Algèbre et Analyse
L2 Mathématiques - Semestre 4

Année 2023 – 2024

Table des matières

Chapitre 1

Quelques rappels d'algèbre linéaire

1 Espaces vectoriels

1.1 Espaces vectoriels

Définition 1.1. Un ensemble E non vide muni d'une loi de composition interne notée $+$ est un espace vectoriel sur $\mathbb{K} = (\mathbb{R}$ ou $\mathbb{C})$ si $(E, +)$ est un groupe commutatif (dont on note l'élément neutre 0_E) muni d'une action de $\mathbb{K}$ distributive.
Les axiomes précis sont les suivants. $\forall (u, v, w) \in E^3\ \forall (\lambda, \mu) \in \mathbb{K}^2$:

1. Associativité : $(u + v) + w = u + (v + w)$.
2. Élément neutre : $u + 0_E = 0_E + u = u$.
3. Existence d'un inverse : $\exists u'$ tel que $u + u' = u' + u = 0_E$ (c'est-à-dire $u' = -u$).
4. Commutativité : $u + v = v + u$.
5. $\lambda(\mu \cdot u) = (\lambda \cdot \mu)u$ et $1 \cdot u = u$.
6. $(\lambda + \mu)u = \lambda u + \mu u$.
7. Distributivité : $\lambda(u + v) = \lambda u + \lambda v$.

Exemples 1.2.

1. $\mathbb{R}$ est un $\mathbb{R}$ espace vectoriel.
2. $\mathbb{C}$ est un $\mathbb{C}$ espace vectoriel et aussi un $\mathbb{R}$ espace vectoriel.
3. $E = \mathbb{K}^n$ est un $\mathbb{K}$ espace vectoriel.
4. $\mathbb{K}^X = \{f : X \to \mathbb{K}\}$ est un espace vectoriel sur $\mathbb{K}$ pour tout ensemble X.
5. $M_{n,p}(\mathbb{K})$ l'ensemble des matrices à n lignes et p colonnes à coefficients dans $\mathbb{K}$ est un espace vectoriel sur $\mathbb{K}$.

1.2 Sous-espaces vectoriels

Définition 1.3. Soit $F \subseteq E$ avec E un $\mathbb{K}$-espace vectoriel. Si, muni des lois de E, F est un $\mathbb{K}$ espace vectoriel, on dit que c'est un sous-espace vectoriel (en abrégé sous-e.v.) de E.

Proposition 1.4. *Si E est un espace vectoriel et $F \subseteq E$ alors F est un sous-espace vectoriel de E si et seulement si :*

1. $0_E \in F$.
2. $\forall (u, v) \in F^2$ *et* $\forall (\lambda, \mu) \in \mathbb{K}^2,\ \lambda u + \mu v \in F$.

Exemples 1.5.

1. $\{0_E\}$ et E sont toujours des sous-espaces vectoriels de E.
2. $\mathbb{R}$ et $i\mathbb{R}$ sont des sous-espaces vectoriels du $\mathbb{R}$-espace vectoriel $\mathbb{C}$.
3. Soit I un intervalle de $\mathbb{R}$, l'espace $C(I,\mathbb{R})$ des fonctions continues $I \to \mathbb{R}$ est un sous-espace vectoriel de $\mathbb{R}^I$. De même pour $C^n(I,\mathbb{R})$ et $C^\infty(I,\mathbb{R})$.

   Attention !

   L'ensemble des bijections de $\mathbb{R}$ dans $\mathbb{R}$ n'est pas un sous-espace vectoriel de $\mathbb{R}^{\mathbb{R}}$. Il ne contient pas la fonction nulle.

4. L'ensemble des polynômes $\mathbb{K}[X]$ est un sous-espace vectoriel de $\mathbb{K}^{\mathbb{N}}$.
   L'ensemble $\mathbb{K}_n[X]$ des polynômes de degré au plus n est un sous-espace vectoriel de $\mathbb{K}[X]$.
5. $\mathrm{Diag}_n(\mathbb{K}) \subset M_n(\mathbb{K})$, l'ensemble des matrices diagonales, est un sous-espace vectoriel de $M_n(\mathbb{K})$. De même, $T_n(\mathbb{K}) \subset M_n(\mathbb{K})$, l'ensemble des matrices triangulaires supérieures, est un sous-espace vectoriel de $M_n(\mathbb{K})$.

   Attention !

   Si $GL_n(\mathbb{K}) \subset M_n(\mathbb{K})$ désigne l'ensemble des matrices inversibles de taille n à coefficients dans $\mathbb{K}$, ce n'est pas un sous-espace vectoriel de $M_n(\mathbb{K})$ car, par exemple, la matrice nulle n'est pas inversible.

1.3 Intersection de sous-espace vectoriels

Proposition 1.6. *Si F_1 et F_2 sont deux sous-espaces vectoriels de E, alors $F_1 \cap F_2$ est encore un sous-espace vectoriel de E.*
De même, si $(F_i)_{i\in I}$ est une famille de sous-espace vectoriels de E, l'ensemble $\bigcap_{i\in I} F_i$ est encore un sous-espace vectoriel de E.

1.4 Sous-espace vectoriel engendré par une partie de E

Définition 1.7. Soit $A \subset E$ une partie d'un espace vectoriel E. Le sous-espace vectoriel engendré par A est le plus petit sous-espace vectoriel de E qui contient A. On le notera $\mathrm{Vect}(A)$. C'est le sous-espace vectoriel qui est l'intersection de tous les sous-espaces vectoriels de E contenant A :

$$\mathrm{Vect}(A) = \bigcap_{A\subseteq F,\, F\ sous-e.v.} F.$$

Proposition 1.8. $\mathrm{Vect}(A)$ *est le sous-espace vectoriel de E formé par les combinaisons linéaires d'éléments de A : on a $v \in \mathrm{Vect}\,A$ ssi il existe $k \in \mathbb{N}$, $\lambda_1, \dots, \lambda_k \in \mathbb{K}$, et $a_1, \dots, a_k \in A$ tels que*

$$v = \lambda_1 a_1 + \cdots + \lambda_k v_k$$

1.5 Somme de sous-espaces vectoriels

Proposition 1.9. *La réunion de deux sous-espace vectoriels F et G d'un espace vectoriel E est un sous-espace vectoriel de E si et seulement si $F \subseteq G$ ou $G \subseteq F$.*

Définition 1.10. On appelle *somme* de F et G, deux sous-espaces vectoriels de E et on note $F + G$ le sous-espace vectoriel de E : $F + G = \{u + v, u \in F, v \in G\}$.

Proposition 1.11. $F + G = \text{Vect}(F \cup G)$.

1.6 Somme directe de sous-espaces vectoriels

Définition 1.12. Soient F_1 et F_2 deux sous-espaces vectoriels de E. On dit que ces sous-espaces vectoriels sont en *somme directe*, et on note $F_1 \oplus F_2$ au lieu de $F_1 + F_2$ leur somme, si $\forall u \in F_1 + F_2$, $\exists!\ (u_1, u_2) \in F_1 \times F_2$ tel que $u = u_1 + u_2$.

Le point clé ici est que l'on requiert l'*unicité* (via le point d'exclamation après le symbole $\exists$) du couple (u_1, u_2), son existence étant garantie par la définition de la somme $F_1 + F_2$.

Proposition 1.13. *F_1 et F_2 sont en somme directe si et seulement si $F_1 \cap F_2 = \{0_E\}$.*

La somme directe de n sous-espace vectoriels est définie comme suit.

Définition 1.14. Si $F_1, F_2, ..., F_n$ sont n sous-espaces vectoriels de E, on dit que la somme $\sum_{i=1}^{n} F_i$ est une somme directe si

$\forall u \in F_1 + F_2 + ... + F_n$, $\exists!\ (u_1, u_2, ..., u_n) \in F_1 \times ... \times F_n$ tel que $u = \sum_{i=1}^{n} u_i$.

On note la somme $\bigoplus_{i=1}^{n} F_i$ si les F_i sont en somme directe.

Caractérisation équivalente :
$\sum_{i=1}^{n} F_i = \bigoplus_{i=1}^{n} F_i$ si pour toute combinaison linéaire nulle $u_1 + u_2 + ... + u_n = 0_E$ avec $u_i \in F_i$, on a nécessairement $u_i = 0_E$ pour tout $i \in \{1, \cdots, n\}$.

Attention !

C'est une condition beaucoup plus forte que de demander seulement que $F_i \cap F_j = \{0_E\}\ \forall i \neq j$.

1.7 Supplémentaire d'un sous-espace vectoriel

Définition 1.15. Si F est un sous-espace vectoriel de E, un sous-espace vectoriel de G de E est un supplémentaire de F dans E si :
– F et G sont en somme directe.
– $F \oplus G = E$.

Attention !

Ne pas confondre le complémentaire $E \setminus F$ (qui n'est jamais un sous-espace vectoriel) et un supplémentaire d'un sous-espace vectoriel.

2 Familles de vecteurs, bases

Définition 1.16. Soit $(u_i)_{i\in I}$ une famille de vecteurs d'un $\mathbb{K}$-espace vectoriel E. On appelle *combinaison* linéaire des vecteurs u_i toute somme de la forme $\sum_{i\in I} \lambda_i u_i$ où $(\lambda_i)_{i\in I} \in \mathbb{K}^I$ est une famille de scalaires presque tous nuls (c'est à dire que seulement un nombre **fini** de coefficients sont non nuls.

Une combinaison linéaire de vecteurs est par définition toujours une somme finie de vecteurs.

2.1 Familles libres

Définition 1.17. On dit qu'une famille de vecteurs $(u_i)_{i\in I}$ d'un $\mathbb{K}$-espace vectoriel E est une famille *libre* si toute combinaison linéaire nulle des vecteurs $(u_i)_{i\in I}$ est à coefficients tous nuls : $\sum_{i\in I} \lambda_i u_i = 0 \Rightarrow \forall i \in I, \lambda_i = 0$.

2.2 Familles génératrices

Définition 1.18. On dit qu'une famille $(u_i)_{i\in I}$ de vecteurs de E est *génératrice* de E si tout élément de E s'écrit comme une combinaison linéaire des vecteurs $(u_i)_{i\in I}$:
$\forall u \in E, \exists\ (\lambda_i)_{i\in I}$ presque tous nuls tels que $u = \sum_{i\in I} \lambda_i u_i$.

2.3 Bases

Définition 1.19. On dit qu'une famille de vecteur d'un $\mathbb{K}$- espace vectoriel, E est une *base* de E si tout élément de E s'écrit de manière unique comme une combinaison linéaire des vecteurs $(u_i)_{i\in I}$:
$\forall u \in E, \exists!\ (\lambda_i)_{i\in I} \in \mathbb{K}^I$, presque tous nuls tel que $u = \sum_{i\in I} \lambda_i u_i$

Proposition 1.20. *Caractérisations équivalentes :*

1. $(u_i)_{i\in I}$ *est libre* $\Leftrightarrow \mathrm{Vect}((u_i)_{i\in I}) = \bigoplus_{i\in I} \mathrm{Vect}(u_i)$.
2. $(u_i)_{i\in I}$ *est génératrice de* $E \Leftrightarrow \mathrm{Vect}((u_i)_{i\in I}) = E$.
3. $(u_i)_{i\in I}$ *est une base de* $E \Leftrightarrow \mathrm{Vect}((u_i)_{i\in I}) = \bigoplus_{i\in I} Vect(u_i) = E$.
4. $(u_i)_{i\in I}$ *est une base* $\Leftrightarrow \begin{cases} (u_i)_{i\in I} \text{ est libre.} \\ (u_i)_{i\in I} \text{ est génératrice de } E. \end{cases}$

Proposition 1.21.
1. *Toute sous-famille d'une famille libre est libre.*
2. *Toute famille contenant une sous-famille liée est liée.*
3. *Toute famille contenant une sous-famille génératrice de E est encore une famille génératrice de E.*

3 Dimension d'un espace vectoriel

3.1 Définition

Définition 1.22. Soit E un $\mathbb{K}$-espace vectoriel, on appelle *dimension* de E la quantité finie ou infinie suivante :

$$\dim_{\mathbb{K}} E = \sup\{n \in \mathbb{N}, \text{il existe une famille libre de n vecteurs dans } E\}$$

Exemples 1.23.

1. $\dim_{\mathbb{K}}\{0\} = 0$.
2. $\dim_{\mathbb{K}} \mathbb{K}^n = n$.
3. $\dim_{\mathbb{K}} \mathbb{K}[X] = +\infty$.
4. $\dim_{\mathbb{K}} \mathbb{K}_n[X] = n + 1$.
5. $\dim_{\mathbb{K}} M_n[\mathbb{K}] = n^2$.
6. $\dim_{\mathbb{K}} D_n(\mathbb{K}) = n$.
7. $\dim_{\mathbb{K}} T_n(\mathbb{K}) = \dfrac{n^2 - n}{2} + n = \dfrac{n(n+1)}{2}$.

3.2 Rang d'un système de vecteurs

Définition 1.24. Soit $(u_i)_{i\in I}$ une famille de vecteurs. On appelle *rang* de la famille $(u_i)_{i\in I}$ et on note $\mathrm{rg}(u_i)_{i\in I}$ la dimension du sous-espace vectoriel engendré par la famille $(u_i)_{i\in I}$: $\mathrm{rg}((u_i)_{i\in I}) = \dim(\mathrm{Vect}(u_i)_{i\in I})$.

Proposition 1.25. *Si* $(u_1, \dots, u_p)$ *sont* p *vecteurs dans* E, $\mathrm{rg}(u_1, u_2, \dots, u_p) = \ell \leq p$. *L'entier* ℓ *représente la taille maximale d'une sous-famille libre de la famille* $(u_1, u_2, \dots, u_p)$. *En particulier :*
— *Il existe une sous-famille* $(u_{i_1}, u_{i_2}, \dots, u_{i_\ell}) \subset (u_1, u_2, \dots, u_p)$ *qui est libre.*

— *Toute sous-famille à $\ell + 1$ vecteurs de $(u_1, ..., u_p)$ est liée (dans le cas $\ell < p$).*

Corollaire 1.26. *Si E est un espace vectoriel de dimension n, toute famille génératrice a au moins n vecteurs.*

Théorème 1.27. *Les assertions suivantes sont équivalentes :*

1. *E est un espace vectoriel de dimension n sur $\mathbb{K}$.*
2. *Toute base de E a exactement n vecteurs.*
3. *Tout système générateur de E a au moins n éléments.*
4. *Tout système générateur à n éléments est libre (donc une base).*
5. *Tout système libre a au plus n éléments.*
6. *Tout système libre à n éléments est générateur (donc une base).*

3.3 Dimension d'un sous-espace vectoriel, théorème de la base incomplète

Proposition 1.28. *Si E est un espace vectoriel de dimension finie et $F \subseteq E$ est un sous-espace vectoriel de E, alors la dimension de F est finie et de dimension $\dim F \leq \dim E$. Si $\dim(F) = \dim(E)$, alors $F = E$.*

Théorème 1.29 (Théorème de la base incomplète)**.** *Soit $(u_1, u_2, ..., u_p)$ une famille libre dans E un espace vectoriel de dimension n, alors si $p < n$, il existe des vecteurs $(u_{p+1}, u_{p+2}, ..., u_n)$ tels que $(u_1, ..., u_n)$ soit une base de E. De plus, si $(v_1, v_2, ..., v_q)$ est une famille génératrice, on peut choisir $u_{p+1}, u_{p+2}, ..., u_n$ parmi les vecteurs de la famille $(v_1, v_2, ..., v_q)$.*

Corollaire 1.30.

1. *Si $(v_1, ..., v_q)$ est une famille génératrice alors il existe une sous-famille $(v_{i_1}, v_{i_2}, ..., v_{i_n})$ qui est une base de E.*
2. *Tout sous-espace vectoriel de E admet au moins un supplémentaire (en fait une infinité).*

3.4 Dimension de la somme de sous-espaces vectoriels

Proposition 1.31. *Soient F_1 et F_2 deux sous-espace vectoriels de dimension finie de E, alors $F_1 + F_2$ est de dimension finie et :*

$$\dim(F_1 + F_2) = \dim(F_1) + \dim(F_2) - \dim(F_1 \cap F_2).$$

Corollaire 1.32. *Si F_1 et F_2 sont en somme directe alors :*

$$\dim(F_1 \oplus F_2) = \dim(F_1) + \dim(F_2).$$

De plus si $(e_1, ..., e_n)$ est une base de F_1 et $(\varepsilon_1, \cdots, \varepsilon_m)$ est une base de F_2, alors $(e_1, \cdots, e_n, \varepsilon_1, \cdots, \varepsilon_m)$ est une base de $F_1 \oplus F_2$.

Cette propriété se généralise à la somme directe de p sous-espace vectoriels de E.

Proposition 1.33. *Si $F_1, F_2, \cdots, F_p$ sont des sous-espaces vectoriels de E, de dimensions finies n_i qui sont en somme directe, et si $\left(\varepsilon_1(i), \cdots, \varepsilon_{n_i}(i)\right)$ est une base de F_i, alors :*

$$\dim(F_1 \oplus F_2 \oplus \cdots \oplus F_p) = \dim(F_1) + \dim(F_2) + ... + \dim(F_p)$$

et $\left(\varepsilon_1(i), ..., \varepsilon_{n_1}(i), \varepsilon_1(2), \cdots, \varepsilon_{n_2}(2), \cdots \varepsilon_1(p), \cdots, \varepsilon_{n_p}(p)\right)$ est une base de $F_1 \oplus F_2 \oplus \cdots \oplus F_p$.

4 Applications linéaires

4.1 Définitions.

Définition 1.34. Soient E et F deux $\mathbb{K}$ espaces vectoriels, on appelle *application linéaire* de E dans F une application $f : E \longrightarrow F$ tel que $f(\lambda u + \mu v) = \lambda f(u) + \mu f(v)$ pour tous $(u, v) \in E^2$ et $(\lambda, \mu) \in \mathbb{K}^2$.
On appelle également *morphisme d'espaces vectoriels* une telle application linéaire.

— Une application linéaire bijective est appelée isomorphisme.
— Une application linéaire d'un espace vectoriel E dans lui-même est appelée endormorphisme.
— Un endomorphisme bijectif est appelé automorphisme.
— Une application linéaire de E dans $\mathbb{K}$ est appelée forme linéaire.

On note $\mathscr{L}(E, F)$ l'ensemble des applications linéaires de E dans F. C'est un sous-espace vectoriel de F^E (l'ensemble de toute les applications de E dans F).

4.2 Exemples

Exemples 1.35.

1. L'application nulle
$$\begin{array}{rccc} 0 : & E & \longrightarrow & F \\ & x & \longmapsto & 0_F \end{array}$$
est un morphisme d'espace vectoriel.

2. L'évaluation en $x \in X$ est, pour tout $x \in X$, une forme linéaire sur $\mathbb{K}^X$:
$$\begin{array}{rccc} ev_x : & \mathbb{K}^X & \longrightarrow & \mathbb{K} \\ & f & \longmapsto & f(x) \end{array}$$

3. Produit de deux fonctions à valeurs dans $\mathbb{K}$: soit $f \in \mathbb{K}^X$: L'application
$$\begin{array}{rccc} m_f : & \mathbb{K}^X & \longrightarrow & \mathbb{K}^X \\ & g & \longmapsto & f \cdot g \end{array}$$
est linéaires, c'est l'application de multiplication par f.

4. Composée d'applications : si $f : X \longrightarrow Y$, l'application
$$\begin{array}{rccc} . \circ f : & \mathbb{K}^Y & \longrightarrow & \mathbb{K}^X \\ & g & \longmapsto & g \circ f \end{array}$$
est une application linéaire de $\mathbb{K}^Y$ dans $\mathbb{K}^X$: on a $(\lambda g + \mu h) \circ f = \lambda g \circ f + \mu h \circ f$.

Attention !

Si $f : X \longrightarrow Y$ n'est pas linéaire (en particulier, si X et Y ne sont pas des espaces vectoriels) alors

$$\begin{array}{ccc} \mathbb{K}^X & \longrightarrow & \mathbb{K}^X \\ g & \longmapsto & f \circ g \end{array}$$

n'est pas linéaire .

Par exemple, soit $X = Y = \mathbb{R}$ et $f(x) = x^2$. Alors

$$f \circ (\lambda g + \mu h) = \lambda^2 g^2 + 2\lambda\mu g h \mu^2 h^2 \neq \lambda g^2 + \mu h^2 = \lambda f \circ g + \mu f \circ h.$$

5. Applications :
   (a) $p_i\,(\mathbb{K}^n \longrightarrow \mathbb{K})\,(a_1, \ldots, a_n) \longmapsto a_i$ est linéaire.
   (b) L'application suivante est linéaire (application de 3)

$$\begin{array}{ccc} \mathbb{K}^{\mathbb{N}} & \longrightarrow & \mathbb{K}^{\mathbb{N}} \\ u & \longmapsto & uv = (u_n v_n)_{n \in \mathbb{N}} \end{array}$$

   (c) Si $Q \in \mathbb{K}[X]$ est fixé, l'application suivante est linéaire :

$$\begin{array}{ccc} \mathbb{K}[X] & \longrightarrow & \mathbb{K}[X] \\ P & \longmapsto & PQ \end{array}$$

6. Soit $B \in M_{p,q}(\mathbb{K})$ alors $\forall n \geq 1$, l'application

$$\begin{array}{ccc} M_{n,p}(\mathbb{K}) & \longrightarrow & M_{n,q}(\mathbb{K}) \\ A & \longmapsto & AB \end{array}$$

   est linéaire et l'application

$$\begin{array}{ccc} M_{q,n} & \longrightarrow & M_{p,n} \\ A & \longmapsto & BA \end{array}$$

   également.

7. Les applications
$\begin{array}{ccc} \mathbb{K}[X] & \longrightarrow & \mathbb{K}[X] \\ P & \longmapsto & P' \end{array}$ et $\begin{array}{ccc} D^1(I) & \longrightarrow & \mathbb{R}^I \\ f & \longmapsto & f' \end{array}$ sont linéaires.

8. Si $f \in C^0([a,b], \mathbb{R})$ l'application suivante est linéaire.

$$\begin{array}{ccc} C^0([a,b], \mathbb{R}) & \longrightarrow & \mathbb{R} \\ f & \longmapsto & \int_a^b f(x)dx \end{array}$$

9. L'application de transposition est un isomorphisme d'espaces vectoriels :

$$\begin{array}{ccc} M_{n,p}(\mathbb{K}) & \longrightarrow & M_{p,n}(\mathbb{K}) \\ A & \longmapsto & {}^t A \end{array}$$

   Si $n = p$, ${}^t : M_n(\mathbb{K}) \longrightarrow M_n(\mathbb{K})$ est donc un automorphisme.

4.3 Opérations sur les applications linéaires

Proposition 1.36. *Si f et g sont des applications linéaires de E dans F et $(\lambda, \mu) \in \mathbb{K}^2$:*
- — *$\lambda f + \mu g$ est encore une application linéaire.*
- — *Si f est un isomorphisme, sa bijection réciproque f^{-1} est également linéaire.*
- — *Si $f \circ g$ est bien définie, alors $f \circ g$ est linéaire.*
- — *Si, de plus, $h : F \longrightarrow G$ est linéaire, alors on a :*
  - — *$(f+g) \circ h = f \circ h + g \circ h$ et $f \circ (g+h) = f \circ g + f \circ h$;*
  - — *$(\lambda f) \circ h = \lambda(f \circ h)$ et $f \circ (\lambda h) = \lambda(f \circ h)$.*
- — *Si f est un endomorphisme, on peut définir f^n pour tout $n \in \mathbb{N}$, par $f^n = \underbrace{f \circ f \circ \ldots.. \circ f}_{n \text{ fois}}$.*

Restriction d'une application linéaire à un sous-espace vectoriel

Si $f : E \longrightarrow G$ est une application linéaire et $F \subseteq E$ un sev de E, alors

$$f_{|F} : \begin{array}{ccc} F & \longrightarrow & G \\ x & \longmapsto & f(x) \end{array}$$

Restriction d'un endomorphisme

Soit $f : E \longrightarrow E$ un endomorphisme et $F \subseteq E$ un sev alors $f_{|F}$ est une application linéaire de F dans E.

Attention !

En général $f_{|F}$ n'est pas un endomorphisme de F. En fait $f_{|F}$ est un endomorphisme de F si et seulement si F est stable par f, c'est-à-dire si et seulement si $f(F) \subset F$.

4.4 Noyau et image d'une application linéaire

Définition 1.37. Soit $f \in \mathcal{L}(E, F)$. On pose : $\mathrm{Im}(f) = \{f(x), x \in E\} = f(E)$ est un sous-espace vectoriel de F.
$\mathrm{Ker}(f) = \{x \in E, f(x) = 0_F\} = f^{-1}(\{0_F\})$ est un sous-espace vectoriel de E.

Proposition 1.38. *Soit $f \in \mathcal{L}(E, F)$, alors*
- — $\mathrm{Im}(f)$ *est un sous-espace vectoriel de F.*
- — $\mathrm{Ker}(f)$ *est un sous-espace vectoriel de E.*

Proposition 1.39. *$f \in \mathcal{L}(E, F)$ est :*
- — *injective si et seulement si* $\mathrm{Ker}\, f = \{0_E\}$;
- — *surjective si et seulement si* $\mathrm{Im}\, f = F$.

4.5 Projecteurs, projections

Définition 1.40. Soit E un espace vectoriel, F et G des sous-espaces vectoriels supplémentaires. On appelle *projection* sur F parallèlement à G l'endomorphisme de E défini par : $p(x) = x_F$ si on a la décomposition $x = x_F + x_G$ dans la somme directe $E = F \oplus G$.

Proposition 1.41.

1. *p est une application linéaire.*
2. *On a* $p(x) = \begin{cases} x \text{ si } x \in F. \\ 0 \text{ si } x \in G. \end{cases}$
3. *On a* $p_{|F} = \mathrm{id}_F$ *et* $P_{|G} = 0 : G \longrightarrow G$.
4. $p \circ p = p$.
5. $\mathrm{id}_E - p$ *est la projection sur G parallèlement à F.*

Définition 1.42. On appelle projecteur tout endomorphisme de E tel que $p \circ p = p$.

Proposition 1.43. *Si p est un projecteur, alors :*
1) $\mathrm{Im}\, p$ et $\mathrm{Ker}\, p$ sont des sous-espaces vectoriels supplémentaires dans E.
2) p est la projection sur $\mathrm{Im}\, p$ parallèlement à $\mathrm{Ker}\, p$.

5 Applications linéaires en dimension finie

5.1 Rang d'une application linéaire

Définition 1.44. Soit $f : E \longrightarrow F$ une application linéaire, on appelle *rang* de f et on note $\mathrm{rg}(f)$ la quantité $\mathrm{rg}(f) = \dim(\mathrm{Im}\, f)$
$\mathrm{Im}\, f = \{y \in F, \exists x \in E \text{ avec } f(x) = y\}$.
Si $(u_i)_{i\in I}$ est une famille génératrice de E, alors $\mathrm{Im}\, f = \mathrm{Vect}(f(u_i))_{i\in I}$.
Le rang de l'application linéaire est le rang de la famille de vecteurs $\big(f(u_i)\big)_{i\in I}$ où $(u_i)_{i\in I}$ est une famille génératrice de E.

On suppose dorénavant E et F de dimensions finies.

Théorème 1.45. *Théorème du rang.*
Soit $f \in \mathscr{L}(E, F)$, alors

$$\dim(E) = \dim(\mathrm{Ker}\, f) + \underset{\substack{\| \\ \dim(\mathrm{Im}\, f)}}{\mathrm{rg}(f)}.$$

Corollaire 1.46.

1. *Si $f \in \mathscr{L}(E, F)$ est injective, alors* $\dim(F) \geq \dim(E)$.
2. *Si $f \in \mathscr{L}(E, F)$ est surjective, alors* $\dim(F) \leq \dim(E)$.
3. *Si $f \in \mathscr{L}(E, F)$ est bijective, alors* $\dim(E) = \dim(F)$.

Théorème 1.47. *Si* $\dim(E) = \dim(F)$ *alors, il existe une application linéaire bijective de E dans F.*

Démonstration. Idée de la démonstration.
Si $\dim(E) = n$, on montre que le théorème est vrai pour $F = \mathbb{K}^n$. Le cas général s'obtient en composant $E \xrightarrow{f} \mathbb{K}^n \xrightarrow{g^{-1}} F$
Comment construire une application linéaire bijective de E dans $\mathbb{K}^n$ (avec $\dim(E) = n$) ?
On choisit une base $\mathscr{B}$ de E, $\mathscr{B} = (u_1, u_2, \ldots, u_n)$.
On appelle $\mathrm{Coeff}_{\mathscr{B}} : E \longrightarrow \mathbb{K}^n$ l'application qui à un vecteur $x \in E$ associe ses coefficients (ou coordonnées) dans la base $\mathscr{B}$. Alors par définition d'une base, $\mathrm{Coeff}_{\mathscr{B}} \in \mathscr{L}(E, \mathbb{K}^n)$ est bijective. □

Proposition 1.48.
L'espace vectoriel $\mathscr{L}(E,F)$ est de dimension finie, égale à $\dim \mathscr{L}(E,F) = \dim E \times \dim F$

5.2 Détermination d'une application linéaire par l'image d'une base

Proposition 1.49. *Pour déterminer une application linéaire $f : E \longrightarrow F$, il suffit de connaître l'image directe d'une base par f.*
Autrement dit, si $\mathscr{B} = (u_1, \ldots., u_n)$ est une base de E, alors $\forall\ (v_1, \ldots, v_n) \in F^n$, il existe une unique application linéaire $f : E \longrightarrow F$ tel que $\forall i \in \{1, \ldots, n\}$, $f(u_i) = v_i$.

5.3 Cas des endomorphismes

Proposition 1.50. $\dim(E) = \dim(\mathrm{Ker}\, f) + \mathrm{rg}(f)$. *En particulier :*
f bijective $\Leftrightarrow$ f est injective $\Leftrightarrow$ f est surjective $\Leftrightarrow$ f est de rang maximal.

Attention !

