mathbf{N}} A_n$ est l'événement qui se réalise si et seulement si au moins l'un des A_n se réalise;
- $\cap_{n \in \mathbf{N}} A_n$ est l'événement qui se réalise si et seulement si tous les A_n se réalisent.

DÉFINITION 1.2. Une famille finie ou dénombrables d'événements $(A_i)_{i \in I}$ forme un *système complet d'événements* (ou *partition* de Ω) si les A_i sont tous non vides, deux à deux incompatibles et recouvrent Ω . Autrement dit,

$$\forall i \in I \ A_i \neq \emptyset; \quad \forall i, j \in I \ (i \neq j \implies A_i \cap A_j = \emptyset); \quad \bigcup_{i \in I} A_i = \Omega.$$

1.3. Probabilités.

DÉFINITION 1.3. Ω étant un ensemble dénombrable, on appelle *probabilité* sur Ω tout application $\mathbf{P}: \mathcal{P}(\Omega) \rightarrow [0; 1]$ telle que $\mathbf{P}(\Omega) = 1$ et, pour toute suite $(A_n)_{n \in \mathbf{N}}$

d'événements deux à deux incompatibles, on ait

$$\mathbf{P}\left(\bigcup_{n \in \mathbf{N}} A_n\right) = \sum_{n=0}^{+\infty} \mathbf{P}(A_n).$$

(on dit alors que \mathbf{P} est σ -additive).

Dans ce cas, on dit que (Ω, \mathbf{P}) est un *espace probabilisé* dénombrable et, pour tout $A \in \mathcal{P}(\Omega)$, le nombre $\mathbf{P}(A)$ est appelé *probabilité* de A .

Dans toute la suite du chapitre, (Ω, \mathbf{P}) désigne un espace probabilisé dénombrable.

PROPOSITION 1.4. *Soient A et B deux événements. Alors :*

- (1) $\mathbf{P}(\emptyset) = 0$;
- (2) $\mathbf{P}(\bar{A}) = 1 - \mathbf{P}(A)$;
- (3) $\mathbf{P}(B \setminus A) = \mathbf{P}(B) - \mathbf{P}(A \cap B)$; en particulier, si $A \subset B$, alors $\mathbf{P}(B \setminus A) = \mathbf{P}(B) - \mathbf{P}(A)$;
- (4) si $A \subset B$, alors $\mathbf{P}(A) \leq \mathbf{P}(B)$;
- (5) $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$;
- (6) Si $\Omega = \{\omega_n \mid n \in \mathbf{N}\}$ et $(p_n)_{n \in \mathbf{N}}$ est une suite de réels positifs tels que $\sum p_n$ soit convergente et de somme 1, alors il existe une unique probabilité \mathbf{P} sur Ω telle que, pour tout $n \in \mathbf{N}$, $\mathbf{P}(\omega_n) = p_n$.

REMARQUE. Si l'univers Ω est infini, il est impossible que les événements élémentaires soient équiprobables.

2. Conditionnement et indépendance

DÉFINITION 2.1. Soient A et B deux événements. Si $\mathbf{P}(B) \neq 0$, on définit la *probabilité conditionnelle* de A sachant B par

$$\mathbf{P}(A \mid B) = \mathbf{P}_B(A) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}.$$

Ceci définit une probabilité \mathbf{P}_B sur Ω .

DÉMONSTRATION. \mathbf{P}_B définit bien une application $\mathcal{P}(\Omega) \rightarrow [0; 1]$ (car $\mathbf{P}(A \cap B) \leq \mathbf{P}(B)$).

De plus, si $(A_i)_{i \in \mathbf{N}}$ est une suite d'événements deux à deux incompatibles, alors $(B \cap A_i)_{i \in \mathbf{N}}$ est aussi une suite d'événements incompatibles, la série $\sum \mathbf{P}(B \cap A_i)$ est donc convergente, de somme $\mathbf{P}(\bigcup_{i \in \mathbf{N}} (B \cap A_i)) = \mathbf{P}(B \cap (\bigcup_{i \in \mathbf{N}} A_i))$.

Autrement dit, en divisant tout par $\mathbf{P}(B)$, on en déduit l'égalité $\mathbf{P}_B(\bigcup_{i \in \mathbf{N}} A_i) = \sum_{i \in \mathbf{N}} \mathbf{P}_B(A_i)$. Ceci démontre que \mathbf{P}_B est une probabilité. \square

PROPOSITION 2.2 (Formule des probabilités composées). *Soit A_1, \dots, A_n une famille d'événements telle que $\mathbf{P}(\bigcap_{i=1}^{n-1} A_i) \neq 0$. Alors*

$$\mathbf{P}\left(\bigcap_{i=1}^n A_i\right) = \mathbf{P}(A_1) \times \mathbf{P}(A_2 \mid A_1) \times \mathbf{P}(A_3 \mid A_2 \cap A_1) \times \cdots \times \mathbf{P}(A_n \mid A_1 \cap \cdots \cap A_{n-1}).$$

DÉMONSTRATION. □

REMARQUE. Si on a $A_n \subset A_{n-1} \subset \cdots \subset A_1$, alors la formule des probabilités composées devient :

$$\mathbf{P}\left(\bigcap_{i=1}^n A_i\right) = \mathbf{P}(A_1) \times \mathbf{P}(A_2 \mid A_1) \times \mathbf{P}(A_3 \mid A_2) \times \cdots \times \mathbf{P}(A_n \mid A_{n-1}).$$

PROPOSITION 2.3 (Formule des probabilités totales). *Soit $(A_i)_{1 \leq i \leq n}$ un système complet d'événements de probabilités non nulles. Alors, si B est un événement,*

$$\mathbf{P}(B) = \sum_{i=1}^n \mathbf{P}_{A_i}(B) \times \mathbf{P}(A_i).$$

En particulier, si $n = 2$,

$$\mathbf{P}(B) = \mathbf{P}_A(B) \times \mathbf{P}(A) + \mathbf{P}_{\bar{A}}(B) \times (1 - \mathbf{P}(A)).$$

DÉMONSTRATION. □

PROPOSITION 2.4 (Formule des probabilités totales, version infinie). *Soit $(A_n)_{n \in \mathbf{N}}$ un système complet d'événements non négligeables. Alors, si B est un événement, $\sum \mathbf{P}(B \cap A_n)$ converge et*

$$\mathbf{P}(B) = \sum_{n=0}^{+\infty} \mathbf{P}(B \cap A_n) = \sum_{n=0}^{+\infty} \mathbf{P}_{A_n}(B) \times \mathbf{P}(A_n).$$

DÉMONSTRATION. On peut écrire

$$B = B \cap \Omega = B \cap \left(\bigcup_{n \in \mathbf{N}} A_n \right) = \bigcup_{n \in \mathbf{N}} (B \cap A_n).$$

Les $B \cap A_n$ sont deux à deux incompatibles car les A_n le sont, donc par définition de la probabilité \mathbf{P} , on a bien $\mathbf{P}(B) = \sum_{n=0}^{+\infty} \mathbf{P}(B \cap A_n)$, on conclut par définition de la probabilité conditionnelle. □

PROPOSITION 2.5 (Formule de Bayes). *Si A et B sont deux événements de probabilité non nulle, alors*

$$\mathbf{P}_B(A) = \frac{\mathbf{P}_A(B) \times \mathbf{P}(A)}{\mathbf{P}(B)} = \frac{\mathbf{P}_A(B) \times \mathbf{P}(A)}{\mathbf{P}_A(B) \times \mathbf{P}(A) + \mathbf{P}_{\bar{A}}(B) \times (1 - \mathbf{P}(A))}$$

DÉMONSTRATION. □

DÉFINITION 2.6. Deux événements A et B sont dits *indépendants* si $\mathbf{P}(A \cap B) = \mathbf{P}(A) \times \mathbf{P}(B)$. On note (parfois) $A \perp B$.

REMARQUE. En d'autres termes, si A et B sont de probabilités non nulles,

$$A \perp B \iff \mathbf{P}_B(A) = \mathbf{P}(A) \iff \mathbf{P}_A(B) = \mathbf{P}(B).$$

PROPOSITION 2.7. *Si A et B sont indépendants, alors \bar{A} et B , \bar{A} et \bar{B} , A et \bar{B} le sont.*

DÉMONSTRATION. □

DÉFINITION 2.8. Les événements A_1, \dots, A_n sont *mutuellement indépendants* si, pour toute partie non vide I de $\llbracket 1; n \rrbracket$, on a

$$\mathbf{P} \left(\bigcap_{i \in I} A_i \right) = \prod_{i \in I} \mathbf{P}(A_i).$$

REMARQUE. On ne définit pas la notion d'indépendance mutuelle pour une famille infinie d'événements.

PROPOSITION 2.9. *Soit $(A_i)_{1 \leq i \leq n}$ une famille d'événements mutuellement indépendants. Alors ;*

- (1) *Toute sous-famille de $(A_i)_{1 \leq i \leq n}$ est une famille d'événements mutuellement indépendants.*
- (2) *Les A_1, \dots, A_n sont deux à deux indépendants.*
- (3) *La propriété d'indépendance mutuelle reste vraie si on remplace certains des A_i par leurs contraires.*

DÉMONSTRATION. □

REMARQUE. Si on a plus de deux événements, l'indépendance deux à deux n'implique pas nécessairement l'indépendance mutuelle.

3. Variables aléatoires discrètes

On rappelle qu'une variable aléatoire (réelle) est une application $X : \Omega \rightarrow \mathbf{R}$.

DÉFINITION 3.1. Une variable aléatoire sur un univers Ω fini ou dénombrable est dite *discrète*.

Dans cette partie, l'univers est supposé fini ou dénombrable et toutes les variables aléatoires sont donc supposées discrètes.

DÉFINITION 3.2. Si $X : \Omega \rightarrow \mathbf{R}$, on appelle *loi (de probabilité)* de X la donnée, pour tout $x \in X(\Omega)$, de la probabilité $\mathbf{P}(X = x)$.

REMARQUE. Ceci définit une probabilité \mathbf{P}_X sur $X(\Omega)$ en posant, pour tout $A \subset X(\Omega)$, $\mathbf{P}_X(A) = \mathbf{P}(X \in A)$.

PROPOSITION 3.3. *Les événements $(X = x)_{x \in X(\Omega)}$ forment un système complet d'événements de Ω .*

DÉMONSTRATION. Si $x \neq y \in X(\Omega)$, alors $(X = x) \cap (X = y) = \emptyset$. De plus, par définition de $X(\Omega)$, on a $\Omega = \bigcup_{x \in X(\Omega)} (X = x)$. \square

DÉFINITION 3.4. Si X est une variable aléatoire, on appelle *fonction de répartition* de X la fonction $F_X: \mathbf{R} \rightarrow \mathbf{R}$ définie par $F_X(x) = \mathbf{P}(X \leq x)$.

PROPOSITION 3.5. (1) F_X est croissante sur \mathbf{R} .

(2) Si $X(\Omega)$ est fini et s'écrit $\{x_1, \dots, x_n\}$ avec $x_1 < \dots < x_n$, alors $\mathbf{P}(X = x_1) = F_X(x_1)$ et, pour tout $i \in \llbracket 2; n \rrbracket$, $\mathbf{P}(X = x_i) = F_X(x_i) - F_X(x_{i-1})$.

(3) Si $X(\Omega)$ est dénombrable et s'écrit $\{x_n \mid n \in \mathbf{N}\}$ avec $x_n < x_{n+1}$ pour tout $n \in \mathbf{N}$, alors $\mathbf{P}(X = x_0) = F(x_0)$ et, pour tout $n \geq 1$, $\mathbf{P}(X = x_n) = F_X(x_n) - F_X(x_{n-1})$.

DÉMONSTRATION. \square

REMARQUE. • Il est possible que $X(\Omega)$ soit dénombrable sans qu'on puisse le mettre sous la forme $\{x_n \mid n \in \mathbf{N}\}$ avec $x_n < x_{n+1}$ pour tout $n \in \mathbf{N}$; par exemple, si $X(\Omega) = \{-n \mid n \in \mathbf{N}\}$ ou $\{\frac{1}{n+1} \mid n \in \mathbf{N}\}$. On peut montrer que, dans ce cas, on a malgré tout $\mathbf{P}(X = x) = F_X(x) - F_X(x^-)$.

• L'écriture explicite d'une variable aléatoire, sa loi ou sa fonction de répartition sont trois manières différentes de présenter la même information.

4. Espérance et variance

On suppose ici Ω dénombrable (le cas fini a été traité au chapitre 3). Soit X une variable aléatoire sur Ω .

DÉFINITION 4.1. On dit que X est *d'espérance finie* si la série $\sum_{x \in X(\Omega)} x \times \mathbf{P}(X = x)$ est absolument convergente. Dans ce cas, on appelle *espérance* de X le nombre

$$\mathbf{E}(X) = \sum_{x \in X(\Omega)} x \times \mathbf{P}(X = x).$$

En particulier, si $X(\Omega) = \{x_n \mid n \in \mathbf{N}\}$ (deux à deux distincts), alors

$$\mathbf{E}(X) = \sum_{n=0}^{+\infty} x_n \times \mathbf{P}(X = x_n).$$

REMARQUE. • On admet que si la série est absolument convergente, alors sa somme ne dépend pas de l'ordre dans lequel on fait l'énumération des éléments x_n (ça ne serait pas le cas pour une série juste convergente).

• Si $X(\Omega)$ est un ensemble fini, alors la série est une somme finie, donc converge absolument; X est donc d'espérance finie.

• Si X est d'espérance finie et que $\mathbf{E}(X) = 0$, alors on dit que X est *centrée*.

PROPOSITION 4.2. (1) Si X est bornée, alors elle est d'espérance finie.

(2) L'ensemble des variables aléatoires d'espérance finie est un sous-espace vectoriel de l'ensemble des variables aléatoires sur Ω , sur lequel l'espérance définit une forme linéaire. Autrement dit, si X et Y sont d'espérance finie et que $\lambda, \mu \in \mathbf{R}$, alors $\lambda X + \mu Y$ est d'espérance finie et $\mathbf{E}(\lambda X + \mu Y) = \lambda \mathbf{E}(X) + \mu \mathbf{E}(Y)$.

(3) Si X est d'espérance finie, alors $X - \mathbf{E}(X)$ est centrée.

(4) Si X est d'espérance finie et positive, alors $\mathbf{E}(X) \geq 0$.

DÉMONSTRATION. □

THÉORÈME 4.3 (Formule de transfert). Si $f: X(\Omega) \rightarrow \mathbf{R}$, alors $f(X)$ est d'espérance finie si et seulement si $\sum_{x \in X(\Omega)} f(x) \times \mathbf{P}(X = x)$ est absolument convergente; dans ce cas,

$$\mathbf{E}(f(X)) = \sum_{x \in X(\Omega)} f(x) \times \mathbf{P}(X = x).$$

En particulier, si $X(\Omega) = \{x_n \mid n \in \mathbf{N}\}$ deux à deux distincts, alors

$$\mathbf{E}(f(X)) = \sum_{n=0}^{+\infty} f(x_n) \times \mathbf{P}(X = x_n).$$

ADMIS. □

LEMME 4.4. Si X^2 est d'espérance finie, alors X est d'espérance finie.