En général, $\mathrm{Ker}\, f$ *et* $\mathrm{Im}\, f$ *ne sont pas en somme directe.*

Exemple 1.51.

$$f : \begin{array}{l} \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \\ (x,y) \longrightarrow (y,0) \end{array}$$

Nous avons $\mathrm{Ker}\, f = \mathrm{Im}\, f$ avec $\mathrm{Ker}\, f = \{(x,0), x \in \mathbb{R}\}$ et $\mathrm{Im}\, f = \{(y,0), y \in \mathbb{R}\}$

Application 1.52. Application de changement de base.
Soit E un ev, $\mathscr{B} = (e_1, \ldots, e_n)$ et $\mathscr{B}' = (\varepsilon_1, \ldots, \varepsilon_n)$ deux bases de E. On définit $f_{\mathscr{B}}^{\mathscr{B}'}$ par $f_{\mathscr{B}}^{\mathscr{B}'}(e_i) = \varepsilon_i$ pour tout $i \in \{1, \ldots, n\}$.
C'est un automorphisme de E que l'on appelle application de changement de base.
$f_{\mathscr{B}}^{\mathscr{B}'} = \mathrm{Coeff}_{\mathscr{B}'}^{-1} \circ \mathrm{Coeff}_{\mathscr{B}}$.
En effet, $\mathrm{Coeff}_{\mathscr{B}}(e_i) = (0,0,0,1,0,\ldots,0) \in \mathbb{K}^n$ et $\mathrm{Coeff}_{\mathscr{B}'}^{-1}(0,0,1,\ldots,0) = \varepsilon_i$.

Proposition 1.53. *$f \in \mathscr{L}(E)$ est bijectif si et seulement si l'image d'une base par f est une base.*

5.4 Rang de l'image d'une famille de vecteurs

Proposition 1.54. *Soit $f \in \mathscr{L}(E,F)$ et $(u_1, u_2, \ldots, u_p)$ une famille de vecteurs. Alors*

$$\mathrm{rg}(f(u_1), f(u_2), \ldots, f(u_p)) \leq \mathrm{rg}(u_1, u_2, \ldots, u_p)$$

Proposition 1.55. *$f \in \mathscr{L}(E,F)$ alors $\mathrm{rg}(f(u_1), \ldots, f(u_p)) = \mathrm{rg}(u_1, \ldots, u_p)$ pour toute famille $(u_1, \ldots, u_p)$ de E si et seulement si f est injective.*

5.5 Rang de la composée de deux applications linéaires

Proposition 1.56. *Si $f \in \mathscr{L}(E,F)$ et $g \in \mathscr{L}(F,G)$, alors* $\mathrm{rg}(g \circ f) \leq \min(\mathrm{rg}(f), \mathrm{rg}(g))$

Démonstration. Soit $\mathscr{B} = (u_1, \ldots, u_n)$ une base de E.
$\mathrm{rg}(g \circ f) = \mathrm{rg}(g \circ f(u_1), \ldots, g \circ f(u_n))$ donc

$$\mathrm{rg}(g \circ f) = \mathrm{rg}(g(f(u_1)), \ldots, g(f(u_n))) \leq \mathrm{rg}(f(u_1), \ldots, f(u_n)) = \mathrm{rg}(f).$$

Par ailleurs, on a $\mathrm{Im}(g \circ f) \subset \mathrm{Im}(g)$, donc

$$\mathrm{rg}(g \circ f) = \dim(\mathrm{Im}(g \circ f)) \leq \dim(\mathrm{Im}\, g) = \mathrm{rg}(g).$$ □

6 Matrices

6.1 Généralités

Définition 1.57. $M_{n,p}(\mathbb{K})$ est l'ensemble des matrices à n lignes et p colonnes à coefficients dans $\mathbb{K}$:

$$M_{n,p}(\mathbb{K}) = \{(a_{i,j})_{1 \leq 1 \leq n \text{ et } 1 \leq j \leq p}, a_{i,j} \in \mathbb{K}\}.$$

Si $A = (a_{i,j})_{i<1 \leq n \text{ et } 1 \leq j \leq p} \in M_{n,p}(\mathbb{K})$, on la note aussi

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,p} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,p} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,p} \end{pmatrix}$$

Proposition 1.58. *$M_{n,p}(\mathbb{K})$ est un espace vectoriel de dimension np. La base canonique de $M_{n,p}(\mathbb{K})$ est la famille $\left(E_{i,j}^{(n,p)}\right)_{1 \leq i \leq n \text{ et } 1 \leq j \leq p}$ où $E_{i,j}^{(n,p)}$ est la matrice à n lignes et p colonnes qui a tous ses coefficients nuls sauf celui situé à l'intersection de la i-ème ligne et de la j-ème colonne qui vaut 1.*
Les coefficients de $(E_{i,j})$ sont $(E_{i,j})_{k,l} = \delta_{i,k}\delta_{l,j}$ où on utilise le symbole de Kronecker $(\delta_{i,j}) = \begin{cases} 0 \text{ si } i \neq j. \\ 1 \text{ si } i = j. \end{cases}$

Vecteurs lignes, vecteurs colonnes

Définition 1.59.
Si $A \in M_{n,p}(\mathbb{K})$, on appelle :
– *vecteurs lignes* les matrices $L_1(A), L_2(A), \cdots L_n(A)$, avec $L_i(A) = (a_{i,1}\ a_{i,2} \cdots a_{i,p}) \in M_{1,p}(\mathbb{K})$;

– *vecteurs colonnes* les matrices $C_1(A), C_2(A), \cdots C_p(A)$, avec

$$C_j(A) = \begin{pmatrix} a_{1,j} \\ a_{2,j} \\ a_{3,j} \\ . \\ . \\ . \\ a_{n,j} \end{pmatrix} \in M_{n,1}(\mathbb{K}).$$

Produit de Matrices

Définition 1.60.
Si $A \in M_{n,p}(\mathbb{K})$ et $B \in M_{p,m}(\mathbb{K})$, on définit la matrice $C = AB \in M_{n,m}(\mathbb{K})$ par

$$c_{i,j} = \sum_{k=1}^{p} a_{i,k} b_{k,j}.$$

Proposition 1.61. *L'ensemble $M_n(\mathbb{K})$ des matrices à n lignes et à n colonnes (matrices carées de taille n) est une algèbre non commutative sur $\mathbb{K}$.*

Transposition d'une matrice

Définition 1.62. Définition et proposition.
L'application linéaire de transposition $^t : M_{n,p}(\mathbb{K}) \longrightarrow M_{p,n}(\mathbb{K})$ définie par :
$E_{i,j}^{(n,p)} = E_{j,i}^{(p,n)}$ est bijective
On obtient ${}^tA$ par l'une des définitions suivantes :

— $C_i({}^tA) = L_i(A)$;.
— $L_j({}^tA) = C_j(A)$;.
— $({}^tA)_{i,j} = a_{j,i}$;

avec $C_j(A)$ la j-ème colonne de la matrice A et $L_i(A)$ la i-ème ligne de la matrice A.

Proposition 1.63. *On a :*
${}^t(AB) = {}^tB \times {}^tA$ *et* ${}^t({}^tA) = A$.

6.2 Opérations élémentaires sur les lignes et les colonnes d'une matrice

Définition 1.64. Les opérations élémentaires sur les lignes d'une matrice sont :

1. La multiplication d'une ligne de A par un scalaire : dans la matrice on remplace la i-ème ligne $L_i(A)$ par $\lambda L_i(A)$, les autres lignes restant inchangées. On appelle cette opération dilatation de la i-ème ligne et on la note $\mathrm{Dil}_i^L(A,\lambda)$ avec $\lambda \in \mathbb{K}$.
2. L'échange de deux lignes de la matrice A : dans la matrice on échange les i-ème et j-ème lignes $L_i(A)$ et $L_j(A)$ de la matrice, les autres lignes restant inchangées. On appelle cette opération échange des lignes i et j et on la note $\mathrm{Ech}_{i,j}^L(A)$.
3. L'addition à une ligne de A du produit d'une autre ligne de A par un scalaire $\lambda \in \mathbb{K}$: dans la matrice on remplace la i-ème ligne $L_i(A)$ par $L_i(A) + \lambda L_j(A)$, les autres lignes restant inchangées. On appelle cette opération ajout à la ligne i et de λ fois la ligne j et on la note $\mathrm{Ajout}_i^L(A,j,\lambda)$.

On définit de manière analogue les opérations élémentaires sur les colonnes :

$$\mathrm{Dil}_i^C(A,\lambda)\ ,\ \mathrm{Ech}_{i,j}^C(A)\ ,\ \mathrm{Ajout}_i^C(A,j,\lambda).$$

Définition 1.65. Matrices élementaires

1. On appelle matrice de dilatation dans $M_n(\mathbb{K})$ une matrice de la forme

$$D_i^n(\lambda) = I_n + (\lambda-1)E_{i,i}^{(n,n)} = \sum_{\substack{k=1\\k\neq i}}^{n} E_{k,k}^{(n,n)} + \lambda E_{i,i}^{(n,n)}$$

pour $i \in \{1,\cdots,n\}$ et $\lambda \in \mathbb{K}$.

2. On appelle matrice de transposition dans $M_n(\mathbb{K})$ une matrice de la forme

$$\pi^n(i,j) = I_n - E_{i,i}^{(n,n)} - E_{j,j}^{(n,n)} + E_{i,j}^{(n,n)} + E_{j,i}^{(n,n)} = \sum_{\substack{k=1\\k\neq i,k\neq j}}^{n} E_{k,k}^{(n,n)} + E_{i,j}^{(n,n)} + E_{j,i}^{(n,n)}$$

avec $i \neq j$ dans $\{1,\cdots,n\}$.

3. On appelle matrice de transvection dans $M_n(\mathbb{K})$ une matrice de la forme

$$T_{i,j}^n(\lambda) = I_n + \lambda E_{i,j}^{(n,n)}$$

avec $i \neq j$ dans $\{1,\cdots,n\}$ et $\lambda \in \mathbb{K}$.

Théorème 1.66. *Les opérations élémentaires sur les lignes d'une matrice A sont décrites par la multiplication de A à gauche par une matrice élémentaire donnée.*
Les opérations élémentaires sur les colonnes d'une matrice A sont décrites par la multiplication de A à droite par une matrice élémentaire donnée.
Plus précisément, si n désigne le nombre de lignes de A et p son nombre de colonnes

1. $Dil_i^L(A,\lambda) = D_i^n(\lambda)A.$
2. $Ech_{i,j}^L(A) = \pi^n(i,j)A.$
3. $Ajout_i^L(A,j,\lambda) = T_{i,j}^n(\lambda)A.$

4. $Dil_i^C(A,\lambda) = AD_i^p(\lambda)$.
5. $Ech_{i,j}^C(A) = A\pi^p(i,j)$.
6. $Ajout_i^C(A,j,\lambda) = AT_{j,i}^p(\lambda)$.

7 Matrices et applications linéaires

7.1 Bijection canonique entre l'ensemble des matrices et celui des applications linéaires

Une matrice $A \in M_{n,p}(\mathbb{K})$ définit une application linéaire de $\mathbb{K}^p$ dans $\mathbb{K}^n$.

Proposition 1.67.
L'application Lin : $M_{n,p}(\mathbb{K}) \longrightarrow \mathscr{L}(\mathbb{K}^p,\mathbb{K}^n)$ définie par $Lin(A)(x) = Ax$ est une application linéaire qui vérifie $Lin(AB) = Lin(A) \circ Lin(B)$. On confond dans cette écriture le vecteur $x \in \mathbb{K}^p$ et le vecteur colonne $x \in M_{p,1}(\mathbb{K})$.

Une application linéaire de $\mathbb{K}^p$ dans $\mathbb{K}^n$ définit une matrice $A \in M_{n,p}(\mathbb{K})$

On note $(e_1, e_2, \cdots e_p)$ la base canonique de $\mathbb{K}^p$ et $(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)$ la base canonique de $\mathbb{K}^n$. Alors si $f \in \mathscr{L}(\mathbb{K}^p,\mathbb{K}^n)$, il existe, pour chaque $j \in \{1,\cdots,p\}$ des coefficients $a_{i,j}$ tels que $f(e_j) = \sum_{i=1}^n a_{i,j}\varepsilon_i$ (ce sont les coefficients du vecteur $f(e_j)$ dans la base canonique de $\mathbb{K}^n$). On note $\mathrm{Mat}(f)$ la matrice de $M_{n,p}(\mathbb{K})$ dont les coefficients sont les $(a_{i,j})_{1\leq i\leq n, 1\leq j\leq p}$.

Proposition 1.68. *L'application Mat : $\mathscr{L}(\mathbb{K}^p,\mathbb{K}^n) \longrightarrow M_{n,p}(\mathbb{K})$ définie ci-dessus est une application linéaire qui vérifie $Mat(f \circ g) = Mat(f)\, Mat(g)$.*
Les applications linéaires Lin et Mat sont des isomorphismes qui sont bijections réciproques l'une de l'autre, qui explicitent un isomorphisme entre $\mathscr{L}(\mathbb{K}^p,\mathbb{K}^n)$ et $M_{n,p}(\mathbb{K})$ et elles sont compatibles avec le produit / la composition.

7.2 Noyau, image, rang d'une matrice

Définition 1.69. Soit $A \in M_{n,p}(\mathbb{K})$.
- On appelle noyau de la matrice A, et on note $\mathrm{Ker}(A)$ l'ensemble :

$$\mathrm{Ker}(A) = \{X \in M_{p,1}(\mathbb{K}), AX = 0\}.$$

- On appelle image de la matrice A, et on note $\mathrm{Ker}(A)$ l'ensemble :

$$\mathrm{Im}(A) = \{Y \in M_{n,1}(\mathbb{K}), \exists X \in M_{p,1}(\mathbb{K}), AX = Y\}.$$

Proposition 1.70. *On a* $\mathrm{Ker}(A) = \mathrm{Ker}(Lin(A))$ *et* $\mathrm{Im}(A) = \mathrm{Im}((Lin(A))$, *à condition d'identifier un vecteur de $\mathbb{K}^p$ et un vecteur colonne dans $M_{p,1}(\mathbb{K})$.*

Définition 1.71. Soit $A \in M_{n,p}(\mathbb{K})$, on appelle rang de la matrice A et on note $\mathrm{rg}(A)$ le rang du système de vecteurs $(C_1(A), C_2(A), \cdots C_p(A))$.

D'après ce qui précède, on a

Proposition 1.72.

1. $\mathrm{rg}(A) = \mathrm{rg}((Lin(A)))$;
2. $\mathrm{rg}({}^tA) = \mathrm{rg}(A)$;
3. $\mathrm{rg}(AB) \leq \min(\mathrm{rg}(A), \mathrm{rg}(B))$;
4. $\mathrm{rg}(A) + \dim(\mathrm{Ker}(A)) = p$.

Définition 1.73. Une matrice $A \in M_{n,p}(\mathbb{K})$ est inversible s'il existe une matrice $B \in M_{p,n}(\mathbb{K})$ telle que $AB = I_n$ et $BA = I_p$.

Proposition 1.74. *La matrice A est inversible si et seulement si l'application $Lin(A)$ est un isomorphisme, c'est à dire, si et seulement si $p = n$ et $\mathrm{Ker}(A) = \{0\} \Leftrightarrow p = n$ et $\mathrm{rg}(A) = n$.*

7.3 Matrice d'une application linéaire entre deux espaces vectoriels, munis chacun d'une base

Soit E un $\mathbb{K}$ espace vectoriel de dimension p, muni d'une base $\mathscr{B}$ et F un espace vectoriel de dimension n, muni d'une base $\mathscr{B}'$. On a vu que l'application $\mathrm{Coeff}_{\mathscr{B}}$ définit un isomorphisme de E dans $\mathbb{K}^p$, qui envoie de manière ordonnée la base $\mathscr{B}$ sur la base canonique de $\mathbb{K}^p$. Ainsi, on a :

$$f \in \mathscr{L}(E, F) \Leftrightarrow \mathrm{Coeff}_{\mathscr{B}'} \circ f \circ \mathrm{Coeff}_{\mathscr{B}}^{-1} \in \mathscr{L}(\mathbb{K}^p, \mathbb{K}^n).$$

Définition 1.75. Soit f une application linéaire entre deux $\mathbb{K}$ espaces vectoriels E et F de dimensions finies, munis respectivement d'une base $\mathscr{B}$ et d'une base $\mathscr{B}'$. On appelle matrice de f dans les bases $\mathscr{B}, \mathscr{B}'$, et on note $\mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(f)$ la matrice :

$$\mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(f) = \mathrm{Mat}(\mathrm{Coeff}_{\mathscr{B}'} \circ f \circ \mathrm{Coeff}_{\mathscr{B}}^{-1}).$$

Pour construire cette matrice, on exprime les coefficients de $f(u_j)$ dans la base $\mathscr{B}'$ (où u_j désigne le j-ième vecteur de la base $\mathscr{B}$) et on range les coefficients en colonne, qui constitue la j-ième colonne de la matrice. Noter que si $p = \dim E$ et $n = \dim F$ alors

$$\mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(f) \in M_{n,p}(\mathbb{K}).$$

Loi de composition

Proposition 1.76. *Si $f \in \mathscr{L}(E, F)$ et $g \in \mathscr{L}(F, G)$ et si l'on note $\mathscr{B}, \mathscr{B}', \mathscr{B}''$ des bases respectives de E, F et G, on a alors :*

$$\mathrm{Mat}_{\mathscr{B}'',\mathscr{B}'}(g)\,\mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(f) = \mathrm{Mat}_{\mathscr{B}'',\mathscr{B}}(g \circ f).$$

Lien avec la base canonique

Faisons le lien avec la construction faite à la section 7.1 d'une bijection canonique entre $\mathscr{L}(\mathbb{K}^p, \mathbb{K}^n)$ et $M_{n,p}(\mathbb{K})$. Lorsque $E = \mathbb{K}^p$ muni de sa base canonique $\mathscr{B}$ et $F = \mathbb{K}^n$ muni de sa base canonique $\mathscr{B}'$, on a que pour tout $f \in \mathscr{L}(E, F)$,

$$\mathrm{Mat}(f) = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(f),$$

et donc si $A \in M_{n,p}(\mathbb{K})$, alors en posant $f = \mathrm{Lin}(A)$ dans la formule ci-dessus on obtient

$$A = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{Lin}(A)).$$

Matrice de changement de base

Soit E un espace vectoriel de dimension n muni de deux bases $\mathscr{B} = (e_1, \cdots e_n)$, $\mathscr{B}' = (\varepsilon_1, \cdots \varepsilon_n)$. Alors il existe une unique application linéaire qui envoie, pour tout $i \in \{1, \cdots, n\}$, e_i sur ε_i. On note $f_{\mathscr{B}}^{\mathscr{B}'}$ cette application.

Définition 1.77. On appelle matrice de passage de la base $\mathscr{B}$ à la base $\mathscr{B}'$ la matrice

$$\mathrm{P}(\mathscr{B}, \mathscr{B}') = \mathrm{Mat}_{\mathscr{B},\mathscr{B}}(f_{\mathscr{B}}^{\mathscr{B}'}).$$

La matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$ est obtenue en écrivant, dans chaque colonne, les coefficients de chaque vecteur de la "nouvelle" base $\mathscr{B}'$ dans "l'ancienne base" $\mathscr{B}$. Plus précisément, si $\mathscr{B} = (e_1, \ldots, e_n)$ et $\mathscr{B}' = (f_1, \ldots, f_n)$, alors si on note $\mathrm{P}(\mathscr{B}, \mathscr{B}') = (p_{i,j})$, on a pour tout $j \in \{1, \ldots, n\}$:

$$f_j = \sum_{i=1}^{n} p_{i,j} e_i.$$

Proposition 1.78. *On a* $\mathrm{Mat}_{\mathscr{B},\mathscr{B}}(f_{\mathscr{B}}^{\mathscr{B}'}) = \mathrm{Mat}_{\mathscr{B},\mathscr{B}'}(\mathrm{id}_E)$. *En particulier* $\mathrm{Mat}_{\mathscr{B},\mathscr{B}}(f_{\mathscr{B}}^{\mathscr{B}'})$ *est inversible, d'inverse*

$$\mathrm{Mat}_{\mathscr{B}',\mathscr{B}'}(f_{\mathscr{B}'}^{\mathscr{B}}) = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(\mathrm{id}_E).$$

On en déduit la formule de changement de base :

Proposition 1.79. *Soient E un $\mathbb{K}$ espace vectoriel de dimension finie, muni de deux bases $\mathscr{B}_1$ et $\mathscr{B}_1'$, et F un $\mathbb{K}$ espace vectoriel de dimension finie, muni de deux bases $\mathscr{B}_2$ et $\mathscr{B}_2'$, et enfin $f \in \mathscr{L}(E, F)$ une application linéaire. Alors, on a*

$$\begin{aligned}\mathrm{Mat}_{\mathscr{B}_2',\mathscr{B}_1'}(f) &= \mathrm{Mat}_{\mathscr{B}_2',\mathscr{B}_2}(\mathrm{id}_F)\,\mathrm{Mat}_{\mathscr{B}_2,\mathscr{B}_1}(f)\,\mathrm{Mat}_{\mathscr{B}_1,\mathscr{B}_1'}(\mathrm{id}_E)\\ &= \mathrm{P}(\mathscr{B}_2', \mathscr{B}_2)\,\mathrm{Mat}_{\mathscr{B}_2,\mathscr{B}_1}(f)\,\mathrm{P}(\mathscr{B}_1, \mathscr{B}_1').\end{aligned}$$

Proposition 1.80. *Toute matrice inversible est une matrice de passage.*

7.4 Matrice d'un endomorphisme dans une base

Définition 1.81. Soit E un $\mathbb{K}$ espace vectoriel de dimension finie, muni d'une base $\mathscr{B}$ et $f \in \mathscr{L}(E)$, un endomorphisme. On appelle matrice de l'endomorphisme f dans la base $\mathscr{B}$, et on note $\mathrm{Mat}_{\mathscr{B}}(f)$ la matrice :

$$\mathrm{Mat}_{\mathscr{B}}(f) = \mathrm{Mat}_{\mathscr{B},\mathscr{B}}(f) = \mathrm{Mat}(\mathrm{Coeff}_{\mathscr{B}} \circ f \circ \mathrm{Coeff}_{\mathscr{B}}^{-1}).$$

On a la formule de changement de base suivante pour les endomorphismes :

Proposition 1.82. *Soient E un $\mathbb{K}$ espace vectoriel de dimension finie, muni de deux bases $\mathcal{B}$ et $\mathcal{B}'$, et $f \in \mathcal{L}(E)$ un endomorphisme. On note $P = \mathrm{Mat}_{\mathcal{B},\mathcal{B}}(f_{\mathcal{B}}^{\mathcal{B}'})$ la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$. Alors, on a*

$$\mathrm{Mat}_{\mathcal{B}'}(f) = \mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{id}_E)\,\mathrm{Mat}_{\mathcal{B}}(f)\,\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(\mathrm{id}_E) = P^{-1}\,\mathrm{Mat}_{\mathcal{B}}(f)P.$$

8 Matrices inversibles

8.1 Définition et Propriétés

Définition 1.83. Une matrice carrée $A \in M_n(\mathbb{K})$ est inversible s'il existe $B \in M_n(\mathbb{K})$ telle que $AB = BA = I_n$.

Proposition 1.84. *Une matrice carrée $A \in M_n(\mathbb{K})$ est inversible si et seulement si l'une des conditions suivantes est vérifiée.*
- *L'application linéaire canoniquement associée à A, $Lin(A)$ est un automorphisme de $\mathbb{K}^n$.*
- $\mathrm{Ker}(A) = \{0\}$.
- $\mathrm{rg}(A) = n$.
- *A est la matrice de passage d'une base à un autre.*
- *Si $A = \mathrm{Mat}_{\mathcal{B}}(f)$, alors f est un automorphisme.*

Théorème 1.85. *Muni de la multiplication, l'ensemble $GL_n(\mathbb{K})$ des matrices inversibles est un groupe non commutatif, engendré par les produits de matrices élémentaires.*

La dernière partie du théorème signifie que toute matrice inversible peut s'écrire $A = M_1 \times M_2 \times \cdots \times M_m$ où les matrices M_i sont des matrices élémentaires (de dilatation, de transposition, de transvection).
La démonstration est exactement la méthode du pivot de Gauss de résolution des systèmes linéaires, en utilisant le Théorème 1.66.

Attention !

$GL_n(\mathbb{K})$ n'est pas un sous-espace vectoriel de $M_n(\mathbb{K})$.

8.2 Calcul de l'inverse

Il existe différentes techniques pour calculer l'inverse d'une matrice.

Interprétation comme matrice de passage

On interprète A comme la matrice de passage de la base canonique de $\mathbb{K}^n$ à une nouvelle base $\mathcal{B}'$ dont la matrice A donne, en colonnes, les coordonnées de chaque vecteur dans la base canonique. Alors il suffit de trouver les coordonnées des vecteurs de la base canonique exprimés dans la nouvelle base $\mathcal{B}'$ pour connaitre la matrice de passage de la base canonique à la nouvelle base $\mathcal{B}'$, qui est exactement l'inverse de la matrice A.

Résolution d'un système linéaire

On utilise le fait que si $A \in M_n(\mathbb{K})$ est inversible, alors pour $X, Y \in M_{n,1}(\mathbb{K})$, on a :

$$AX = Y \Leftrightarrow X = A^{-1}Y.$$

Déterminer A^{-1} revient donc à résoudre le système linéaire de n équations à n inconnues (avec second membre) $AX = Y$.

Méthode parallèle

Principe : on réalise des opérations élémentaires successives simultanément sur les lignes de la matrice A et sur celles de la matrice identité I_n. Lorsque l'on est arrivé à la matrice identité en partant de A, on a l'expression de A^{-1}, en prenant la matrice obtenue en effectuant les mêmes opérations à partir de la matrice I_n.
On écrit en fait $\underbrace{M_pM_{p-1}\cdots M_2M_1}_{B}A = I_n$, où les M_i sont des matrices élémentaires, et on obtient ainsi que

$$B = A^{-1} = M_pM_{p-1}\cdots M_2M_1 = M_pM_{p-1}\cdots M_2M_1I_n.$$

Attention !

On peut choisir de faire des opérations sur les colonnes plutôt que sur les lignes, mais on ne peut mélanger les opérations sur les lignes et les colonnes pour obtenir A^{-1} par cette méthode.

9 Matrices carrées

9.1 Matrices semblables

Définition 1.86. Deux matrices carrées $A, B \in M_n(\mathbb{K})$ sont dites semblables, s'il existe une matrice inversible $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $B = P^{-1}AP$.

Théorème 1.87. *Deux matrices sont semblables si et seulement si elles représentent le même endomorphisme dans des bases différentes.*

9.2 Matrices carrées particulières

Matrices diagonales

Définition 1.88. On dit qu'une matrice carrée $A \in M_n(\mathbb{K})$ est *diagonale* si tous ses coefficients en dehors de la diagonale sont nuls, c'est à dire si $A \in \mathrm{Vect}(E_{i,i})_{1\leq i\leq n}$.
On note $\mathrm{Diag}_n(\mathbb{K})$ ou $D_n(\mathbb{K})$ l'ensemble des matrices diagonales et on écrit $D = \mathrm{diag}(a_1, a_2, \cdots, a_n)$ pour désigner la matrice diagonale dont les coefficients sont $d_{i,j} = 0$ si $i \neq j$ et $d_{i,i} = a_i$.

Exemples 1.89.

1. La matrice nulle est diagonale.
2. La matrice identité I_n est diagonale.
3. Les matrices d'homothéties λI_n sont diagonales.

Proposition 1.90. *$D_n(\mathbb{K})$ est une sous-algèbre commutative unitaire de $M_n(\mathbb{K})$, de dimension n.*

Cela signifie que :
— $D_n(\mathbb{K})$ est un sous-espace vectoriel de $M_n(\mathbb{K})$ de dimension n.
— La matrice identité $I_n \in D_n(\mathbb{K})$.
— Si A et B sont diagonales, AB est diagonale.
— Si A et B sont diagonales, $AB = BA$.

Proposition 1.91.

1. *Le rang d'une matrice diagonale est donné par le nombre de coefficients non nuls.*
2. *En particulier, une matrice diagonale $D = \mathrm{Diag}(d_1, d_2, \cdots d_n)$ est inversible si et seulement si $\forall i \in \{1, \cdots, n\}, d_i \neq 0$. Dans ce cas, l'inverse de D est la matrice*

$$D^{-1} = \mathrm{Diag}(d_1^{-1}, d_2^{-1}, \cdots d_n^{-1}).$$

Définition 1.92. Soit E un $\mathbb{K}$ espace vectoriel et $\lambda \in \mathbb{K}$. On appelle *homothétie* sur E de rapport λ l'endomorphisme h_λ défini par : $h_\lambda(x) = \lambda x$.

Proposition 1.93. *Dans toute base $\mathcal{B}$ de E, on a $\mathrm{Mat}_{\mathcal{B}}(h_\lambda) = \lambda I_n$.*

Proposition 1.94. *Si $A \in M_n(\mathbb{K})$ est une matrice qui commute à toute autre matrice $B \in M_n(\mathbb{K})$ (c'est-à-dire que $AB = BA$), alors A est une matrice d'homothétie.*

Matrices triangulaires

Définition 1.95. Soit $T \in M_n(\mathbb{K})$ une matrice. On dit que :
— T est *triangulaire supérieure* si $t_{i,j} = 0$ dès que $i > j$.
— T est *triangulaire supérieure stricte* si $t_{i,j} = 0$ dès que $i \geq j$.
On note $T_n(\mathbb{K})$ l'ensemble des matrices triangulaires supérieures et $T_n^s(\mathbb{K})$ l'ensemble des matrices triangulaires supérieures strictes.

Proposition 1.96. *Les ensembles $T_n(\mathbb{K})$ et $T_n^s(\mathbb{K})$ sont des sous-algèbres de $M_n(\mathbb{K})$, de dimensions respectives $\frac{n(n+1)}{2}$ et $\frac{n(n-1)}{2}$, et on a $T_n(\mathbb{K}) = D_n(\mathbb{K}) \oplus T_n^s(\mathbb{K})$.*

Proposition 1.97. *Si $N \in T_n^s(\mathbb{K})$, alors $N^n = 0$, c'est à dire que N est une matrice nilpotente.*

Pour montrer cela, on montre par exemple par récurrence que N^k a des coefficients $n_{i,j}(k)$ qui vérifient $n_{i,j}(k) = 0$ dès que $i + k \geq j$.

9.3 Trace d'une matrice carrée

Définition 1.98. On appelle trace d'une matrice carrée $A \in M_n(\mathbb{K})$, et on note $\mathrm{Tr}(A)$ la somme de ses coefficients diagonaux : $\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{i,i}$.

Proposition 1.99. *La trace vérifie les assertions suivantes.*

1. *L'application* Tr *est une application linéaire* $M_n(\mathbb{K}) \longrightarrow \mathbb{K}$.
2. *Pour toutes matrices* $A, B \in M_n(\mathbb{K})$, *on a* $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.
3. *Deux matrices semblables ont la même trace.*

Corollaire 1.100. *Si* $f \in \mathscr{L}(E)$ *est un endomorphisme, on peut définir* $\mathrm{Tr}(f) = \mathrm{Tr}(\mathrm{Mat}_{\mathscr{B}}(f))$, *où* $\mathscr{B}$ *est une base quelconque de* E.

9.4 Produit par blocs

Proposition 1.101. *Si* A *et* B *sont deux matrices de* $M_n(\mathbb{K})$ *qui s'écrivent*

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix} \text{ et } B = \begin{pmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{2,2} \end{pmatrix},$$

avec $A_{1,1}, B_{1,1} \in M_r(\mathbb{K})$, $A_{2,2}, B_{2,2} \in M_{n-r}(\mathbb{K})$, $A_{1,2}, B_{1,2} \in M_{r,n-r}(\mathbb{K})$ *et* $A_{2,1}, B_{2,1} \in M_{n-r,r}(\mathbb{K})$. *Alors, le produit des matrices* AB *s'écrit :*

$$AB = \begin{pmatrix} A_{1,1}B_{1,1} + A_{1,2}B_{2,1} & A_{1,1}B_{1,2} + A_{1,2}B_{2,2} \\ A_{2,1}B_{1,1} + A_{2,2}B_{2,1} & A_{2,1}B_{1,2} + A_{2,2}B_{2,2} \end{pmatrix}.$$

Attention !

L'ordre dans l'écriture est essentiel car, en général, $A_{1,1}B_{1,1} \neq B_{1,1}A_{1,1}$. On en déduit en particulier que le produit de matrices diagonales par blocs est encore diagonal par blocs et que le produit de matrices triangulaires par blocs est encore triangulaire par blocs.

10 Quelques exemples pratiques

Dimension et base d'un sous-e.v. donné par une famille génératrice.

Soit F le sous-e.v. de $\mathbb{R}^4$ engendré par la famille de vecteurs (u_1, u_2, u_3, u_4), définis ci-dessous :

$$u_1 = \begin{pmatrix} 1 \\ -1 \\ 3 \\ 2 \end{pmatrix} \qquad u_2 = \begin{pmatrix} 2 \\ 0 \\ -3 \\ -1 \end{pmatrix} \qquad u_3 = \begin{pmatrix} 1 \\ 1 \\ -6 \\ -3 \end{pmatrix} \qquad u_4 = \begin{pmatrix} 5 \\ 1 \\ -12 \\ -5 \end{pmatrix}$$

Déterminer la dimension de F. En donner une base et représenter F par un système d'équations cartésiennes.

Eléments de correction 1.102. La dimension de F est aussi le rang de la matrice A dont les colonnes sont les 4 vecteurs $u_1, \ldots, u_4$.

Le rang d'une matrice est invariant lorsqu'on effectue des manipulations élémentaires sur les lignes et les colonnes. Ici, puisqu'on cherche aussi à déterminer une base de F, on va se limiter uniquement à des manipulations élémentaires sur les *colonnes* jusqu'à arriver à isoler les pivots. Les vecteurs colonnes non nuls qu'on obtiendra alors seront des vecteurs d'une base de F.

$$A = \begin{pmatrix} 1 & 2 & 1 & 5 \\ -1 & 0 & 1 & 1 \\ 3 & -3 & -6 & -12 \\ 2 & -1 & -3 & -5 \end{pmatrix} \xrightarrow{\substack{C_2 \longleftarrow C_2 - 2C_1 \\ C_3 \longleftarrow C_3 - C_1 \\ C_4 \longleftarrow C_4 - 5C_1}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & 2 & 6 \\ 3 & -9 & -9 & -27 \\ 2 & -5 & -5 & -15 \end{pmatrix}$$

$$\xrightarrow{\substack{C_3 \longleftarrow C_3 - C_2 \\ C_4 \longleftarrow C_4 - 3C_2}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 3 & -9 & 0 & 0 \\ 2 & -5 & 0 & 0 \end{pmatrix}$$

La dimension de F est 2, et une base est donnée par (v_1, v_2), où v_1 et v_2 sont les deux premières colonnes de la dernière matrice ci-dessus. Une représentation paramétrique de F (non demandée dans l'énoncé) est donc :

$$F = \left\{ \begin{pmatrix} \lambda \\ -\lambda+2\mu \\ 3\lambda-9\mu \\ 2\lambda-5\mu \end{pmatrix} \ : \ (\lambda, \mu) \in \mathbb{R}^2 \right\}$$

Pour obtenir un système d'équations cartésiennes représentant F, on élimine les paramètres λ et μ dans le système suivant :

$$\begin{cases} x = \lambda \\ y = -\lambda + 2\mu \\ z = 3\lambda - 9\mu \\ t = 2\lambda - 5\mu \end{cases} \iff \begin{cases} x = \lambda \\ y + x = 2\mu \\ z - 3x = -9\mu \\ t - 2x = -5\mu \end{cases} \iff \begin{cases} x = \lambda \\ y + x = 2\mu \\ 2z - 6x = -9(y + x) \\ 2(t - 2x) = -5(y + x) \end{cases}$$

On obtient donc la représentation suivante de F sous forme de système d'équations cartésiennes :

$$\begin{cases} 3x + 9y + 2z = 0 \\ x + 5y + 2t = 0 \end{cases}$$

Matrice d'une projection.

Dans $\mathbb{R}^3$, soit P le plan d'équation cartésienne $2x + y - z = 0$, et soit D le sous-e.v. donné par le système d'équations cartésiennes suivantes :

$$D \ : \ \begin{cases} x - z = 0 \\ 2x - y = 0 \end{cases}$$

Montrer que D est une droite vectorielle et que $\mathbb{R}^3 = P \oplus D$. Déterminer la matrice dans la base canonique de $\mathbb{R}^3$ de la projection sur P parallèlement à D.

Eléments de correction 1.103. Dans la représentation de D, on peut prendre x comme paramètre libre, ce qui montre immédiatement que D est la droite vectorielle engendrée par le vecteur $u = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$.

Montrons que la somme $P + D$ est directe. Il suffit de vérifier que $P \cap D = \{0\}$. Ici, comme les éléments de D sont tous multiplies de u, c'est-à-dire de la forme tu avec $t \in \mathbb{R}$, il suffit de vérifier que $tu \in P \implies t = 0$. En remplaçant dans l'équation de P, on trouve $t(2 + 2 - 1) = 0$, ce qui implique bien que $t = 0$. Donc la somme est directe. Donc sa dimension est égale à la somme des dimensions de P et de D, c'est-à-dire 3, et comme $\dim \mathbb{R}^3 = 3$ il s'ensuit que $P \oplus D = \mathbb{R}^3$.

Soit p la projection sur P parallèlement à D, et soit A la matrice de p dans la base canonique. On calcule directement A en déterminant une expression pour la projection d'un vecteur. Étant donné un vecteur $a = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, on cherche $t \in \mathbb{R}$ et $w \in P$ tels que $a = tu + w$, et on sait que le couple (t, w) est unique. Il est équivalent de chercher d'abord $t \in \mathbb{R}$ tel que $a - tu \in P$, c'est-à-dire que $a - tu$ vérifie l'équation du plan P :

$$\begin{aligned} & 2(x - t) + (y - 2t) - (z - t) = 0 \\ \iff & t = (2x + y - z)/3 \end{aligned}$$

Le vecteur $w = p(a)$ est alors obtenu par $w = a - tu$, soit :

$$w = \begin{pmatrix} x - t \\ y - 2t \\ z - t \end{pmatrix} = \begin{pmatrix} 1/3x - 1/3y + 1/3z \\ -4/3x + 1/3y + 2/3z \\ -2/3x - 1/3y + 4/3z \end{pmatrix}$$

On lit directement les coefficients de la matrice A cherchée sur cette dernière expression puisqu'on a $w = A \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, soit :

$$A = \begin{pmatrix} 1/3 & -1/3 & 1/3 \\ -4/3 & 1/3 & 2/3 \\ -2/3 & -1/3 & 4/3 \end{pmatrix}$$

Application linéaire et matrice.

On considère l'application linéaire $f : \mathbb{R}^4 \to \mathbb{R}^3$ donnée par la matrice suivante :

$$A = \begin{pmatrix} -11 & 7 & 0 & 3 \\ 0 & 1 & 11 & 2 \\ 1 & 0 & 7 & 1 \end{pmatrix}$$

Déterminer le rang de f, et donner une base de $\ker f$ et une base de $\operatorname{Im} f$. Donner une représentation de l'image sous forme de systèmes d'équations cartésiennes.

Eléments de correction 1.104. Le rang de f est égal au rang de la matrice A. On va le déterminer en effectuant des opérations élémentaires uniquement sur les *colonnes* de A. Ainsi, les vecteurs correspondants aux pivots non nuls seront des vecteurs de base de l'image de f.

$$\begin{pmatrix} -11 & 7 & 0 & 3 \\ 0 & 1 & 11 & 2 \\ 1 & 0 & 7 & 1 \end{pmatrix} \xrightarrow[]{\substack{C_2 \longleftarrow C_2 + 7/11 C_1 \\ C_4 \longleftarrow C_4 + 3/11 C_1}} \begin{pmatrix} -11 & 0 & 0 & 0 \\ 0 & 1 & 11 & 2 \\ 1 & 7/11 & 7 & 14/11 \end{pmatrix}$$

$$\xrightarrow[]{\substack{C_3 \longleftarrow C_3 - 11 C_2 \\ C_4 \longleftarrow C_4 - 2 C_2}} \begin{pmatrix} -11 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 7/11 & 0 & 0 \end{pmatrix}$$

L'application f est donc de rang 2, et une base de son image est donnée par $\left(\begin{pmatrix} -11 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 7/11 \end{pmatrix}\right)$. Pour trouver une équation de l'image, on élimine les paramètres λ et μ dans le système suivant :

$$\begin{cases} x = -11\lambda \\ y = \mu \\ z = \lambda + 7/11\mu \end{cases} \iff x - 7y + 11z = 0.$$

Pour trouver une base du noyau, on revient à la matrice M de départ, en effectuant cette fois des manipulations élémentaires uniquement sur les *lignes*, qui correspondent exactement à la résolution du système $f\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = 0$.

$$\begin{pmatrix} -11 & 7 & 0 & 3 \\ 0 & 1 & 11 & 2 \\ 1 & 0 & 7 & 1 \end{pmatrix} \xrightarrow{L_3 \longleftarrow L_3 + 1/11 L_1} \begin{pmatrix} -11 & 7 & 0 & 3 \\ 0 & 1 & 11 & 2 \\ 0 & 7/11 & 7 & 14/11 \end{pmatrix}$$

$$\xrightarrow{L_3 \longleftarrow L_3 - 7/11 L_2} \begin{pmatrix} -11 & 7 & 0 & 3 \\ 0 & 1 & 11 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

On retrouve bien sûr que f est de rang 2. Nous avons maintenant un système linéaire *échelonné* qui décrit son noyau :

$$\begin{cases} -11x + 7y + 3t = 0 \\ \quad y + 11z + 2t = 0 \end{cases}$$

Comme ce système est échelonné, on peut prendre z et t comme paramètres libres et obtenir une base du noyau en faisant $(z,t) = (1,0)$ puis $(z,t) = (0,1)$. On obtient donc $\left(\begin{pmatrix} -7 \\ -11 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -2 \\ 0 \\ 1 \end{pmatrix}\right)$ comme base de $\ker f$.

Chapitre 2

Formes linéaires et dualité

1 Formes linéaires et hyperplans

Formes linéaires et crochet de dualité. Espace dual.

Rappelons que $\mathscr{L}(E, F)$ est un $\mathbb{R}$-e.v. pour l'addition et la multiplication par les scalaires naturelles, rappelées dans le chapitre précédent. Les *formes linéaires* correspondent au cas où $F = \mathbb{R}$.

Définition 2.1. Soit E un $\mathbb{R}$-e.v.. On appelle *forme linéaire sur E* toute application linéaire de E à valeurs dans $\mathbb{R}$. On note E^* l'espace vectoriel $\mathscr{L}(E, \mathbb{R})$ des formes linéaires sur E, que l'on appelle *espace dual de E*.

On note souvent $\langle \varphi, x \rangle$ pour $\varphi(x)$ lorsque $\varphi \in E^*$ et $x \in E$. On appelle cette notation le *crochet de dualité*.

Exemple 2.2. Prenons $E = \mathbb{R}^3$, et choisissons (a_1, a_2, a_3) des coefficients réels. Posons :

$$f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = a_1 x_1 + a_2 x_2 + a_3 x_3.$$

Il est clair que $f : \mathbb{R}^3 \to \mathbb{R}$ ainsi définie est une forme linéaire. Cet exemple se généralise de façon évidente à $\mathbb{R}^n$ au lieu de $\mathbb{R}^3$. Nous verrons que toutes les formes linéaires sur $\mathbb{R}^n$ sont de cette forme, et que les coefficients $(a_1, \ldots, a_n)$ qui apparaissent correspondent aux coordonnées de f dans la base dite *duale canonique*. □

L'image d'une forme linéaire est un sous-e.v. de $\mathbb{R}$, donc est $\{0\}$ ou $\mathbb{R}$. Ce n'est $\{0\}$ que dans le cas de la forme linéaire nulle, et donc toute forme linéaire non nulle est de rang 1.

Hyperplans.

Définition 2.3. Soit E un $\mathbb{R}$-e.v.. On appelle *hyperplan* de E tout noyau d'une forme linéaire non nulle sur E.

La notion d'hyperplan est définie même si E n'est pas de dimension finie. On peut caractériser les hyperplans de la façon suivante.

Proposition 2.4. *Soit E un $\mathbb{R}$-e.v.*

1. *Un sous-e.v. H de E est un hyperplan si et seulement si il existe un vecteur $x_0 \neq 0$ tel que $E = H \oplus \mathbb{R}x_0$.*
2. *Si de plus E est de dimension finie égale à $n \geq 1$, alors les hyperplans de E sont tous ses sous-e.v. de dimension $n - 1$.*

Dans le premier point, le vecteur x_0 n'est pas unique, même pas à un scalaire près. En effet, la démonstration suivante montre que n'importe quel vecteur non élément de H convient. De plus, le premier point est valide y compris si E n'est pas de dimension finie.

Démonstration. 1. Supposons que H est un hyperplan, alors $H = \ker f$ avec $f \in E^*$ et $f \neq 0$. Il existe $x_0 \notin H$, car sinon f serait identiquement nulle, contredisant notre hypothèse. Vérifions que $E = H \oplus \mathbb{R}x_0$. Il est clair d'une part que $H \cap \mathbb{R}x_0 = \{0\}$. Et d'autre part, pour montrer que $E = H + \mathbb{R}x_0$, on choisit $x \in E$ arbitraire et on pose $t = f(x)/f(x_0)$ ce qui est loisible puisque $f(x_0) \neq 0$. En posant maintenant $y = x - tx_0$, nous avons visiblement $f(y) = 0$ et donc $x = y + tx_0 \in H + \mathbb{R}x_0$, ce qu'on voulait montrer.

   Réciproquement, supposons que $E = H \oplus \mathbb{R}x_0$. Soit $p : E \longrightarrow \mathbb{R}x_0$ la projection sur $\mathbb{R}x_0$ parallèlement à H. Alors pour chaque $x \in E$, il existe un unique réel $f(x)$ tel que $p(x) = f(x)x_0$, et il est clair que $f : E \longrightarrow \mathbb{R}$ est linéaire puisque p est linéaire. Donc $f \in E^*$, et $f \neq 0$ puisque $f(x_0) = 1$. De plus, nous avons $\ker f = \ker p = H$, donc H est bien un hyperplan puisque c'est le noyau de f.

2. On suppose maintenant que $\dim E = n$. Si H est un hyperplan de E, alors $H = \ker f$ avec $\operatorname{rg} f = 1$ donc $\dim H = n - 1$ d'après le théorème du rang. Réciproquement, si $\dim H = n - 1$, choisissons un supplémentaire de H, qui est nécessairement de dimension 1 donc de la forme $\mathbb{R}x_0$ pour un certain $x_0 \neq 0$. Alors $E = H \oplus \mathbb{R}x_0$ donc H est un hyperplan de E d'après le point précédent. □

On a vu que toute forme linéaire non nulle, disons f sur E, définit un hyperplan H de E par la formule $H = \ker f$. Deux formes linéaires différentes peuvent-elles définir le même hyperplan ? La proposition suivante apporte une réponse à cette question.

Proposition 2.5. *Soit E un $\mathbb{R}$-e.v., et soit $f, g \in E^*$ deux formes linéaires toutes les deux non nulles, d'hyperplans associés $H = \ker f$ et $G = \ker g$. Alors $H = G$ si et seulement si f et g sont proportionnelles.*

Démonstration. Il est clair que si $f = \lambda g$, avec λ nécessairement non nul, alors $\ker f = \ker g$. Réciproquement, supposons que $\ker f = \ker g$. Comme f est non nulle, la proposition précédente fournit $x_0 \in E$ non nul tel que $E = \ker f \oplus \mathbb{R}x_0$, en particulier $x_0 \notin \ker f$. Posons $\lambda = g(x_0)/f(x_0)$, ce qui est loisible puisque $x \notin \ker f$ et donc $f(x_0) \neq 0$. Montrons que $g = \lambda f$. Comme H et $\mathbb{R}x_0$ sont supplémentaires, il suffit de le vérifier sur H et sur $\mathbb{R}x_0$. Or c'est évident sur H puisque f et g y sont toutes les deux nulles, et c'est vérifié également sur $\mathbb{R}x_0$ par définition de λ, d'où le résultat. □

2 Bases duales

Dimension du dual.

On se fixe un $\mathbb{R}$-e.v. E de dimension finie, égale à n. On sait déjà que E^* est de dimension finie égale à n : c'est un cas particulier de $\dim \mathscr{L}(E, F) = (\dim E) \cdot (\dim F)$ avec $F = \mathbb{R}$.

Bases duales.

On va s'intéresser à des bases de E^* qu'on peut construire à partir de bases de E.

Théorème 2.6. *Soit E un $\mathbb{R}$-e.v. de dimension finie égale à n, et muni d'une base $\mathscr{B} = (e_1, \dots, e_n)$. Alors il existe une unique base de E^*, notée $\mathscr{B}^* = (e_1^*, \dots, e_n^*)$, vérifiant l'une des deux conditions ci-dessous* (2.1)–(2.2)*, qui sont équivalentes :*

$$\forall x \in E \qquad x = \langle e_1^*, x\rangle e_1 + \cdots + \langle e_n^*, x\rangle e_n \tag{2.1}$$

$$\forall i, j = 1, \dots, n \qquad \langle e_i^*, e_j\rangle = \delta_i^j, \tag{2.2}$$

où δ_i^j est le symbole de Kronecker.

Démonstration. L'équivalence (2.1) $\Longleftrightarrow$ (2.2) est laissée en exercice. L'unicité découle immédiatement de la forme (2.1). Pour l'existence : pour tout vecteur $x \in E$, il existe un unique n-uplet de scalaires $(\lambda_1, \dots, \lambda_n)$ tels que $x = \lambda_1 e_1 + \cdots + \lambda_n e_n$. Posons $e_i^*(x) = \lambda_i$ pour tout $i = 1, \dots, n$ et pour tout vecteur $x \in E$. Il suffit alors de remarquer que les e_i^* ainsi définies sont linéaires, donc des éléments de E^*, qui vérifient (2.1) par construction.

Montrons que la famille $e^* = (e_i^*, \dots, e_n^*)$ forme une base de E^*. Soit $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ tels que :

$$\alpha_1 e_1^* + \cdots + \alpha_n e_n^* = 0,$$

on applique cette égalité successivement aux vecteurs $e_1, \dots, e_n$ pour obtenir $\alpha_1 = \cdots = \alpha_n = 0$, ce qui montre que la famille e^* est libre. Pour voir que e^* est génératrice, soit φ une forme linéaire sur E, et soit $\psi \in E^*$ la forme linéaire définie sur E par $\psi = \langle \varphi, e_1\rangle e_1^* + \cdots + \langle \varphi, e_n^*\rangle$. Alors il est clair que φ et ψ coïncident sur les vecteurs de base $e_1, \dots, e_n$, donc $\varphi = \psi$, ce qui montre que e^* est génératrice. □

Définition 2.7. Soit E un $\mathbb{R}$-e.v. de dimension finie muni d'une base $\mathscr{B} = (e_1, \dots, e_n)$. La base $\mathscr{B}^*$ définie comme dans le Théorème 2.6 est appelée *base duale de e*.

La question de savoir si toute base de E^* peut être obtenue comme base duale d'une base de E est naturelle. Elle sera traitée en 3.

Exemples. Base duale canonique.

Soit $\mathscr{B}_c = (e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$. Alors la base duale de $\mathscr{B}_c$ est appelée *base duale canonique de* $\mathbb{R}^n$. Elle est formée des formes linéaires $(e_1^*, \dots, e_n^*)$ telles que :

$$\langle e_1^*, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\rangle = x_1, \qquad \langle e_2^*, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\rangle = x_2, \qquad \dots \qquad \langle e_n^*, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\rangle = x_n.$$

Soit maintenant f une forme linéaire sur $\mathbb{R}^n$. Soit $(\lambda_1, \dots, \lambda_n)$ les coefficients de f dans la base duale canonique, donc $f = \lambda_1 e_1^* + \cdots + \lambda_n e_n^*$. Les coefficients λ_i sont ceux qui interviennent dans l'expression de f "en coordonnées", c'est-à-dire :

$$\langle f, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\rangle = \lambda_1 x_1 + \cdots + \lambda_n x_n. \tag{2.3}$$

On remarque que les coefficients λ_i s'obtiennent en évaluant la forme linéaire f sur les vecteurs de la base canonique :

$$\lambda_1 = \langle f, e_1 \rangle \qquad \lambda_2 = \langle f, e_2 \rangle \qquad \ldots \qquad \lambda_n = \langle f, e_n \rangle.$$

La proposition suivante généralise cette observation à tout couple $(\mathscr{B}, \mathscr{B}^*)$ de bases, même si $\mathscr{B}$ n'est pas la base canonique de $\mathbb{R}^n$.

Proposition 2.8. *Soit E un $\mathbb{R}$-e.v. de dimension finie muni d'une base $\mathscr{B} = (e_1, \ldots, e_n)$, et soit $\mathscr{B}^* = (e_1^*, \ldots, e_n^*)$ la base duale de E. Alors, pour tout $\varphi \in E^*$, les scalaires λ_i tels que $\varphi = \lambda_1 e_1^* + \cdots + \lambda_n e_n^*$ s'obtiennent en évaluant φ sur les vecteurs de la base $\mathscr{B}$:*

$$\lambda_1 = \langle \varphi, e_1 \rangle \qquad \lambda_2 = \langle \varphi, e_2 \rangle \qquad \ldots \qquad \lambda_n = \langle \varphi, e_n \rangle.$$

Démonstration. On a $\langle e_i^*, e_j \rangle = \delta_i^j$ par définition de la base duale. Les formules $\lambda_i = \langle \varphi, e_i \rangle$ découlent donc immédiatement de l'écriture $\varphi = \lambda_1 e_1^* + \cdots + \lambda_n e_n^*$. □

Représentation matricielle.

Supposons données une base $\mathscr{B}$ de E et sa base duale $\mathscr{B}^*$. Soit $x \in E$ et $\varphi \in E^*$, dont les décompositions dans les bases $\mathscr{B}$ et $\mathscr{B}^*$ sont respectivement :

$$x = \sum_i x_i e_i, \qquad \varphi = \sum_i y_i e_i^*.$$

Évaluons le crochet de dualité :

$$\begin{aligned} \langle \varphi, x \rangle &= \sum_i y_i \langle e_i^*, \sum_j x_j e_j \rangle \\ &= \sum_i y_i x_i \qquad \text{puisque } \langle e_i^*, e_j \rangle = \delta_i^j. \end{aligned}$$

Considérons les deux vecteurs *colonnes* $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$. On constate que le crochet de dualité coïncide avec le produit matriciel suivant :

$$\langle \varphi, x \rangle = {}^tY \cdot X$$

On obtient donc : *à condition d'exprimer les coordonnées des vecteurs et des formes linéaires dans une base et sa base duale, le crochet de dualité correspond au produit scalaire usuel dans $\mathbb{R}^n$.*

Pratique de la détermination d'une base duale.

En pratique, déterminer la base duale d'une base $\mathscr{B} = (u_1, \ldots, u_n)$ de $\mathbb{R}^n$ revient souvent à déterminer les coordonnées de chaque forme linéaire u_i^* dans la base duale canonique.

Soit $\mathscr{B}_c = (e_1, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$, et soit $\mathscr{B}_c^* = (e_1^*, \ldots, e_n^*)$ la base duale canonique. Supposons que les vecteurs u_i sont exprimés par leurs coordonnées dans la base canonique de $\mathbb{R}^n$, et soit $A = (a_{i,j})$ la matrice de leurs coefficients en colonnes :

$$u_j = \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{n,j} \end{pmatrix}.$$

La matrice A est donc la matrice de passage de la base $\mathscr{B}_c$ vers la base $\mathscr{B}$: $A = \mathrm{P}(\mathscr{B}_c, \mathscr{B})$ (cf. Définition 1.77). On cherche à déterminer la matrice de passage de $\mathscr{B}_c*$ vers $\mathscr{B}*$, disons $P = \mathrm{P}(\mathscr{B}_c^*, \mathscr{B}^*) = (p_{i,j})_{i,j}$. Les n^2 relations $\langle u_i^*, u_j \rangle = \delta_i^j$ correspondent à la relation matricielle :

$$ {}^tP \cdot A = I_n . $$

Donc une façon d'obtenir les $(u_i^*)_i$ est d'inverser la matrice A, puis de lire en ligne dans A^{-1} les coordonnés des u_i^* dans la base duale canonique.

Exemple 2.9. Soit (u_1, u_2) la base de $\mathbb{R}^2$ définie par $u_1 = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ et $u_2 = \left(\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}\right)$. La matrice de passage de la base canonique de $\mathbb{R}^2$ vers (u_1, u_2) est $A = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 1 \end{smallmatrix}\right)$, d'inverse $A^{-1} = P = \left(\begin{smallmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{smallmatrix}\right)$. On lit les coordonnées de la base duale en ligne sur cette matrice :

$$ u_1^* = \frac{1}{2} e_1^* + \frac{1}{2} e_2^* \qquad u_2^* = -\frac{1}{2} e_1^* + \frac{1}{2} e_2^* , $$

c'est-à-dire :

$$ u_1^*(\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)) = \frac{x}{2} + \frac{y}{2} \qquad u_2^*(\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)) = -\frac{x}{2} + \frac{y}{2} . $$

3 Bases antéduales

Existence des bases antéduales.

On a vu la construction de bases duales de E^* à partir de bases de E. Réciproquement, toute base de E^* peut-elle être obtenue comme une base duale ? Le résultat suivant répond à cette question par l'affirmative.

Définition 2.10. Soit E un $\mathbb{R}$-e.v. de dimension finie égale à n, et soit f une base de E^*. Une base e de E est *antéduale* de f si $e^* = f$.

Théorème 2.11. *Soit E un $\mathbb{R}$-e.v. de dimension finie égale à n. Soit $f = (f_1, \ldots, f_n)$ une base de E^*. Alors il existe une unique base e de E telle que $e^* = f$.*

Pour la démonstration du théorème, on utilisera le lemme suivant, qui est utile en lui-même.

Lemme 2.12 (lemme de séparation). *Soit E un $\mathbb{R}$-e.v. de dimension finie égale à n. Soit $(f_1, \ldots, f_n)$ une base de E^* et soit $x \in E$. Si $f_i(x) = 0$ pour tout $i = 1, \ldots, n$, alors $x = 0$.*

Démonstration du lemme 2.12. Comme les f_i forment une base de E^* par hypothèse, il s'ensuit que $f(x) = 0$ pour tout $f \in E^*$. Maintenant, supposons par l'absurde que $x \neq 0$, et soit H un supplémentaire de $\mathbb{R}x$. Alors $\dim H = n - 1$ donc H est un hyperplan d'après la proposition 2.4, c'est-à-dire $H = \ker f$ avec $f \in E^*$ et $f \neq 0$. Comme la somme $H \oplus \mathbb{R}x$ est directe, l'intersection $H \cap \mathbb{R}x$ est réduite à $\{0\}$, donc $x \notin \ker f$, contradiction. □

Démonstration du théorème 2.11. Unicité. Soit $e = (e_1, \ldots, e_n)$ et $e' = (e'_1, \ldots, e'_n)$ deux bases de E telles que $e^* = (e')^* = f$. Alors en particulier $f_i(e_j) = f_i(e'_j)$ pour tous i et j, donc $f_i(e_j - e'_j) = 0$. D'après le lemme précédent, cela implique $e_j = e'_j$, et ceci pour tout j donc $e = e'$.

Existence. On propose deux démonstrations de l'existence de la base antéduale.

Première démonstration. Introduisons le bidual E^{**} de E, défini par $E^{**} = (E^*)^*$. C'est un $\mathbb{R}$-e.v. de même dimension que E^*, et donc que E. Considérons l'application linéaire $J : E \longrightarrow E^{**}$ définie par $u \in E \mapsto J_u \in E^{**}$ avec $J_u(\varphi) = \varphi(u)$ pour $u \in E$ et $\varphi \in E^*$. Cette application est appelée *évaluation*.

Vérifions que J est injective, c'est-à-dire que $\ker J = \{0_E\}$: c'est une conséquence immédiate du lemme précédent. Comme $\dim E = \dim E^{**}$, il s'ensuit que J est un isomorphisme de E sur E^{**}.

Considérons maintenant la base $f^* = (f_1^*, \ldots, f_n^*)$ de E^{**}, duale de f. Puisque J est un isomorphisme, les vecteurs $(e_1, \ldots, e_n)$ de E tels que $J(e_j) = f_j^*$ forment une base de E. Ils forment la base cherchée. En effet, on a :

$$\begin{aligned}
f_i(e_j) &= J_{e_j}(f_i) && \text{par définition de } J, \\
&= f_j^*(f_i) && \text{par définition de } (e_1, \ldots, e_n), \\
&= \delta_i^j && \text{par définition de la base duale } f^*.
\end{aligned}$$

On a montré que $e^* = f$, ce qu'on voulait.

Deuxième démonstration. On part toujours d'une base $f = (f_1, \ldots, f_n)$ de E^* dont on cherche une base antéduale e. Choisissons une base $u = (u_1, \ldots, u_n)$ de E, et soit $u^* = (u_1^*, \ldots, u_n^*)$ sa base duale. Soit P la matrice de passage de u^* vers f, qui exprime donc dans la colonne j les coordonnées de f_j dans la base $(u_1^*, \ldots, u_n^*)$.