DÉMONSTRATION. Le lemme est trivial si X est une variable aléatoire finie. Supposons donc $X(\Omega)$ dénombrable et notons-le $X(\Omega) = \{x_n\}_{n \in \mathbf{N}}$ avec des x_n deux à deux distincts. La variable aléatoire X^2 étant d'espérance finie, d'après le théorème de transfert, $\sum x_n^2 \mathbf{P}(X = x_n)$ est absolument convergente.

Or, pour tout $x \in \mathbf{R}$, $x^2 - 2|x| + 1 = (|x| - 1)^2 \geq 0$ donc $|x| \leq \frac{1}{2}(x^2 + 1)$. Alors, pour tout $n \in \mathbf{N}$, $|x_n \mathbf{P}(X = x_n)| \leq \frac{1}{2}(x_n^2 + 1) \mathbf{P}(X = x_n)$. Comme $\sum x_n^2 \mathbf{P}(X = x_n)$ et $\sum \mathbf{P}(X = x_n)$ convergent, on en déduit (théorème de comparaison) que $\sum |x_n \mathbf{P}(X = x_n)|$ est convergente. □

DÉFINITION 4.5. Si X^2 est d'espérance finie, on appelle *variance* de X le réel $\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2$.

PROPOSITION 4.6. $\mathbf{V}(X) \geq 0$. On appelle *écart-type* de X le nombre $\sigma(X) = \sqrt{\mathbf{V}(X)}$.

DÉMONSTRATION.

$$\begin{aligned} \mathbf{V}(X) &= \mathbf{E}(X^2) - \mathbf{E}(X)^2 = \mathbf{E}(X^2) - 2\mathbf{E}(X)\mathbf{E}(X) + \mathbf{E}(X)^2 \\ &= \mathbf{E}(X^2 - 2\mathbf{E}(X)X + \mathbf{E}(X)^2) = \mathbf{E}((X - \mathbf{E}(X))^2) \geq 0. \end{aligned}$$

□

PROPOSITION 4.7. Pour tous $a, b \in \mathbf{R}$, on a $\mathbf{V}(aX + b) = a^2 \mathbf{V}(X)$.

DÉMONSTRATION. $\mathbf{V}(aX + b) = \mathbf{E}(a^2 X^2 + 2abX + b^2) - (a\mathbf{E}(X) + b)^2 = a^2 \mathbf{E}(X^2) + 2ab\mathbf{E}(X) + b^2 - a^2 \mathbf{E}(X)^2 - 2ab\mathbf{E}(X) - b^2 = a^2 \mathbf{V}(X)$. □

THÉORÈME 4.8 (Inégalité de Bienaymé-Tchebychev). Pour tout $\varepsilon > 0$,

$$\mathbf{P}(|X - \mathbf{E}(X)| \geq \varepsilon) \leq \frac{\mathbf{V}(X)}{\varepsilon^2}.$$

DÉMONSTRATION. On pose $X(\Omega) = \{x_n\}_{k \in K}$ (2 à 2 distincts), où K est un ensemble fini ou dénombrable. On note $K^+ = \{k \in K \mid |x_k - \mathbf{E}(X)| \geq \varepsilon\}$. Alors

$$\begin{aligned} \mathbf{P}(|X - \mathbf{E}(X)| \geq \varepsilon) &= \sum_{k \in K^+} \mathbf{P}(X = x_k) = \frac{1}{\varepsilon^2} \sum_{k \in K^+} \varepsilon^2 \mathbf{P}(X = x_k) \\ &\leq \frac{1}{\varepsilon^2} \sum_{k \in K^+} (x_k - \mathbf{E}(X))^2 \mathbf{P}(X = x_k) \\ &\leq \frac{1}{\varepsilon^2} \sum_{k \in K} (x_k - \mathbf{E}(X))^2 \mathbf{P}(X = x_k) \\ &\leq \frac{1}{\varepsilon^2} \mathbf{E}((X - \mathbf{E}(X))^2) = \frac{1}{\varepsilon^2} \mathbf{V}(X). \quad \square \end{aligned}$$

REMARQUE. • Si X admet une variance (autrement dit, si X^2 est d'espérance finie), on dit que X est *réduite* si $\mathbf{V}(X) = 1$. Si X admet une variance non nulle, alors $\frac{X - \mathbf{E}(X)}{\sigma(X)}$ est centrée et réduite.

• On utilise souvent l'inégalité de Bienaymé-Tchebychev avec l'événement contraire :

$$\mathbf{P}(|X - \mathbf{E}(X)| < \varepsilon) = 1 - \mathbf{P}(|X - \mathbf{E}(X)| \geq \varepsilon) \geq 1 - \frac{\mathbf{V}(X)}{\varepsilon^2}.$$

5. Lois usuelles

DÉFINITION 5.1. On dit que X suit une *loi géométrique* de paramètre $p \in]0; 1[$, et on note $X \sim \mathcal{G}(p)$, si $X(\Omega) = \mathbf{N}^*$ et, pour tout $k \in \mathbf{N}^*$,

$$\mathbf{P}(X = k) = p(1 - p)^{k-1} = pq^{k-1}.$$

Alors $\mathbf{E}(X) = \frac{1}{p}$ et $\mathbf{V}(X) = \frac{1-p}{p^2} = \frac{q}{p^2}$.

DÉMONSTRATION. □

EXEMPLE(S). On considère une suite infinie d'épreuves de Bernoulli mutuellement indépendantes de paramètre p . Alors, si on appelle X le rang du premier succès, X suit une loi géométrique de paramètre p .

DÉFINITION 5.2. On dit que X suit une *loi de Poisson* de paramètre $\lambda > 0$, et on note $X \sim \mathcal{P}(\lambda)$, si $X(\Omega) = \mathbf{N}$ et, pour tout $k \in \mathbf{N}$,

$$\mathbf{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Alors $\mathbf{E}(X) = \mathbf{V}(X) = \lambda$.

EXEMPLE(S). La loi de Poisson est utilisée pour modéliser le nombre d'occurrences d'un événement sur un intervalle de temps donné lorsque les occurrences sont isolées les unes des autres, par exemple :

- désintégration d'atomes radioactifs
- nombre de clients se présentant à une caisse de supermarché
- nombre d'erreurs typographiques sur une page d'un livre
- nombre de morts par coup de pied de cheval dans un corps d'armée prussien

THÉORÈME 5.3 (Approximation de la loi binomiale par la loi de Poisson). *Si (X_n) est une suite de variables aléatoires réelles sur Ω telles que, pour tout $n \in \mathbf{N}$, X_n suit une loi binomiale de paramètres (n, p_n) avec $\lim_{n \rightarrow +\infty} np_n = \lambda$, alors, pour tout $k \in \mathbf{N}$,*

$$\lim_{n \rightarrow +\infty} \mathbf{P}(X_n = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

REMARQUE. En pratique, cela signifie que pour de faibles valeurs de p et de grandes valeurs de n (typiquement $p \leq 0,1$, $n \geq 30$), on peut approcher la loi binomiale $\mathcal{B}(n, p)$ par la loi de Poisson $\mathcal{P}(np)$.

Séries entières

« [Grandi] also argued that since the sum $[1 - 1 + 1 - 1 + \dots]$ was both 0 and $\frac{1}{2}$, he had proved that the world could be created out of nothing. »

M. KLINE, *Mathematical thought from ancient to modern times* (1983)

1. Convergence d'une série entière

DÉFINITION 1.1. On appelle *série entière* toute série de la forme $\sum a_n z^n$, où $(a_n)_{n \in \mathbf{N}}$ est une suite réelle ou complexe de *coefficients*, et $z \in \mathbf{C}$ une *variable*.

REMARQUE. On peut se restreindre au cas où la variable est réelle, et dans ce cas on la note traditionnellement x plutôt que z .

Pour une même suite $(a_n)_{n \in \mathbf{N}}$, la série $\sum a_n z^n$ peut être convergente ou divergente suivant la valeur de z .

EXEMPLE(S). Étudier la convergence des séries entières

$$\sum z^n; \quad \sum \frac{z^n}{n!}.$$

LEMME 1.2 (Lemme d'Abel). Soit $\sum a_n z^n$ une série entière. On suppose qu'il existe $r > 0$ tel que la suite $(a_n r^n)_{n \in \mathbf{N}}$ soit bornée. Alors, pour tout $z \in \mathbf{C}$, si $|z| < r$, alors $\sum a_n z^n$ est absolument convergente.

DÉMONSTRATION. Soit $z \in \mathbf{C}$, $|z| < r$. Notons M un majorant de la suite $(|a_n r^n|)_{n \in \mathbf{N}}$. Alors, pour tout $n \in \mathbf{N}$,

$$0 \leq |a_n z^n| = |a_n r^n| \times \left(\frac{|z|}{r}\right)^n \leq M \left(\frac{|z|}{r}\right)^n,$$

mais $|z| < r$ donc la série $\sum (|z|/r)^n$ est géométrique de raison $0 < \frac{|z|}{r} < 1$, elle est donc convergente. D'après le théorème de comparaison des séries à termes positifs, la série $\sum |a_n z^n|$ est donc convergente, autrement dit la série $\sum a_n z^n$ est absolument convergente. \square

DÉFINITION 1.3. Pour toute série entière $\sum a_n z^n$, il existe un nombre $R \in \mathbf{R}_+ \cup \{+\infty\}$ tel que, pour tout $z \in \mathbf{C}$:

- si $|z| < R$, alors $\sum a_n z^n$ est absolument convergente ;
- si $|z| > R$, alors $\sum a_n z^n$ est grossièrement divergente.

Le nombre R est appelé *rayon de convergence* de la série entière. Le disque $D(0, R) = \{z \in \mathbf{C} \mid |z| < R\}$ (resp. l'intervalle $] -R; R[$) est appelé *disque ouvert de convergence* (resp. *intervalle ouvert de convergence*).

DÉMONSTRATION. Il suffit de considérer l'ensemble $\{r \in \mathbf{R}_+ \mid (a_n r^n)_{n \in \mathbf{N}} \text{ bornée}\}$, R est alors sa borne supérieure. En effet, pour tout $z \in \mathbf{C}$, si $|z| > R$, alors $(a_n |z|^n)_{n \in \mathbf{N}}$ n'est pas bornée donc ne tend pas vers 0. C'est donc également le cas de la suite $(a_n z^n)$. Ainsi, $\sum a_n z^n$ est grossièrement divergente. \square

EXEMPLE(S). Déterminer le rayon de convergence des deux séries entières étudiées ci-dessus. Déterminer le rayon de convergence de $\sum \frac{z^n}{n}$ et $\sum n! z^n$.

REMARQUE. Si $|z| = R$, on ne peut rien dire en général et il faut étudier au cas par cas.

EXEMPLE(S). Étudier la convergence de $\sum \frac{z^n}{n}$ pour $z = 1$ et $z = -1$.

THÉORÈME 1.4 (Théorème de comparaison des séries entières). Soient $\sum a_n z^n$ et $\sum b_n z^n$ deux séries entières.

- si, pour tout $n \in \mathbf{N}$, $|a_n| \leq |b_n|$, alors le rayon de convergence de $\sum a_n z^n$ est supérieur à celui de $\sum b_n z^n$;
- si $|a_n| \sim |b_n|$, alors $\sum a_n z^n$ et $\sum b_n z^n$ ont le même rayon de convergence.

DÉMONSTRATION. Il suffit d'appliquer les divers théorèmes de comparaison des séries à termes positifs aux séries $\sum |a_n z^n|$ et $\sum |b_n z^n|$. \square

EXEMPLE(S). Déterminer le rayon de convergence des séries entières suivantes :

$$\sum \frac{z^n}{1+n^2}; \quad \sum \frac{z^n}{1+n \sin^2 n}$$

PROPOSITION 1.5. Soit $(a_n)_{n \in \mathbf{N}}$ une suite réelle ou complexe. Alors les séries entières $\sum a_n z^n$ et $\sum n a_n z^n$ ont le même rayon de convergence.

ADMIS. \square

EXEMPLE(S). Comparer les rayons de convergence des séries $\sum z^n$, $\sum \frac{z^n}{n}$, $\sum \frac{z^n}{n^2}$. Comparer leurs comportements sur le cercle $|z| = R$.

2. Somme d'une série entière

Dans toute cette partie, on se restreint au cas d'une variable x réelle.

DÉFINITION 2.1. Soit $\sum a_n x^n$ une série entière de rayon de convergence $R > 0$. On appelle *fonction somme* de la série entière la fonction $S: x \mapsto \sum_{n=0}^{+\infty} a_n z^n$, définie au moins sur l'intervalle $] -R; R[$.

PROPOSITION 2.2. Si $\sum a_n x^n$ est une série entière de rayon $R > 0$, alors sa fonction somme est continue sur $] -R; R[$.

ADMIS. □

DÉFINITION 2.3. Soit $S = \sum_{n=0}^{+\infty} a_n x^n$ une série entière. On appelle *série dérivée* de S la série entière

$$\sum_{n=1}^{+\infty} n a_n x^{n-1} = \sum_{n=0}^{+\infty} (n+1) a_{n+1} x^n.$$

PROPOSITION 2.4 (Dérivation terme à terme). *Soit $S = \sum a_n x^n$ une série entière de rayon $R > 0$. Alors la fonction somme de S est de classe \mathcal{C}^∞ . De plus, la série dérivée de S est également de rayon R , et la fonction somme de la série dérivée est égale à la dérivée de la fonction somme.*

ADMIS. □

COROLLAIRE 2.5 (Intégration terme à terme). *Si $\sum a_n x^n$ est une série entière de rayon $R > 0$, alors $\sum a_n \frac{x^{n+1}}{n+1}$ est également de rayon R , et sa somme est une primitive de la somme de $\sum a_n x^n$. Plus précisément,*

$$\int_0^x \left(\sum_{n=0}^{+\infty} a_n t^n \right) dt = \sum_{n=0}^{+\infty} a_n \frac{x^{n+1}}{n+1}.$$

EXEMPLE(S). Déterminer le rayon de convergence et calculer la somme des séries entières $\sum \frac{x^n}{n}$, $\sum n x^n$, $\sum \frac{x^n}{n!}$.

3. Fonctions développables en séries entières

DÉFINITION 3.1. Soit f une fonction définie au voisinage de 0. On dit que f est *développable en série entière* en 0 s'il existe une série entière $\sum a_n x^n$ de rayon de convergence $R > 0$ telle que, pour tout $x \in]-R; R[$,

$$f(x) = \sum_{n=0}^{+\infty} a_n x^n.$$

Dans ce cas, la série entière $\sum a_n x^n$ est appelée un *développement en série entière* de f .

PROPOSITION 3.2. *Si une fonction f est développable en série entière, alors elle est de classe \mathcal{C}^∞ au voisinage de 0. De plus, pour tout $n \in \mathbf{N}$, le n -ème terme du développement est donné par*

$$a_n = \frac{f^{(n)}(0)}{n!}$$

REMARQUE. Autrement dit, si f est développable en série entière, on peut écrire, au voisinage de 0 :

$$f(x) = \sum_{n=0}^{+\infty} f^{(n)}(0) \frac{x^n}{n!}$$

On reconnaît les termes de la formule de Taylor (on parle de *série de Taylor*).

Attention : toutes les fonctions de classe \mathcal{C}^∞ ne sont pas développables en série entière !