Raisonnons par analyse-synthèse. Si la base antéduale e existe, la matrice de passage $A = \mathrm{P}(u, e)$ doit vérifier ${}^tP \cdot A = I_n$, ce qui correspond aux n^2 relations $\langle f_i, e_j \rangle = \delta_i^j$. En effet :

$$\begin{aligned}
\langle f_i, e_j \rangle &= \left\langle \sum_k P_{k,i} u_k^*, \sum_k A_{k,j} u_k \right\rangle \\
&= \sum_k P_{k,i} A_{k,j} && \text{car } u^* \text{ est la base duale de } u \\
&= ({}^tP \cdot A)_{i,j}
\end{aligned}$$

Il suffit donc de poser $A = ({}^tP)^{-1}$, et de consider pour e_i le vecteur dont les coordonnées dans la base $(u_1, \ldots, u_n)$ sont données par la i^{e} colonne de A. Le calcul précédent nous dit alors que f est bien duale de e. □

Détermination pratique d'une base antéduale.

La deuxième démonstration de l'existence de la base antéduale dans la démonstration du théorème 2.11 fournit une méthode pratique pour obtenir les coordonnées d'une base antéduale.

Exemple 2.13. Soit $f = (f_1, f_2)$ la base de $(\mathbb{R}^2)^*$ définie par $f_1 = x + y$, $f_2 = 2x - y$. Cherchons les coefficients $(a_{i,j})$ de la base antéduale de f, avec $(a_{i,1})_{i=1,2}$ les coefficients de e_1 et $(a_{i,2})_{i=1,2}$ les coefficients de e_2. La matrice $A = (a_{i,j})$ vérifie : $\left(\begin{smallmatrix} 1 & 1 \\ 2 & -1 \end{smallmatrix}\right) \cdot A = I_2$. Après inversion : $A = \begin{pmatrix} 1/3 & 1/3 \\ 2/3 & -1/3 \end{pmatrix}$, et on y lit les coordonnées de e_1 et e_2 en colonnes : $e_1 = \begin{pmatrix} 1/3 \\ 2/3 \end{pmatrix}$ et $e_2 = \begin{pmatrix} 1/3 \\ -1/3 \end{pmatrix}$.

4 Application aux systèmes linéaires

Systèmes linéaires et hyperplans.

Un système linéaire homogène, c'est-à-dire, avec second membre nul, s'interprète comme une intersection d'hyperplans. Par exemple, dans $\mathbb{R}^4$, le système :

$$\begin{cases} 2x - y + z + t = 0 \\ -x - y + 2z - t = 0 \end{cases} \tag{2.4}$$

est l'intersection des deux hyperplans de $\mathbb{R}^4$ d'équations $2x - y + z + t = 0$ et $-x - y + 2z - t = 0$.

Dimension d'une intersection d'hyperplans.

Cette interprétation permet de retrouver des résultats bien connus.

Proposition 2.14. *Soit E un $\mathbb{R}$-e.v. de dimension finie égale à n. Soit $(\varphi_1, \ldots, \varphi_k)$ des formes linéaires sur E, et soit V le sous-e.v. de E défini par :*

$$V = \bigcap_{i=1}^{k} \ker \varphi_i \, .$$

Si le système $(\varphi_1, \ldots, \varphi_k)$ est libre, alors $\dim V = n - k$. *Plus généralement, si* $k' = \dim \mathrm{Vect}\langle \varphi_1, \ldots, \varphi_k \rangle$, *alors* $\dim V = n - k'$.

Démonstration. Traitons d'abord le cas où les formes linéaires sont libres. D'après le théorème de la base incomplète, on choisit $\varphi_{k+1}, \ldots, \varphi_n$ de sorte que $(\varphi_1, \ldots, \varphi_n)$ soit une base de E^*. Considérons l'application linéaire $F : E \longrightarrow \mathbb{R}^n$ définie comme suit :

$$F(x) = \begin{pmatrix} \varphi_1(x) \\ \vdots \\ \varphi_n(x) \end{pmatrix}.$$

Montrons que F est un isomorphisme. Il suffit de voir que $\ker F = \{0\}$ puisque E et $\mathbb{R}^n$ ont même dimension. Or cela découle immédiatement du lemme de séparation 2.12.

Maintenant, on voit que $V = F^{-1}(U)$, où U est le sous-e.v. de $\mathbb{R}^n$ défini par :

$$U = \{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \ : \ x_1 = 0, \ldots, x_k = 0 \}.$$

Comme F est un isomorphisme d'une part, et comme il est clair que $\dim U = n - k$ d'autre part, on en déduit que $\dim V = n - k$, ce qu'on voulait montrer.

Dans le cas général, choisissons $\varphi'_1, \ldots, \varphi'_{k'}$ une base du sous-e.v. $\mathrm{Vect}\langle \varphi_1, \ldots, \varphi_k \rangle$, et posons :

$$V' = \bigcap_{i=1}^{k'} \ker \varphi'_i.$$

Alors, d'une part, nous avons $\dim V' = n - k'$ d'après le cas précédent puisque $(\varphi'_1, \dots, \varphi'_{k'})$ est libre par construction. Et d'autre part, il est immédiat de vérifier que $V = V'$ en utilisant $\mathrm{Vect}\langle \varphi_1, \dots, \varphi_k \rangle = \mathrm{Vect}\langle \varphi'_1, \dots, \varphi'_{k'} \rangle$. Le résultat en découle. □

Remarque 2.15. Si on se donne des formes linéaires de $\mathbb{R}^n$ par leurs expressions en coordonnées, on lit immédiatement leurs coefficients dans la base duale canonique de $\mathbb{R}^n$. On peut donc calculer le rang du système qu'elles forment en considérant la matrice par lignes de leurs coefficients, et en évaluant le rang de cette matrice.

Par exemple, pour le système (2.4) ci-dessus, on forme la matrice :

$$A = \begin{pmatrix} 2 & -1 & 1 & 1 \\ -1 & -1 & 2 & -1 \end{pmatrix} \xrightarrow{L_2 \leftarrow L_2 + 1/2 L_1} \begin{pmatrix} 2 & -1 & 1 & 1 \\ 0 & -3/2 & 5/2 & -1/2 \end{pmatrix}$$

qui est de rang 2. On en déduit que le sous-e.v. des solutions du système est de dimension $4 - 2 = 2$. Les manipulations uniquement sur les lignes correspondent aux manipulations sur les équations, on a donc trouvé un système échelonné équivalent. On en déduit une base (u_1, u_2) des solutions en faisant $(z, t) = (1, 0)$ puis $(z, t) = (0, 1)$:

$$\begin{cases} 2x = y - z - t \\ y = 5/3z - 1/3t \end{cases} \qquad u_1 = \begin{pmatrix} 1/3 \\ 5/3 \\ 1 \\ 0 \end{pmatrix} \quad u_2 = \begin{pmatrix} -2/3 \\ -1/3 \\ 0 \\ 1 \end{pmatrix}$$

Une application.

Le résultat précédent fournit une preuve de l'égalité des rangs par lignes et par colonnes d'un système de vecteurs.

Théorème 2.16. *Soit M une matrice de taille $m \times n$. Soit p le rang du système des vecteurs ligne de M, et soit q le rang du système des vecteurs colonne de M. Alors $p = q$.*

Démonstration. Posons $E = \mathbb{R}^n$, et soit $(f_1, \dots, f_m)$ les m formes linéaires sur E correspondant aux vecteurs ligne de M. Soit également $F : \mathbb{R}^n \to \mathbb{R}^m$ l'application linéaire définie par :

$$F(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}.$$

Posons finalement :

$$V = \ker F = \bigcap_{j=1}^{m} \ker f_j.$$

Alors nous avons $q = \mathrm{rg}\, F$ par définition du rang d'une application linéaire, et $p = \mathrm{rg}(f_1, \dots, f_m)$. Appliquons la proposition 2.14 : on trouve $\dim V = n - p$. D'après le théorème du rang, on a aussi $\dim V = n - \mathrm{rg}\, F = n - q$, d'où $p = q$. □

5 Applications linéaires transposées

Définition 2.17. Soit E et F deux $\mathbb{R}$-e.v., et soit $g : E \to F$ une application linéaire. La *transposée* de l'application g est l'application linéaire notée

$$g^* : F^* \to E^*$$

définie par $g^*(\varphi) = \varphi \circ g$ pour $\varphi \in F^*$. De manière équivalente, $g^*(\varphi)$ est caractérisée par :

$$\langle g^*(\varphi), x\rangle = \langle \varphi, g(x)\rangle.$$

On peut visualiser $g^*(\varphi)$ sur le diagramme suivant :

$$\begin{tikzcd} E \arrow[r, "g"] \arrow[dr, "g^*(\varphi)"'] & F \arrow[d, "\varphi"] \\ & \mathbb{R} \end{tikzcd}$$

Proposition 2.18. *Soit E, F et G des $\mathbb{R}$-e.v., et soit des applications linéaires*

$$\begin{tikzcd} E \arrow[r, "g"] \arrow[rr, bend right, "h\circ g"'] & F \arrow[r, "h"] & G \end{tikzcd}.$$

Alors $(h \circ g)^* = g^* \circ h^*$. De plus, $(\mathrm{Id}_E)^* = \mathrm{Id}_{E^*}$.

Démonstration. Il suffit de vérifier la définition. Pour $\varphi \in G^*$ et $x \in E$, on a :

$$\begin{aligned}
\langle (h\circ g)^*(\varphi), x\rangle &= \langle \varphi, h\circ g(x)\rangle && \text{par définition de } (h\circ g)^* \\
&= \langle \varphi, h\big(g(x)\big)\rangle && \text{par définition de la composition} \\
&= \langle h^*(\varphi), g(x)\rangle && \text{par définition de } h^* \\
&= \langle g^*\big(h^*(\varphi)\big), x\rangle && \text{par définition de } g^* \\
&= \langle g^*\circ h^*(\varphi), x\rangle && \text{par définition de la composition.}
\end{aligned}$$

L'identité $(\mathrm{Id}_E)^* = \mathrm{Id}_{E^*}$ découle d'un calcul similaire laissé au lectorat. □

Proposition 2.19. *Soit E et F deux $\mathbb{R}$-e.v. tous deux de dimension finie, et soit $g : E \to F$ une application linéaire. Soit e une base de E, soit f une base de F, et posons :*

$$M = \mathrm{Mat}(g, e, f), \qquad M' = \mathrm{Mat}(g^*, f^*, e^*).$$

Alors M' est la transposée de M.

Démonstration. Posons $e = (e_1, \ldots, e_n)$ et $f = (f_1, \ldots, f_m)$. Évaluons la forme linéaire $g^*(f_j^*)$. D'une part, par définition de la matrice M', nous avons :

$$g^*(f_j^*) = \sum_k M'_{k,j} e_k^*. \tag{2.5}$$

D'autre part, par définition de g^*, nous avons :

$$\langle g^*(f_j^*), x\rangle = \langle f_j^*, g(x)\rangle$$

Or x se décompose dans la base e par $x = \sum_k e_k^*(x)e_k$, d'où :

$$\langle g^*(f_j^*), x\rangle = \sum_k e_k^*(x)\langle f_j^*, g(e_k)\rangle$$

ce qui s'écrit aussi :

$$\begin{aligned} g^*(f_j^*) &= \sum_k \langle f_j^*, g(e_k)\rangle e_k^* \\ &= \sum_k M_{j,k} e_k^* \qquad \text{par définition de la matrice } M. \end{aligned}$$

En comparant avec (2.5), on en déduit donc $M'_{j,k} = M_{k,j}$, ce qu'on voulait. □

Corollaire 2.20. *Soit E et F deux $\mathbb{R}$-e.v. de tous deux de dimension finie, et soit $f : E \to F$ une application linéaire. Alors* $\operatorname{rg} f = \operatorname{rg} f^*$.

Démonstration. En effet, cela découle de la proposition 2.18, puisqu'une matrice a même rang que sa transposée, d'après le résultat rappelé au théorème 2.16. □

6 Retour sur les bases duales

On interprète la recherche de bases duales et antéduales dans $\mathbb{R}^n$ à l'aide de la notion de transposée.

Soit $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ la base canonique de $\mathbb{R}^n$, et supposons donnée une autre base $e = (e_1, \dots, e_n)$ de $\mathbb{R}^n$. En écrivant en colonnes les coordonnées des e_i dans la base ε, on obtient la matrice de passage de ε vers e :

$$P = \mathrm{P}(\varepsilon, e) = \mathrm{Mat}(\mathrm{Id}_{\mathbb{R}^n}, e, \varepsilon),$$

c'est-à-dire la matrice de l'application identité de $\mathbb{R}^n$, avec e comme base de départ et ε comme base d'arrivée. D'après la proposition 2.18, la transposée de l'identité de $\mathbb{R}^n$ est l'identité de $(\mathbb{R}^n)^*$. En appliquant la proposition 2.19, on a donc :

$${}^tP = \mathrm{Mat}(\mathrm{Id}_{(\mathbb{R}^n)^*}, \varepsilon^*, e^*), \quad \text{et donc} \quad {}^t(P^{-1}) = \mathrm{Mat}(\mathrm{Id}_{(\mathbb{R}^n)^*}, e^*, \varepsilon^*).$$

Or ε^* est la base duale canonique. Donc les coordonnées des vecteurs de e^* s'expriment en colonnes avec les coefficients de P^{-1}. Autrement dit, vues en lignes, ces coordonnées s'obtiennent en calculant l'inverse de la matrice de passage P. C'est bien la même opération décrite plus haut au paragraphe 2.

Chapitre 3

Espaces euclidiens

1 Produit scalaire et norme

Produits scalaires.

La notion de produit scalaire a déjà été vue dans les cours précédents, en tout cas à travers des exemples. En voici une définition générale.

Définition 3.1. Soit E un $\mathbb{R}$-e.v.. On se fixe une fonction $P : E \times E \longrightarrow \mathbb{R}$, et on note $\langle x, y\rangle = P(x, y)$. La fonction $P = \langle \cdot, \cdot \rangle$ est un *produit scalaire* sur E si :

1. (Symétrie) $\langle y, x\rangle = \langle x, y\rangle$ pour tous $x, y \in E$.
2. (Bilinéarité) $\langle x + \lambda y, z\rangle = \langle x, z\rangle + \lambda\langle y, z\rangle$ pour tous $x, y, z \in E$ et $\lambda \in \mathbb{R}$.
3. (Positivité) $\langle x, x\rangle \geq 0$ et $\langle x, x\rangle = 0 \iff x = 0$ pour tout $x \in E$.

Exemples.

On retrouve d'abord l'exemple du produit scalaire canonique sur $\mathbb{R}^n$, défini pour $x = (x_i)_i$ et $y = (y_i)_i$ par :

$$\langle x, y\rangle = \sum_i x_i y_i$$

Étant un $\mathbb{R}$-e.v. de dimension finie E, il n'y *a priori* pas de produit scalaire "canonique" sur E. Mais si on se fixe un isomorphisme $\Phi : \mathbb{R}^n \longrightarrow E$, alors on définit un produit scalaire sur E en posant $P(x, y) = \langle \Phi^{-1}(x), \Phi^{-1}(y)\rangle$.

En particulier, si $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ est un isomorphisme, la formule $P(x, y) = \langle f(x), f(y)\rangle$ définit un produit scalaire sur $\mathbb{R}^n$, différent *a priori* du produit scalaire canonique.

Un exemple en dimension infinie : prenons $E = C([0, 1])$, l'espace des fonctions continues sur $[0, 1]$ et à valeurs dans $\mathbb{R}$. Alors, pour $f, g \in E$, on pose :

$$\langle f, g\rangle = \int_0^1 f(t)g(t)dt,$$

ce qui définit un produit scalaire sur E. Remarquez que la positivité provient du résultat affirmant que l'intégrale d'une fonction *continue et positive* est nulle si et seulement si la fonction est nulle sur l'intervalle d'intégration.

Définition 3.2. On appelle *espace euclidien* un $\mathbb{R}$-e.v. E de dimension finie et muni d'un produit scalaire. Le plus souvent, on notera $\langle x, y\rangle$ pour le produit scalaire.

Identités remarquables.

La bilinéarité et la symétrie du produit scalaire fournissent des identités remarquables semblables aux identités habituelles :

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

Norme d'un produit scalaire.

L'objectif est maintenant de définir une notion de "longueur" des vecteurs, qu'on appelle leur norme, à partir d'un produit scalaire qu'on s'est fixé. L'ingrédient essentiel est le résultat suivant.

Proposition 3.3 (inégalité de Cauchy-Schwarz). *Soit E un espace euclidien. Alors pour tous vecteurs $x, y \in E$, l'inégalité suivante a lieu :*

$$\langle x,y\rangle^2 \le \langle x,x\rangle\langle y,y\rangle, \tag{3.1}$$

et l'égalité a lieu si et seulement si la famille (x, y) est liée.

Démonstration. Soit $x, y \in E$. Le cas $y = 0$ étant trivial, on peut supposer sans perte de généralité que $y \neq 0$. Considérons la fonction réelle de la variable réelle t définie par $P(t) = \langle x+ty, x+ty\rangle$. D'une part, la positivité du produit scalaire montre que $P(t) \ge 0$ pour tout $t \in \mathbb{R}$. D'autre part, développons en utilisant l'identité remarquable :

$$P(t) = \langle x,x\rangle + 2t\langle x,y\rangle + t^2\langle y,y\rangle.$$

La fonction P est donc un polynôme du second degré, puisque $y \neq 0$. Comme ce polynôme ne change pas de signe, son discriminant réduit est négatif, soit :

$$\Delta = \langle x,y\rangle^2 - \langle x,x\rangle\langle y,y\rangle \le 0, \tag{3.2}$$

ce qui est exactement l'inégalité cherchée (3.1). Puisque $y \neq 0$, la famille (x, y) est liée si et seulement si il existe $\lambda \in \mathbb{R}$ tel que $y = \lambda x$. Si c'est le cas, l'égalité dans (3.1) est évidente. Et réciproquement, s'il y a égalité dans (3.1), alors le discriminant réduit (3.2) est nul, donc le polynôme admet une racine, disons $P(t_0) = 0$. D'après la positivité du produit scalaire, ceci implique $x + t_0 y = 0$, donc la famille (x, y) est liée. □

La notion intuitive de longueur des vecteurs est formalisée par la notion de norme, définie ci-dessous.

Définition 3.4. Soit E un $\mathbb{R}$-e.v.Une fonction $N : E \to \mathbb{R}$, qu'on note souvent $N(x) = \|x\|$, est appelée une *norme sur E* si elle vérifie les trois propriétés suivantes :

1. (Positivité) $\|x\| \ge 0$ et $\|x\| = 0 \iff x = 0$ pour tout $x \in E$.
2. (Homogénéité) $\|\lambda x\| = |\lambda| \cdot \|x\|$ pour tout $x \in E$ et tout $\lambda \in \mathbb{R}$.
3. (Inégalité triangulaire) $\|x + y\| \le \|x\| + \|y\|$ pour tous $x, y \in E$.

Théorème 3.5. *Soit E un $\mathbb{R}$-e.v. et soit $\langle \cdot, \cdot \rangle$ un produit scalaire sur E. Alors la fonction $\|\cdot\| : E \longrightarrow \mathbb{R}$ définie par $\|x\| = \sqrt{\langle x, x \rangle}$ est une norme sur E.*

Démonstration. La positivité de $\|\cdot\|$ provient de la positivité du produit scalaire. Vérifions l'homogénéité en évaluant le carré de la norme :

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

En passant aux racines carrées on obtient bien $\|\lambda x\| = |\lambda| \cdot \|x\|$. Il reste à vérifier l'inégalité triangulaire. À nouveau, en évaluant les carrés des normes, il est équivalent de vérifier : $\big(\|x\| + \|y\|\big)^2 - \|x+y\|^2 \geq 0$. Développons les carrés *via* les identités remarquables, on obtient :

$$\begin{aligned}\big(\|x\| + \|y\|\big)^2 - \|x+y\|^2 &= \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| - (\|x\|^2 + \|y\|^2 + 2\langle x, y \rangle) \\ &= 2\|x\| \cdot \|y\| - 2\langle x, y \rangle,\end{aligned}$$

qui est bien positif d'après l'inégalité de Cauchy-Schwartz. □

Définition 3.6. Si E est un $\mathbb{R}$-e.v. muni d'un produit scalaire $\langle \cdot, \cdot \rangle$, la norme définie par $\|x\| = \sqrt{\langle x, x \rangle}$ est la norme *associée au produit scalaire.*

Angle non orienté de vecteurs.

La notion de produit scalaire fournit immédiatement une notion d'angle non orienté entre vecteurs.

Définition 3.7. Soit E un espace euclidien, et soit x et y deux vecteurs non nuls. *L'angle non orienté entre x et y* est le réel défini par :

$$\arccos \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}.$$

D'après ce qu'on sait sur la fonction arccosinus, c'est donc un réel compris entre 0 et π au sens large.

La notion d'angle *orienté* de vecteurs sera abordée plus loin.

Distance.

Si on voit les vecteurs de E comme des *points*, on souhaite mesurer la distance entre les points. La notion de norme donne une mesure de la longueur des vecteurs. On en déduit une notion de distance en posant $d(x, y) = \|x - y\|$ pour $x, y \in E$, et on appelle $d : E \times E \longrightarrow \mathbb{R}$ la *fonction distance sur E.*

La fonction distance a les propriétés suivantes :

1. (Positivité) $d(x, y) \geq 0$ et $d(x, y) = 0 \iff x = y$ pour tous $x, y \in E$.
2. (Symétrie) $d(x, y) = d(y, x)$ pour tous $x, y \in E$.
3. (Inégalité triangulaire) $d(x, z) \leq d(x, y) + d(y, z)$ pour tous $x, y, z \in E$.
4. (Homogénéité) $d(\lambda x, \lambda y) = |\lambda| \cdot d(x, y)$ pour tous $x, y \in E$.

Identités remarquables. Inégalité triangulaire inversée. Vecteurs normés.

Avec la notion de norme, les identités remarquables s'écrivent comme suit :

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

L'inégalité suivante est vérifiée pour tous vecteurs $x, y \in E$. En fait, elle est équivalente à l'inégalité triangulaire :

$$\|x-y\| \geq \|x\| - \|y\|.$$

En effet appliquons l'inégalité triangulaire à $x = (x-y)+y$, on trouve $\|x\| \leq \|x-y\|+\|y\|$, d'où l'inégalité cherchée.

Remarque 3.8. Il ne faut pas se laisser abuser par les signes dans les deux inégalités triangulaires ! En effet, d'après l'homogénéité, on peut très bien écrire l'inégalité triangulaire $\|x-y\| \leq \|x\| + \|y\|$ et l'inégalité triangulaire inversée sous la forme $\|x+y\| \geq \|x\| - \|y\|$.

On choisit ces signes pour énoncer les inégalités pour des raisons de "mnémotechnique intuitive". La première inégalité peut être comprise comme "si x et y sont petits, alors $x+y$ est petit". Et la deuxième peut être comprise comme : "si $x-y$ est petit, alors x et y ont des longueurs proches".

On dit qu'un vecteur x est *normé* s'il est de norme 1 : $\|x\| = 1$. Si $x \neq 0$, il existe exactement deux vecteurs y_1 et y_2 qui lui sont colinéaires et normés, ce sont $y_1 = x/\|x\|$ et $y_2 = -y_1$.

2 Orthogonalité

Vecteurs orthogonaux.

Soit E un espace euclidien. Deux vecteurs $x, y \in E$ sont dits *orthogonaux* si $\langle x, y\rangle = 0$. On le note :

$$x \perp y.$$

Par linéarité, on voit que 0 est orthogonal à tout vecteur de E, et que 0 est le seul vecteur de E qui est orthogonal à lui-même.

Avec la notion d'angle non orienté introduite plus haut, on voit que deux vecteurs *non nuls* sont orthogonaux si et seulement ils font un angle de $\pi/2$.

Familles orthogonales.

Une famille de vecteurs $(x_1, \ldots, x_k)$ est *orthogonal* si $x_i \perp x_j$ pour tous $i \neq j$. Deux sous-e.v. U et V de E sont *orthogonaux* si $x \perp y$ pour tous vecteurs $x \in U$ et $y \in V$, ce qu'on notera $U \perp V$.

Proposition 3.9 (théorème de Pythagore). *Soit E un $\mathbb{R}$-e.v. muni d'un produit scalaire. Soit $(x_1, \ldots, x_k)$ une famille orthogonal de E, et soit $x = x_1 + \cdots + x_k$. Alors :*

$$\|x\|^2 = \|x_1\|^2 + \cdots + \|x_k\|^2.$$

Démonstration. Le résultat est clair pour $k = 2$ d'après l'identité remarquable $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle$. On montre le résultat pour $k \geq 2$ par récurrence sur k. □

Familles orthonormées et bases orthonormées.

Une famille $(x_i)_i$ est *orthonormée* si elle est orthogonale et si tous les x_i sont de norme 1. En particulier on parle de bases orthogonales et de bases orthonormées. Ainsi, une base $e = (e_1, \ldots, e_n)$ est orthonormée si et seulement si :

$$\forall i, j = 1, \ldots, n \qquad \langle e_i, e_j \rangle = \delta_i^j.$$

Les bases orthonormées sont extrêmement agréables à utiliser car les coordonnés des vecteurs s'y expriment facilement, comme le montre le résultat suivant.

Proposition 3.10. *Soit E un espace euclidien muni d'une base $e = (e_1, \ldots, e_n)$ orthonormée. Alors tout vecteur $x \in E$ admet la décomposition suivante dans la base e :*

$$x = \sum_i \langle e_i, x \rangle e_i. \tag{3.3}$$

Démonstration. Puisque e est une base, il existe des scalaires $\lambda_1, \ldots, \lambda_n$ tels que $x = \sum_i \lambda_i e_i$. En effectuant le produit scalaire avec un vecteur de base e_j dans cette égalité, et en utilisant que $\langle e_i, e_j \rangle = \delta_i^j$, on trouve $\lambda_j = \langle e_j, x \rangle$. □

Le résultat suivant est élémentaire, mais très utile.

Proposition 3.11. *Soit E un $\mathbb{R}$-e.v. muni d'un produit scalaire. Alors toute famille orthogonale de vecteurs non nuls est libre. En particulier, toute famille orthonormé est libre.*

Démonstration. Soit $(x_1, \ldots, x_n)$ une famille orthogonale de vecteurs non nuls, et soit $\lambda_1, \ldots, \lambda_n$ des réels tels que $\sum_i \lambda_i x_i = 0$. Effectuons le produit scalaire avec x_j, on trouve par linéarité $\sum_i \lambda \langle x_j, x_i \rangle = 0$. Or comme la famille est orthogonale, tous les termes pour $i \neq j$ dans cette somme sont nuls, d'où finalement $\lambda_j \langle x_j, x_j \rangle = 0$. Comme $x_j \neq 0$, il s'ensuit que $\lambda_j = 0$, et ceci quel que soit j, donc la famille est libre. □

Expression du produit scalaire dans une base orthonormée.

Dans $\mathbb{R}^n$ muni du produit scalaire usuel, la base canonique est orthonormée.

Soit E euclidien muni d'une base orthonormée $e = (e_1, \ldots, e_n)$. La formule (3.3) montre comment calculer les coordonnées d'un vecteur dans la base e. On peut aussi calculer facilement le produit scalaire de deux vecteurs donnés par leurs coordonnées dans la base e.

Proposition 3.12. *Soit E un espace euclidien muni d'une base orthonormée $e = (e_1, \ldots, e_n)$. Si $x = \sum_{i=1}^n x_i e_i$ et $y = \sum_{i=1}^n y_i e_i$, alors :*

$$\langle x, y \rangle = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \langle e_i, e_j \rangle = \sum_{i=1}^n x_i y_i.$$

Autrement dit, étant donnée une base orthonormée, on se ramène à l'espace euclidien $\mathbb{R}^n$ muni de son produit scalaire canonique.

Orthonormalisation de Gram-Schmidt.

Ce procédé est un algorithme qui construit des familles orthonormées. On part d'un $\mathbb{R}$-e.v. E muni d'un produit scalaire $\langle \cdot, \cdot \rangle$. Soit $(x_1, \dots, x_k)$ une famille libre. On définit $(u_1, \dots, u_k)$ par récurrence. On pose $u_1 = x_1/\|x_1\|$. Puis par induction, si $(u_1, \dots, u_i)$ ont été construits avec $i < k$, on pose :

$$u'_{i+1} = x_{i+1} - \big(\langle x_{i+1}, u_1 \rangle u_1 + \dots + \langle x_{i+1}, u_i \rangle u_i\big) \qquad\qquad u_{i+1} = \frac{u'_{i+1}}{\|u'_{i+1}\|}$$

On verra plus tard que cette formule est motivée par le fait que l'on obtient u'_{i+1} en lui retirant la projection orthogonale de x_{i+1} sur l'espace vectoriel engendré par $(u_1, \dots, u_i)$.

Théorème 3.13. *Soit $(x_1, \dots, x_k)$ une famille libre, et $(u_1, \dots, u_k)$ la famille de vecteurs construits par le procédé d'orthonormalisation de Gram-Schmidt. Alors :*

1. *La famille $(u_1, \dots, u_k)$ est orthonormé.*
2. *Posons $U_i = \mathrm{Vect}\langle u_1, \dots, u_i \rangle$ et $V_i = \mathrm{Vect}\langle x_1, \dots, x_i \rangle$ pour $i = 1, \dots, k$. Alors $U_i = V_i$ pour tout $i = 1, \dots, k$.*

En particulier, $(u_1, \dots, u_k)$ est une base orthonormée de $\mathrm{Vect}\langle x_1, \dots, x_k \rangle$.

Démonstration. La suite $(u_1, \dots, u_k)$ étant construite par récurrence, il est naturel de prouver le résultat énoncé par récurrence sur i : on va montrer que pour tout $i \in \{1, \dots, k\}$ on a $(u_1, \dots, u_i)$ orthonormé et $U_i = V_i$. Pour $i = 1$, il n'y a presque rien à faire : u_1 et x_1 étant colinéaires, on a bien $V_1 = U_1$, et comme $\|u_1\| = \left\| \frac{1}{\|x_1\|} x_1 \right\| = \frac{\|x_1\|}{\|x_1\|} = 1$, on a également que (u_1) est une base orthonormée de $U_1 = V_1$.

Supposons le résultat établi pour $i \geq 1$ (avec $i < k$), établissons le pour $i+1$. Il s'agit de prouver que $(u_1, \dots, u_{i+1})$ est orthonormé et que $V_{i+1} = U_{i+1}$.

Notons $u = \sum_{j=1}^{i} \langle x_{i+1}, u_j \rangle u_j \in \mathrm{Vect}(u_1, \dots, u_i) = U_i$. Tout d'abord, comme $x_{i+1} \notin U_i = V_i$, on a que $u'_{i+1} = x_{i+1} + u \notin U_i$, en particulier u'_{i+1} est non nul et la définition de u_{i+1} est donc licite. Par construction $u_{i+1} \in \mathrm{Vect}(u_1, \dots, u_i, x_{i+1}) = \mathrm{Vect}(x_1, \dots, x_i, x_{i+1})$ car $V_i = U_i$, donc $U_{i+1} \subseteq V_{i+1}$. De plus $x_{i+1} \notin V_i$ donc $u'_{i+1} \notin V_i$, et ainsi $u_{i+1} \notin V_i$. Il suit que $(u_{i+1}, x_1, \dots, x_i)$ est libre, donc $\dim U_{i+1} = i + 1 = \dim V_{i+1}$ et donc $U_{i+1} = V_{i+1}$. Ceci termine d'établir le point 2.

Concernant le premier point, on remarque tout d'abord que comme $(u_1, ..., u_i)$ est orthonormée, il suffit de montrer que u_{i+1} est de norme 1 et qu'il est orthogonal à u_j pour tout $j \in \{1, \dots i\}$. Comme $u_{i+1} = \frac{1}{\|u'_{i+1}\|} u'_{i+1}$, il est de norme 1, et pour tout $j \in \{1, \dots, i\}$, on a

$$\begin{aligned}
\langle u'_{i+1}, u_j \rangle &= \left\langle \Big(\langle x_{i+1} - \sum_{k=1}^{i} \langle x_{i+1}, u_k \rangle u_k \Big), u_j \right\rangle \\
&= \langle x_{i+1}, u_j \rangle - \sum_{k=1}^{i} \langle x_{i+1}, u_k \rangle \langle u_k, u_j \rangle \\
&= \langle x_{i+1}, u_j \rangle - \sum_{k=1}^{i} \langle x_{i+1}, u_k \rangle \delta_{k,j} \\
&= \langle x_{i+1}, u_j \rangle - \langle x_{i+1}, u_j \rangle \\
&= 0.
\end{aligned}$$

Comme $u_{i+1} = \frac{1}{\|u'_{i+1}\|} u'_{i+1}$, on a $\langle u_{i+1}, u_j \rangle = \frac{1}{\|u'_{i+1}\|} \langle u'_{i+1}, u_j \rangle = 0$, ce qui achève la preuve de la récurrence et donc du théorème. □

Comme tout espace vectoriel de dimension finie admet une base finie, on déduit le résultat suivant en appliquant le théorème à une base quelconque.

Corollaire 3.14. *Tout espace euclidien admet des bases orthonormées.*

Projections orthogonales.

Commençons par énoncer le théorème de la projection orthogonale.

Théorème 3.15. *Soit E un espace euclidien, et soit U un sous-e.v. de E. Alors :*

1. *Pour tout vecteur $x \in E$, il existe un unique vecteur $y \in U$ tel que $(x - y) \perp U$.*
   Le vecteur y est appelé projeté orthogonal de x sur U.
2. *L'application qui associe à $x \in E$ son projeté $p(x) \in U$ est un projecteur, d'image U et de noyau* $\ker p = \{z \in E \ : \ z \perp U\}$.
   L'application p est appelée projection orthogonale sur U.

Démonstration.

1. (a) *Unicité.* Choisissons $(e_1, \ldots, e_k)$ une base orthonormée de U, qui existe d'après le corollaire 3.14.
   Supposons que $y \in U$ vérifie $(x - y) \perp U$. Alors en particulier $\langle x - y, e_i \rangle = 0$ et donc $\langle y, e_i \rangle = \langle x, e_i \rangle$ pour tout $i = 1, \ldots, k$. Il s'ensuit, d'après la proposition 3.10 qui s'applique puisque e a été choisie orthonormée, que y se décompose comme suit :
   $$y = \sum_{i=1}^{k} \langle x, e_i \rangle e_i. \tag{3.4}$$
   Si y existe, il est donc nécessairement égal au vecteur du membre de droite ci-dessus. Ceci prouve l'unicité.

   (b) *Existence.* Soit y le vecteur de U défini par (3.4). Vérifions que y convient. Puisque e est orthonormée, il est évident sur l'expression ci-dessus que $\langle x - y, e_i \rangle = 0$ pour tout $i = 1, \ldots, k$. Et comme e engendre U, on en déduit que $(x - y) \perp U$.
2. Il est clair sur la formule (3.4) que p est linéaire et vérifie $p^2 = p$. C'est donc un projecteur, et visiblement $\operatorname{Im} p = U$. Pour son noyau, on a :
   $$\ker p = \{x \in E \ : \ \langle x, e_i \rangle \text{ pour } i = 1, \ldots, k\} = \{x \in E \ : \ x \perp U\},$$
   ce qu'on voulait.

□

Remarque 3.16. On peut montrer que le projeté orthogonal y de x sur U est l'unique vecteur de U qui minimise la distance de x à U :
$$d(x, y) = \min\{d(x, z) \ : \ z \in U\}.$$

Remarque 3.17. Dans le procédé d'orthonormalisation de Gram-Schmidt, on reconnaît l'utilisation de la projection orthogonale. En effet, avec les notations utilisées précédemment, la construction du vecteur u_{i+1} s'interprète comme :

$$u_{i+1} = u'_{i+1}/\|u'_{i+1}\| \qquad\qquad u'_{i+1} = x_{i+1} - p(x_{i+1})$$

où p est la projection orthogonale sur $V_i = \mathrm{Vect}\langle u_1, \ldots, u_i\rangle = \mathrm{Vect}\langle x_1, \ldots, x_i\rangle$.

Supplémentaires orthogonaux.

Définition 3.18. Deux sous-e.v. U et V d'un espace euclidien E sont des *supplémentaires orthogonaux*, ce qu'on note $E = U \oplus^{\perp} V$, si $E = U \oplus V$ et $U \perp V$.

L'*orthogonal* d'un sous-e.v. U de E est le sous-e.v. noté $U^{\perp}$ et défini par :

$$U^{\perp} = \{x \in E \ : \ x \perp U\}.$$

En comparant avec l'énoncé du théorème 3.15, point 2, on voit que $U^{\perp}$ est aussi le noyau de la projection orthogonale sur U.

Théorème 3.19. *Soit U un sous-e.v. d'un espace euclidien E. Alors :*

1. $E = U \oplus^{\perp} U^{\perp}$.
2. $(U^{\perp})^{\perp} = U$.

Démonstration. Soit p la projection orthogonale sur U. Alors $E = U \oplus U^{\perp}$ puisque $\mathrm{Im}\, p = U$ et $\ker p = U^{\perp}$. Comme il est clair que $U \perp U^{\perp}$, on en déduit que U et $U^{\perp}$ sont des supplémentaires orthogonaux, d'où le point 1.

Pour le point 2, observons d'abord que l'inclusion $U \subseteq (U^{\perp})^{\perp}$ est évidente. On en déduit l'égalité pour des raisons de dimension. □

On en déduit une représentation des hyperplans d'un espace euclidien.

Corollaire 3.20. *Soit H un hyperplan d'un espace euclidien. Alors il existe un vecteur $x_0 \in E$ tel que $H = x_0^{\perp}$. Tout autre vecteur x tel que $H = x^{\perp}$ est colinéaire à x_0.*

Démonstration. Soit H un hyperplan de E. Alors $\dim H = n - 1$ si $\dim E = n$, donc $H^{\perp}$ est de dimension 1, donc $H^{\perp} = \mathbb{R}x_0$ pour un certain $x_0 \neq 0$. D'après le théorème 3.19, on a donc $H = (\mathbb{R}x_0)^{\perp} = x_0^{\perp}$.

Supposons que x est un vecteur tel que $H = x^{\perp}$. Alors $H = (\mathbb{R}x)^{\perp} = (\mathbb{R}x_0)^{\perp}$. En appliquant à nouveau le théorème 3.19, on obtient que $\mathbb{R}x = \mathbb{R}x_0$, donc que x et x_0 sont colinéaires. □

Dual d'un espace euclidien.

On a vu que tout $\mathbb{R}$-e.v. de dimension finie est isomorphe à son dual, mais l'isomorphisme n'est défini qu'à travers le choix d'une base. Dans le cas d'un espace euclidien E, il existe une identification naturelle entre E et E^*, qu'on va décrire maintenant.

Ce théorème est important car il donne une *représentation des formes linéaires*.

Théorème 3.21. *Soit E un espace euclidien. À tout vecteur $x \in E$, on associe la forme linéaire $j_E(x)$ définie par $j_E(x)(y) = \langle x, y\rangle$. Alors l'application $j_E : E \to E^*$ est un isomorphisme.*

Autrement dit, pour tout forme linéaire $f \in E^*$ il existe un unique vecteur $x \in E$ tel que :

$$\forall y \in E \qquad f(y) = \langle x, y\rangle.$$

Démonstration. L'application j_E est visiblement linéaire. Si $x \in E$ est tel que $\langle x, y\rangle = 0$ pour tout $y \in E$, alors en particulier $\langle x, x\rangle = 0$ donc $x = 0$. Donc $\ker j_E = \{0\}$, et ceci suffit à montrer que j_E est un isomorphisme puisque $\dim E = \dim E^*$. □

Remarque 3.22. L'identification $j_E : E \to E^*$ est intrinsèque au sens où elle ne fait pas intervenir de choix d'une base. Elle explique qu'on puisse choisir sans confusion la même notation pour le crochet de dualité $\langle x, y\rangle$ avec $(x, y) \in E^* \times E$ dans le cas général, et le produit scalaire $\langle x, y\rangle$ avec $(x, y) \in E \times E$ dans le cas d'un espace euclidien.

3 Produit scalaire pour un espace vectoriel complexe : Espaces Hermitiens

Lorsque qu'on travaille dans un espace vectoriel complexe, il faut modifier la notion de produit scalaire.

Définition 3.23. Soit E un $\mathbb{C}$ espace vectoriel. On se fixe une fonction $P : E \times E \to \mathbb{C}$, et on note $\langle x, y\rangle = P(x, y)$. La fonction $P = \langle\cdot,\cdot\rangle$ est un *produit scalaire (complexe)* sur E si elle vérifie :

1. (Hermitienne) $\langle y, x\rangle = \overline{\langle x, y\rangle}$ pour tous $x, y \in E$.
2. (Sesquilinéarité) $\langle x + \lambda y, z\rangle = \langle x, z\rangle + \overline{\lambda}\langle y, z\rangle$ et $\langle x, z + \lambda y\rangle = \langle x, z\rangle + \lambda\langle x, y\rangle$ pour tous $x, y, z \in E$ et $\lambda \in \mathbb{C}$.
3. (Positivité) $\langle x, x\rangle \geq 0$ et $\langle x, x\rangle = 0 \iff x = 0$ pour tout $x \in E$.

Exemple 3.24. Produit scalaire canonique sur $\mathbb{C}^n$ Pour $x = (x_i)_i \in \mathbb{C}^n$ et $y = (y_i)_i \in \mathbb{C}^n$, l'application définie par :

$$\langle x, y\rangle = \sum_i \overline{x_i} y_i$$

est un produit scalaire sur $\mathbb{C}^n$.

Définition 3.25. On appelle *espace hermitien* un $\mathbb{C}$- espace vectoriel E de dimension finie et muni d'un produit scalaire complexe.

Nous n'étudierons pas en détail les espaces hermitiens dans ce cours, mais l'essentiel des propriétés et définitions (norme, orthogonalité, familles orthogonales et orthonormales, théorème de Pythagore, représentation matricielle, ...) vues pour les espace euclidiens restent valables dans le cas hermitien, à condition d'adapter les formules. Seule la notion d'angle non orienté perd sa signification dans ce cadre.
Par exemple la proposition 3.12 se transforme en

Proposition 3.26. *Soit E un espace hermitien muni d'une base orthonormée $e = (e_1, \ldots, e_n)$. Si $x = \sum_{i=1}^{n} x_i e_i$ et $y = \sum_{i=1}^{n} y_i e_i$, alors :*

$$\langle x, y\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \overline{x_i} y_j \langle e_i, e_j\rangle = \sum_{i=1}^{n} \overline{x_i} i y_i.$$

Remarque 3.27. La démonstration de l'inégalité de Cauchy-Schwarz, qui permet de montrer que l'application $x \mapsto \sqrt{\langle x, x\rangle}$ est également légèrement différente dans le cas complexe.

En voici une démonstration.

Proposition 3.28 (Inégalité de Cauchy-Schwarz)**.** *Soit E un espace hermitien. Alors pour tous vecteurs $x, y \in E$, l'inégalité suivante a lieu :*

$$\left|\langle x, y\rangle\right|^2 \leq \langle x, x\rangle\langle y, y\rangle, \tag{3.5}$$

et l'égalité a lieu si et seulement si la famille (x, y) est liée.

Démonstration. Fixons x et y et posons $\mu = \overline{\langle x, y\rangle}$. On considère maintenant, pour $\lambda \in \mathbb{R}$, l'expression $P(\lambda) = \langle \lambda x + \mu y, \lambda x + \mu y\rangle$. On sait que c'est une quantité positive et en calculant, on obtient :

$$P(\lambda) = \langle \lambda x + \mu y, \lambda x + \mu y\rangle = \lambda^2\langle x, x\rangle + \lambda\mu\langle x, y\rangle + \lambda\overline{\mu}\langle y, x\rangle + \overline{\mu}\mu\langle y, y\rangle.$$

Par définition de μ, on obtient ainsi

$$P(\lambda) = \lambda^2\langle x, x\rangle + 2\lambda\left|\mu\right|^2 + \left|\mu\right|^2\langle y, y\rangle$$

et en utilisant la même technique que dans le cas euclidien, on voit que cela implique l'inégalité $\Delta = \left|\mu\right|^4 - \left|\mu\right|^2\langle x, x\rangle\langle y, y\rangle \leq 0$ qui implique immédiatement l'inégalité recherchée :

$$\left|\mu\right|^2 \leq \langle x, x\rangle\langle y, y\rangle. \qquad \square$$

Chapitre 4

Endomorphismes des espaces euclidiens

1 Représentation matricielle d'un endomorphisme dans une base orthonormée

Proposition 4.1.
Soit E un espace euclidien muni d'une base $e = (e_1, \dots, e_n)$ orthonormée.
Si $M = (M_{i,j})$ est la matrice dans la base e d'un endomorphisme $u \in \mathscr{L}(E)$, alors :

$$M_{i,j} = \langle e_i, u(e_j) \rangle. \tag{4.1}$$

Démonstration. La relation (4.1) découle de la proposition 3.10 car le j^{e} vecteur colonne $(M_{i,j})_i$ collecte les coordonnées dans la base e de l'image du vecteur de base $u(e_j)$. □

2 Applications adjointes.

La notion d'application adjointe est liée à celle d'application transposée, en utilisant l'isomorphisme naturel entre un espace vectoriel et son dual dans un espace euclidien.

Proposition 4.2. *Soit E, F deux espaces euclidiens, et soit $u : E \longrightarrow F$ une application linéaire. Alors il existe une unique application linéaire $v : F \longrightarrow E$ ayant la propriété suivante :*

$$\forall x \in E \quad \forall y \in F \qquad \langle u(x), y \rangle = \langle x, v(y) \rangle. \tag{4.2}$$

Démonstration. Fixons $y \in F$. Alors $\ell_y(x) = \langle u(x), y \rangle$ est clairement une forme linéaire sur E. D'après le théorème de représentation des formes linéaires (théorème 3.21) il existe un unique vecteur, notons-le $v(y) \in E$, tel que $\ell_y = \langle v(y), \cdot \rangle$.

Ceci montre l'unicité de v en tant qu'application $v : F \longrightarrow E$ avec la propriété (4.2). Il reste à voir que v est linéaire. Pour $y, y' \in F$ et $\lambda \in \mathbb{R}$, on a :

$$\forall x \in E \quad \langle u(x), y + \lambda y' \rangle = \langle u(x), y \rangle + \lambda \langle u(x), y' \rangle = \langle x, v(y) + \lambda v(y') \rangle.$$

D'après l'unicité déjà démontrée, on a donc $v(y + \lambda y') = v(y) + \lambda v(y')$. □

Définition 4.3. L'application $v : F \longrightarrow E$ de la proposition 4.2 est appelé *application adjointe de u*. On la note $v = u^*$.

Remarque 4.4. La notation entre l'adjoint de u est la même que pour la transposée vue en 5. En effet, dans le cas d'un espace euclidien E, les deux se correspondent *via* les identifications intrinsèques $j_E : E \longrightarrow E^*$ et $j_F : F \longrightarrow F^*$ du théorème 3.21, comme illustré dans le diagramme ci-dessous.

$$\begin{tikzcd} E \arrow[d, "j_E"'] & F \arrow[l, "\text{adjoint de } u"'] \arrow[d, "j_F"] \\ E^* & F^* \arrow[l, "\text{transposée de } u"] \end{tikzcd}$$

3 Représentation matricielle de l'adjoint.

On se place dans le cadre des endomorphismes, bien qu'une généralisation aux applications linéaires entre espaces euclidiens ne présente pas de difficulté supplémentaire.

Proposition 4.5. *Soit E un espace euclidien muni d'une base orthonormée e. Soit $u \in \mathscr{L}(E)$, d'adjoint $u^* \in \mathscr{L}(E)$, et soit $M = \mathrm{Mat}(u, e)$ et $M' = \mathrm{Mat}(u^*, e)$. Alors M' est la transposée de M.*

Démonstration. Puisque e est orthonormée on a :

$$\begin{aligned} M'_{i,j} &= \langle u^*(e_j), e_i \rangle && \text{d'après la proposition 4.1} \\ &= \langle e_j, u(e_i) \rangle && \text{par définition de } u^* \\ &= M_{j,i} && \text{d'après la proposition 4.1.} \end{aligned}$$

Ceci montre que M et M' sont transposées l'une de l'autre.

On peut aussi donner une autre démonstration, laissée en exercice, basée sur l'identification de u^* avec la transposée de u, et en appliquant la proposition 2.19 et le deuxième point du théorème 3.21. □

Remarque 4.6. Le résultat n'est pas vrai en général si e n'est pas orthonormée, de même que le résultat de la proposition 2.19 n'est vrai que pour des couples de bases duales, ce qui correspond bien à considérer une base orthonormée comme on l'a remarqué au théorème 3.21.

4 Endomorphismes auto-adjoints.

Définition 4.7. Soit E un espace euclidien. Un endomorphisme $u \in \mathscr{L}(E)$ est *auto-adjoint* si $u = u^*$.

Deux propriétés importantes des endomorphismes auto-adjoints sont les suivantes.

Proposition 4.8. *Soit E un espace euclidien. Alors :*

1. *Pour tout $u \in \mathscr{L}(E)$, et pour toute base orthonormée e de E, u est auto-adjoint si et seulement si* $\mathrm{Mat}(u, e)$ *est symétrique.*
2. *Soit $u \in \mathscr{L}(E)$ auto-adjoint, et soit F un sous-e.v. de E. Si F est stable par u, alors $F^\perp$ est aussi stable par u.*

Démonstration. La première partie de l'énoncé découle directement de la proposition 4.5.

Soit maintenant $u \in \mathscr{L}(E)$ auto-adjoint, et soit F stable par u. Pour montrer que $F^\perp$ est stable par u, soit $x \in F^\perp$, montrons que $u(x) \in F^\perp$. Pour cela, considérons $y \in F$ arbitraire, et calculons :

$$\langle u(x), y\rangle = \langle x, u(y)\rangle = 0$$

puisque $x \in F^\perp$ et $u(y) \in F$, d'où le résultat. □

Remarque 4.9. Considérons une matrice carrée symétrique M de taille $n \times n$. Alors, en munissant $\mathbb{R}^n$ de son produit scalaire usuel, la matrice M représente un endomorphisme auto-adjoint.

5 Isométries dans l'espace euclidien

Définitions des isométries.

Définition 4.10. Soit E un espace euclidien. Un endomorphisme $f \in \mathscr{L}(E)$ est une *isométrie* si $\|f(x)\| = \|x\|$ pour tout $x \in E$.

Autrement dit, les isométries de E sont les endomorphismes qui conservent la norme. Plusieurs propriétés équivalentes caractérisent les isométries.

Proposition 4.11. *Soit E un espace euclidien, et soit $f \in \mathscr{L}(E)$. Alors les propriétés suivantes sont équivalentes :*

1. *f est une isométrie.*
2. *f conserve le produit scalaire, c'est-à-dire :*
$$\forall x, y \in E \qquad \langle f(x), f(y)\rangle = \langle x, y\rangle. \tag{4.3}$$
3. *f envoie une base orthonormée de E sur une base orthonormée de E.*
4. *f envoie toute base orthonormée de E sur une base orthonormée de E.*

Démonstration. $1 \Longrightarrow 2$. D'après les identités remarquables, on a :

$$\langle x, y\rangle = \frac{\|x+y\|^2 - \|x-y\|^2}{4} \qquad \langle f(x), f(y)\rangle = \frac{\|f(x+y)\|^2 - \|f(x-y)\|^2}{4}$$

d'où $\langle f(x), f(y)\rangle = \langle x, y\rangle$.

$2 \Longrightarrow 3$. Si $e = (e_1, \ldots, e_n)$ est orthonormée, alors $\langle f(e_i), f(e_j)\rangle = \langle e_i, e_j\rangle = \delta_i^j$ donc $f(e)$ est orthonormée.

$3 \Longrightarrow 1$. Soit e une base orthonormée telle que $f(e)$ est aussi orthonormée. Alors, pour $x \in E$ de décomposition $x = \sum_i x_i e_i$, on a :

$$\|x\|^2 = \sum_i x_i^2 \qquad f(x) = \sum_i x_i f(e_i) \qquad \|f(x)\|^2 = \sum_i x_i^2 = \|x\|^2$$

donc f est une isométrie.

$1 \Longrightarrow 4$. En appliquant le même raisonnement à une base arbitraire qu'à la base e dans la preuve de $2 \Longrightarrow 3$, et puisqu'on sait déjà que f conserve le produit scalaire.

$4 \Longrightarrow 1$. On sait qu'il existe au moins une base orthonormée, on peut donc appliquer $3 \Longrightarrow 1$. □

Groupe des isométries

Les propriétés de composition des isométries découlent presque immédiatement de la définition.

Proposition 4.12. *Soit E un espace euclidien. Alors toute isométrie de E est un isomorphisme de E, et son inverse est aussi une isométrie. La composée de deux isométries de E est encore une isométrie de E.*

Démonstration. Soit f une isométrie de E. Alors $x \in \ker f \implies f(x) = 0 \implies \|x\| = 0 \implies x = 0$, donc f est injective et est donc un isomorphisme. Et $\|x\| = \|f(f^{-1}(x))\| = \|f^{-1}(x)\|$ donc f^{-1} est une isométrie. Finalement, si f et g sont deux isométries de E, alors $\|g \circ f(x)\| = \|f(x)\| = \|x\|$ donc $g \circ f$ est une isométrie. □

Comme il est clair que Id_E est une isométrie, l'ensemble des isométries de E forme un *sous-groupe* de $GL(E)$, puisque stable par composition et passage à l'inverse. On le note $O(E)$.

Exemple : rotation dans $\mathbb{R}^2$.

Soit r la rotation d'angle θ dans $\mathbb{R}^2$, dont la matrice dans la base canonique est :

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

Alors R_θ est une isométrie. Montrons-le directement en montrant que R_θ conserve la norme :

$$\|R_\theta \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)\|^2 = \| \left(\begin{smallmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{smallmatrix}\right) \|^2 = (x\cos\theta - y\sin\theta)^2 + (x\sin\theta + y\cos\theta)^2 = \| \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \|^2$$

Alternativement, examinons l'image par R_θ de la base canonique, c'est-à-dire les vecteurs colonne de la matrice. Ce sont deux vecteurs visiblement normés et orthogonaux, donc R_θ est une isométrie d'après le point 3 de la proposition 4.11.

Représentation matricielle.

Pour les matrices, on introduit une définition supplémentaire.

> **Définition 4.13.** Une matrice carrée M de taille n est *orthogonale* si les colonnes de M forment une famille orthonormée pour le produit scalaire canonique de $\mathbb{R}^n$.

À condition de se placer dans une base orthonormée, les matrices orthogonales caractérisent exactement les isométries comme le montre le résultat suivant.

Théorème 4.14. *Soit E un espace euclidien, et soit e une base orthonormée de E.*

Soit u un endomorphisme de E, et soit $M = \mathrm{Mat}(u, e)$. Alors les propriétés suivantes sont équivalentes :

(i) u est une isométrie.

(ii) M est une matrice orthogonale.

Adjoint d'une isométrie.

Les isométries sont caractérisées par la propriété suivante de leur adjoint.

Théorème 4.15. *Soit E un espace euclidien. Alors un endomorphisme $u \in \mathscr{L}(E)$ est une isométrie si et seulement si $u^* = u^{-1}$.*

Démonstration. Supposons que u est une isométrie. C'est donc un isomorphisme, et u^{-1} conserve le produit scalaire d'après les propositions 4.11 et 4.12. Pour $x, y \in E$, on a donc :

$$\langle u(x), y\rangle = \langle x, u^{-1}(y)\rangle.$$

Or cette relation caractérise l'adjoint, donc $u^* = u^{-1}$.

Réciproquement, soit $u \in \mathscr{L}(E)$ tel que $u^* = u^{-1}$, et montrons que u conserve le produit scalaire.

$$\langle u(x), u(y)\rangle = \langle x, u^* \circ u(y)\rangle = \langle x, y\rangle.$$

Donc u est une isométrie. □

Corollaire 4.16. *Une matrice M est orthogonale si et seulement si ${}^tM \cdot M = I_n$.*

Reprenons l'exemple précédent de la rotation R_θ dans $\mathbb{R}^2$. On vérifie immédiatement que ${}^tR_\theta \cdot R_\theta = I_2$. On en déduit donc que R_θ est la matrice d'une isométrie de $\mathbb{R}^2$. On constate aussi que $R_\theta^{-1} = {}^tR_\theta = R_{-\theta}$.

Valeurs propres et déterminant des isométries.

Proposition 4.17. *Soit E un espace euclidien, et soit $u \in \mathscr{L}(E)$ une isométrie de E. Alors les seules valeurs propres possibles de u sont ± 1, et $\det u = \pm 1$.*

Démonstration. Soit λ une valeur propre de u, et soit x un vecteur propre de u associé à λ. Alors $\|u(x)\| = |\lambda| \cdot \|x\| = \|x\|$ donc $|\lambda| = 1$ c'est-à-dire $\lambda = \pm 1$.

Soit e une base orthonormée de E, et soit $M = \mathrm{Mat}(u, e)$. Alors M est orthogonale, et puisque $\det {}^tM = \det M$ on en déduit $(\det M)^2 = 1$, et donc $\det u = \det M = \pm 1$. □

Remarque 4.18. Les vecteurs propres associés à la valeur propre 1 sont appelés *points fixes* de l'isométrie.

Exemple : symétries orthogonales.

Définition 4.19. Soit E un espace euclidien, et soit F un sous-e.v. de E. Considérons la projection orthogonale sur F, notons-la p. La *symétrie orthogonale* par rapport à F est l'endomorphisme $s : E \to E$ défini par :

$$s(x) = 2p(x) - x.$$

Si F est un hyperplan de E, alors s est appelée une *réflexion*.

Soit s une symétrie orthogonale. Alors s est une isométrie. En effet, comme $p(x)-x \perp F$ et $p(x) \in F$, on a d'après le théorème de Pythagore : $\|s(x)\|^2 = \|p(x)\|^2 + \|x-p(x)\|^2 = \|x\|^2$, donc s est une isométrie.

Les points fixes de s sont exactement les éléments de F. En effet, si $s(x) = x$, alors $p(x) = x$ et donc $x \in F$. Réciproquement, si $x \in F$, alors $p(x) = x$ donc $s(x) = x$.

Par ailleurs, s est une involution, c'est-à-dire $s^2 = \mathrm{Id}$. En effet, puisque $x - p(x) \perp F$ d'une part, et puisque $s(x) = p(x) + \big(p(x) - x\big)$, on en déduit que $s\big(p(x)\big) = p(x)$ donc $s^2(x) = 2p(x) - s(x) = x$.

On en déduit que s est à la fois auto-adjoint et isométrique. En effet, on a $s^{-1} = s^*$ puisque s est une isométrie, et $s^{-1} = s$ puisque $s^2 = \mathrm{Id}$, d'où $s^* = s$.

Supposons que $F \neq \{0\}$, alors s a au moins un point fixe, donc 1 est valeur propre de s. Et si $F \neq E$ alors $F^\perp \neq \{0\}$ donc pour au moins un vecteur $x \neq 0$ on a $s(x) = -x$ donc -1 est valeur propre de s. Donc si $\{0\} \subsetneq F \subsetneq E$, la symétrie orthogonale s a pour valeurs propres les deux seules valeurs possibles 1 et -1.

Considérons en particulier le cas où F est un hyperplan de E. Alors $F = x_0^\perp$ pour un certain vecteur x_0 qu'on peut choisir normé. Considérons $(x_1, \dots, x_{n-1})$ une base orthonormée de F, la matrice de la symétrie orthogonale par rapport à F a la forme suivante :

$$M = \begin{pmatrix} -1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & I_{n-1} & \\ 0 & & & \end{pmatrix}$$

Donc $\det s = \det M = -1$.

Décomposition des isométries comme produits de réflexions.

Nous avons vu que les symétries orthogonales sont des exemples d'isométries, et évidemment les produits de symétries orthogonales restent donc des isométries. En fait, on peut toujours obtenir une isométrie comme produit de symétries orthogonales, comme le précise le résultat élémentaire ci-dessous.

Proposition 4.20. *Soit E un espace euclidien de dimension n, et soit f une isométrie de E. Alors f peut s'écrire comme un produit $f = s_k \circ \cdots \circ s_1$ de réflexions, avec de plus $k \leq n$.*

Énonçons et démontrons d'abord deux lemmes.

Lemme 4.21. *Soit E un espace euclidien, et soit x_1 et x_2 deux vecteurs distincts et de même norme. Alors la réflexion par rapport à l'hyperplan $H = (x_1 - x_2)^\perp$ échange x_1 et x_2.*

Démonstration. Comme $x_1 - x_2 \neq 0$, il est clair que H est bien un hyperplan. Soit s la symétrie par rapport à H, et soit p la projection orthogonale sur H.