CONTRE-EXEMPLE. On considère la fonction $f: \mathbf{R} \rightarrow \mathbf{R}$ définie par $f(x) = e^{-1/x^2}$ si $x > 0$ et 0 si $x \leq 0$. Montrer que f est de classe \mathcal{C}^∞ mais qu'elle n'est pas développable en série entière.

THÉORÈME 3.3. *Les fonctions suivantes sont développables en série entière en 0 :*

- $\frac{1}{1-x} = \sum_{n=0}^{+\infty} x^n$, de rayon 1 ;
- $\ln(1+x) = \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} x^n$, de rayon 1 ;
- $e^x = \sum_{n=0}^{+\infty} \frac{x^n}{n!}$, de rayon infini ;
- $\cos x = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n}}{(2n)!}$, de rayon infini ;
- $\sin x = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$, de rayon infini ;
- $(1+x)^\alpha = \sum_{n=0}^{+\infty} \binom{\alpha}{n} x^n$, de rayon 1.

Espaces préhilbertiens

« Xeno pointed to the far end of the table, where a glum-looking, heavy-drinking man was trying to determine the angle between two bread rolls. »

T. PRATCHETT, *Pyramids* (1989)

Dans tout le chapitre, E désigne un \mathbf{R} -espace vectoriel.

1. Définitions

DÉFINITION 1.1. On appelle *forme bilinéaire* sur E toute application $\varphi: E \times E \rightarrow \mathbf{R}$ telle que, pour tous $x_0, y_0 \in E$, les applications $x \mapsto \varphi(x, y_0)$ et $y \mapsto \varphi(x_0, y)$ soient linéaires. Autrement dit, pour tous $x_1, x_2, y_1, y_2 \in E$ et $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbf{R}$, on a

$$\varphi(\lambda_1 x_1 + \lambda_2 x_2, \mu_1 y_1 + \mu_2 y_2) = \lambda_1 \mu_1 \varphi(x_1, y_1) + \lambda_1 \mu_2 \varphi(x_1, y_2) + \lambda_2 \mu_1 \varphi(x_2, y_1) + \lambda_2 \mu_2 \varphi(x_2, y_2).$$

REMARQUE. Si φ est une forme bilinéaire sur E , alors pour tous $x, y \in E$, on a $\varphi(x, 0) = \varphi(0, y) = 0$.

DÉFINITION 1.2. Une forme bilinéaire φ sur E est dite :

- *symétrique* si, pour tous $x, y \in E$, $\varphi(x, y) = \varphi(y, x)$;
- *positive* si, pour tout $x \in E$, $\varphi(x, x) \geq 0$;
- *définie* si, pour tout $x \in E$, si $\varphi(x, x) = 0$ alors $x = 0$.

EXEMPLE(S). Soient $a, b \in \mathbf{R}$. Pour tous $P, Q \in \mathbf{R}[X]$, on pose $\varphi(P, Q) = P(a) \times Q(b)$. Vérifier que φ est une forme bilinéaire. Est-elle symétrique ? Positive ? Définie ?

DÉFINITION 1.3. Une forme bilinéaire symétrique définie positive sur E est appelée un *produit scalaire* sur E .

On appelle *espace préhilbertien* tout \mathbf{R} -espace vectoriel muni d'un produit scalaire.

REMARQUE. Notations alternatives pour $\varphi(x, y) : \langle x, y \rangle, (x | y), x \cdot y, \dots$

EXEMPLE(S) (à connaître). Montrer que les \mathbf{R} -ev ci-dessous sont préhilbertiens :

- $E = \mathbf{R}^n$ avec le produit scalaire $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$.
- $E = \mathcal{C}([a; b], \mathbf{R})$ avec le produit scalaire $(f | g) = \int_a^b f(t)g(t) dt$.
- $E = \mathbf{R}[X]$ avec le produit scalaire $P \cdot Q = \int_0^{+\infty} P(t)Q(t)e^{-t} dt$.

2. Norme associée à un produit scalaire

THÉORÈME 2.1 (Inégalité de Cauchy-Schwarz). *Soit E un espace préhilbertien. Alors, pour tous $x, y \in E$,*

$$\langle x, y \rangle^2 \leq \langle x, x \rangle \times \langle y, y \rangle,$$

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

DÉMONSTRATION. On pose, pour tout $\lambda \in \mathbf{R}$, $f(\lambda) = \langle x + \lambda y, x + \lambda y \rangle$. On a, pour tout $\lambda \in \mathbf{R}$, $f(\lambda) = \langle x, x \rangle + 2\lambda \langle x, y \rangle + \lambda^2 \langle y, y \rangle$; c'est un polynôme du second degré (sauf si $y = 0$) qui est toujours positif (car $\langle \cdot, \cdot \rangle$ est une forme bilinéaire positive); on en déduit que son discriminant est négatif ou nul, d'où le résultat.

Le cas d'égalité correspond à un discriminant nul, c'est-à-dire qu'il existe $\lambda \in \mathbf{R}$ tel que $\langle x + \lambda y, x + \lambda y \rangle = 0$. La forme bilinéaire $\langle \cdot, \cdot \rangle$ étant définie, cela implique que $x + \lambda y = 0$, donc que x et y sont colinéaires. \square

THÉORÈME 2.2. *Pour tout $x \in E$, on pose $\|x\| = \sqrt{\langle x, x \rangle}$. Alors $\|\cdot\|$ définit une norme sur E .*

DÉMONSTRATION. Il faut vérifier les quatre propriétés d'une norme :

- $\|\cdot\|$ est bien définie sur E car, pour tout $x \in E$, $\langle x, x \rangle \geq 0$ (positivité de la forme bilinéaire);
- pour tout $x \in E$, $\|x\| = 0 \iff \sqrt{\langle x, x \rangle} = 0 \iff x = 0$ (définition de la forme bilinéaire);
- pour tout $x \in E$ et $\lambda \in \mathbf{R}$, $\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x \rangle} = \sqrt{\lambda^2 \langle x, x \rangle} = |\lambda| \|x\|$;
- pour tous $x, y \in E$,

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &= \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| \\ &\leq (\|x\| + \|y\|)^2. \end{aligned} \quad \square$$

REMARQUE. • L'inégalité de Cauchy-Schwarz peut se réécrire $|\langle x, y \rangle| \leq \|x\|\|y\|$.

- Application aux exemples précédents :

$$\sum_{i=1}^n x_i y_i \leq \sqrt{\sum_{i=1}^n x_i^2} \times \sqrt{\sum_{i=1}^n y_i^2}$$

$$\left| \int_a^b f(t)g(t) dt \right| \leq \sqrt{\int_a^b f^2(t) dt} \times \sqrt{\int_a^b g^2(t) dt}$$

- On rappelle qu'à toute norme est associée une distance, donc à tout produit scalaire est associé une distance : $d(x, y) = \sqrt{\langle x - y, x - y \rangle}$.

PROPOSITION 2.3 (Identités de polarisation). *Pour tous $x, y \in E$:*

$$\begin{aligned}\langle x, y \rangle &= \frac{1}{2} \left(\|x + y\|^2 - \|x\|^2 - \|y\|^2 \right) \\ &= \frac{1}{2} \left(\|x\|^2 + \|y\|^2 - \|x - y\|^2 \right) \\ &= \frac{1}{4} \left(\|x + y\|^2 - \|x - y\|^2 \right).\end{aligned}$$

DÉMONSTRATION.

$$\begin{aligned}\|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &= \|x\|^2 + 2\langle x, y \rangle + \|y\|^2;\end{aligned}$$

De même,

$$\|x - y\|^2 = \|x\|^2 - 2\langle x, y \rangle + \|y\|^2. \quad \square$$

PROPOSITION 2.4 (Identité du parallélogramme). *Pour tous $x, y \in E$, alors*

$$\|x + y\|^2 + \|x - y\|^2 = 2 \left(\|x\|^2 + \|y\|^2 \right).$$

REMARQUE. Interprétation géométrique : la somme des carrés des côtés d'un parallélogramme est égale à la somme des carrés des diagonales.

3. Orthogonalité

Soit E un espace préhilbertien.

DÉFINITION 3.1. • On dit que deux vecteurs $x, y \in E$ sont *orthogonaux* si $\langle x, y \rangle = 0$.

• Si A est une partie de E , on appelle *orthogonal* de A la partie :

$$A^\perp = \{x \in E \mid \forall y \in A \langle x, y \rangle = 0\}$$

• Deux parties A et B sont dites *orthogonales* si tout élément de A est orthogonal à tout élément de B .

• Une famille $(x_i)_{i \in I}$ de vecteurs de E est dite *orthogonale* si, pour tous $i, j \in I$, si $i \neq j$, alors $\langle x_i, x_j \rangle = 0$.

• Une famille $(x_i)_{i \in I}$ de vecteurs de E est dite *orthonormale* si, pour tous $i, j \in I$, $\langle x_i, x_j \rangle = \delta_{i,j}$.

REMARQUE. A et B sont orthogonales si et seulement si $A \subset B^\perp$

PROPOSITION 3.2. *Soit A une partie de E . Alors :*

- (1) A^\perp est une sous-espace vectoriel de E ;
- (2) $A^\perp = \text{Vect}(A)^\perp$;
- (3) A et A^\perp sont deux parties orthogonales de E ;
- (4) $A \subset (A^\perp)^\perp$ (il n'y a pas égalité en général) ;
- (5) $\{0_E\}^\perp = E$ et $E^\perp = \{0_E\}$;
- (6) Toute famille orthogonale ne contenant pas 0_E est libre ;

- (7) Toute famille orthonormale est libre ;
 (8) Si F est un sev de E , alors F et F^\perp sont en somme directe (pas nécessairement supplémentaires) ;
 (9) Si F_1, \dots, F_p sont des sev de E deux à deux orthogonaux, alors ils sont en somme directe.

DÉMONSTRATION. (1) On a évidemment $A^\perp \subset E$ et $0 \in A^\perp$ (car $\langle 0, x \rangle = 0$ pour tout $x \in A$) ; de plus, si $x_1, x_2 \in A^\perp$ et $\lambda_1, \lambda_2 \in \mathbf{R}$, alors, pour tout $y \in A$,

$$\langle \lambda_1 x_1 + \lambda_2 x_2, y \rangle = \lambda_1 \langle x_1, y \rangle + \lambda_2 \langle x_2, y \rangle = 0$$

donc $\lambda_1 x_1 + \lambda_2 x_2 \in A^\perp$. On en déduit que A^\perp est bien un sous-ev.

- (2) $A \subset \text{Vect}(A)$ donc $\text{Vect}(A)^\perp \subset A^\perp$. De plus, pour tout $y \in \text{Vect}(A)$, il existe $\lambda_1, \dots, \lambda_n \in \mathbf{R}$ et $y_1, \dots, y_n \in A$ tels que $y = \lambda_1 y_1 + \dots + \lambda_n y_n$, et alors, pour tout $x \in A^\perp$, $\langle x, y \rangle = \sum_{k=1}^n \lambda_k \langle x, y_k \rangle = 0$ donc $x \in \text{Vect}(A)^\perp$. Ainsi $A \subset \text{Vect}(A)^\perp$; il y a donc égalité.
 (3), (4) Pour tous $x \in A^\perp$ et $y \in A$, on a $\langle x, y \rangle = 0$ donc $y \in (A^\perp)^\perp$.
 (5) Pour tout $x \in E$, $\langle x, 0 \rangle = 0$.
 (6) Soit \mathcal{V} une famille orthogonale ne contenant pas 0. Soient $v_1, \dots, v_n \in \mathcal{C}$ et $\lambda_1, \dots, \lambda_n \in \mathbf{R}$. On suppose que $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$. Mais alors, en prenant le produit scalaire de $\lambda_1 v_1 + \dots + \lambda_n v_n$ avec v_1 , on obtient $\lambda_1 \|v_1\|^2 = 0$, donc $\lambda_1 = 0$. De même, $\lambda_2 = \dots = \lambda_n = 0$. La famille est donc libre.
 (7) C'est une conséquence du point précédent, puisqu'une famille orthonormale est orthogonale et ne peut pas contenir le vecteur nul.
 (8) si $x \in F \cap F^\perp$, alors $\langle x, x \rangle = 0$ donc $x = 0$.
 (9) Soit $x \in F_1 + \dots + F_p$. On suppose qu'on a deux écritures $x = x_1 + \dots + x_n = y_1 + \dots + y_n$, avec $x_1, y_1 \in F_1, \dots, x_p, y_p \in F_p$. Alors $\sum_{i=1}^p (x_i - y_i) = 0$ donc, pour tout $j \in \llbracket 0; p \rrbracket$, $0 = \langle x_j - y_j; \sum_{i=1}^p (x_i - y_i) \rangle = \|x_j - y_j\|^2$, donc $x_j = y_j$; ceci étant vrai pour tout $j \in \llbracket 0; p \rrbracket$, on en déduit que la décomposition de x est unique. \square

THÉORÈME 3.3 (Théorème de Pythagore). Pour tous $x, y \in E$,

$$\langle x, y \rangle = 0 \iff \|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

DÉMONSTRATION. $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle$. \square

PROPOSITION 3.4 (Généralisation du théorème de Pythagore). Si (x_1, \dots, x_p) est une famille orthogonale de E , alors

$$\sum_{i=1}^p \|x_i\|^2 = \left\| \sum_{i=1}^p x_i \right\|^2$$

DÉMONSTRATION. On procède par récurrence sur p . Le cas $p = 1$ est trivial, le cas $p = 2$ a déjà été prouvé. Soit donc $p \geq 1$, on suppose le résultat vrai au rang p . Soit (x_1, \dots, x_{p+1}) une famille orthogonale. Alors

$$\left\| \sum_{i=1}^p x_i \right\|^2 = \left\| x_1 + \sum_{i=2}^p x_i \right\|^2 = \|x_1\|^2 + \left\| \sum_{i=2}^p x_i \right\|^2 = \sum_{i=1}^p \|x_i\|^2$$

car x_1 et $\sum_{i=2}^p x_i$ sont orthogonaux, puis par hypothèse de récurrence. \square

4. Familles orthonormales

THÉORÈME 4.1 (Procédé d'orthogonalisation de Gram-Schmidt). *Soit v_1, \dots, x_n une famille libre de E . Alors il existe une unique famille orthogonale (e_1, \dots, e_n) telle que $e_1 = v_1$ et, pour tout $k \in \llbracket 2; n \rrbracket$, $e_k - v_k \in \text{Vect}(e_1, \dots, e_{k-1})$.*

En particulier, la famille e_1, \dots, e_n est libre et, pour tout $k \in \llbracket 1; n \rrbracket$, $\text{Vect}(e_1, \dots, e_k) = \text{Vect}(v_1, \dots, v_k)$.

EXEMPLE(S). Dans $E = \mathbf{R}^4$, orthogonaliser la famille (v_1, v_2, v_3) suivante :

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

REMARQUE. La famille obtenue par le procédé de Gram-Schmidt est orthogonale et libre ; en particulier, elle ne contient pas le vecteur 0. On peut donc la rendre orthonormale en divisant chacun de ses éléments par sa norme. On parle dans ce cas de procédé d'*orthonormalisation* de Gram-Schmidt

EXEMPLE(S). Orthonormaliser la famille de l'exemple précédent.