Montrons que $x_1 + x_2 \in H$. En effet, d'après les identités remarquables, on a $\langle x_1 + x_2, x_1 - x_2\rangle = \|x_1\|^2 - \|x_2\|^2 = 0$. D'après la décomposition :

$$x_1 = \underbrace{\frac{x_1 + x_2}{2}}_{\in H} + \underbrace{\frac{x_1 - x_2}{2}}_{\in H^\perp}$$

on déduit que $p(x_1) = (x_1 + x_2)/2$, et donc $s(x_1) = 2p(x_1) - x_1 = x_2$. Comme s est une involution, il s'ensuit aussi que $s(x_2) = x_1$. □

Lemme 4.22. *Soit E un espace euclidien, et soit $f \in \mathscr{L}(E)$ une isométrie. Si F est un sous-e.v. de E stable par f, alors $F^\perp$ est aussi stable par f.*

Démonstration. Soit $x \in F^\perp$, montrons que $f(x) \in F^\perp$. Pour tout $y \in F$, on a $\langle f(x), y\rangle = \langle x, f^{-1}(y)\rangle$. Or $f(F) \subseteq F$ et f est une isométrie donc en particulier un isomorphisme, donc $f(F) = F$ et $f^{-1}(F) = F$, d'où $f^{-1}(y) \in F$ et donc $\langle f(x), y\rangle = 0$, ce qu'on voulait montrer. □

Démonstration de la proposition 4.20. La démonstration est par récurrence sur la dimension de E. Si $n = 1$, le résultat est clair car $f(x) = \det(f)x$, avec $\det(f) = \pm 1$ donc f est ou bien l'identité ou bien la symétrie orthogonale par rapport à $\{0\}$, donnée par $f(x) = -x$.

Supposons le résultat établi pour tous les $\mathbb{R}$-e.v. de dimension au plus $n-1$ avec $n \geq 1$. Soit E de dimension n, et soit f une isométrie de E. Choisissons $x_1 \in E$, et posons $x_2 = f(x_1)$. Alors x_1 et x_2 ont même norme, donc d'après le lemme 4.21, on choisit une symétrie orthogonale par rapport à un hyperplan échangeant x_1 et x_2. Soit s cette symétrie orthogonale.

Posons $g = s \circ f$. Alors $g(x_1) = s(x_2) = x_1$, donc x_1 est un point fixe de g, et donc $\mathbb{R}x_1$ est stable par g. D'après le lemme 4.22, il s'ensuit que $(\mathbb{R}x_1)^\perp$ est aussi stable par g. Alors sa restriction g' à $(\mathbb{R}x_1)^\perp$ est une isométrie produit d'au plus $n-1$ symétries orthogonales par rapport à des hyperplans de $(\mathbb{R}x_1)^\perp$, $g' = s'_k \circ \cdots \circ s'_1$ avec $k \leq n-1$. Disons que s'_i est la symétrie orthogonale par rapport à F'_i. Considérons la symétrie orthogonale s_i de E par rapport à $F_i = F'_i \oplus \mathbb{R}x_1$, qui est un hyperplan de E. Alors s_i prolonge s'_i à E tout entier, et laisse $\mathbb{R}x_1$ fixe. On en déduit que $g = s_k \circ \cdots \circ s_1$, et donc finalement le résultat cherché pour $f = s \circ s_k \circ \cdots \circ s_1$. □

Remarque 4.23. Le nombre de symétries orthogonales par rapport à des hyperplans telles que $f = s_k \circ \cdots \circ s_1$ n'est pas unique, de même que les s_i ne sont pas uniques. Cependant la *parité* de k ne dépend que de f, puisque $\det f = \det s_k \cdot \ldots \cdot \det s_1 = (-1)^k$.

6 Réduction des endomorphismes auto-adjoints.

Le but de cette section est de montrer que tout endomorphisme auto-adjoint d'un espace euclidien est diagonalisable. On en déduira donc le fait remarquable que tout matrice symétrique réelle est diagonalisable sur $\mathbb{R}$.

Lemme 4.24. *Soit E un espace euclidien de dimension n, et soit $u \in \mathscr{L}(E)$ auto-adjoint. Alors u possède n valeurs propres réelles (non nécessairement distinctes).*

Pour démontrer le lemme, on a besoin de se placer dans un espace vectoriel où on est sûr que u admet une valeur propre. C'est le cas si on regarde E comme un espace vectoriel complexe, plutôt que réel.

On peut considérer la matrice M de u dans une base orthonormée de E. D'après ce qui précède, il s'agit d'une matrice symétrique réelle et on va montrer que ses valeurs propres, considérée comme une matrice complexe, sont toutes réelles. On a besoin pour cela d'utiliser la structure de produit scalaire sur un espace vectoriel complexe, introduite à la fin du chapitre précédent et rappelée ci-dessous.

Utilisons le produit scalaire usuel de de $\mathbb{C}^n$ qui est donné, pour $X, Y \in \mathbb{C}^n$ par :

$$\langle X, Y\rangle = \sum_i \overline{X_i} Y_i = {}^t\overline{X} \cdot Y. \tag{4.4}$$

Ce produit scalaire complexe est bien une forme sesquilinéaire hermitienne définie positive car il vérifie les propriétés suivantes :

1. $\langle X, Y\rangle = \overline{\langle Y, X\rangle}$.
2. $\langle X, \lambda Y + Z\rangle = \lambda\langle X, Y\rangle + \langle X, Z\rangle$ et $\langle \lambda X + Z, Y\rangle = \overline{\lambda}\langle X, Y\rangle + \langle Z, Y\rangle$ pour tout $\lambda \in \mathbb{C}$ et $X, Y \in \mathbb{C}^n$.
3. $\langle X, X\rangle \in \mathbb{R}$ pour tout $X \in \mathbb{C}^n$ et $\langle X, X\rangle > 0$ sauf si $X = 0$.

Dans le cadre complexe, la matrice adjointe est donnée comme suit :

Proposition 4.25. *Soit $A \in M_n(\mathbb{C})$ alors il existe une unique matrice $A^* \in M_n(\mathbb{C})$ telle que $\langle A^*X, Y\rangle = \langle X, AY\rangle$ pour tous $X, Y \in \mathbb{C}^n$, et on a $A^* = {}^t\overline{A}$.*

En effet, en utilisant l'expression $\langle X, Y\rangle = {}^t\overline{X} \cdot Y$ on obtient :

$$\begin{aligned}
\langle {}^t\overline{A}X, Y\rangle &= {}^t(\overline{{}^t\overline{A}X}) \cdot Y \\
&= {}^t\overline{X} \cdot A \cdot Y \\
&= \langle X, A \cdot Y\rangle
\end{aligned}$$

Démonstration du lemme. Comme expliqué ci-dessus, il revient au même de se fixer une matrice symétrique réelle M, de considérer M comme une matrice à coefficients complexes et de montrer que toute valeur propre de M, *a priori* complexe, est en fait réelle. Soit donc $\lambda \in \mathbb{C}$ une valeur propre de M.
Choisissons Z un vecteur de $\mathbb{C}^n$, vecteur propre de M pour la valeur propre λ.
Comme M est une matrice symétrique réelle, on a $M^* = M$, ce qui implique que $\langle MX, Y\rangle = \langle X, MY\rangle$ pour tous $X, Y \in \mathbb{C}^n$.
Calculons $\langle MZ, Z\rangle$ de deux façons. D'une part, puisque $MZ = \lambda Z$, on a $\langle MZ, Z\rangle = \langle \lambda Z, Z\rangle = \overline{\lambda}\langle Z, Z\rangle$, en utilisant le point 2 ci-dessus. Mais d'après la proposition 4.25 ci-dessus, on a aussi $\langle MZ, Z\rangle = \langle Z, MZ\rangle = \langle Z, \lambda Z\rangle = \lambda\langle Z, Z\rangle$ en utilisant le point 2 ci-dessus. Comme $\langle Z, Z\rangle \neq 0$ d'après le point 3 ci-dessus, on en déduit $\lambda = \overline{\lambda}$, donc $\lambda \in \mathbb{R}$, ce qu'on voulait démontrer. □

Théorème 4.26 (diagonalisation des endomorphismes auto-adjoints)**.** *Soit E un espace euclidien et soit $u \in \mathscr{L}(E)$. Si u est auto-adjoint, alors u est diagonalisable dans une base orthonormée.*

En particulier, les sous-espace propres de u sont deux à deux orthogonaux.

Remarque 4.27. Le théorème dit *plus* que la seule diagonalisation de u, puisqu'il affirme que u est diagonalisable *dans une base orthonormée*. Autrement dit, l'espace E possède une base orthonormée de vecteurs propres pour u.

Démonstration. Démontrons d'abord que les sous-espaces propres de u sont deux à deux orthogonaux. Soit donc F et F' deux sous-espaces propres distincts, donc associés à des valeurs propres λ et λ' avec $\lambda \neq \lambda'$. Alors, pour $x \in F$ et $x' \in F'$, on a :

$$\langle u(x), y\rangle = \lambda\langle x, y\rangle \qquad \text{et d'autre part } \langle u(x), y\rangle = \langle x, u(y)\rangle = \lambda'\langle x, y\rangle.$$

Donc $\lambda\langle x, y\rangle = \lambda'\langle x, y\rangle$, d'où $\langle x, y\rangle = 0$ puisque $\lambda \neq \lambda'$.

Soit maintenant $F_1, \ldots, F_k$ tous les espaces propres de u, associés à toutes les valeurs propres $\lambda_1, \ldots, \lambda_k$, et soit F leur somme directe, qui est de plus orthogonale d'après ce qu'on vient de voir : $F = F_1 \oplus^\perp \cdots \oplus^\perp F_k$. Pour montrer que u est diagonalisable, il suffit de montrer que $E = F$. Supposons par l'absurde que $F \neq E$. Alors, comme F est évidemment stable par u, il découle de la proposition 4.8, point 2, que $F^\perp$ est stable par u.

Soit v la restriction de u à $G = F^\perp$. C'est donc un endomorphisme auto-adjoint de G. D'après le lemme 4.24, cet endomorphisme possède une valeur propre réelle, et donc un vecteur propre dans G, qui est aussi un vecteur propre de u, contredisant la définition de F.

Donc $F_1 \oplus^\perp \cdots \oplus^\perp F_k = E$ et u est bien diagonalisable. En choisissant une base orthonormée de chaque sous-espace propre $F_1, \ldots, F_k$, et en réunissant les bases orthonormées ainsi obtenues, on obtient une base orthonormée de E formée de vecteurs propres de u. □

Énonçons la version matricielle du théorème.

Corollaire 4.28. *Toute matrice symétrique réelle est diagonalisable et admet une base de vecteurs propres qui est orthonormée pour le produit scalaire usuel de $\mathbb{R}^n$.*

En particulier la matrice de passage de la base canonique de $\mathbb{R}^n$ à la base diagonalisation est une matrice P orthogonale : on a $P^{-1} = {}^t P$ et ${}^t P P = \mathrm{I}_n$.

Chapitre 5

Isométries et orientation des espaces euclidiens en dimensions 2 et 3

1 Orientation du plan et classification des isométries du plan

Appliquons la proposition 4.20 à $\mathbb{R}^2$ muni du produit scalaire usuel. Les hyperplans de $\mathbb{R}^2$ sont les droites vectorielles. Les isométries de $\mathbb{R}^2$ sont donc le produit de $k = 0$, 1 ou 2 symétries orthogonales par rapport à des droites vectorielles de $\mathbb{R}^2$. Le cas $k = 0$ correspond à l'identité, le cas $k = 1$ correspond aux symétries orthogonales par rapport à des droites vectorielles.

Il reste à examiner le cas $k = 2$. On va voir qu'on obtient dans ce cas les rotations vectorielles. Introduisons d'abord une définition.

Définition 5.1. Soit E un plan euclidien. Une *rotation* est une isométrie dont la matrice dans au moins une base orthonormée de E est de la forme :

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

pour un certain réel θ.

Il faut remarquer que le réel θ peut dépendre *a priori* de la base orthonormée qu'on a choisie. Il se pourrait également que la matrice d'une isométrie ait une forme M_θ dans une certaine base orthonormée, mais pas dans une autre. On verra qu'il n'en est rien, mais cependant le réel θ peut changer de signe avec la base orthonormée qu'on choisit.

Lemme 5.2. *Soit E un plan euclidien et soit e une base orthonormée de e. Pour tout réel θ, soit r_θ l'endomorphisme de E tel que $M_\theta = \mathrm{Mat}(r_\theta, e)$. Alors :*

1. *r_θ est une isométrie et $\det r_\theta = 1$.*
2. *Pour tous $\theta, \theta' \in \mathbb{R}$, on a $M_\theta \cdot M_{\theta'} = M_{\theta+\theta'}$.*

Démonstration. Le point 1 est évident, et le point 2 découle des formules d'addition du sinus et du cosinus. □

Théorème 5.3. *Soit E un plan euclidien, et soit f une isométrie de E. Alors :*

1. *Si $\det f = 1$, alors f est une rotation et sa matrice dans toute base orthonormée est de la forme M_θ pour un certain réel θ (qui dépend* a priori *de la base orthonormée choisie).*
2. *Si $\det f = -1$, alors f est une symétrie orthogonale par rapport à une droite de E.*

Démonstration. Soit $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ la matrice de f dans une base orthonormée. Les vecteurs colonne de M forment une base orthonormée, donc les coefficients vérifient :

$$a^2 + c^2 = 1 \qquad b^2 + d^2 = 1 \qquad ab + cd = 0.$$

Il existe donc $\alpha \in [0, 2\pi[$ et $\theta \in [0, 2\pi[$ tels que $a = \cos\theta$, $c = \sin\theta$, $b = \cos\alpha$ et $d = \sin\alpha$. La dernière équation devient $\cos\theta\cos\alpha + \sin\theta\sin\alpha = 0$, soit $\cos(\theta - \alpha) = 0$, donc $\theta = \alpha \pm \pi/2 \mod (2\pi)$. Examinons les deux cas possibles :

$$\text{Si } \alpha = \theta + \pi/2 \mod (2\pi) : M = \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right) \qquad \text{et alors } \det M = 1$$
$$\text{Si } \alpha = \theta - \pi/2 \mod (2\pi) : M = \left(\begin{smallmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{smallmatrix}\right) \qquad \text{et alors } \det M = -1$$

Dans le premier cas, f est une rotation. Dans le second cas, on pourrait vérifier à la main que M est la matrice d'une symétrie orthogonale, en déterminant ses vecteurs propres. Mais on peut conclure plus rapidement puisqu'on sait que f est le produit de $k \leq 2$ symétries orthogonales, et que le cas $k = 2$ est incompatible avec $\det M = -1$. Donc f est nécessairement une symétrie orthogonale. □

Remarque 5.4. La démonstration précédente montre qu'une rotation a toujours la forme $\left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$ pour un certain réel θ dans toute base orthonormée. Il reste à déterminer comment θ varie lorsqu'on change de base orthonormée. On le comprend mieux en introduisant la notion d'orientation du plan.

Orientations du plan euclidien.

Soit e et e' deux bases orthonormées d'un plan euclidien E. Il existe alors une unique application linéaire, disons f, qui envoie e et sur e', et alors f est une isométrie d'après la proposition 4.11. On dit que la paire (e, e') est *de la même orientation* si $\det f > 0$, c'est-à-dire si $\det f = 1$. On le notera $e\mathbb{R}e'$.

Cette relation entre les paires de bases orthonormées définit une relation d'équivalences, c'est-à-dire : $e\mathbb{R}e$ (réflexivité), $e\mathbb{R}e' \iff e'\mathbb{R}e$ (symétrie), et $(e\mathbb{R}e' \wedge e'\mathbb{R}e'') \implies e\mathbb{R}e''$ (transitivité).

Il y a exactement deux classes d'équivalence de bases orthonormées de E. En effet, si $e_0 = (u, v)$ est une base orthonormée, alors $e_1 = (v, u)$ est une base orthonormée qui n'est pas dans la même classe que e_0, et toute base orthonormée est soit dans la classe de e_0 soit dans celle de e_1.

Par définition, choisir une orientation, c'est se fixer une de ces deux classes d'équivalence. On le fait en choisissant une base orthonormée qu'on décrète d'orientation *directe*. Toutes les bases orthonormées de la même classe sont dites directe, toutes celles de l'autre classe sont dite indirectes.

Isométries directes et indirectes du plan euclidien.

Rappelons que le déterminant d'un endomorphisme $f \in \mathscr{L}(E)$ d'un $\mathbb{R}$-e.v. E de dimension finie peut se calculer comme $\det \operatorname{Mat}(f, e)$, où e est une base quelconque de E.

Définition 5.5. Une isométrie f d'un plan euclidien E est dite *directe* (ou *positive*) si $\det f = 1$, et *indirecte* si $\det f = -1$.

Lemme 5.6. *Soit f une isométrie d'un plan euclidien E, et soit e une base orthonormée de E. Alors sont équivalentes :*

1. *e et $f(e)$ ont la même orientation.*

2. *Pour toute base orthonormée e' de E, e' et $f(e')$ ont la même orientation.*
3. $\det f = 1$.

Démonstration. On a $e\mathcal{R}f(e) \iff \det f = 1$. Le même raisonnement s'applique pour toute autre base orthonormée. □

On en déduit la forme des matrices des rotations du plan euclidien.

Proposition 5.7. *Soit E un plan euclidien orienté, et soit r une rotation de E. Alors il existe un unique réel $\theta \in]-\pi, \pi]$ tel que :*

1. *Dans toute base orthonormée directe e,* $\mathrm{Mat}(r, e) = M_\theta$.
2. *Dans toute base orthonormée indirecte e,* $\mathrm{Mat}(r, e) = M_{-\theta}$.

Démonstration. Soit e une base orthonormée directe, alors d'après le point 1 du théorème 5.3, il existe un réel $\theta \in]-\pi, \pi]$ tel que $\mathrm{Mat}(r, e) = M_\theta$, et ce réel est évidemment unique.

Soit maintenant e' une base orthonormée directe arbitraire, et soit P la matrice de passage de e à e'. La matrice P exprime les coordonnées de e' dans la base e, c'est donc elle-même la matrice d'une isométrie, qui est de plus directe puisque $e\mathcal{R}e'$ par hypothèse. P est donc de la forme $P = M_\alpha$ et donc $P^{-1} = M_{-\alpha}$. Posons $M' = \mathrm{Mat}(r, e')$. Appliquons la formule de changement de base en utilisant le lemme 5.2, on obtient :

$$M' = M_{-\alpha} \cdot M_\theta \cdot M_\alpha = M_\theta,$$

ce qu'on voulait montrer.

Soit maintenant e_1 la base indirecte obtenue en échangeant les deux vecteurs de base de e. Il est clair que $\mathrm{Mat}(r, e_1) = {}^tM_\theta = M_{-\theta}$. Appliquons le résultat précédent à l'orientation opposée du plan, on obtient que la matrice de r dans toutes base indirecte est $M_{-\theta}$. □

Définition 5.8. Soit E un plan euclidien *orienté*. On appelle *angle* d'une rotation r l'unique réel θ défini modulo 2π tel que la matrice de r dans une base orthonormée directe soit M_θ.

Remarque 5.9. D'après la proposition 5.7, l'angle de la rotation r est changé en son opposé lorsqu'on change l'orientation du plan. L'angle de r est donc conservé par changement d'orientation si et seulement si $r = \pm \mathrm{Id}$.

Mesure des angles orientés et non orientés entre deux vecteurs.

On laisse la démonstration du résultat suivant en exercice.

Lemme 5.10. *Soit E un plan euclidien, et soit x et y deux vecteurs normés. Alors il existe une unique rotation r de E telle que $r(x) = y$.*

On peut donc introduire la définition suivante. Remarquez l'importance de l'orientation du plan dans cette définition.

Définition 5.11. Soit E un plan euclidien orienté, et soit x et y deux vecteurs non nuls. La *mesure de l'angle orienté* $\widehat{(x,y)}$ est le réel défini modulo 2π qui est l'angle dans une base orthonormée directe de l'unique rotation envoyant $x/\|x\|$ sur $y/\|y\|$.

Faisons le lien avec la notion d'angle non orienté entre deux vecteurs vue au chapitre 3. On l'a défini comme l'unique réel $\alpha \in [0,\pi]$ tel que

$$\cos(\alpha) = \frac{\langle x, y\rangle}{\|x\| \cdot \|y\|}.$$

Supposons que E est un espace euclidien, sans orientation définie, et fixons une orientation arbitraire. Si l'on choisit la mesure de l'angle $\widehat{(x,y)}$ comme étant représenté par un nombre θ compris dans l'intervalle $]-\pi,\pi]$, alors l'angle non orienté formé par x et y est donné par $\alpha = |\theta|$. Evidemment, si l'on change d'orientation, on change θ en $-\theta$ et ainsi l'angle non orienté entre x et y ne change pas.

On peut également utiliser la définition précédente pour définir la notion d'angle orienté entre deux droites vectorielles :

Définition 5.12. Soit E un plan euclidien orienté , D et D' deux droites et soit $x \in D$ et $y \in D'$ deux vecteurs non nuls. La *mesure de l'angle orienté* entre D et D' est $\widehat{(D,D')} = \widehat{(x,y)} \mod \pi$.

On vérifie que la définition ne dépend pas du choix de x et y

$$\widehat{(x,-y)} = \widehat{(x,y)} + \pi \mod 2\pi \text{ et } \widehat{(-x,y)} = \pi + \widehat{(x,y)} \mod 2\pi$$

On en déduit l'égalité suivante modulo π :

$$\widehat{(x,-y)} \mod \pi = \widehat{(-x,y)} \mod \pi = \widehat{(-x,-y)} \mod \pi = \widehat{(x,y)} \mod \pi.$$

2 Produit vectoriel dans l'espace euclidien en dimension 3

Orientation.

Soit E un espace euclidien de dimension 3. Comme dans le cas du plan, on considère la relation d'équivalence entre bases orthonormées de E définie par $e\mathbb{R}e' \iff \det(f) = 1$, où f est l'unique isométrie envoyant e sur e'.

On choisit une des deux classes qu'on décrète être la classe des bases *directes*, les autres sont *indirectes*.

Une isométrie est dite *directe* si son déterminant vaut 1, négative si son déterminant vaut -1. Cette notion est indépendante de l'orientation choisie.

Produit mixte

Soit E un espace euclidien de dimension 3. Considérons une base orthonormée e de E. On peut calculer le déterminant d'un système (u, v, w) de vecteurs dans la base e, noté $det_e(u, v, w)$. Cette valeur dépend *a priori* de la base e, puisque :

$$\det_e(x, y, z) = \left(\det \mathrm{P}(e, e')\right) \cdot \det_{e'}(x, y, z).$$

Cependant, si e et e' ont la même orientation, alors $\det \mathrm{P}(e, e') = 1$. On introduit donc la définition suivante.

Définition 5.13. Soit E un espace euclidien de dimension 3, muni d'une orientation. le *produit mixte* d'un système (x, y, z) de trois vecteurs est le déterminant $\det_e(x, y, z)$, calculé dans une base orthonormée directe quelconque. On le note $[x, y, z]$.

Produit vectoriel.

Soit E un espace euclidien de dimension 3 orienté et muni d'une base orthonormée directe e. Soit u et v deux vecteurs de E. Alors l'application $x \in E \mapsto \varphi(x) = [u, v, x]$ est une forme linéaire. D'après le théorème de représentation 3.21, il existe un unique vecteur w de E tel que $\varphi(x) = \langle w, x \rangle$ pour tout $x \in E$.

Définition 5.14 (définitive). Soit E un espace euclidien de dimension 3 orienté. Soit u et v deux vecteurs de E. Le *produit vectoriel de u et v*, noté $u \wedge v$, est l'unique vecteur w de E vérifiant la propriété suivante :

$$\forall x \in E \quad [u, v, x] = \langle u \wedge v, x \rangle. \tag{5.1}$$

Le théorème suivant regroupe quelques propriétés à connaître du produit vectoriel.

Théorème 5.15. *Soit E un espace euclidien de dimension 3.*

1. *Le produit vectoriel est linéaire en chacune de ses composantes et anti-symétrique, c-à-d :*

$$v \wedge u = -(u \wedge v) \qquad\qquad (\lambda u + v) \wedge w = \lambda(u \wedge w) + (v \wedge w).$$

2. *$u \wedge v = 0$ si et seulement si (u, v) est un système lié.*
3. *Le vecteur $u \wedge v$ est orthogonal à u et à v.*
4. *Si (u, v) est un système orthonormé, alors $(u, v, u \wedge v)$ est un système orthonormé direct.*
5. *Dans $\mathbb{R}^3$ muni de la base canonique et de l'orientation directe canonique, si $u = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ et $v = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ alors :*

$$u \wedge v = \begin{pmatrix} yc - zb \\ za - xc \\ xb - ya \end{pmatrix}$$

6. *Soit θ l'angle non orienté entre deux vecteurs u et v non nuls, c'est-à-dire l'unique réel $\theta \in [0, \pi]$ tel que $\langle u, v\rangle = \|u\| \cdot \|v\| \cdot \cos\theta$. Alors :*
$$\|u \wedge v\| = \|u\| \cdot \|v\| \cdot \sin\theta.$$

Démonstration. On fixe une base orthonormée directe e, et on calcule les produits mixtes comme des déterminants dans cette base.

Pour le point 1, on utilise le plus possible la caractérisation (5.1). Commençons pas montrer l'antisymétrie. Pour $x, y, u \in E$ on calcule le déterminant dans une base e orthonormée directe :

$$\begin{aligned}
\det_e(v, u, x) &= -\det_e(u, v, x) && \text{d'après les propriétés du déterminant}\\
&= -\langle u \wedge v, x\rangle && \text{d'après la définition de } u \wedge v\\
&= \langle -(u \wedge v), x\rangle
\end{aligned}$$

D'après la caractérisation du produit vectoriel, on a donc $v \wedge u = -(u \wedge v) =$. La bilinéarité se montre de la même manière.

De même, pour le point 2, on a $u \wedge v = 0$ si et seulement si la forme linéaire $\det_e(u, v, \cdot)$ est nulle, ce qui est le cas si et seulement si (u, v) est lié.

Pour montrer que u et v sont orthogonaux à $u \wedge v$ (point 3 du théorème), on utilise encore la caractérisation du produit vectoriel :
$$\langle u \wedge v, u\rangle = \det_e(u, v, u) = 0,$$
et idem pour v.

Pour le point 4, il nous reste seulement à montrer que, si (u, v) est orthonormé, alors $u \wedge v$ est de norme 1 et $\det_e(u, v, u \wedge v) = 1$. Soit z de norme 1 et tel que (u, v, z) est un système orthonormé. Alors $\langle z, u\rangle = \langle z, v\rangle = 0$ et $\langle z, z\rangle = 1$. Mais on a aussi, en posant $w = u \wedge v$: $\langle w, u\rangle = \langle w, v\rangle = 0$ et $\langle w, z\rangle = \det_e(u, v, z) = 1$, donc $w = z$, ce qu'on voulait montrer.

Pour le point 5, on calcule le déterminant $\det_e(u, v, w)$ avec $w = \left(\begin{smallmatrix}\alpha\\ \beta\\ \gamma\end{smallmatrix}\right)$, et on cherche à l'exprimer comme forme linéaire en (α, β, γ) :
$$\begin{vmatrix} x & a & \alpha\\ y & b & \beta\\ z & c & \gamma\end{vmatrix} = \alpha(yc - zb) - \beta(xc - za) + \gamma(xb - ya) = \langle \begin{pmatrix}\alpha\\ \beta\\ \gamma\end{pmatrix}, \begin{pmatrix} yc - zb\\ za - xc\\ xb - ya\end{pmatrix}\rangle$$

On montre maintenant le point 6. Si $\theta = 0$ ou $\theta = \pi$, alors le résultat est vrai car les deux vecteurs u et v sont liés, donc $u \wedge v = 0$ et $\sin\theta = 0$. On suppose donc que u et v ne sont pas liés. Par bilinéarité du produit vectoriel, on peut supposer que u et v sont tous les deux normés. On calcule maintenant le produit vectoriel en coordonnées, en se plaçant dans une base orthonormée directe (e_1, e_2, e_3) avec $e_1 = u$ et $\mathrm{Vect}\langle e_1, e_2\rangle = \mathrm{Vect}\langle u, v\rangle$. En identifiant les vecteurs avec leurs coorodnnées, on a :
$$u = \begin{pmatrix}1\\0\\0\end{pmatrix} \qquad v = \begin{pmatrix}\cos\theta\\ \varepsilon\sin\theta\\ 0\end{pmatrix} \qquad u \wedge v = \begin{pmatrix}0\\0\\ \lambda\end{pmatrix}$$

Et d'après la formule précédente, on a $\lambda = \varepsilon\sin\theta$, d'où $\|u \wedge v\| = \sin\theta$; ce qu'on voulait montrer. □

Remarque 5.16. 1. Le produit vectoriel fournit une méthode rapide pour trouver un vecteur orthogonal à deux vecteurs de $\mathbb{R}^3$. Si les deux vecteurs, disons u et v, dont on part forment un système orthonormé, on est aussi garanti de trouver le vecteur manquant pour former l'unique système orthonormé direct commençant par (u, v).

2. Le produit vectoriel dans $\mathbb{R}^3$ généralise à la dimension 3 l'astuce élémentaire consistant à considérer $\left(\begin{smallmatrix} -b \\ a \end{smallmatrix}\right)$ pour obtenir un vecteur directement orthognal à $\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$.

3 Rotations de l'espace euclidien de dimension 3

Généralités sur les isométries de l'espace euclidien de dimension 3.

Soit u une isométrie d'un espace eucliden E de dimension 3. Le polynôme caractéristique de u étant de degré impair, il s'annule au moins une fois, ce qui montre que u a au moins une valeur propre. Rappelons que, puisque u est une isométrie, les seules valeurs propres possibles de u sont ± 1.

Par ailleurs, si x est un vecteur propre de u, alors $\mathbb{R}x$ est stable par u. Comme u est une isométrie, on en déduit que le plan $F = (\mathbb{R}x)^\perp$ est également stable par u. Donc la restriction $u|_F$ est une isométrie en dimension 2, à laquelle on peut appliquer les résultats précédents sur la classification des isométries en dimension 2.

Isométries directes de l'espace euclidien de dimension 3.

On rappelle que la notion d'isométrie directe ne nécessite pas d'orienter l'espace pour être bien définie : une isométrie u est dite directe si $\det(u) = 1$.

Lemme 5.17. *Toute isométrie directe d'un espace euclidien de dimension 3 admet au moins un point fixe.*

Démonstration. Soit u une isométrie de E euclidien de dimension 3. Raisonnons par l'absurde, et supposons que u n'a pas de point fixe. Soit x un vecteur propre de u pour la valeur propre -1, soit $F = (\mathbb{R}x)^\perp$ et soit $v = u|_F$. Alors $\det u = -\det v$, donc $\det v = -1$, c'est-à-dire que v est une isométrie indirecte de F. Mais alors v est une symetrie orthogonale par rapport à une droite, ce qui entraîne que v, et donc u, a un point fixe, contradiction. □

Définition 5.18. Soit E un espace euclidien de dimension 3. Soit Δ une droite vectorielle de E et soit $F = \Delta^\perp$. On appelle *rotation d'axe* Δ toute isométrie u de E telle que $u|_\Delta = \mathrm{Id}|_\Delta$, et $u|_F$ est une rotation de F.

Soit u un vecteur $\neq 0$ de E, et soit $\theta \in \mathbb{R}$. On définit $R_{u,\theta}$ comme la rotation d'axe $\Delta = \mathbb{R}u$, et dont la restriction à $\Delta^\perp$ est la rotation plane d'angle θ, lorsque $\Delta^\perp$ a été orienté par une base orthonormée (e_1, e_2) faisant de (e_1, e_2, u) une base directe.

Les rotations de la forme $R_{u,\pi}$ sont appelées des *retournements*.

Remarque 5.19. 1. Toute rotation de E s'écrit sous la forme $R_{u,\theta}$ mais attention ! l'écriture n'est pas unique. En effet $R_{-u,-\theta} = R_{u,\theta}$.

2. On montrera en exercice la *formule de Rodrigues*, valable dans un espace orienté :

$$R_{u,\theta}(x) = (\cos\theta)\cdot x + (1-\cos\theta)\cdot\langle u, x\rangle\cdot u + (\sin\theta)\cdot u\wedge x.$$

3. Si $e = (e_1, e_2, e_3)$ est une base orthonormée directe telle que $e_1 = u$, alors :

$$\operatorname{Mat}(R_{u,\theta}, e) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

Théorème 5.20. *Toute isométrie directe d'un espace euclidien de dimension* 3 *est une rotation.*

Démonstration. Soit u une isométrie directe de E euclidien avec $\dim(E) = 3$. Alors u admet au moins un point fixe, disons x normé, d'après le lemme 5.17. Soit $F = x^\perp$. Alors $u\big|_F$ est une isométrie positive de F, c'est donc une rotation de F. Donc u est une rotation d'axe $\Delta = \mathbb{R}x$. □

Détermination pratique des éléments d'une rotation dans $\mathbb{R}^3$

On explique sur un exemple comment déterminer l'axe et l'angle de rotation d'une rotation de $\mathbb{R}^3$ donnée sous forme matricielle.

On considère la matrice suivante :

$$M = \begin{pmatrix} 9/25 & 12/25 & -4/5 \\ 12/25 & 16/25 & 3/5 \\ 4/5 & -3/5 & 0 \end{pmatrix}$$

Question : *Montrer que M est la matrice d'une rotation r dans $\mathbb{R}^3$ muni de son produit scalaire canonique, et déterminer ses éléments. En particulier, trouver une base orthonormée directe de $\mathbb{R}^3$ et l'unique réel θ tel que :*

$$\operatorname{Mat}(r, e) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

Pour vérifier que M est orthogonale et de déterminant 1, on peut soit :

1. vérifier ${}^tM\cdot M = I_3$ et $\det(M) = 1$;
2. vérifier que les deux premières colonnes sont normées et orthogonales et que le produit vectoriel des deux premières colonnes donne la troisième colonne.

On trouve l'axe de la rotation en résolvant $M\cdot x = x$. On choisit un vecteur normé pour l'axe, on peut donc prendre :

$$u = \begin{pmatrix} 3/5 \\ 4/5 \\ 0 \end{pmatrix}$$

L'unique autre choix possible aurait été $-u$. Maintenant que u a été choisi, il existe un unique réel θ tel que $r = R_{u,\theta}$. Le cosinus de θ est déterminé par sa trace : $1+2\cos\theta = 1$ donc $\theta = \pm\pi/2$, mais il reste à déterminer le signe de θ.

Choisissons $e_1 = u$, puis e_2 un vecteur normé et orthogonal à e_1, par exemple $e_2 = \begin{pmatrix} -4/5 \\ 3/5 \\ 0 \end{pmatrix}$. Alors l'unique vecteur e_3 tel que (e_1, e_2, e_3) est orthonormée directe nous est donné par le produit vectoriel :

$$e_3 = e_1 \wedge e_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Exprimons Me_2 en fonction de e_2 et de e_3 :

$$Me_2 = \begin{pmatrix} 9/25 & 12/25 & -4/5 \\ 12/25 & 16/25 & 3/5 \\ 4/5 & -3/5 & 0 \end{pmatrix} \cdot \begin{pmatrix} -4/5 \\ 3/5 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} = -e_3.$$

Donc $\theta = -\pi/2$. Dans la base orthonormée directe $e = (e_1, e_2, e_3)$, on a donc :

$$\mathrm{Mat}(r, e) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

4 Isométries indirectes de l'espace eucliden de dimension 3

Définition 5.21. Soit E un espace euclidien de dimension 3, soit F un plan de E, soit r une rotation d'axe $F^\perp$ et soit s la réflexion par rapport à F. Alors $r \circ s = s \circ r$, et cette isométrie est une isométrie indirecte de E appelée *rotation-réflexion* d'axe $F^\perp$.

Remarque 5.22. En considérant l'identité comme une rotation, les réflexions sont elles-mêmes des rotation-réflexions.

Pour justifier les affirmations de la définition ci-dessus, posons $\Delta = F^\perp$. On vérifie l'égalité $r \circ s = s \circ r$ sur F et sur Δ, et on la déduit sur E. Enfin, $r \circ s$ est une isométrie comme composée de deux isométries, et de déterminant -1 puisque $\det(r) = 1$ et $\det(s) = -1$.

Les rotation-réflexions sont donc une classe d'isométries indirectes. Le résultat suivant nous dit qu'il n'y en a pas d'autres.

Proposition 5.23. *Soit E un espace euclidien de dimension 3. Alors toute isométrie indirecte de E est une rotation-réflexion.*

Démonstration. Soit f une isométrie indirecte de E. On a vu que f a au moins un vecteur propre, disons u, de valeur propre $\lambda = \pm 1$. Posons $F = u^\perp$, et soit s la réflexion par rapport à F.

Posons $r = s \circ f$. Alors r est une isométrie directe de E, et $r(u) = \lambda s(u) = -\lambda u$.

Donc, si $\lambda = -1$, alors r est une rotation d'axe $\mathbb{R}u$, et alors $f = s \circ r$ est bien une rotation-réflexion.

Si $\lambda = 1$, alors $f\big|_F$ est une isométrie indirecte de F, donc une réflexion de F. Soit (i, j) une base orthonormée de F telle que $f(i) = -i$ et $f(j) = j$. Alors, comme $f(u) = u$, on voit que f est la réflexion par rapport au plan $\mathrm{Vect}(j, k)$, et est donc bien une rotation-réflexion d'après la remarque 5.22. □

Chapitre 6

Formes bilinéaires et formes quadratiques

La notion de forme bilinéaire généralise celle de produit scalaire, en supprimant l'hypothèse de positivité. De nouveaux phénomènes intéressants apparaissent dans ce cadre.

1 Formes bilinéaires

Définition 6.1. Soit E un $\mathbb{R}$-e.v.. Une *forme bilinéaire sur $E \times E$* est une application $\varphi : E \times E \longrightarrow \mathbb{R}$ telle que :

1. Pour tout $x_0 \in E$, l'application $y \in E \mapsto \varphi(x_0, y)$ est linéaire.
2. Pour tout $y_0 \in E$, l'application $x \in E \mapsto \varphi(x, y_0)$ est linéaire.

Exemples.

1. L'application φ définie par $\varphi(x, y) = xy$ est une forme bilinéaire sur $\mathbb{R} \times \mathbb{R}$.
2. Plus généralement, si ℓ_1 et ℓ_2 sont deux formes linéaires sur E, alors $\varphi(x, y) = \ell_1(x) \cdot \ell_2(y)$ définit une forme bilinéaire sur $E \times E$.
3. Soit E l'espace des fonctions réelles et continues sur $[0, 1]$. Alors on définit une forme bilinéaire φ sur $E \times E$ en posant :
$$\varphi(f, g) = \int_0^1 f(t)g(t)dt.$$
4. Tout produit scalaire sur E définit une forme bilinéaire sur $E \times E$.
5. On définit une forme bilinéaire sur $\mathbb{R}^2 \times \mathbb{R}^2$ qui *n'est pas* un produit scalaire en posant :
$$\varphi(\left(\begin{smallmatrix} x_1 \\ y_1 \end{smallmatrix}\right), \left(\begin{smallmatrix} x_2 \\ y_2 \end{smallmatrix}\right)) = x_1x_2 - y_1y_2.$$
6. Soit $u : E \longrightarrow E^*$ une application linéaire. On définit une forme bilinéaire sur $E \times E$ en posant :
$$\varphi(x, y) = \big(u(x)\big)(y).$$

L'exemple 6 ci-dessus est en fait général car toute forme bilinéaire φ peut se mettre sous cette forme. En effet, étant donnée une forme bilinéaire φ sur $E \times E$, définissons $u_g : E \longrightarrow E^*$ en posant $\langle u_g(x), y\rangle = \varphi(x, y)$. Alors la bilinéarité de φ fait que u_g est linéaire.

On peut inverser les rôles de x et y en définissant $u_d : E \longrightarrow E^*$, toujours pour φ une forme bilinéaire sur $E \times E$ fixée, par $\langle u_d(x), y\rangle = \varphi(y, x)$. On a donc deux identifications possibles entre formes bilinéaires sur $E \times E$ et applications linéaires $E \longrightarrow E^*$.

Formes symétriques.

Définition 6.2. Une forme bilinéaire $\varphi : E\times E \to \mathbb{R}$ est *symétrique* si $\varphi(x, y) = \varphi(y, x)$ pour tout $(x, y) \in E \times E$. On note $\mathscr{B}(E)$ le $\mathbb{R}$-e.v. des formes bilinéaires symétriques sur $E \times E$.

Exemple 6.3. Tout produit scalaire est par définition symétrique. La forme sur $\mathbb{R}^2\times\mathbb{R}^2$ de l'exemple 5 ci-dessus est symétrique.

Dans l'optique des identifications entre formes bilinéaires sur $E \times E$ et applications linéaires $E \to E^*$, la situation est plus simple dans le cas des formes bilinéaires symétriques car alors les deux applications $u_g, u_d : E \to E^*$ sont les mêmes. On notera $D_\varphi : E \to E^*$ cette unique application linéaire.

Définition 6.4. Si $\varphi : E \times E \to \mathbb{R}$ est une forme bilinéaire symétrique sur $E \times E$, on note $D_\varphi : E \to E^*$ l'application linéaire définie par $\langle D_\varphi(x), y\rangle = \varphi(x, y)$. L'application linéaire $D_\varphi : E \to E^*$ est *l'application linéaire canonique* associée à φ.

Représentation matricielle.

Supposons E de dimension finie, et soit $e = (e_1, \ldots, e_n)$ une base de E. Soit φ une forme bilinéaire sur $E \times E$. Alors, pour $x = \sum_i x_i e_i$ et $y = \sum_j y_j e_j$, on a par bilinéarité :

$$\varphi(x, y) = \sum_{i,j} x_i y_j \varphi(e_i, e_j).$$

Si X et Y sont les vecteurs colonnes avec les coordonnées de x et de y, on a donc l'écriture matricielle :

$$\varphi(x, y) = {}^tX \cdot M \cdot Y \qquad \text{avec } M = (M_{i,j})_{i,j} \text{ et } M_{i,j} = \varphi(e_i, e_j).$$

Définition 6.5. Soit E un $\mathbb{R}$-e.v. de dimension finie, et soit $e = (e_1, \ldots, e_n)$ une base de E. Soit φ une forme bilinéaire sur $E\times E$. La *matrice de φ dans la base e* est la matrice carrée $M = (M_{i,j})_{1\leq i,j\leq n}$ définie par $M_{i,j} = \varphi(e_i, e_j)$, aussi notée $M = \mathrm{Mat}(\varphi, e)$.

Proposition 6.6. *Soit E un $\mathbb{R}$-e.v. de dimension finie muni d'une base e, et soit φ une forme bilinéaire sur $E \times E$. Alors φ est symétrique si et seulement la matrice $M = \mathrm{Mat}(\varphi, e)$ est symétrique. Dans ce cas, la matrice M est aussi la matrice de l'application linéaire canonique $D_\varphi : E \to E^*$, dans les bases e et e^*.*

Démonstration. L'écriture $\varphi(x, y) = {}^tX \cdot M \cdot Y$ vue plus haut montre l'équivalence entre φ symétrique et M symétrique. Supposons que φ est symétrique. Alors pour chaque vecteur de base e_i, on a :

$$\langle D_\varphi(e_i), y\rangle = \varphi(e_i, y) = \sum_j y_j \varphi(e_i, e_j) = \sum_j e_j^*(y)\varphi(e_i, e_j),$$

ce qui s'écrit aussi :

$$D_\varphi(e_i) = \sum_j \varphi(e_i, e_j) e_j^*.$$

Les coordonnées de $D_\varphi(e_i)$ dans la base duale e^* coïncident avec la j^e colonne de la matrice M, c'est donc que $M = \mathrm{Mat}(D_\varphi, e, e^*)$. □

Changement de base

La formule de changement de base pour les formes bilinéaires *n'est pas la même que pour les applications linéaires.*

Proposition 6.7. *Soit E un espace vectoriel de dimension finie, soit e et e' deux bases de E, et posons $P = \mathrm{P}(e, e')$ la matrice de passage de e vers la nouvelle base e'. Soit φ une forme bilinéaire sur $E \times E$, et soit $M = \mathrm{Mat}(\varphi, e)$ et $M' = \mathrm{Mat}(\varphi, e')$. Alors M et M' sont reliées par la relation matricielle suivante :*

$$M' = {}^tP \cdot M \cdot P.$$

Démonstration. Soit $x, y \in E$, dont les coordonnées dans les bases e et e' sont respectivement X et X' et Y et Y'. Alors $X = P \cdot X'$ et $Y = P \cdot Y'$. On a donc :

$$\varphi(x, y) = {}^tX \cdot M \cdot Y = {}^tX' \cdot ({}^tP \cdot M \cdot P) \cdot Y' = {}^tX' \cdot M' \cdot Y'.$$

Posons $M'' = {}^tP \cdot M \cdot P$. Il s'ensuit de l'égalité précédente que la formule ${}^tX \cdot (M'' - M') \cdot Y' = 0$ est vraie pour tous vecteurs X' et Y' de $\mathbb{R}^n$. Cela implique que $M' = M''$; il suffit pour le voir de prendre $X' = u_i$ et $Y' = u_j$ des vecteurs de la base canonique de $\mathbb{R}^n$, pour obtenir l'égalité des coefficients (i, j) des deux matrices. □

Rang et noyau d'une forme bilinéaire symétrique.

Grâce à l'application linéaire $D_\varphi : E \to E^*$ associée à une forme bilinéaire symétrique φ, on va pouvoir utiliser les notions d'algèbre *linéaire* pour étudier les *formes bilinéaires*.

Définition 6.8. Soit E un $\mathbb{R}$-e.v. de dimension finie, et soit φ une forme bilinéaire symétrique sur $E \times E$. Soit $D_\varphi : E \to E^*$ l'application linéaire canonique associée à φ.

1. Le *noyau* de φ est $\ker \varphi = \ker D_\varphi$.
2. Le *rang* de φ est $\mathrm{rg}\, \varphi = \mathrm{rg}\, D_\varphi$.
3. φ est *non-dégénérée* si $\ker \varphi = \{0\}$.

Autrement dit, φ est non dégénérée si et seulement si le seul vecteur $x \in E$ tel que : $\forall y \in E \quad \varphi(x, y) = 0$ est le vecteur $x = 0$.

Exemple 6.9.
1. Un produit scalaire $b : E \times E \to \mathbb{R}$ est une forme bilinéaire symétrique non dégénérée. En effet, soit $x \in \ker b$, c'est-à-dire tel que $b(x, y) = 0$ pour tout $y \in E$. Alors en particulier pour $y = x$ on obtient $\|x\|^2 = 0$ et donc $x = 0$.

2. Soit φ la forme bilinéaire symétrique définie par $\varphi(\binom{x_1}{x_2}, \binom{y_1}{y_2}) = x_1y_1 - x_2y_2$. Alors nous avons déjà remarqué que φ n'est pas un produit scalaire. Pour autant, φ n'est pas dégénérée.

   On peut par exemple le voir en déterminant le rang de sa matrice, qui est $M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Puisque M est de rang 2, φ est non dégénérée.

   Cet exemple révèle que l'argument utilisé dans l'exemple précédent n'est pas forcément concluant. En effet, on avait utilisé plus haut la propriété $b(x, x) = 0 \implies x = 0$. Mais ici, la propriété $\varphi(x, x) = 0 \implies x = 0$ *n'est pas* vérifiée. En effet on a par exemple $\varphi(\binom{1}{1}, \binom{1}{1}) = 0$.

   Il faut donc se garder de transposer directement l'intuition héritée des produits scalaires aux formes bilinéaires symétriques plus générales. On y reviendra ci-dessous dans l'étude des vecteurs isotropes.
3. Sur $\mathbb{R}^2 \times \mathbb{R}^2$, la forme bilinéaire symétrique $\varphi(\binom{x_1}{x_2}, \binom{y_1}{y_2}) = x_1y_1$ est dégénérée. En effet $\varphi(\binom{0}{1}, \binom{y_1}{y_2}) = 0$ pour tout $\binom{y_1}{y_2}$ donc $\binom{0}{1} \in \ker\varphi$. On voit ici que $\operatorname{rg}\varphi = 1$.

Orthogonalité.

On adapte la notion d'orthogonalité vue précédemment pour les produits scalaires au cadre plus général des formes bilinéaires symétriques.

Définition 6.10. Soit φ une forme bilinéaire symétrique sur un $\mathbb{R}$-e.v. E.

1. Deux vecteurs $x, y \in E$ sont *orthogonaux* si $\varphi(x, y) = 0$
2. L'*orthogonal* d'une partie $X \subseteq E$ est le sous-e.v. noté $H^\perp$ et défini par $H^\perp = \{z \in E \ : \ \forall x \in X \quad x \perp z\}$.
3. Un vecteur $x \in E$ est *isotrope* s'il est orthogonal à lui-même, c'est-à-dire si $x \perp x$, c'est-à-dire si $\varphi(x, x) = 0$.

Remarque 6.11.
1. L'ensemble des vecteurs isotropes a une structure de *cône*, c'est-à-dire que si $x \in E$ est isotrope, alors λx est isotrope quel que soit $\lambda \in \mathbb{R}$. En général, le cône des vecteurs isotropes *n'a pas* la structure d'un sous-espace.
2. Comme l'exemple 1 ci-dessus l'a montré (Exemple 6.9), une forme bilinéaire sans vecteur isotrope est nécessairement non dégénérée.
3. Mais comme l'exemple 2 ci-dessus l'a montré (Exemple 6.9), une forme bilinéaire non dégénérée peut très bien avoir des vecteurs isotropes.

Exemple 6.12.
1. On voit immédiatement que pour toute partie A, on a $A^\perp = (\operatorname{Vect} A)^\perp$. En particulier, si F est un sous-e.v. de dimension finie et muni d'une base $e = (e_1, \ldots, e_k)$, alors :

$$F^\perp = \{x \in E \ : \ \varphi(x, e_i) = 0 \text{ pour } i = 1, \ldots, k\}.$$

2. On a $E^\perp = \ker D_\varphi$. Donc φ est non dégénérée si et seulement si $E^\perp$ est réduit à $\{0\}$.

Dans le cas d'une forme bilinéaire symétrique non dégénérée, on peut facilement calculer la dimension de l'orthogonal d'un sous-e.v., comme le montre la proposition suivante.

Proposition 6.13. *Soit E un $\mathbb{R}$-e.v. de dimension finie égale à n. Soit φ une forme bilinéaire symétrique sur $E \times E$, et soit F un sous-e.v. de E.*

1. *On suppose qu'aucun vecteur non nul de F n'est orthogonal à l'espace E tout entier. Alors :*

$$\dim(F^{\perp}) = n - \dim F.$$

2. *En particulier, le résultat s'applique à tout sous-e.v. F de E si φ est non dégénérée.*

Démonstration. Il est clair que le point 2 découle du point 1, il suffit donc de montrer ce dernier. Soit $k = \dim F$, et soit $(e_1, \dots, e_k)$ une base de F. Comme on l'a vu ci-dessus, les vecteurs de $F^{\perp}$ sont ceux qui sont orthogonaux à tous les e_i. En posant $\ell_i = \varphi(e_i, \cdot)$, on a donc :

$$F^{\perp} = \bigcap_{i=1}^{k} \ker \ell_i.$$

Il suffit donc de voir que les ℓ_i sont libres dans E^* pour avoir le résultat $\dim(F^{\perp}) = n - k$ d'après la proposition 2.14.

Soit donc $\lambda_1, \dots, \lambda_k$ des réels tels que $\lambda_1\ell_1 + \dots + \lambda_k\ell_k = 0$. Cette égalité qui a lieu dans E^* se ré-écrit :

$$\forall y \in E \quad \varphi\Big(\sum_{i=1}^{k} \lambda_i e_i, y\Big) = 0.$$

Or l'hypothèse est que 0 est le seul vecteur de $F \cap E^{\perp}$, c'est donc que $\lambda_1 e_1 + \dots + \lambda_k e_k = 0$. Comme $(e_1, \dots, e_k)$ est une base de F, il s'ensuit que $\lambda_1 = \dots = \lambda_k = 0$, ce qu'on voulait démontrer. □

Corollaire 6.14. *Si φ est une forme bilinéaire symétrique non dégénérée sur $E \times E$, avec E de dimension finie, alors $(F^{\perp})^{\perp} = F$ pour tout sous-e.v. F de E.*

Démonstration. On a $\dim(F^{\perp})^{\perp} = \dim F$ d'après la proposition précédente. Or il est clair que $F \subseteq (F^{\perp})^{\perp}$, d'où l'égalité. □

Restriction d'une forme bilinéaire symétrique.

Si b est un produit scalaire sur E, il est clair que la restriction de b à tout sous-e.v. F de E est encore un produit scalaire sur F. En revanche, l'exemple suivant montre qu'une forme bilinéaire symétrique non dégénérée peut avoir des restrictions qui sont dégénérées.

Exemple 6.15. Reprenons $\varphi(\binom{x_1}{x_2}, \binom{y_1}{y_2}) = x_1y_1 - x_2y_2$, et considérons le vecteur $x_0 = \binom{1}{1}$ qui est isotrope pour φ. Soit $F = \mathbb{R}x_0$. Alors la restriction de φ à F est identiquement nulle, et elle est donc dégénérée.

Proposition 6.16. *Soit E un $\mathbb{R}$-e.v., et soit φ une forme bilinéaire symétrique sur $E \times E$. Soit F un sous-e.v. de E. Alors les propositions suivantes sont équivalentes :*

1. *La restriction de φ à $F \times F$ est non dégénérée.*
2. *$F \cap F^{\perp} = \{0\}$.*

3. $E = F \oplus F^{\perp}$.

Démonstration. Montrons d'abord l'équivalence 1 $\Longleftrightarrow$ 2. Soit $\psi = \varphi\big|_F$, et $H = \ker \psi$. Alors ψ est non dégénérée si et seulement si $H = \{0\}$. Il est clair que $\ker \psi = F \cap F^{\perp}$, d'où le résultat.

2 $\Longrightarrow$ 3. Puisque $F \cap F^{\perp} = \{0\}$, pour obtenir $E = F \oplus F^{\perp}$ il suffit d'avoir $\dim(F^{\perp}) = \dim E - \dim F$. Or nous pouvons appliquer le point 1 de la proposition 6.13 si aucun vecteur non nul $x \in F$ n'est orthogonal à E tout entier, ce qui est bien le cas *a fortiori* puisque $F \cap F^{\perp} = \{0\}$.

3 $\Longrightarrow$ 2. Évident. □

2 Formes quadratiques

Dans le cadre des produits scalaires, on a vu que la fonction "carré de la norme" avait des propriétés remarquables, en particulier vis-à-vis de l'orthogonalité grâce au théorème de Pythagore. La généralisation de cette fonction dans le cadre des formes bilinéaires conduit à la notion de forme quadratique.

Définition 6.17. Soit E un $\mathbb{R}$-e.v.. Une fonction $q : E \longrightarrow \mathbb{R}$ est une *forme quadratique* s'il existe une forme bilinéaire φ sur $E \times E$ (non symétrique *a priori*) tel que $q(x) = \varphi(x, x)$ pour tout $x \in E$.

Remarque 6.18. Toute forme quadratique q vérifie $q(\lambda x) = \lambda^2 q(x)$, pour tout vecteur x et pour tout scalaire λ. En particulier $q(0) = 0$.

Forme polaire d'une forme quadratique.

Bien que φ ne soit pas supposée *a priori* symétrique dans la définition ci-dessus, on peut en fait toujours se ramener au cas d'une forme bilinéaire symétrique, comme le montre la proposition suivante.

Proposition 6.19. *Soit $q : E \longrightarrow \mathbb{R}$ une forme quadratique. Alors il existe une unique forme bilinéaire* symétrique *sur $E \times E$ telle que $q(x) = \varphi(x, x)$ pour tout $x \in E$.*

Démonstration. Supposons que φ est une forme bilinéaire symétrique telle que $\varphi(x, x) = q(x)$ pour tout $x \in E$. Alors on a nécessairement :

$$q(x + y) = q(x) + q(y) + 2\varphi(x, y)$$

donc $\varphi(x, y) = \big(q(x+y) - q(x) - q(y)\big)/2$, ce qui montre l'unicité. Pour l'existence, il suffit de poser $\varphi(x, y) = \big(q(x+y) - q(x) - q(y)\big)/2$ et de vérifier que φ est bilinéaire symétrique, puis que $q(x) = \varphi(x, x)$. Or tout ceci découle facilement de l'écriture $q(x, y) = \psi(x, y)$ pour une certaine forme bilinéaire ψ (non symétrique *a priori*). □

Définition 6.20. Soit $q : E \longrightarrow \mathbb{R}$ une forme quadratique. La *forme polaire* de q est l'unique forme bilinéaire symétrique φ sur $E \times E$ telle que $q(x) = \varphi(x, x)$ pour tout $x \in E$.

Pratique de la détermination de la forme polaire.

La démonstration de la proposition 6.19 fournit une formule explicite pour exprimer φ en fonction de q. Il y en a aussi d'autres :

$$\varphi(x,y) = \frac{1}{2}\big(q(x+y) - q(x) - q(y)\big) \qquad \varphi(x,y) = \frac{1}{4}\big(q(x+y) - q(x-y)\big)$$

En pratique, il est cependant maladroit de passer par ces formules pour trouver la forme polaire d'une forme quadratique donnée explicitement. Il est plus aisé de procéder directement, comme suit.

Exemple 6.21. Sur $\mathbb{R}^3$, soit $q(\left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right)) = x^2 + y^2 - 2z^2 + 2xy - xz$. Alors q est une forme quadratique, dont la forme polaire est :

$$\varphi(\begin{pmatrix} x \\ y \\ z \end{pmatrix}, \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}) = xx' + yy' - 2zz' + (xy' + x'y) - \frac{1}{2}(xz' + x'z).$$

La méthode consiste donc à reporter chaque terme carré, tel que x^2 dans la forme quadratique, comme un terme comportant les deux instance (ici x et x') de la même composante ; et chaque terme rectangle, tel que $2xy$, est distribué avec un facteur $1/2$ dans chaque composante du produit en croix qui lui correspond (ici $xy' + x'y$), ceci afin de maintenir la symétrie de la forme polaire.

Représentation matricielle.

Soit E un $\mathbb{R}$-e.v. de dimension finie muni d'une base $e = (e_1, \dots, e_n)$ et d'une forme quadratique q. Soit φ la forme polaire de q, et soit $M = \mathrm{Mat}(\varphi, e)$. Alors pour tout vecteur $x \in E$, si $X \in \mathbb{R}^n$ est le vecteur des coordonnées de x dans la base e, on a :

$$q(x) = \varphi(x,x) = {}^tX \cdot M \cdot X.$$

Cette écriture justifie la définition suivante.

Définition 6.22. Soit E un $\mathbb{R}$-e.v. de dimension finie muni d'une base e, et soit $q : E \longrightarrow \mathbb{R}$ une forme quadratique sur E. La *matrice de q dans la base e* est $M = \mathrm{Mat}(\varphi, e)$, où φ est la forme polaire de q. Le *rang* de q est $\mathrm{rg}\, q = \mathrm{rg}\, \varphi = \mathrm{rg}\, M$, et le *noyau* de q est $\ker q = \ker \varphi = \ker D_\varphi$, où $D_\varphi : E \longrightarrow E^*$ est l'application linéaire canonique associée à φ.

Évidemment, la formule de changement de base $M' = {}^tP \cdot M \cdot P$ est la même pour les formes quadratiques que pour les formes bilinéaires.

Bases q-orthogonales.

L'orthogonalité de vecteurs et de sous-espaces par rapport à une forme quadratique a est définie comme l'orthogonalité par rapport à la forme polaire de q (*cf.* définition 6.10). On parle de *q-orthogonalité* pour souligner la dépendance par rapport à q. De même, un vecteur x est q-isotrope si $q(x) = 0$, c'est-à-dire si x est isotrope pour la forme polaire de q.

Remarque 6.23. Attention à ne pas confondre les vecteurs isotropes de q et ceux de son noyau : $q(x) = 0$ est une condition nécessaire mais en général *non suffisante* pour que x soit dans le noyau de q.

Définition 6.24. Soit E un $\mathbb{R}$-e.v. de dimension finie, et soit $q : E \longrightarrow \mathbb{R}$ une forme quadratique sur E. Une base $e = (e_1, \dots, e_n)$ de E est *q-orthogonale* si les vecteurs de la base e sont deux à deux q-orthogonaux

Remarque 6.25. Une base e de E est q-orthogonale si et seulement si la matrice $\mathrm{Mat}(q, e)$ est diagonale.

Combinaisons linéaires de carrés de formes linéaires.

Un exemple de forme bilinéaire symétrique est $\varphi(x, y) = \ell(x)\ell(y)$, avec $\ell \in E^*$, pour $x, y \in E$. La forme quadratique associée est $q(x) = \big(\ell(x)\big)^2$.