COROLLAIRE 4.2. *Tout espace préhilbertien de dimension finie admet une base orthonormée ; plus précisément, dans un espace préhilbertien fini, toute famille orthonormée peut être complétée en une base orthonormée.*

DÉMONSTRATION. Toute famille orthonormée est libre ; on peut donc la compléter en une base, puis orthonormaliser cette base. On obtient une famille orthonormée de cardinal égal à la dimension de l'espace, donc une base. Enfin, les vecteurs d'origine (avant complétion) formant déjà une famille orthonormée, ils ne sont pas modifiés par le procédé de Gram-Schmidt. La base orthonormée obtenue est donc bien une complétion de la famille de départ. \square

REMARQUE. Si (e_1, \dots, e_n) est une base orthonormée de E , alors pour tout $x \in E$, on peut écrire

$$x = \sum_{i=1}^n \langle x, e_i \rangle e_i.$$

Si on note x_i le résultat du produit scalaire $\langle x, e_i \rangle$, alors pour tout $y \in E$, on a

$$\langle x, y \rangle = \left\langle \sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j \right\rangle = \sum_{(i,j) \in \llbracket 1; n \rrbracket^2} x_i y_j \langle e_i, e_j \rangle = \sum_{i=1}^n x_i y_i$$

En particulier,

$$\|x\| = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{{}^t X X}, \quad \text{où } X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

DÉFINITION 4.3. Un espace préhilbertien de dimension finie est un espace *euclidien*.

5. Sous-espaces de dimension finie et orthogonalité

THÉORÈME 5.1. *Si F est un sous-espace vectoriel de dimension finie de E , alors $E = F \oplus F^\perp$.*

DÉMONSTRATION. On a montré (prop. 3.2 (8)) que F et F^\perp sont en somme directe.

Soit (e_1, \dots, e_n) une base orthonormale de F (existe d'après le corollaire 4.2). Soit x un vecteur quelconque de E . On pose $u = \sum_{i=1}^n \langle e_i, x \rangle e_i$ et $v = x - u$. On a alors :

- $x = u + v$;
- $v \in F^\perp$;
- pour tout $j \in \llbracket 1; n \rrbracket$, $\langle e_j, v \rangle = \langle e_j, x \rangle - \sum_{i=1}^n \langle e_i, x \rangle \langle e_j, e_i \rangle = \langle e_j, x \rangle - \langle e_j, x \rangle = 0$, donc $v \in F^\perp$.

On a donc $E = F + F^\perp$, d'où le résultat. \square

COROLLAIRE 5.2. *Si E est un espace euclidien et F un sev de E , alors $\dim F^\perp = \dim E - \dim F$.*

DÉFINITION 5.3. Soit F un sev de E de dimension finie. On appelle *projection orthogonale* sur F la projection sur F parallèlement à F^\perp . De même, on appelle *symétrie orthogonale* par rapport à F la symétrie par rapport à F parallèlement à F^\perp .

PROPOSITION 5.4. *Soit F un sev de E de dimension finie et (e_1, \dots, e_n) une base orthonormée de F . On note p la projection orthogonale sur F . Alors, pour tout $x \in E$, on a*

$$p(x) = \sum_{i=1}^n \langle x, e_i \rangle e_i.$$

DÉMONSTRATION. Démontré au passage pendant la démonstration du théorème 5.1. \square

REMARQUE. Si on ne dispose pas d'une base orthonormée de F , une piste pour rechercher $p(x)$ est d'exploiter le fait que $x - p(x)$ est un élément de F^\perp , donc qu'il est orthogonal à tous les vecteurs d'une partie génératrice de F .

On rappelle que si $x \in E$ et $A \subset E$, la *distance* de x à A est le nombre

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

THÉORÈME 5.5 (Théorème des projections). *Soit F un sev de dimension finie de E . On note p la projection orthogonale sur F . Alors, pour tout $x \in E$,*

$$d(x, F)^2 = \|x\|^2 - \|p(x)\|^2$$

De plus, $p(x)$ est l'unique élément $y \in F$ tel que $d(x, F) = \|x - y\|$.

DÉMONSTRATION. Soit $x \in E$. Pour tout $y \in F$,

$$\|x - y\|^2 = \|x - p(x)\|^2 + \|p(x) - y\|^2 \geq \|x - p(x)\|^2,$$

avec égalité si et seulement si $y = p(x)$.

En particulier, $d(x, F) \geq \|x - p(x)\|$. Or $p(x) \in F$ donc $d(x, F) \leq \|x - p(x)\|$; on en déduit l'égalité. \square

Séries de Fourier

« M. Fourier avait l'opinion que le but principal des mathématiques était l'utilité publique et l'explication des phénomènes naturels ; mais un philosophe comme lui aurait dû savoir que le but unique de la science, c'est l'honneur de l'esprit humain. »

C. JACOBI, *lettre à A. Legendre* (1830)

Dans tout le chapitre, on T désigne un réel strictement positif, et on pose $\omega = \frac{2\pi}{T}$.

1. Un espace préhilbertien

THÉORÈME 1.1. On pose $\mathcal{C}_T(\mathbf{R}) = \{f: \mathbf{R} \rightarrow \mathbf{R} \mid f \text{ continue et } T\text{-périodique}\}$ et, pour tous $f, g \in \mathcal{C}_T(\mathbf{R})$,

$$\varphi(f, g) = \frac{1}{T} \int_0^T f(t)g(t) dt.$$

Alors $(\mathcal{C}_T(\mathbf{R}), \varphi)$ est un espace préhilbertien.

DÉMONSTRATION. • $\mathcal{C}_T(\mathbf{R})$ est un sous-espace vectoriel de $\mathcal{C}(\mathbf{R}, \mathbf{R})$; (exercice)
 • φ définit clairement une forme bilinéaire symétrique sur $\mathcal{C}_T(\mathbf{R})$;
 • si $f \in \mathcal{C}_T(\mathbf{R})$, $\varphi(f, f) \geq 0$ car c'est l'intégrale de la fonction positive f^2 sur un segment dont les bornes sont dans le « bon » sens. De plus, f^2 étant continue et positive et $0 < T$, si $\varphi(f, f) = 0$, alors f est nulle sur l'intervalle $[0; T]$, donc nulle sur \mathbf{R} par périodicité. \square

DÉFINITION 1.2. Une fonction $f: I \rightarrow \mathbf{R}$ est dite *continue par morceaux* si on peut décomposer I en intervalles sur lesquels f est continue et prolongeable par continuité aux bornes.

EXEMPLE(S). Une fonction créneau, une fonction en dents de scie, etc. sont continues par morceaux. Par contraste, une fonction comme la fonction tangente n'est pas continue par morceaux car on ne peut pas la prolonger par continuité aux bornes des intervalles $]n\pi - \frac{\pi}{2}; n\pi + \frac{\pi}{2}[$.

PROPOSITION 1.3. Si $f: \mathbf{R} \rightarrow \mathbf{R}$ est continue par morceaux et T -périodique, alors, pour tout $\alpha \in \mathbf{R}$,

$$\int_{\alpha}^{\alpha+T} f(t) dt = \int_0^T f(t) dt.$$

DÉMONSTRATION. On applique la relation de Chasles, puis le changement de variables affine $u = t - T$:

$$\begin{aligned} \int_{\alpha}^{\alpha+T} f(t) dt &= \int_{\alpha}^0 f(t) dt + \int_0^T f(t) dt + \int_T^{\alpha+T} f(t) dt \\ &= \int_{\alpha}^0 f(t) dt + \int_0^T f(t) dt + \int_0^{\alpha} \underbrace{f(u+T)}_{=f(u)} du \\ &= \int_0^T f(t) dt. \end{aligned} \quad \square$$

THÉORÈME 1.4. *La famille $\{t \mapsto \cos(n\omega t) \mid n \in \mathbf{N}\} \cup \{t \mapsto \sin(n\omega t) \mid n \in \mathbf{N}^*\}$ est une famille orthogonale de $\mathcal{C}_T(\mathbf{R})$.*

DÉMONSTRATION. Il s'agit de vérifier que tous les produits scalaires $\varphi(\cos(n\omega t), \cos(p\omega t))$, $\varphi(\cos(n\omega t), \sin(p\omega t))$ et $\varphi(\sin(n\omega t), \sin(p\omega t))$ sont nuls.

Soient $n, p \in \mathbf{N}$.

Si $n = p = 0$, alors $\frac{1}{T} \int_0^T \cos(n\omega t) \cos(p\omega t) dt = 1$.

Si $n = p > 0$, alors $\frac{1}{T} \int_0^T \cos(n\omega t) \cos(p\omega t) dt = \frac{1}{2T} \int_0^T 1 + \cos(2n\omega t) dt = \frac{1}{2}$.

Si $n \neq p$, alors

$$\begin{aligned} \frac{1}{T} \int_0^T \cos(n\omega t) \cos(p\omega t) dt &= \frac{1}{2T} \int_0^T \cos((n+p)\omega t) + \cos((n-p)\omega t) dt \\ &= \frac{1}{2T} \left[\frac{\sin((n+p)\omega t)}{(n+p)\omega} + \frac{\sin((n-p)\omega t)}{(n-p)\omega} \right]_0^T \\ &= 0. \end{aligned}$$

Si $n = p$, alors $\frac{1}{T} \int_0^T \sin(n\omega t) \sin(p\omega t) dt = \frac{1}{2T} \int_0^T 1 - \cos(2n\omega t) dt = \frac{1}{2}$.

Si $n \neq p$, alors

$$\begin{aligned} \frac{1}{T} \int_0^T \sin(n\omega t) \sin(p\omega t) dt &= \frac{1}{2T} \int_0^T \cos((n-p)\omega t) - \cos((n+p)\omega t) dt \\ &= \frac{1}{2T} \left[\frac{\sin((n-p)\omega t)}{(n-p)\omega} - \frac{\sin((n+p)\omega t)}{(n+p)\omega} \right]_0^T \\ &= 0 \end{aligned}$$

Si $n = p$, alors $\frac{1}{T} \int_0^T \cos(n\omega t) \sin(p\omega t) dt = \frac{1}{2T} \int_0^T \sin(2n\omega t) dt = 0$.

Si $n \neq p$, alors

$$\begin{aligned} \frac{1}{T} \int_0^T \cos(n\omega t) \sin(p\omega t) dt &= \frac{1}{2T} \int_0^T \sin((n+p)\omega t) - \sin((n-p)\omega t) dt \\ &= \frac{1}{2T} \left[\frac{\cos((n-p)\omega t)}{(n-p)\omega} - \frac{\cos((n+p)\omega t)}{(n+p)\omega} \right]_0^T \\ &= 0. \end{aligned} \quad \square$$

2. Coefficients de Fourier

On peut remarquer que toutes les fonctions de $\mathcal{C}_T(\mathbf{R})$ ne sont pas des combinaisons linéaires des $\cos(n\omega t)$ et $\sin(n\omega t)$. En effet, ces combinaisons linéaires sont toutes de classe \mathcal{C}^∞ , alors qu'il existe des fonctions continues T -périodiques qui ne soient pas \mathcal{C}^∞ (on peut par exemple penser à une fonction en « dents de scie »).

Une question que l'on peut en revanche se poser est si n'importe quelle fonction de $\mathcal{C}_T(\mathbf{R})$ peut être « approchée » par une combinaison linéaire des $\cos(n\omega t)$ et $\sin(n\omega t)$.

Plus précisément, si on considère $f \in \mathcal{C}_T(\mathbf{R})$ et, pour tout $p \in \mathbf{N}^*$, F_p l'espace vectoriel engendré par les fonctions $t \mapsto \cos(n\omega t)$, $n \in \llbracket 0; p \rrbracket$ et $t \mapsto \sin(n\omega t)$, $n \in \llbracket 1; p \rrbracket$ et $S_p(f)$ la projection orthogonale de f sur F_p , peut-on dire que la suite $(S_p(f))_{p \in \mathbf{N}^*}$ converge vers f ?

Plus précisément, a-t-on :

- $\|f - S_p(f)\| \longrightarrow 0$?
- $\forall x \in \mathbf{R} \ (S_p(f))(x) \longrightarrow f(x)$?

On va introduire ici quelques définitions et faire quelques remarques avant de s'intéresser pleinement à ces questions dans la partie 3.

DÉFINITION 2.1. Soit $f: \mathbf{R} \rightarrow \mathbf{R}$ T -périodique et continue par morceaux. On appelle *coefficients de Fourier* de f les nombres :

$$\begin{aligned} a_0(f) &= \frac{1}{T} \int_0^T f(t) dt; \\ \forall n \in \mathbf{N}^* \ a_n(f) &= \frac{2}{T} \int_0^T f(t) \cos(n\omega t) dt; \\ \forall n \in \mathbf{N}^* \ b_n(f) &= \frac{2}{T} \int_0^T f(t) \sin(n\omega t) dt. \end{aligned}$$

REMARQUE. Pour tout $p \in \mathbf{N}^*$, on a

$$(S_p(f))(x) = a_0 + \sum_{n=1}^p (a_n \cos(n\omega t) + b_n \sin(n\omega t));$$

en effet, une conséquence de la démonstration du théorème 1.3 est que la famille

$$\{t \mapsto 1\} \cup \{t \mapsto \sqrt{2} \cos(n\omega t) \mid n \in \mathbf{N}^*\} \cup \{t \mapsto \sqrt{2} \sin(n\omega t) \mid n \in \mathbf{N}^*\}$$

est une base orthonormée de F_p .

PROPOSITION 2.2. Si f est paire, alors, pour tout $n \in \mathbf{N}^*$, $b_n(f) = 0$.

Si f est impaire, alors, pour tout $n \in \mathbf{N}$, $a_n(f) = 0$.

DÉMONSTRATION. Si f est paire, alors, pour tout $n \in \mathbf{N}^*$, la fonction $t \mapsto f(t) \sin(n\omega t)$ est impaire, donc $\int_0^T f(t) \sin(n\omega t) dt = \int_{-T/2}^{T/2} f(t) \sin(n\omega t) dt = 0$. On raisonne de même si f est impaire. \square

REMARQUE. • Si f est paire, alors, pour tout $n \geq 1$, la fonction $t \mapsto f(t) \cos(n\omega t)$ est également paire; on peut alors écrire

$$a_n(f) = \frac{1}{T} \int_0^T f(t) \cos(n\omega t) dt = \frac{1}{T} \int_{-T/2}^{T/2} f(t) \cos(n\omega t) dt = \frac{2}{T} \int_0^{T/2} f(t) \cos(n\omega t) dt.$$

- On peut calculer en une seule intégrale les coefficients $a_n(f)$ et $b_n(f)$ en remarquant que, si $n \geq 1$, $a_n(f) + ib_n(f) = \frac{2}{T} \int_0^T f(t) e^{in\omega t} dt$.

PROPOSITION 2.3. Si f est une application de \mathbf{R} dans \mathbf{R} T -périodique et continue par morceaux, vérifiant $\forall x \in \mathbf{R} f(x + \frac{T}{2}) = -f(x)$, alors tous les coefficients de Fourier d'indice pair de f sont nuls. Autrement dit, $a_0(f) = 0$ et, pour tout $n \in \mathbf{N}^*$, $a_{2n}(f) = b_{2n}(f) = 0$.

DÉMONSTRATION. Soit $n \in \mathbf{N}$. Alors

$$\begin{aligned} \int_{T/2}^T f(t) e^{2in\omega t} dt &= \int_0^{T/2} f(u + T/2) e^{i2n\omega(u+T/2)} du \\ &= - \int_0^{T/2} f(u) e^{i2n\omega u + i2n\pi} du \\ &= - \int_0^{T/2} f(u) e^{i2n\omega u} du, \end{aligned}$$

donc $\int_0^T f(t) e^{i2n\omega t} dt = 0$, d'où $a_{2n}(f) = b_{2n}(f) = 0$. \square

3. Série de Fourier

DÉFINITION 3.1. Si $f: \mathbf{R} \rightarrow \mathbf{R}$ continue par morceaux et T -périodique, alors on appelle *série de Fourier* de f la série de fonctions

$$t \mapsto a_0(f) + \sum_{n \geq 1} (a_n(f) \cos(n\omega t) + b_n(f) \sin(n\omega t)).$$

REMARQUE. On n'a aucune information a priori sur la convergence de cette série.

THÉORÈME 3.2 (Formule de Parseval). Si f est continue par morceaux et T -périodique, alors

$$\frac{1}{T} \int_0^T f^2(t) dt = a_0^2 + \frac{1}{2} \sum_{n=1}^{\infty} (a_n^2(f) + b_n^2(f)).$$

ADMIS. \square

REMARQUE. • La formule de Parseval implique que la série $\sum(a_n^2(f) + b_n^2(f))$ converge, donc que les séries $\sum a_n^2(f)$ et $\sum b_n^2(f)$ convergent; on en déduit en particulier que $a_n(f)$ et $b_n(f)$ tendent vers 0 pour $n \rightarrow +\infty$.

• Si $f \in \mathcal{C}_T(\mathbf{R})$, alors pour tout $p \geq 1$,

$$\|S_p(f)\|^2 = a_0^2 + \frac{1}{2} \sum_{n=1}^p (a_n^2(f) + b_n^2(f))$$

donc

$$\|f - S_p(f)\|^2 = \|f\|^2 - \|S_p(f)\|^2 = \frac{1}{T} \int_0^T f^2(t) dt - a_0^2 - \frac{1}{2} \sum_{n=1}^{\infty} (a_n^2(f) + b_n^2(f)),$$

qui tend vers 0 lorsque p tend vers $+\infty$ d'après la formule de Parseval. La fonction $S_p(f)$ « converge » donc vers f au sens de la norme de $\mathcal{C}_T(\mathbf{R})$ lorsque p tend vers l'infini.

DÉFINITION 3.3. Si f est une fonction continue par morceaux sur \mathbf{R} et $t \in \mathbf{R}$, on note $f(t^+)$ (resp. $f(t^-)$) la limite à droite (resp. à gauche) de $f(x)$ lorsque $x \rightarrow t$.

On appelle *régularisée* de f la fonction $t \mapsto \frac{1}{2} (f(t^+) + f(t^-))$.

REMARQUE. En particulier, si f est continue, on a, pour tout $t \in \mathbf{R}$, $f(t^+) = f(t^-) = f(t)$ donc f est égale à sa régularisée.

THÉORÈME 3.4 (Théorème de Dirichlet). *si f est T -périodique et \mathcal{C}^1 par morceaux, alors, pour tout $t \in \mathbf{R}$, on a*

$$a_0(f) + \sum_{n=1}^{\infty} (a_n(f) \cos(n\omega t) + b_n(f) \sin(n\omega t)) = \frac{1}{2} (f(t^+) + f(t^-)).$$

REMARQUE. Le théorème de Dirichlet donne une réponse partielle à la deuxième question de convergence de la deuxième partie : si f est continue et \mathcal{C}^1 par morceaux, alors, pour tout $t \in \mathbf{R}$, $S_p(f)(x)$ converge vers $f(x)$.

EXEMPLE(S). Soit f la fonction 2π -périodique définie pour $x \in [0; 2\pi[$ par $f(x) = x^2$. Déterminer la série de Fourier de f , puis en déduire la valeur des séries de terme général $\frac{1}{n^2}$ et $\frac{1}{n^4}$.

Isométries des espaces euclidiens

« Ἀγεωμέτρητος μηδεὶς εἰσὶτω » (Que nul n'entre ici s'il n'est géomètre.)

devise de l'Académie de Platon, v. 387 av. J.-C.

1. Rappels

DÉFINITION 1.1. Un *espace vectoriel euclidien* est un espace préhilbertien de dimension finie; autrement dit, c'est un \mathbf{R} -espace vectoriel de dimension finie muni d'un produit scalaire.

Dans tout le chapitre, E désignera un espace euclidien et on notera $n = \dim E$.

PROPOSITION 1.2. Si (e_1, \dots, e_n) est une base orthonormée de E , si $x = \sum_{i=1}^n x_i e_i$ et

$$u = \sum_{i=1}^n y_i e_i, \text{ et si on note } X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ et } Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \text{ alors}$$

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i = {}^t X Y; \quad \|x\|^2 = \sum_{i=1}^n x_i^2 = {}^t X X; \quad x = \sum_{i=1}^n \langle x, e_i \rangle \cdot e_i.$$

PROPOSITION 1.3. Si F est un sev de E , alors $E = F \oplus F^\perp$.

2. Isométries vectorielles de E

DÉFINITION 2.1. On appelle *isométrie (vectorielle)* de E toute application linéaire $f: E \rightarrow E$ préservant la norme, c'est-à-dire telle que, pour tout $x \in E$, $\|f(x)\| = \|x\|$. On note $\mathcal{O}(E)$ l'ensemble des isométries de E et on l'appelle *groupe orthogonal* de E .

THÉORÈME 2.2. f est une isométrie vectorielle si et seulement si, pour tous $x, y \in E$, on a $\langle f(x), f(y) \rangle = \langle x, y \rangle$.

DÉMONSTRATION. $\boxed{\Leftarrow}$ Soit $x \in E$, alors $\|f(x)\|^2 = \langle f(x), f(x) \rangle = \langle x, x \rangle = \|x\|^2$ donc $\|f(x)\| = \|x\|$.

\Rightarrow Soient $x, y \in E$. Alors

$$\begin{aligned} \langle f(x), f(y) \rangle &= \frac{1}{2}(\|f(x) + f(y)\|^2 - \|f(x)\|^2 - \|f(y)\|^2) \\ &= \frac{1}{2}(\|f(x + y)\|^2 - \|f(x)\|^2 - \|f(y)\|^2) \\ &= \frac{1}{2}(\|x + y\|^2 - \|x\|^2 - \|y\|^2) \\ &= \langle x, y \rangle. \end{aligned} \quad \square$$

PROPOSITION 2.3. *Si f est une isométrie de E , alors f est un automorphisme de E .*

DÉMONSTRATION. L'application f étant un endomorphisme de E qui est de dimension finie, il suffit de montrer que f est injective, donc que $\text{Ker } f = \{0\}$.

Soit donc $x \in \text{Ker } f$. Alors $f(x) = 0$ donc $\|f(x)\| = 0$, mais $\|f(x)\| = \|x\|$ car f est une isométrie donc $\|x\| = 0$, autrement dit $x = 0$. L'application f est donc un automorphisme de E . \square

REMARQUE. Attention : la notion d'isométrie dépend évidemment du choix de produit scalaire sur E . Une même application peut être isométrique pour un certain produit scalaire, mais non isométrique pour un autre.

Les isométries préservent l'orthogonalité, c'est-à-dire que si x et y sont deux vecteurs orthogonaux et f une isométrie, alors $f(x)$ et $f(y)$ sont orthogonaux (en effet, $\langle f(x), f(y) \rangle = \langle x, y \rangle$).

PROPOSITION 2.4. *Si $f, g \in \mathcal{O}(E)$, alors $f \circ g \in \mathcal{O}(E)$. Si $f \in \mathcal{O}(E)$, alors $f^{-1} \in \mathcal{O}(E)$.*

DÉMONSTRATION. Soient $f, g \in \mathcal{O}(E)$. Soit $x \in E$. Alors $\|(g \circ f)(x)\| = \|g(f(x))\| = \|f(x)\| = \|x\|$ donc $g \circ f \in \mathcal{O}(E)$.

Soient $f \in \mathcal{O}(E)$ et $x \in E$. Alors $(f \circ f^{-1})(x) = x$ donc $\|(f \circ f^{-1})(x)\| = \|x\|$. Mais par ailleurs $f \in \mathcal{O}(E)$ donc $\|(f \circ f^{-1})(x)\| = \|f(f^{-1}(x))\| = \|f^{-1}(x)\|$, donc $\|f^{-1}(x)\| = \|x\|$; ainsi $f^{-1} \in \mathcal{O}(E)$. \square

PROPOSITION 2.5. (1) *Soit \mathcal{B} une base orthonormée de E . Soit $f \in \mathcal{L}(E)$. Alors $f \in \mathcal{O}(E) \iff f(\mathcal{B})$ est une base orthonormée de E .*

(2) *Soient $f \in \mathcal{O}(E)$ et F un sev de E stable par f (autrement dit, tel que $f(F) \subset F$). Alors F^\perp est également stable par f (autrement dit, $f(F^\perp) \subset F^\perp$).*

(3) *Soit $f \in \mathcal{O}(E)$; alors $\text{Sp}f \subset \{-1; 1\}$ et $\text{Ker}(f - \text{id}_E) \perp \text{Ker}(f + \text{id}_E)$.*

DÉMONSTRATION. (1) On note $\mathcal{B} = (e_1, \dots, e_n)$.

Si on suppose que $f \in \mathcal{O}(E)$, alors pour tous $i, j \in \llbracket 1; n \rrbracket$, on a $\langle f(e_i), f(e_j) \rangle = \langle e_i, e_j \rangle = \delta_{i,j}$. La famille $f(\mathcal{B})$ est donc une famille orthonormée de n vecteurs de E ; autrement dit, c'est une base orthonormée.

Réciproquement, si on suppose que $f(\mathcal{B})$ est orthonormée, soit $x = x_1 e_1 + \dots + x_n e_n \in E$. Alors $f(x) = x_1 f(e_1) + \dots + x_n f(e_n)$, mais $f(\mathcal{B})$ est orthonormée donc $\|f(x)\|^2 = x_1^2 + \dots + x_n^2 = \|x\|^2$; on en déduit que f est une isométrie.

- (2) Préliminaire : $f(F) \subset F$, mais f étant un isomorphisme, $\dim f(F) = \dim F$; on en déduit que $F = f(F)$ (ou, autrement dit, que $F = f^{-1}(F)$).

Soit $x \in F^\perp$; il s'agit de montrer que $f(x) \in F^\perp$. Soit donc $y \in F$. $\langle f(x), y \rangle = \langle f(x), f(f^{-1}(y)) \rangle = \langle x, f^{-1}(y) \rangle$. Mais comme $F = f^{-1}(F)$, $f^{-1}(y) \in F$ donc $\langle x, f^{-1}(y) \rangle = 0$. Ceci étant vrai pour tout $y \in F$, on en déduit que $f(x) \in F^\perp$, ce qui montre que F^\perp est stable par f .

- (3) Soit $\lambda \in \text{Sp}(f)$; il existe donc $x \in E \setminus \{0\}$ tel que $f(x) = \lambda x$. L'application f étant une isométrie, on a $\|f(x)\| = \|x\|$, c'est-à-dire $|\lambda|\|x\| = \|x\|$. Le vecteur x n'étant pas nul, on en déduit que $|\lambda| = 1$ donc $\lambda = \pm 1$.

De plus, si $x \in \text{Ker}(f - \text{id})$ et $y \in \text{Ker}(f + \text{id})$ (autrement dit $f(x) = x$ et $f(y) = -y$), alors (comme f est une isométrie) $\langle x, y \rangle = \langle f(x), f(y) \rangle = \langle x, -y \rangle$; on en déduit que $\langle x, y \rangle = 0$. \square

PROPOSITION 2.6. *Soit F un sev de E . Alors s la symétrie orthogonale par rapport à F (c'est-à-dire la symétrie par rapport à F parallèlement à F^\perp) est une isométrie.*

DÉMONSTRATION. Soit $x \in E$. Il existe une unique écriture $x = y + z$ avec $y \in F$ et $z \in F^\perp$ et, par définition d'une symétrie, $s(x) = y - z$.

Mais alors $\|s(x)\|^2 = \|y - z\|^2 = \|y\|^2 + \|z\|^2 = \|y + z\|^2 = \|x\|^2$ (d'après le théorème de Pythagore, car y et z sont orthogonaux). La symétrie s est donc bien une isométrie. \square

REMARQUE. Une symétrie orthogonale par rapport à un hyperplan est appelé *réflexion*.

On rappelle que si s est la symétrie orthogonale par rapport à F , p la projection orthogonale sur F et p' la projection orthogonale sur F^\perp , alors

$$s = p - p'; \quad p = \frac{1}{2}(s + \text{id}); \quad p' = \frac{1}{2}(\text{id} - s).$$

3. Matrices orthogonales

DÉFINITION 3.1. Une matrice $A \in \mathcal{M}_n(\mathbf{R})$ est dite *orthogonale* si ses vecteurs colonnes forment une base orthonormée de \mathbf{R}^n (pour le produit scalaire usuel). On note $\mathcal{O}_n(\mathbf{R})$ ou $\mathcal{O}(n)$ l'ensemble des matrices orthogonales de taille n .

EXEMPLE(S). Si $\theta \in \mathbf{R}$, $R(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \in \mathcal{O}(2)$.

PROPOSITION 3.2. *Les propositions suivantes sont équivalentes :*

- (1) A est orthogonale ;
- (2) ${}^tAA = I_n$;
- (3) A est inversible et $A^{-1} = A^T$;
- (4) les lignes de A forment une base orthonormée de \mathbf{R}^n ;

DÉMONSTRATION. Je dis qu'il suffit de montrer l'équivalence entre (1) et (2). Pour cela, notons $A = (a_{i,j})_{1 \leq i,j \leq n}$ et écrivons le produit $A^T A$:

$$A^T A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}$$

On remarque que le coefficient sur la i -ème ligne et la j -ème colonne de $A^T A$ est donné par le produit scalaire des i -ème et j -ème colonnes de A . Les colonnes forment donc une base orthonormée si et seulement si $A^T A = I_n$, ce qui montre que (1) \iff (2).

L'équivalence (2) \iff (3) est immédiate.

Enfin, remarquons que $A^T A = I_n \iff AA^T = I_n$. On se convainc aisément que les coefficients de AA^T sont donnés par le produit scalaire des lignes de A entre elles, ce qui montre l'équivalence entre (2) et (3). \square

REMARQUE. Le critère (2) peut sembler être une façon pratique de vérifier si une matrice est orthogonale; en réalité, calculer $A^T A$ demande davantage de calcul que de vérifier que la base des vecteurs colonnes est orthonormée!

PROPOSITION 3.3. Soient \mathcal{B} une base orthonormée de E et $f \in \mathcal{L}(E)$. Alors $f \in \mathcal{O}(E)$ si et seulement si $\text{Mat}_{\mathcal{B}}(f) \in \mathcal{O}_n(\mathbf{R})$.

DÉMONSTRATION. Notons $M = \text{Mat}_{\mathcal{B}}(f)$.

On suppose $M \in \mathcal{O}_n(\mathbf{R})$. Soit $x \in E$ et notons X le vecteur de ses coordonnées dans la base \mathcal{B} . Alors MX est le vecteur des coordonnées de $f(x)$. Mais comme \mathcal{B} est orthonormée, $\|f(x)\|^2 = {}^t(MX)MX = X^T M^T M X = X^T X = \|x\|^2$, donc f est une isométrie.

Réciproquement, supposons que f est une isométrie. Soient $X, Y \in \mathbf{R}^n$ et notons x, y les vecteurs de E dont les coordonnées dans \mathcal{B} sont X et Y . Dans \mathbf{R}^n , on peut écrire

$$\langle X, (M^T M - I_n)Y \rangle = X^T (M^T M - I_n)Y = (MX)^T MY - X^T Y = \langle f(x), f(y) \rangle - \langle x, y \rangle = 0$$

car f est une isométrie. Ceci étant vrai pour tout $X \in \mathbf{R}^n$, on en déduit que $(M^T M - I_n)Y \in \mathbf{R}^\perp$, autrement dit $(M^T M - I_n)Y = 0$. Ceci étant vrai pour tout $Y \in \mathbf{R}^n$, on en déduit que $M^T M - I_n = 0$, donc $M^T M = I_n$, ce qui montre que M est une matrice orthogonale. \square

PROPOSITION 3.4. Le produit de deux matrices orthogonales est une matrice orthogonale; l'inverse d'une matrice orthogonale est une matrice orthogonale.

DÉMONSTRATION. Soient $M, N \in \mathcal{O}(n)$. Alors $(MN)^T MN = N^T M^T MN = N^T N = I_n$. De plus, $(M^{-1})^T M^{-1} = (M^T)^{-1} \times M^T = I_n$ d'après le point (3) de la prop. 3.2. \square

PROPOSITION 3.5. *Soit \mathcal{B} une base orthonormée de E et \mathcal{B}' une autre base de E . Alors \mathcal{B}' est orthonormée si et seulement si la matrice de passage de \mathcal{B} à \mathcal{B}' est orthogonale.*

DÉMONSTRATION. Par définition, la matrice de passage de \mathcal{B} à \mathcal{B}' a pour colonnes les coordonnées des vecteurs de \mathcal{B}' dans la base \mathcal{B} . Mais la base \mathcal{B} étant orthonormée, on peut calculer le produit scalaire des vecteurs de \mathcal{B}' à partir de leurs coordonnées dans \mathcal{B} ; on en déduit que \mathcal{B}' est orthonormée si et seulement si cette matrice est orthogonale. \square

THÉORÈME 3.6. *Pour tout $M \in \mathcal{O}_n(\mathbf{R})$, $\det M \in \{\pm 1\}$. L'ensemble des matrices orthogonales de déterminant 1 est appelé groupe spécial orthogonal, noté $\mathcal{SO}_n(\mathbf{R})$.*

DÉMONSTRATION. Pour tout $M \in \mathcal{O}_n(\mathbf{R})$, on a $M^T M = I_n$, donc $\det(M^T) \det M = 1$, autrement dit $(\det M)^2 = 1$, d'où $\det M = \pm 1$. \square

REMARQUE. Pour tout $n \geq 1$, il existe des matrices orthogonales de déterminant 1 (I_n par exemple) et de déterminant -1 (par exemple la matrice diagonale dont tous les coefficients valent 1 sauf le premier qui vaut -1).

Le produit de deux matrices, l'inverse d'une matrice du groupe spécial orthogonal sont toujours dans le groupe spécial orthogonal.

COROLLAIRE 3.7. *Pour toute isométrie f de E , on a $\det f \in \{\pm 1\}$. L'ensemble des isométries de E de déterminant 1 est appelé groupe spécial orthogonal de E , noté $\mathcal{SO}(E)$. Les éléments de $\mathcal{SO}(E)$ sont appelé isométries directes, ceux de $\mathcal{O}(E) \setminus \mathcal{SO}(E)$ isométries indirectes.*

Application : orientation d'un espace euclidien. Orienter un espace euclidien E , c'est en choisir une base orthonormée \mathcal{B}_0 de référence. D'après le paragraphe précédent, pour toute autre base orthonormée \mathcal{B} de E , la matrice de passage de \mathcal{B}_0 à \mathcal{B} est orthogonale, donc de déterminant ± 1 .

DÉFINITION 3.8. Une base orthonormée \mathcal{B} de E est dite *directe* si la matrice de passage de \mathcal{B}_0 à \mathcal{B} est de déterminant 1 (sinon, elle est dite *indirecte*).

PROPOSITION 3.9. *Si \mathcal{B} est une base orthonormée de E et f une isométrie de E , alors la base orthonormée $f(\mathcal{B})$ est de même orientation que \mathcal{B} si et seulement si l'isométrie f est directe.*

4. Diagonalisation des matrices symétriques réelles

THÉORÈME 4.1. *Les sous-espaces propres d'une matrice symétrique réelle sont deux à deux orthogonaux.*

DÉMONSTRATION. Soit M une matrice symétrique réelle et $\lambda, \mu \in \mathbf{R}$ deux valeurs propres distinctes de M . Alors, pour tout $X \in \text{Ker}(M - \lambda I_n)$ et pour tout $Y \in \text{Ker}(M - \mu I_n)$, on a $(MX)^T Y = (\lambda X)^T Y = \lambda \langle X, Y \rangle$. Mais, par ailleurs, M étant symétrique,

$M^T = M$ donc $(MX)^TY = X^TMY = \mu X^TY = \mu \langle X, Y \rangle$. On a donc $\lambda \langle X, Y \rangle = \mu \langle X, Y \rangle$, mais comme $\lambda \neq \mu$, on en déduit que $\langle X, Y \rangle = 0$. Ceci étant vrai pour tout vecteur propre X associé à λ et pour tout vecteur propre Y associé à μ , on en déduit que les deux sous-espaces propres sont orthogonaux. \square

THÉORÈME 4.2 (Théorème spectral). *Toute matrice symétrique réelle est diagonalisable en base orthonormée. Autrement dit, pour tout $A \in \mathcal{M}_n(\mathbf{R})$, si A est symétrique, alors il existe $P \in \mathcal{O}_n(\mathbf{R})$ et $D \in \mathcal{M}_n(\mathbf{R})$ diagonale telles que $P^{-1}AP = P^TAP = D$.*

ADMIS. \square

5. Classification des isométries du plan et de l'espace

Dans cette partie, on cherche à décrire tous les éléments des groupes $\mathcal{O}(2)$ et $\mathcal{O}(3)$.

5.1. Isométries du plan.

THÉORÈME 5.1. *Soit $M \in \mathcal{O}_2(\mathbf{R})$. Alors il existe $\theta \in \mathbf{R}$ tel que $M = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ (si $\det M = 1$) ou $M = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$ (si $\det M = -1$).*

DÉMONSTRATION. Notons $M = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. Si $M \in \mathcal{O}_2(\mathbf{R})$, alors $tM = M^{-1}$, autrement dit $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{\varepsilon} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$, où $\varepsilon = ad - bc = \pm 1$. On en déduit les trois équations $d = \varepsilon a$, $c = -\varepsilon b$ et $a^2 + b^2 = 1$; il existe donc $\theta \in \mathbf{R}$ tel que $a = \cos \theta$ et $b = \sin \theta$, et on a donc $d = \varepsilon \cos \theta$ et $c = -\varepsilon \sin \theta$. \square

REMARQUE. La réciproque est clairement vraie : toutes les matrices d'une des formes ci-dessus sont bien dans $\mathcal{O}_2(\mathbf{R})$.

THÉORÈME 5.2. *Soit E un espace euclidien orienté de dimension 2, soit $\mathcal{B} = (i, j)$ une base orthonormée directe de E . Alors, pour tout $f \in \mathcal{O}(E)$, la matrice de f dans la base \mathcal{B} est de la forme $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ (si f est directe) ou $\begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$ (si f est indirecte).*

Dans le premier cas, f est la rotation vectorielle d'angle θ ; dans le second cas, f est la réflexion par rapport à la droite vectorielle formant un angle de $\frac{\theta}{2}$ avec l'axe des abscisses.

DÉMONSTRATION. D'après le théorème précédent la matrice de f dans la base \mathcal{B} est bien d'une des deux formes proposées.

Supposons que la matrice de f est $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$. Soit $u \in E$, notons $\begin{pmatrix} x \\ y \end{pmatrix}$ les coordonnées de u dans la base \mathcal{B} . Alors :

$$\cos(\widehat{u, f(u)}) = \frac{\langle u, f(u) \rangle}{\|u\| \cdot \|f(u)\|} = \frac{\begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} x \cos \theta - y \sin \theta \\ x \sin \theta + y \cos \theta \end{pmatrix}}{\|u\|^2} = \frac{x^2 \cos \theta + y^2 \cos \theta}{x^2 + y^2} = \cos \theta;$$

$$\sin(\widehat{u, f(u)}) = \frac{\det_{\mathcal{B}}(u, f(u))}{\|u\| \cdot \|f(u)\|} = \frac{\begin{vmatrix} x & x \cos \theta - y \sin \theta \\ y & x \sin \theta + y \cos \theta \end{vmatrix}}{x^2 + y^2} = \frac{x^2 \sin \theta + y^2 \sin \theta}{x^2 + y^2} = \sin \theta.$$

Ainsi, $(\widehat{u, f(u)}) = \theta$ donc f est bien la rotation vectorielle d'angle θ .

Supposons maintenant que la matrice de f est $M = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$. Remarquons que $M^2 = I_2$ donc f est une symétrie.

Le polynôme caractéristique de M est $\chi_M(X) = \begin{vmatrix} X - \cos \theta & -\sin \theta \\ -\sin \theta & X + \cos \theta \end{vmatrix} = X^2 - \cos^2 \theta - \sin^2 \theta = (X - 1)(X + 1)$

Posons $F = \text{Ker}(f - \text{id})$ et $G = \text{Ker}(f + \text{id})$. Ces deux sous-espaces propres sont de dimension 1 (car les valeurs propres sont simples) et orthogonaux entre eux car M est symétrique réelle. L'isométrie f est donc une symétrie orthogonale, plus précisément la réflexion d'axe la droite F . Reste à déterminer un vecteur directeur de F .

Soit $\varphi \in \mathbf{R}$. On considère le vecteur u de coordonnées $\begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}$ dans la base \mathcal{B} . Alors

$$u \in F \iff \begin{cases} \cos \theta \cos \varphi + \sin \theta \sin \varphi = \cos \varphi \\ \sin \theta \cos \varphi - \cos \theta \sin \varphi = \sin \varphi \end{cases} \iff \begin{cases} \cos(\theta - \varphi) = \cos \varphi \\ \sin(\theta - \varphi) = \sin \varphi \end{cases}$$

donc ssi $\theta - \varphi = \varphi$, autrement dit $\varphi = \frac{\theta}{2}$.

L'axe de symétrie forme donc un angle de $\frac{\theta}{2}$ avec l'horizontale. \square

REMARQUE. En dimension 2, les isométries (vectorielles) directes sont donc les rotations et les isométries (vectorielles) indirectes les réflexions.

5.2. Isométries de l'espace. Soient E un espace euclidien de dimension 3 et f une isométrie de E .

Remarquons que le polynôme caractéristique χ_f de f est de degré 3 et à coefficients réels; il admet nécessairement une racine réelle. On sait de plus que $\text{Sp}f \subset \{-1; 1\}$, donc cette racine réelle est ± 1 .

Considérons donc les sous-espaces $E_1 = \text{Ker}(f - \text{id})$ et $E_{-1} = \text{Ker}(f + \text{id})$. Au moins l'un de ces deux espaces n'est pas réduit à $\{0\}$. On va séparer les cas suivant la dimension de E_{-1} .

5.2.1. *Si* $\dim E_{-1} = 0$. Alors $\dim E_1 \geq 1$. Considérons donc k un vecteur de E_1 de norme 1, et (i, j) deux vecteurs de E tels que (i, j, k) soit une BOND de E .

Comme k est un vecteur propre de f , $\text{Vect}(k)$ est stable par f , donc $\text{Vect}(k)^\perp$ est également stable par f ; autrement dit, $\text{Vect}(i, j)$ est stable par f . L'application f induit donc une isométrie \tilde{f} du sous-espace $\text{Vect}(i, j)$, qui est de dimension 2; d'après le paragraphe précédent, \tilde{f} est soit une rotation, soit une réflexion. Mais comme -1 n'est pas valeur propre de f , on en déduit que $\det \tilde{f} = 1$, donc il s'agit d'une rotation.

Bilan : il existe $\theta \in \mathbf{R}$ tel que $\text{Mat}_{(i,j,k)} f = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$; l'isométrie f est la rotation d'axe k et d'angle θ .

5.2.2. *Si* $\dim E_{-1} = 1$. On considère un vecteur unitaire $k \in E_{-1}$ et deux vecteurs i, j tels que (i, j, k) soit une BOND.

De la même façon que précédemment, $\text{Vect}(k)$ est stable par f , qui induit donc une isométrie \tilde{f} sur $\text{Vect}(i, j)$. On remarque que -1 n'est pas valeur propre de \tilde{f} , qui est donc encore une rotation.

Bilan : il existe $\theta \in \mathbf{R}$ tel que $\text{Mat}_{(i,j,k)} f = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & -1 \end{pmatrix}$; l'isométrie f est la composée de la rotation d'axe k et d'angle θ avec la réflexion par rapport à $\text{Vect}(k)^\perp$.

5.2.3. *Si* $\dim E_{-1} = 2$. On pose (i, j) une base orthonormée de E_{-1} que l'on complète par $k = i \wedge j$ de telle sorte à avoir une BOND (i, j, k) de E . Comme précédemment, E_{-1} est stable par f , donc $E_{-1}^\perp = \text{Vect}(k)$ est stable par f ; autrement dit, k est un vecteur propre de f . Mais comme $k \notin E_{-1}$, nécessairement $k \in E_1$ donc $f(k) = k$.

Bilan : $\text{Mat}_{(i,j,k)} f = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$; l'isométrie f est la rotation d'axe k et d'angle π .

5.2.4. *Si* $\dim E_{-1} = 3$. Dans ce cas, $f = -\text{id}$; dans n'importe quelle base de E , on aura $\text{Mat} f = -I_3$; l'isométrie f est la symétrie centrale. C'est aussi, pour tout vecteur k unitaire, la composée de la rotation d'axe k et d'angle π avec la réflexion d'axe $\text{Vect}(k)^\perp$.

THÉORÈME 5.3. *Si* E est de dimension 3, les isométries de E sont les rotations vectorielles (isométries directes) et les composées d'une rotation et d'une réflexion par rapport au plan orthogonal à l'axe de rotation (isométries indirectes).

REMARQUE. (1) en dimension 3, pour décrire correctement une rotation, il est important d'orienter son axe (en donner un vecteur directeur); en effet, remplacer k par $-k$ a pour conséquence de changer l'angle de θ en $-\theta$.

(2) plusieurs isométries peuvent sembler absentes de notre description mais sont en fait prises en compte : les réflexions (mais ce sont les composées de réflexions avec

une rotation d'angle 0) et les symétries orthogonales par rapport à une droite (mais ce sont les rotations d'axe cette droite et d'angle π).

Équations différentielles

« Le savant n'étudie pas la nature parce que cela est utile ; il l'étudie parce qu'il y prend plaisir et il y prend plaisir parce qu'elle est belle. »

H. POINCARÉ, *Science et Méthode* (1908)

Dans tout le chapitre, I est un intervalle de \mathbf{R} de longueur non nulle, et $\mathbf{K} = \mathbf{R}$ ou \mathbf{C} .

1. Équations différentielles linéaires du premier ordre

Ce sont les équations du type $(E): y' + a(t)y = b(t)$, où a et b sont deux applications continues $I \rightarrow \mathbf{K}$. L'équation $(E_0): y' + a(t)y = 0$ est appelée *équation homogène* associée à (E) et b est appelé *second membre* de (E) .

On effectue la résolution en deux parties :

(1) On détermine une primitive A de a . En effet,

$$y' + a(t)y = 0 \iff \exists \lambda \in \mathbf{K} \quad \forall t \in I \quad y(x) = \lambda e^{-A(t)}.$$

(2) On résout (E_0) à l'aide du changement de fonction inconnue défini par $y = \lambda(t)e^{-A(t)}$ (technique dite de *variation de la constante*). On a alors

$$\begin{aligned} y' + a(t)y = b(t) &\iff \lambda'(t)e^{-A(t)} - \lambda(t)A'(t)e^{-A(t)} + a(t)\lambda(t)e^{-A(t)} = b(t) \\ &\iff \lambda'(t)e^{-A(t)} = b(t) \\ &\iff \lambda'(t) = b(t)e^{A(t)}. \end{aligned}$$

On effectue alors un calcul de primitive de $b(t)e^{A(t)}$ pour déterminer $\lambda(t)$ (à une constante additive près), on en déduit alors toutes les solutions $y = (\lambda(t)+c)e^{-A(t)}$, $c \in \mathbf{K}$.

REMARQUE. (1) L'ensemble des solutions de (E_0) est un \mathbf{K} -ev de dimension 1.

(2) Si φ est une solution de (E) , alors l'ensemble des solutions de (E) s'écrit $\{y + \varphi \mid y \text{ solution de } (E_0)\}$. En d'autres termes, si on connaît une « solution particulière » de (E) , on obtient toutes les solutions de (E) en ajoutant cette solution particulière à toutes les solutions de l'équation homogène associée.

(3) Pour tout $t_0 \in I$ et tout $a \in \mathbf{K}$, il existe une *unique* solution y de (E) telle que $y(t_0) = a$ (théorème de Cauchy).

- (4) Il est important de toujours résoudre les équations différentielles sur des *intervalles*. En effet, une fonction dont la dérivée est nulle n'est pas forcément constante si elle n'est pas définie sur un intervalle.

2. Équations différentielles linéaires du second ordre

On s'intéresse à l'équation $(E): y'' + a(t)y' + b(t)y = c(t)$, où $a, b, c \in \mathcal{C}(I, \mathbf{K})$. L'équation $(E_0): y'' + a(t)y' + b(t)y = 0$ est l'équation homogène associée à (E) et c est le *second membre*.

THÉORÈME 2.1 (Théorème de Cauchy). *Pour tous $t_0 \in I$ et $\alpha, \beta \in \mathbf{K}$, il existe une unique solution de (E) sur I telle que $y(t_0) = \alpha$ et $y'(t_0) = \beta$.*

ADMIS. □

COROLLAIRE 2.2. *L'ensemble des solutions de (E_0) est un \mathbf{K} -ev de dimension 2.*

DÉMONSTRATION. Notons \mathcal{S}_0 l'ensemble des solutions de (E_0) . Montrons que c'est un sous-ev de $\mathcal{C}^2(I, \mathbf{K})$ de dimension 2.

Il est clair que $0 \in \mathcal{S}_0$, ainsi $\mathcal{S}_0 \neq \emptyset$.

Si $y_1, y_2 \in \mathcal{S}_0$ et $\lambda_1, \lambda_2 \in \mathbf{K}$, alors $(\lambda_1 y_1 + \lambda_2 y_2)'' + a(t)(\lambda_1 y_1 + \lambda_2 y_2)' + b(t)(\lambda_1 y_1 + \lambda_2 y_2) = 0$ donc $\lambda_1 y_1 + \lambda_2 y_2 \in \mathcal{S}_0$. Ceci démontre que \mathcal{S}_0 est un sous-ev de $\mathcal{C}^2(I, \mathbf{K})$.

Enfin, si on fixe $t_0 \in I$, d'après le théorème de Cauchy, l'application $\varphi: \mathcal{S}_0 \rightarrow \mathbf{K}^2$ définie par $\varphi(f) = (f(t_0), f'(t_0))$ est un isomorphisme. On a donc $\dim \mathcal{S}_0 = \dim \mathbf{K}^2 = 2$. □

PROPOSITION 2.3 (Principe de superposition). *Soient $a, b, c_1, c_2: I \rightarrow \mathbf{K}$ continues. Si φ_1 est une solution de $y'' + a(t)y' + b(t)y = c_1(t)$ et φ_2 est une solution de $y'' + a(t)y' + b(t)y = c_2(t)$, alors $\varphi_1 + \varphi_2$ est solution de $y'' + a(t)y' + b(t)y = c_1(t) + c_2(t)$.*

DÉMONSTRATION.

$$\begin{aligned} (\varphi_1 + \varphi_2)'' + a(t)(\varphi_1 + \varphi_2)' + b(t)(\varphi_1 + \varphi_2) &= \varphi_1'' + a(t)\varphi_1' + b(t)\varphi_1 + \varphi_2'' + a(t)\varphi_2' + b(t)\varphi_2 \\ &= c_1(t) + c_2(t). \end{aligned} \quad \square$$

REMARQUE. Si φ est une solution de (E) , alors l'ensemble des solutions de (E) s'écrit $\{y + \varphi \mid y \text{ solution de } (E_0)\}$. En d'autres termes, si on connaît une « solution particulière » de (E) , on obtient toutes les solutions de (E) en ajoutant cette solution particulière à toutes les solutions de l'équation homogène associée.

Cas particulier. Si on connaît une solution y_1 de (E_0) ne s'annulant pas sur I , alors on cherche y sous la forme $y = zy_1$. En effet, dans ce cas,

$$\begin{aligned} y'' + a(t)y' + b(t)y = c(t) &\iff z''y_0 + 2z'y_0' + zy_0'' + a(t)z'y_0 + a(t)zy_0' + b(t)zy_0 = c(t) \\ &\iff z''y_0 + 2z'y_0' + a(t)z'y_0 = c(t), \end{aligned}$$

et on s'est ainsi ramené à une équation différentielle d'ordre 1, d'inconnue z' .

On peut obtenir y_0 à partir d'une indication de l'énoncé, comme solution évidente, comme solution développable en série entière, etc. On peut éventuellement garantir la condition de non-annulation en restreignant l'intervalle d'étude I (il faut alors gérer les éventuels problèmes de recollement par la suite).

Cette idée est particulièrement pratique pour terminer la résolution de l'équation homogène car on a alors $c(t) = 0$.

Rappel : cas où les coefficients sont constants. On se place dans le cas où les fonctions a et b sont constantes. Dans ce cas, on résout l'équation caractéristique $r^2 + ar + b = 0$.

Si cette équation possède une unique solution r , alors les solutions de l'équation homogène (E_0) sont les fonctions $t \mapsto (\lambda t + \mu)e^{rt}$, $\lambda, \mu \in \mathbf{K}$.

Si cette équation possède deux solutions r_1 et r_2 (réelles ou complexes), alors les solutions de (E_0) sont les fonctions $t \mapsto \lambda e^{r_1 t} + \mu e^{r_2 t}$, $\lambda, \mu \in \mathbf{K}$.

Dans le cas particulier d'une équation à coefficients a, b réels, si les racines r_1 et r_2 sont complexes, elles sont nécessairement conjuguées et s'écrivent $r = k \pm i\omega$. Les solutions *réelles* de (E_0) sont alors les fonctions $t \mapsto (\lambda \cos \omega t + \mu \sin \omega t)e^{kt}$, $\lambda, \mu \in \mathbf{R}$, ou encore $t \mapsto \lambda \cos(\omega t + \varphi)e^{kt}$, $\lambda, \varphi \in \mathbf{R}$.

3. Systèmes différentiels linéaires

DÉFINITION 3.1. Un système différentiel linéaire du premier ordre est un système d'équations différentielles de la forme $X' = A(t)X + B(t)$, où

$$X: t \mapsto \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}, \quad A: t \mapsto \begin{pmatrix} a_{1,1}(t) & \cdots & a_{1,n}(t) \\ \vdots & & \vdots \\ a_{n,1}(t) & \cdots & a_{n,n}(t) \end{pmatrix} \quad \text{et} \quad B: t \mapsto \begin{pmatrix} b_1(t) \\ \vdots \\ b_n(t) \end{pmatrix},$$

les fonctions $a_{i,j}$ et b_i étant toutes définies et continues sur un intervalle I ; les fonctions x_i étant des fonctions dérivables sur I à déterminer.

EXEMPLE(S). Cas de deux masses couplées avec trois ressorts :

$$\begin{cases} m_1 x_1'' = -k_1 x_1 + k_{1,2}(x_2 - x_1) \\ m_2 x_2'' = -k_2 x_2 - k_{1,2}(x_2 - x_1) \end{cases}$$

donc, en posant $x_3 = x_1'$ et $x_4 = x_2'$,

$$\begin{cases} x_1' = x_3 \\ x_2' = x_4 \\ x_3' = \frac{1}{m_1} (-k_1 x_1 + k_{1,2}(x_2 - x_1)) \\ x_4' = \frac{1}{m_2} (-k_2 x_2 - k_{1,2}(x_2 - x_1)). \end{cases}$$

3.1. Cas d'un système homogène à coefficients constants. Il s'agit du cas où le système est de la forme $X' = AX$ pour $A \in \mathcal{M}_n(\mathbf{K})$; on a alors $I = \mathbf{R}$.

THÉORÈME 3.2 (Théorème de Cauchy). *Pour tout $t_0 \in \mathbf{R}$ et pour tout $X_0 \in \mathbf{R}^n$, il existe une unique solution de l'équation $X' = AX$ telle que $X(t_0) = X_0$.*

ADMIS. □

COROLLAIRE 3.3. *L'ensemble des solutions de $X' = AX$ est un \mathbf{K} -ev de dimension n .*

DÉMONSTRATION. Notons \mathcal{S} l'ensemble des solutions. C'est un sous-espace vectoriel de l'ensemble des fonctions de \mathbf{R} dans \mathbf{K}^n (il contient 0 et il est stable par combinaison linéaire).

De plus, si $t_0 \in \mathbf{R}$, alors l'application $\varphi: \mathcal{S} \rightarrow \mathbf{K}^n$ définie par $\varphi(X) = X(t_0)$ est un isomorphisme d'après le théorème de Cauchy, on en déduit que $\dim \mathcal{S} = \dim \mathbf{K}^n = n$. □

THÉORÈME 3.4. *Soient $\lambda \in \mathbf{K}$ et $v_0 \in \mathbf{K}^n \setminus \{0\}$. On considère l'application $\varphi: t \mapsto e^{\lambda t}v_0$. Alors φ est solution de $X' = AX$ si et seulement si λ est une valeur propre de A et v_0 un vecteur propre associé à λ .*

DÉMONSTRATION.

$$\begin{aligned} \varphi \text{ solution de } X' = AX &\iff \forall t \in \mathbf{R} \quad \varphi'(t) = A\varphi(t) \\ &\iff \forall t \in \mathbf{R} \quad \lambda e^{\lambda t}v_0 = Ae^{\lambda t}v_0 \\ &\iff \lambda v_0 = Av_0. \end{aligned} \quad \square$$

THÉORÈME 3.5 (Cas où A est diagonalisable). *Si A est diagonalisable, soit (v_1, \dots, v_n) une base de \mathbf{R}^n constituée de vecteurs propres de A et $\lambda_1, \dots, \lambda_n$ les valeurs propres correspondantes. Alors l'ensemble des solutions du système $X' = AX$ est $\mathcal{S} = \text{Vect}\{t \mapsto e^{\lambda_i t}v_i \mid i \in \llbracket 1; n \rrbracket\}$.*

DÉMONSTRATION. On sait d'après le théorème de Cauchy que \mathcal{S} est un \mathbf{R} -ev de dimension n . Or le théorème ci-dessus nous donne n solutions $t \mapsto e^{\lambda_i t}v_i$ pour tout $i \in \llbracket 1; n \rrbracket$. Il suffit de montrer que ces solutions forment une famille libre.

Soient $\alpha_1, \dots, \alpha_n \in \mathbf{R}$. On suppose que la fonction $t \mapsto \alpha_1 e^{\lambda_1 t}v_1 + \dots + \alpha_n e^{\lambda_n t}v_n$ est constamment nulle. Mais alors, pour tout $t \in \mathbf{R}$, $\alpha_1 e^{\lambda_1 t}v_1 + \dots + \alpha_n e^{\lambda_n t}v_n = 0$, mais comme la famille (v_1, \dots, v_n) est libre, alors $\alpha_1 e^{\lambda_1 t} = \dots = \alpha_n e^{\lambda_n t} = 0$, d'où $\alpha_1 = \dots = \alpha_n = 0$. □

EXEMPLE(S). Résoudre l'équation du second ordre $y'' - 3y' + 2y = 0$.

Dans le cas, moins favorable, où A n'est que trigonalisable : il existe une matrice inversible P et une matrice triangulaire T telle que $A = PTP^{-1}$. Mais alors

$$X' = AX \iff X' = PTP^{-1}X \iff P^{-1}X' = TP^{-1}X \iff Y' = TY,$$

om $Y = P^{-1}X$. On résoud alors le système pour Y et on en déduit les solutions pour X via le calcul $X = PY$. La résolution du système $Y' = TY$ est facile car il existe une ligne où une seule fonction inconnue apparaît, qu'on peut résoudre et ensuite injecter dans les autres lignes... et ainsi résoudre le système de proche en proche.

EXEMPLE(S). Résoudre l'équation du second ordre $y'' - 2y' + y = 0$.

Enfin, si A est une matrice réelle non trigonalisable, on sait qu'elle est trigonalisable sur \mathbf{C} car tous les polynômes à coefficients complexes sont scindés. On utilise alors le paragraphe précédent pour trouver toutes les solutions *complexes* de l'équation différentielle. On conclut alors à l'aide du lemme suivant :

LEMME 3.6. *Si X est une solution à valeurs complexes du système différentiel $X' = AX$ où A est à coefficients réels, alors $\operatorname{Re}(X)$ et $\operatorname{Im}(X)$ sont des solutions du même système.*

DÉMONSTRATION. A est à coefficients réels donc $\operatorname{Re}(AX) = A \operatorname{Re}(X)$; par ailleurs $\operatorname{Re}(X') = \operatorname{Re}(X)'$, donc $\operatorname{Re}(X)' = A \operatorname{Re}(X)$; de même pour $\operatorname{Im}(X)$. \square

EXEMPLE(S). Résoudre l'équation du second ordre $y'' + y = 0$.

THÉORÈME 3.7 (Comportement asymptotique des solutions). *Si, pour tout $\lambda \in \operatorname{Sp}(A)$, $\operatorname{Re}(\lambda) < 0$, alors toutes les solutions du système différentiel $X' = AX$ tendent vers 0 en $+\infty$.*

DÉMONSTRATION. Si A est diagonalisable, les solutions de $X' = AX$ sont des combinaisons linéaires de $t \mapsto e^{\lambda t}v$ avec $\lambda \in \operatorname{Sp}(A)$ et $v \in \mathbf{K}_n$. Or, pour tout $t \in \mathbf{R}$, $|e^{\lambda t}| = e^{\operatorname{Re}(\lambda)t} \rightarrow 0$, d'où le résultat.

On admet que le théorème reste vrai dans le cas où A n'est pas diagonalisable. \square

3.2. Lien avec les équations différentielles linéaires d'ordre supérieur.

REMARQUE. L'équation différentielle $y^{(n)} + \sum_{k=0}^{n-1} a_k(t)y^{(k)} = 0$ est équivalente au système linéaire $X' = A(t)X$, avec

$$A(t) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \\ -a_0(t) & \cdots & \cdots & \cdots & -a_{n-1}(t) \end{pmatrix}$$

COROLLAIRE 3.8. *L'ensemble des solutions de $y^{(n)} + \sum_{k=0}^{n-1} a_k y^{(k)}$ est un \mathbf{K} -ev de dimension n .*

En pratique, pour résoudre l'équation $y^{(n)} + \sum_{k=0}^{n-1} a_k y^{(k)}$, on doit réduire la matrice A donc déterminer les racines de son polynôme caractéristique. Cela revient à résoudre

l'équation $x^n + \sum_{k=0}^{n-1} a_k x^k = 0$, qu'on appelle *équation caractéristique* de l'équation différentielle.

En particulier, si l'équation admet n racines distinctes r_1, \dots, r_n , la matrice A est diagonalisable et les solutions sont alors les fonctions $t \mapsto a_1 e^{r_1 t} + \dots + a_n e^{r_n t}$, $a_1, \dots, a_n \in \mathbf{K}$.

Fonctions de plusieurs variables réelles

$$\frac{\partial u}{\partial t} = \alpha \Delta u$$

J. FOURIER, *Théorie analytique de la chaleur* (1822)

Dans tout le chapitre, on fixe n un entier naturel supérieur ou égal à 2 (en pratique, $n = 2$ ou 3); l'espace \mathbf{R}^n est muni de sa structure euclidienne canonique.

1. Éléments de topologie de \mathbf{R}^n

On rappelle que si $a, b \in \mathbf{R}^n$, alors la distance entre a et b est $d(a, b) = \|a - b\|$.

DÉFINITION 1.1. Soient $a \in \mathbf{R}^n$ et $r \in \mathbf{R}_+$. On appelle :

- *boule ouverte* de centre a et de rayon r l'ensemble $B(a, r) = \{x \in \mathbf{R}^n \mid \|x - a\| < r\}$;
- *boule fermée* de centre a et de rayon r l'ensemble $\bar{B}(a, r) = \{x \in \mathbf{R}^n \mid \|x - a\| \leq r\}$.

REMARQUE. Les boules ouvertes de rayon nul sont vides; les boules fermées de rayon nul sont les singletons.

DÉFINITION 1.2. Une partie $A \subset \mathbf{R}^n$ est dite *bornée* s'il existe $M \geq 0$ tel que, pour tout $a \in A$, $\|a\| \leq M$; autrement dit, tel que $A \subset \bar{B}(0, M)$.

DÉFINITION 1.3. Soit $U \subset \mathbf{R}^n$. On dit que U est une partie *ouverte* de \mathbf{R}^n si, pour tout $a \in U$, il existe $r > 0$ tel que $B(a, r) \subset U$.

Une partie de \mathbf{R}^n est dite *fermée* si la partie complémentaire est ouverte.

REMARQUE. Attention : *ouvert* et *fermé* ne sont pas des contraires; il est possible pour une partie d'être ni ouverte ni fermée, ou d'être à la fois ouverte et fermée.

DÉFINITION 1.4. Soit $A \subset \mathbf{R}^n$. On appelle :

- *intérieur* de A le plus grand ouvert $\overset{\circ}{A}$ contenu à l'intérieur de A ;
- *adhérence* de A le plus petit fermé \bar{A} contenant A ;
- *frontière* de A l'ensemble $\bar{A} \setminus \overset{\circ}{A}$.

2. Limites et continuité

DÉFINITION 2.1. Soient A une partie de \mathbf{R}^n et $f: A \rightarrow \mathbf{R}$. Soit $a \in \bar{A}$.

On dit que f admet une limite finie $\ell \in \mathbf{R}$ pour x tendant vers a si, pour tout $\varepsilon > 0$, il existe $r > 0$ tel que, pour tout $x \in A$, si $\|x - a\| \leq r$, alors $|f(x) - \ell| \leq \varepsilon$.

On dit que f tend vers $+\infty$ (resp. $-\infty$) lorsque x tend vers A si, pour tout $M \in \mathbf{R}$, il existe $r > 0$ tel que, pour tout $x \in A$, si $\|x - a\| \leq r$, alors $f(x) \geq M$ (resp. $f(x) \leq M$).

EXEMPLE(S). Soit f la fonction définie sur $\mathbf{R}^2 \setminus \{(0, 0)\}$ par $f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}$. Montrer que f n'admet pas de limite en $(0, 0)$.

DÉFINITION 2.2. Soient $U \subset \mathbf{R}^n$ et $f: U \rightarrow \mathbf{R}$.

- Soit $a \in U$. On dit que f est *continue* en a si elle admet une limite finie en a .
- On dit que f est *continue* sur U si elle est continue en a pour tout $a \in U$.

PROPOSITION 2.3 (Théorème de composition). *Une somme, une différence, un produit, une composée de fonctions continues est continue. L'inverse d'une fonction continue ne s'annulant pas est continue.*

ADMIS. □

EXEMPLE(S). La fonction f définie par $f(x, y) = \frac{\cos(x) + \sin(xy)}{x^2 + e^y}$ est continue sur \mathbf{R}^2 par composition.

THÉORÈME 2.4 (Théorème de Weierstrass). *Une fonction continue sur une partie fermée et bornée de \mathbf{R}^n est bornée et « atteint ses bornes » (autrement dit, elle admet un maximum et un minimum).*

ADMIS. □

3. Dérivées partielles

DÉFINITION 3.1. Soient U un ouvert de \mathbf{R}^n et $f: U \rightarrow \mathbf{R}$. Soit $a = (a_1, \dots, a_n) \in U$. On dit que f admet une *dérivée partielle* par rapport à la i -ème variable au point a si la fonction $\varphi: x \mapsto f(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_n)$ est dérivable en a_i . On note alors

$$\varphi'(a_i) = \partial_i f(a) = \frac{\partial f}{\partial x_i}(a).$$

EXEMPLE(S). Calculer les dérivées partielles de la fonction $f: (x, y, z) \mapsto \frac{e^{x+y}}{1+z^2}$.

DÉFINITION 3.2. Soient U un ouvert de \mathbf{R}^n , $f: U \rightarrow \mathbf{R}$ et $a \in U$. On suppose que f admet des dérivées partielles par rapport à toutes les variables au point a . On appelle *gradient* de f en a le vecteur

$$\nabla f(a) = \begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_n f(a) \end{pmatrix} \in \mathbf{R}^n.$$

EXEMPLE(S). Calculer le gradient de la fonction f de l'exemple précédent.

REMARQUE. Les formules usuelles de dérivation de la somme, du produit, etc. restent valables.

PROPOSITION 3.3. Soient x et y deux fonctions définies sur un intervalle I , et $f: \mathbf{R}^2 \rightarrow \mathbf{R}$. On note, pour tout $t \in I$, $F(t) = f(x(t), y(t))$. Alors, sous réserve d'existence

$$F'(t) = \frac{\partial f}{\partial x}(x(t), y(t)) \times x'(t) + \frac{\partial f}{\partial y}(x(t), y(t)) \times y'(t).$$

PROPOSITION 3.4. Soient f, g, h trois fonctions $\mathbf{R}^2 \rightarrow \mathbf{R}$. On note $F: (u, v) \mapsto h(f(u, v), g(u, v))$. Alors, sous réserve d'existence,

$$\begin{aligned} \frac{\partial F}{\partial u}(u, v) &= \frac{\partial h}{\partial f}(f(u, v), g(u, v)) \times \frac{\partial f}{\partial u}(u, v) + \frac{\partial h}{\partial g}(f(u, v), g(u, v)) \times \frac{\partial g}{\partial u}(u, v); \\ \frac{\partial F}{\partial v}(u, v) &= \frac{\partial h}{\partial f}(f(u, v), g(u, v)) \times \frac{\partial f}{\partial v}(u, v) + \frac{\partial h}{\partial g}(f(u, v), g(u, v)) \times \frac{\partial g}{\partial v}(u, v). \end{aligned}$$

REMARQUE. La généralisation au cas de fonction de plus de 2 variables est immédiate (mais pénible à écrire).

DÉFINITION 3.5. Soit U un ouvert de \mathbf{R}^n et $f: U \rightarrow \mathbf{R}$. On dit que f est de classe \mathcal{C}^1 sur U si elle admet des dérivées partielles en tout point de U et que les fonctions dérivées partielles sont toutes continues sur U .

PROPOSITION 3.6. Soit U un ouvert de \mathbf{R}^n et f une fonction de classe \mathcal{C}^1 sur U . Soit $a = (a_1, \dots, a_n) \in U$. Alors, pour tout $x \in U$ tendant vers a ,

$$\begin{aligned} f(x) &= f(a) + \sum_{i=1}^n (x_i - a_i) \frac{\partial f}{\partial x_i}(a) + o(\|x - a\|) \\ &= f(a) + (x - a) \cdot \nabla f(a) + o(\|x - a\|). \end{aligned}$$

DÉFINITION 3.7. Soit f une fonction définie sur un ouvert $U \subset \mathbf{R}^n$ et à valeurs dans \mathbf{R} . Alors, pour tous $(i, j) \in \llbracket 1; n \rrbracket^2$ et sous réserve d'existence, on appelle *dérivée partielle seconde* d'ordres (i, j) la fonction

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right).$$

Si $i = j$, on note $\frac{\partial^2 f}{\partial x_i^2}$.

Si toutes les dérivées partielles secondes de f existent et sont continues, on dit que f est de classe \mathcal{C}^2 .

THÉORÈME 3.8 (Théorème de Schwarz). Si f est de classe \mathcal{C}^2 , alors pour tous $(i, j) \in \llbracket 1; n \rrbracket^2$,

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

4. Extrema des fonctions de plusieurs variables

DÉFINITION 4.1. Soient $U \subset \mathbf{R}^n$ et $f: U \rightarrow \mathbf{R}$.

- si $a \in U$, on dit que f admet un *maximum global* en a si, pour tout $x \in U$, $f(x) \leq f(a)$;

- si $a \in \overset{\circ}{U}$, on dit que f admet un *maximum local* en a s'il existe $r > 0$ tel que $B(a, r) \subset U$ et, pour tout $x \in B(a, r)$, $f(x) \leq f(a)$.

On définit de manière analogue les notions de minimum global et local. Le terme générique pour désigner un maximum ou un minimum est *extremum*.

REMARQUE. Il est possible d'être un extremum local sans être global, il est également possible d'être un extremum global sans être local (pour un point au bord de U).

DÉFINITION 4.2. Soit f une fonction définie sur un ouvert U de \mathbf{R}^n dont toutes les dérivées partielles existent. On appelle *point critique* de f un point a tel que toutes les dérivées partielles de f soient nulles en a (ou, autrement dit, tel que $\nabla f(a) = 0$).

PROPOSITION 4.3. Si f est définie sur un ouvert et que toutes ses dérivées partielles existent, alors tous les extrema de f sont des points critiques.

REMARQUE. La réciproque est fautive, par exemple le point $(0, 0)$ est un point critique non extrémal de $(x, y) \mapsto xy$.

Protocole de recherche des extrema d'une fonction de plusieurs variables.

- (1) Justification de l'existence des extrema. Le plus souvent, f est continue sur une partie fermée bornée et on peut invoquer le théorème de Weierstrass.
- (2) Étude sur l'intérieur du domaine de définition. C'est un ouvert, sur lequel les points extrémaux sont des points critiques. On dresse donc la liste des points critiques, ce qui fournit une liste de « candidats » au titre d'extremum.
- (3) Étude sur les bords du domaine de définition. L'expression des bords permet souvent d'y écrire f en fonction d'une seule variable, ce qui permet de trouver relativement facilement les valeurs extrémales prises par f (par les moyens d'étude classique d'une fonction d'une variable réelle). Ces « extrema sur les bords » rejoignent la liste des candidats au titre d'extremum global.
- (4) Bilan. On compare tous les candidats entre eux et on en déduit les véritables extrema globaux.

EXEMPLE(S). Déterminer les extrema globaux de $f : (x, y) \mapsto x^2 + y^2 + 2xy - 6x$ sur le carré $[-4; 4]^2$.

Dans le cas d'une fonction de deux variables (le plus fréquent), la proposition suivante peut être utile :

PROPOSITION 4.4 (Déterminant hessien). Soit f définie et de classe \mathcal{C}^2 sur un ouvert U de \mathbf{R}^2 . Soit $a \in U$ un point critique de f . On note $\Delta = \frac{\partial^2 f}{\partial x^2}(a) \times \frac{\partial^2 f}{\partial y^2}(a) - \left(\frac{\partial^2 f}{\partial x \partial y}(a)\right)^2$.

- si $\Delta > 0$, alors a est un extremum local de f (un maximum si $\frac{\partial^2 f}{\partial x^2}(a) < 0$, un minimum sinon) ;
- si $\Delta < 0$, alors a n'est pas un extremum local de f ; on dit que c'est un point col ou selle.

REMARQUE. • si $\Delta = 0$, on ne peut pas conclure (a peut être, ou ne pas être, extrémal).

- le nombre Δ est le déterminant de la *matrice hessienne*

$$Hf(a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(a) & \frac{\partial^2 f}{\partial x \partial y}(a) \\ \frac{\partial^2 f}{\partial y \partial x}(a) & \frac{\partial^2 f}{\partial y^2}(a) \end{pmatrix}.$$

5. Équations aux dérivées partielles

EXEMPLE(S). • Trouver toutes les fonctions $f: \mathbf{R}^2 \rightarrow \mathbf{R}$ de classe \mathcal{C}^1 vérifiant $\frac{\partial f}{\partial x} = 0$.

- Trouver toutes les fonctions $f: \mathbf{R}^2 \rightarrow \mathbf{R}$ de classe \mathcal{C}^1 vérifiant $\frac{\partial f}{\partial x} = f$.
- Trouver toutes les fonctions $f: \mathbf{R}^2 \rightarrow \mathbf{R}$ de classe \mathcal{C}^1 vérifiant $\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} = 2$. On pourra faire appel au changement de variables $u = x + y$, $v = x - y$.
- Trouver toutes les fonctions $f: \mathbf{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbf{R}$ de classe \mathcal{C}^1 vérifiant $x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} = 0$. On pourra faire appel à un passage en coordonnées polaires.