Les formes quadratiques qu'on obtient ainsi sont toutes de rang 1 (exercice). Plus généralement, si $\ell_1, \dots, \ell_k$ sont k formes linéaires sur E, alors on définit une forme quadratique q en posant :

$$q(x) = c_1\big(\ell_1(x)\big)^2 + \cdots + c_k\big(\ell_k(x)\big)^2, \tag{6.1}$$

où $c_1, \dots, c_k$ sont des réels. La forme q ainsi définie est une combinaison linéaire de carrés de formes linéaires.

Réciproquement, il est naturel de se demander si toutes forme quadratique peut être exprimée sous cette forme. La réponse est affirmative, et la décomposition de Gauss est un algorithme fournissant une décomposition effective.

Avant d'en venir à la décomposition de Gauss, remarquons que si les formes linéaires ℓ_i entrant dans la décomposition (6.1) sont indépendantes, alors l'expression (6.1) fournit directement des renseignements sur la forme quadratique.

Proposition 6.26. *Soit E un $\mathbb{R}$-e.v. de dimension finie, soit $(\ell_1, \dots, \ell_k)$ un système libre de formes linéaires sur E, et soit $c_1, \dots, c_k$ des réels tous non nuls. Alors :*

1. *La forme quadratique* (6.1) *est de rang k, et est donc non dégénérée si et seulement si $k = \dim E$.*
2. *Soit $\ell_{k+1}, \dots, \ell_n$ des formes linéaires telles que $(\ell_1, \dots, \ell_n)$ soit une base de E^*, et soit e sa base préduale. Alors* $\mathrm{Mat}(q, e)$ *est diagonale.*

Remarque 6.27. Réciproquement au point 2 de la proposition ci-dessus, on peut remarquer que si e est une base de E telle que $\mathrm{Mat}(q, e)$ est diagonale, et si $e^* = (e_1^*, \dots, e_n^*)$ en est la base duale, alors q est une combinaison linéaire des carrés $(e_1^*)^2, \dots, (e_n^*)^2$.

Démonstration. Complétons le système libre $(\ell_1, \dots, \ell_k)$ en une base $(\ell_1, \dots, \ell_n)$ de E^*, et soit $e = (e_1, \dots, e_n)$ sa base préduale. Puisque $(\ell_1, \dots, \ell_n)$ est alors la base duale de e, la forme (6.1) nous donne directement $M = \mathrm{Mat}(q, e)$:

$$M = \begin{pmatrix} c_1 & 0 & & \dots & 0 \\ & & & & \\ 0 & & c_k & \dots & 0 \\ 0 & & & 0 & \vdots \\ 0 & & & \dots & 0 \end{pmatrix}$$

d'où les deux résultats, puisque $\operatorname{rg} q = \operatorname{rg} M$. □

Dans le cas des *applications linéaires* dont la matrice est écrite sous forme diagonale dans une certaine base, nous sommes habitués à ce que les termes diagonaux, qui sont les valeurs propres de l'application linéaire, soient invariants avec la base. Dans le cas des matrices de formes quadratiques, nous verrons ci-dessous que seul le signe des coefficients diagonaux est invariant.

Pratique de la décomposition de Gauss.

On va expliquer la décomposition de Gauss sur des exemples dans $\mathbb{R}^3$. On note u les vecteurs de $\mathbb{R}^3$, et x, y et z les coordonnées dans la base canonique.

Cas d'une forme quadratique comportant au moins un terme carré. C'est le cas le plus simple. Par exemple :

$$q(u) = x^2 - 2y^2 + xy - xz. \tag{6.2}$$

On traite d'abord le terme x^2. On le regroupe avec tous les termes rectangle en x, ici xy et $-xz$ en débutant le développement d'un carré, et on soustrait les termes carré *et* rectangle qu'on a rajoutés.

$$\begin{aligned} q(u) &= \Big((x + \frac{1}{2}y - \frac{1}{2}z)^2 - \frac{1}{4}y^2 - \frac{1}{4}z^2 + \frac{1}{2}yz\Big) - 2y^2 \\ &= (x + \frac{1}{2}y - \frac{1}{2}z)^2 - \frac{9}{4}y^2 - \frac{1}{4}z^2 + \frac{1}{2}yz \end{aligned}$$

On pose $\ell_1(u) = x + \frac{1}{2}y - \frac{1}{2}z$ et on recommence avec le terme en y^2 associé aux termes rectangle comportant y, c'est-à-dire ici $\frac{1}{2}yz$. Le plus simple est de factoriser le coefficient $-\frac{9}{4}$ de y^2, ce qui donne :

$$\begin{aligned} q(u) &= \big(\ell_1(u)\big)^2 - \frac{9}{4}(y^2 - \frac{2}{9}yz) - \frac{1}{4}z^2 \\ &= \big(\ell_1(u)\big)^2 - \frac{9}{4}(y - \frac{1}{9}z)^2 - \frac{2}{9}z^2 \end{aligned}$$

En posant :

$$\ell_2(u) = y - \frac{1}{9}z \qquad\qquad \ell_3(u) = z,$$

on a donc exprimé $q(u)$ comme la combinaison linéaire :

$$q = \ell_1^2 - \frac{9}{4}\ell_2^2 - \frac{2}{9}\ell_3^2. \tag{6.3}$$

Pour trouver une base de $\mathbb{R}^3$ qui soit q-orthogonale, il suffit donc de trouver la base préduale, appelons-la e, de $\ell = (\ell_1, \ell_2, \ell_3)$. Les coordonnés des vecteurs de $e = (e_1, e_2, e_3)$ se lisent en colonnes dans l'inverse de la matrice M des coefficients de ℓ mis en ligne, ce qui donne :

$$M = \begin{pmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & -\frac{1}{9} \\ 0 & 0 & 1 \end{pmatrix} \qquad M^{-1} = \begin{pmatrix} 1 & -\frac{1}{2} & \frac{4}{9} \\ 0 & 1 & \frac{1}{9} \\ 0 & 0 & 1 \end{pmatrix} \qquad e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad e_2 = \begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix} \qquad e_3 = \begin{pmatrix} \frac{4}{9} \\ \frac{1}{9} \\ 1 \end{pmatrix}$$

En écrivant la forme polaire $\varphi(u,v) = xx' - 2yy' + \frac{1}{2}(xy' + x'y) - \frac{1}{2}(xz' + x'z)$, on vérifie que les vecteurs e_1, e_2, e_3 sont deux à deux orthogonaux. On lit directement sur l'expression (6.3) la forme de la matrice de q dans la base e préduale de ℓ :

$$\mathrm{Mat}(q,e) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{9}{4} & 0 \\ 0 & 0 & -\frac{2}{9} \end{pmatrix}$$

Cas d'une forme quadratique ne comportant aucun terme carré. Ce cas peut aussi se rencontrer dans le cadre précédent, dans l'une des étapes intermédiaires. La technique précédente ne s'applique plus. Considérons par exemple la forme quadratique q définie sur $\mathbb{R}^3$ par :

$$q(u) = xy.$$

L'idée est d'exprimer un terme rectangle comme une différence de deux carrés, ce qui donne :

$$q(u) = \frac{(x+y)^2 - (x-y)^2}{4}.$$

Considérons les deux formes linéaires définies par :

$$\ell_1 = x + y \qquad\qquad \ell_2 = x - y.$$

On complète (ℓ_1, ℓ_2) en une base (ℓ_1, ℓ_2, ℓ_3) de $(\mathbb{R}^3)^*$ en posant par exemple $\ell_3(u) = z$, dont la base préduale est $e = (e_1, e_2, e_3)$ donnée par :

$$e_1 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix} \qquad e_2 = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{pmatrix} \qquad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

La matrice de q dans la base e est :

$$\mathrm{Mat}(q,e) = \begin{pmatrix} \frac{1}{4} & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Exemple 6.28. Réduction de Gauss de $q(x_1,x_2,x_3,x_4) = x_1x_2 + x_2x_3 + 2x_1x_3 + 3x_1x_4 - x_3x_4$ dans $\mathbb{R}^4$.

On doit **isoler deux variables** qui apparaissent dans un terme rectangle :

$$\begin{aligned}
q(x_1,x_2,x_3,x_4) &= x_1x_2 + x_2x_3 + 2x_1x_3 + 3x_1x_4 - x_3x_4 \\
&= x_1x_2 + x_2x_3 + x_1(2x_3 + 3x_4) - x_3x_4 \\
&= [x_1 + x_3][x_2 + (2x_3 + 3x_4)] - x_3(2x_3 + 3x_4) - x_3x_4 \\
&= \frac{1}{4}[(x_1 + x_3) + (x_2 + (2x_3 + 3x_4))]^2 \\
&\quad -\frac{1}{4}[(x_1 + x_3) - (x_2 + (2x_3 + 3x_4))]^2 - 2x_3^2 - 4x_3x_4 \\
&= \frac{1}{4}\ell_1^2(x_1,x_2,x_3,x_4) - \frac{1}{4}\ell_2^2(x_1,x_2,x_3,x_4) - 2(x_3^2 + 2x_3x_4) \\
&= \frac{1}{4}\ell_1^2(x_1,x_2,x_3,x_4) - \frac{1}{4}\ell_2^2(x_1,x_2,x_3,x_4) - 2(x_3 + x_4)^2 + 2x_4^2
\end{aligned}$$

Donc avec

$$\begin{array}{rcl} \ell_1(x_1,x_2,x_3,x_4) & = & (x_1+x_3)+(x_2+(2x_3+3x_4)) = x_1+x_2+3x_3+3x_4 \\ \ell_2(x_1,x_2,x_3,x_4) & = & (x_1+x_3)-(x_2+(2x_3+3x_4)) = x_1-x_2-x_3-3x_4 \\ \ell_3(x_1,x_2,x_3,x_4) & = & x_3+x_4 \\ \ell_4(x_1,x_2,x_3,x_4) & = & x_4 \end{array}$$

On a bien construit une base de $(R^4)^*$ dont la base antéduale est une base q-orthogonale et

$$q = \frac{1}{4}\ell_1^2 - \frac{1}{4}\ell_2^2 - 2\ell_3^2 + 2\ell_4^2$$

Les deux techniques précédentes s'appliquent au cas d'une forme quadratique générale. On peut à la fin séparer les carrés des formes linéaires suivant le signe de leurs coefficients. On obtient alors le résultat suivant.

Théorème 6.29. *Soit E un $\mathbb{R}$-e.v. de dimension finie. Soit q une forme quadratique sur E. Alors il existe des formes linéaires linéairement indépendantes $\ell_1,\dots,\ell_p$, $\ell_{p+1},\dots,\ell_{p+r}$ telles que :*

$$q = \ell_1^2 + \dots + \ell_p^2 - \ell_{p+1}^2 - \dots - \ell_{p+r}^2.$$

Dans le premier des deux exemples précédents, on pose $\ell_1' = \ell_1$, $\ell_2' = \frac{3}{2}\ell_2$ et $\ell_3' = \frac{\sqrt{2}}{3}\ell_3$ pour obtenir $q = (\ell_1')^2 - (\ell_2')^2 - (\ell_3')^2$.

Signature d'une forme quadratique.

Les coefficients des carrés de formes linéaires dans la décomposition d'une forme quadratique *ne sont pas* invariants en général lorsqu'on change de base. Mais on va montrer que les nombres (p,r) intervenant dans l'énoncé de la proposition 6.29, eux, sont invariants.

Pour cela, on introduit la notion de forme positive et de forme négative.

Définition 6.30. Une forme quadratique q définie sur E est *positive* si $q(x) \geq 0$ pour tout $x \in E$, et *négative* si $q(x) \leq 0$ pour tout $x \in E$. La forme q est *définie positive* si q est positive et si de plus $q(x) = 0 \implies x = 0$. De même, q est *définie négative* si q est négative et si $q(x) = 0 \implies x = 0$. Les mêmes définitions s'appliquent aux formes bilinéaires symétriques.

Lemme 6.31. *Soit q une forme quadratique sur E de dimension finie, et soit $(e_1,\dots,e_n)$ une base q-orthogonale.*

1. *Si $q(e_i) \geq 0$ pour tout i, alors q est positive.*
2. *Si $q(e_i) > 0$ pour tout i, alors q est définie positive.*

Démonstration. Le premier point est évident d'après l'expression de $q(x)$ dans une base q-orthogonale. Pour le deuxième point, vérifions que q est de plus définie sous l'hypothèse $q(e_i) > 0$ pour tout i. Soit x un vecteur isotrope, et soit $x = \sum_i x_i e_i$ la décomposition de x dans la base q-orthogonale $(e_1, \dots, e_n)$. Alors :

$$q(x) = \sum_i x_i^2 q(e_i).$$

Comme cette somme est nulle et formée de nombres ≥ 0, tous les termes sont en fait nuls et donc $x_i = 0$ pour tout i, d'où $x = 0$. □

Théorème 6.32 (loi d'inertie de Sylvester). *Soit E un $\mathbb{R}$-e.v. de dimension finie, soit q une forme quadratique sur E. Alors le couple (p, r) d'entiers tels que la forme q s'exprime par :*

$$q = \ell_1^2 + \cdots + \ell_p^2 - \ell_{p+1}^2 - \cdots - \ell_{p+r}^2$$

pour un certain système libre $(\ell_1, \dots, \ell_{p+r})$ de formes linéaires sur E, ne dépend pas du système libre $(\ell_1, \dots, \ell_{p+r})$.

Remarque 6.33. Autrement dit, si $e = (e_1, \dots, e_n)$ est une base q-orthogonale, le nombre de vecteurs de base tels que $q(e_i) > 0$, le nombre de vecteurs de base tels que $q(e_i) < 0$, et le nombre de vecteurs de bases tels que $q(e_i) = 0$, sont trois entiers indépendants de la base q-orthogonale choisie.

Démonstration. Soit $(e_1, \dots, e_n)$ une base q-orthogonale, et supposons que $q(e_i) = 1$ pour $i = 1, \dots, p$ $q(e_i) = -1$ pour $i = p+1, \dots, p+r$, et $q(e_i) = 0$ pour $i > p + r$. On va montrer que le couple (p, r) ne dépend pas de la base q-orthogonale choisie, ce qui impliquera le résultat.

Pour cela, soit $F = \mathrm{Vect}(e_1, \dots, e_p)$ et $G = \mathrm{Vect}(e_{p+1}, \dots, e_n)$. Alors q est définie positive sur F et négative sur G d'après le lemme 6.31.

Soit $e' = (e'_1, \dots, e'_n)$ une autre base q-orthogonale, et soit (p', r') le couple d'entiers associé à q dans cette base. Soit $F' = \mathrm{Vect}(e'_1, \dots, e'_{p'})$ et $G' = \mathrm{Vect}(e'_{p'+1}, \dots, e'_n)$. La forme q est négative sur G'. Donc $F \cap G' = \{0\}$, et donc $p + (n - p') \leq n$ puisque $\dim F = p$ et $\dim G' = n - p'$. On a donc $p \leq p'$. En raisonnant de la même manière en échangeant les rôles de e et de e', on obtient aussi $p' \leq p$ donc $p = p'$.

Le raisonnement est identique pour montrer que $r = r'$, d'où le résultat. □

La proposition précédente justifie l'introduction de la définition suivante.

Définition 6.34. La *signature* d'une forme quadratique q définie sur un $\mathbb{R}$-e.v. de dimension finie est le couple (p, r) d'entiers tel que, dans une certaine base q-orthogonale $(e_1, \dots, e_n)$ de base duale $(\ell_1, \dots, \ell_n)$, on ait $q = \ell_1^2 + \cdots + \ell_p^2 - \ell_{p+1}^2 - \cdots - \ell_{p+r}^2$. La signature (p, r) ne dépend pas de la base q-orthogonale choisie.

Exemple 6.35. Un produit scalaire sur E de dimension n est de signature $(n, 0)$. La forme quadratique $q(\binom{x}{y}) = xy = \frac{1}{4}\big((x+y)^2 - (x-y)^2\big)$ est de signature $(1, 1)$.

Chapitre 7

Séries de fonctions

1 Suites de fonctions

Définition 7.1. Soit I un intervalle dans $\mathbb{R}$ et $E = \mathbb{R}$ ou $\mathbb{C}$, on appelle suite de fonctions à valeurs dans E une application $\mathbb{N} \longrightarrow \mathcal{F}(I, E)$ avec $\mathcal{F}(I, E)$ l'ensemble des applications $f : I \longrightarrow E$. On note $(f_n)_{n\in\mathbb{N}}$ une telle suite de fonctions avec $\begin{aligned} f_n : I &\longrightarrow E \\ x &\longmapsto f_n(x) \end{aligned}$

Remarque 7.2. On peut facilement généraliser cette notion :

1. En remplaçant E par un espace vectoriel normé (complet).
2. En remplaçant I par un disque dans $\mathbb{C}$. (Les théorèmes concernant la dérivation ou l'intégration ne sont plus forcément applicables tels quels dans ce cadre.)

Dans la suite on se placera sur une partie U du plan complexe telle que U est un disque dans $\mathbb{C}$, ou un intervalle de $\mathbb{R}$, qu'on notera I dans ce cas.

Définition 7.3. Soit U un disque dans $\mathbb{C}$ ou intervalle dans $\mathbb{R}$ et $E = \mathbb{R}$ ou $\mathbb{C}$, on appelle suite de fonctions à valeurs dans E une application $\mathbb{N} \longrightarrow \mathcal{F}(U, E)$ avec $\mathcal{F}(U, E)$ l'ensemble des applications $f : U \longrightarrow E$. On note $(f_n)_{n\in\mathbb{N}}$ une telle suite de fonctions avec $\begin{aligned} f_n : U &\longrightarrow E \\ x &\longmapsto f_n(x) \end{aligned}$

Exemples 7.4.

1. $\begin{aligned} f_n : [0, +\infty[&\longrightarrow \mathbb{R} \\ x &\longmapsto x^n \end{aligned}$
2. $\begin{aligned} f_n : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{1}{x^2+n^2} \end{aligned}$
3. $\begin{aligned} f_n : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \cos(nx) \end{aligned}$

1.1 Convergence simple

Définition 7.5. Soit U une partie du plan complexe telle que U est un disque dans $\mathbb{C}$, ou un intervalle de $\mathbb{R}$. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions définies sur U. On dit que $(f_n)_{n\in\mathbb{N}}$ est simplement convergente au point $x_0 \in U$ si la suite numérique $(f_n(x_0))_{n\in\mathbb{N}}$ converge vers une valeur notée $f(x_0)$ dans E.

Remarque 7.6. Cette propriété est locale.

Exemples 7.7. 1. Posons $I = [0, 1]$, la suite de fonctions $(f_n)_{n\in\mathbb{N}}$ donnée par $f_n(x) = x^n$ converge simplement vers la fonction f définie par $\left| \begin{array}{l} f(x) = 0 \text{ si } x < 1 \\ f(x) = 1 \text{ si } x = 1 \end{array} \right.$

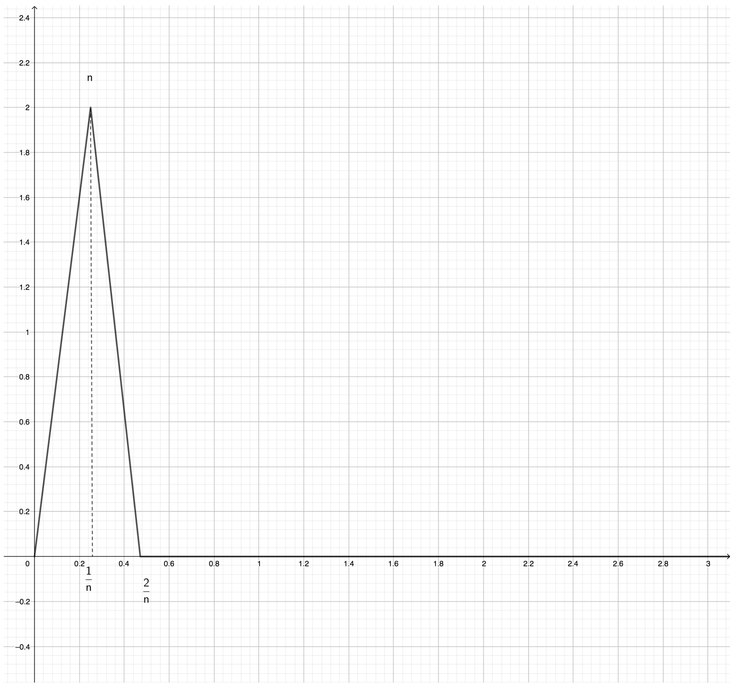
2. Posons $I =]0, +\infty[$ et $\begin{array}{rl} f_n : &]0, +\infty[\longrightarrow \mathbb{R} \\ & x \longmapsto \frac{nx^2}{1+nx} \end{array}$

   $(f_n)_{n\in\mathbb{N}}$ converge simplement sur I vers la fonction $f(x) = x$.

1.2 Convergence uniforme

Pourquoi la convergence simple ne suffit pas ?

1. $f_n : x \mapsto x^n$ est continue sur $[0, 1]$ mais $\lim\limits_{n\to+\infty}(f_n)$ ne l'est pas.
2. Soit $I = [0, 1[$ et f_n la fonction donnée par $f_n(x) = \left| \begin{array}{l} n^2x \text{ si } x \in [0, \frac{1}{n} \\ 2n - n^2x \text{ si } x \in [\frac{1}{n}, \frac{2}{n}] \\ 0 \text{ si } x \geq \frac{2}{n} \end{array} \right.$



   $\int_0^1 f_n(x)dx$ ne tend pas vers $\int_0^1 \lim\limits_{n\to+\infty} f_n(x)dx$

   $f_n(x)$ converge simplement sur $[0, 1]$ vers la fonction $f(x) = 0$.

   En effet $\left| \begin{array}{l} f(0) = 0 \\ \text{Soit } x > 0,\ f_n(x) = 0 \text{ si } n > \frac{2}{x} \end{array} \right.$

   $\int_0^1 f(x)dx = 0$ et $\int_0^1 f_n(x)dx = 1$

Définition 7.8. Soit U une partie du plan complexe telle que U est un disque dans $\mathbb{C}$, ou un intervalle de $\mathbb{R}$.
Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions définies sur U. On dit que la suite des fonctions $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers la fonction f sur U si $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ tel que si $n \geq N$ alors $\forall x \in U, |f_n(x) - f(x)| < \varepsilon$.

Remarque 7.9. La convergence uniforme est une propriété globale : elle dépend de manière essentielle de U.

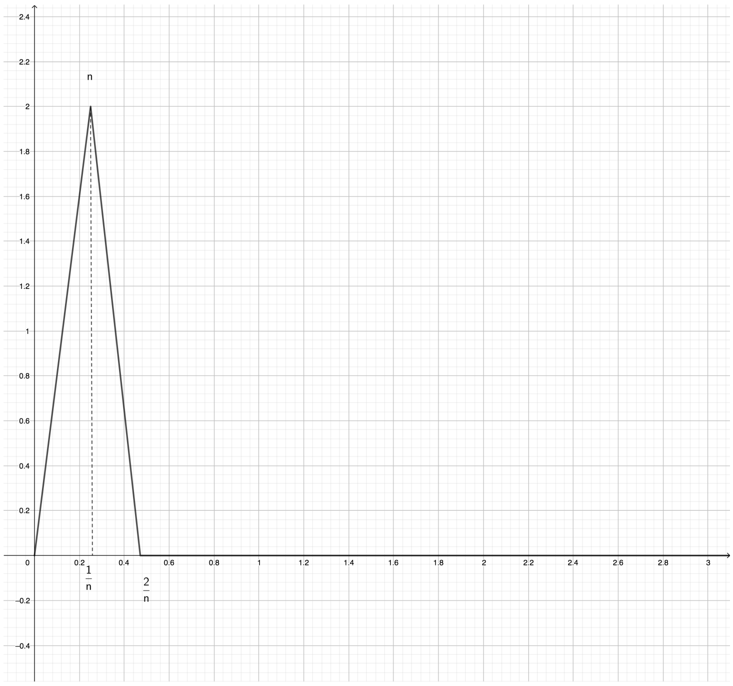
Autre exemple de propriété globale : être bornée.

Soit $\begin{aligned} f :]0,\ 1] &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{1}{x} \end{aligned}$

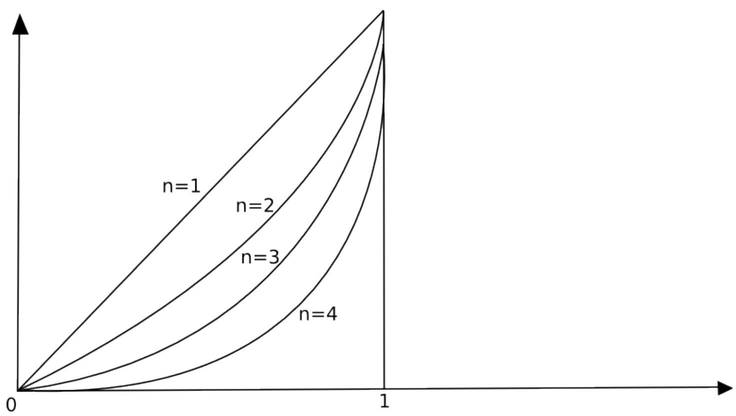
Sur tout intervalle $[a,\ b[\subset]0,\ 1]$ f est bornée mais f n'est pas bornée sur $]0,\ 1]$.

Exemples 7.10.

1. La suite de fonctions $(f_n)_{n\in\mathbb{N}}$ définie dans l'exemple 2 et représentée graphiquement ne converge pas uniformément vers 0 sur $[0,\ 1]$ car $f_n(\frac{1}{n}) = n \longrightarrow +\infty$



2. $x \longmapsto x^n$ ne converge pas uniformément sur $[0,\ 1]$ vers $f : x \longmapsto \left| \begin{array}{l} 0 \text{ si } x < 1 \\ 1 \text{ si } x = 1 \end{array} \right.$



3. $\begin{aligned} f_n : [0,\ 1] &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{1}{n+1}x^{n+1} \end{aligned}$ converge uniformément vers 0 sur l'intervalle $[0,\ 1]$.

$\left\|f_n\right\|_\infty^{[0,1]} = \sup_{x\in[0,1]} |f_n(x)| = \frac{1}{n+1} \longrightarrow_{n\to+\infty} 0$

Proposition 7.11. *Si $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers une fonction f sur U alors $(f_n)_{n\in\mathbb{N}}$ converge simplement vers f sur U.*

Démonstration. $\forall \varepsilon > 0$, $\exists N$ tel que si $n \geq N$, $\forall x \in U$, $|f_n(x) - f(x)| < \varepsilon$. On en déduit donc que pour $x \in U$ fixé, $\displaystyle f_n(x) \underset{n \to +\infty}{\longrightarrow} f(x)$ alors $(f_n)_{n\in\mathbb{N}}$ converge simplement vers f au point x. □

Soit U une partie du plan complexe telle que U est un disque dans $\mathbb{C}$, ou un intervalle de $\mathbb{R}$.
On pose $B(U, E) = \left\{f \in \mathcal{F}(U, E),\ \text{f est bornée sur } U\right\}$
$B(U, E)$ est un sous-espace vectoriel de $\mathcal{F}(U, E)$.

Définition 7.12. Soit U une partie du plan complexe telle que U est un disque dans $\mathbb{C}$, ou un intervalle de $\mathbb{R}$. On appelle norme infinie sur U et on note $\|\ \|_\infty^U$ l'application
$$\begin{aligned} \|\ \|_\infty^U :\ & B(U,E) \longrightarrow \mathbb{R}_+ \\ & f \longmapsto \sup_{x\in U} |f(x)| \end{aligned}$$

Proposition 7.13. *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de U dans E, alors $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur U si et seulement si*
- *$f_n - f$ est bornée sur U à partir d'un certain rang.*
- $\left\|f_n - f\right\|_\infty^U \to_{n\to+\infty} 0$

Démonstration.
- Si $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur U alors $\forall\varepsilon > 0$, $\exists N$ tel que si $n \geq N$, $|f_n(x) - f(x)| < \varepsilon$ pour $x \in U$.
- En particulier si $n \geq N$, $|f_n - f|$ est bornée par ε sur U.
- Si $n \geq N$, alors $||f_n - f||_\infty^U = \sup_{x\in U} |f_n(x) - f(x)| < \varepsilon$ donc $||f_n - f||_\infty^U \to_{n\to+\infty} 0$

Réciproquement, si $||f_n - f||_\infty^U \to_{n\to+\infty} 0$, alors $\forall\varepsilon > 0$, $\exists N$ tel que si $n \geq N$, $||f_n - f||_\infty^U < \varepsilon$.
ie $\sup_{x\in U} |f_n(x) - f(x)| < \varepsilon$, cela implique pour $n \geq N$, $\forall x \in U$, $|f_n(x) - f(x)| \leq ||f_n - f||_\infty^U < \varepsilon$ donc $(f_n)_{n\in\mathbb{N}}$ converge vers f uniformément sur U. □

1.3 Suites uniformément de Cauchy sur une partie U de $\mathbb{C}$

Définition 7.14. Soit U une partie du plan complexe telle que U est un disque dans $\mathbb{C}$, ou un intervalle de $\mathbb{R}$. On dit qu'une suite de fonctions $(f_n)_{n\in\mathbb{N}}$ de U dans E est uniformément de Cauchy sur U si $\forall\varepsilon > 0, \exists N \in \mathbb{N}$ tel que si q et $p \geq N$ alors $\forall x \in U$, $|f_p(x) - f_q(x)| < \varepsilon$.

Théorème 7.15. *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de U dans E.*
Alors $(f_n)_{n\in\mathbb{N}}$ converge uniformément sur U si et seulement si $(f_n)_{n\in\mathbb{N}}$ est uniformément de Cauchy sur U.

Démonstration. $(f_n)_{n\in\mathbb{N}}$ est uniformément de Cauchy sur U si $\forall\varepsilon > 0, \exists N \in \mathbb{N}$ tel que si q et $p \geq N$ alors $||f_p - f_q||_\infty^U < \varepsilon$.

— On montre que $(f_n)_{n\in\mathbb{N}}$ converge simplement vers une fonction f sur U en utilisant que $(f_n(x))_{n\in\mathbb{N}}$ est une suite de Cauchy dans E.
— On montre ensuite que $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur U.

□

2 Propriétés de la limite uniforme sur U d'une suite de fonctions

Soit U une partie du plan complexe telle que U est un disque dans $\mathbb{C}$, ou un intervalle de $\mathbb{R}$. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions à valeurs dans E, et convergeant uniformément vers f sur U.

2.1 Limite uniforme de fonctions bornées

Proposition 7.16. *Soit U une partie du plan complexe telle que U est un disque dans $\mathbb{C}$, ou un intervalle de $\mathbb{R}$. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions à valeurs dans E, et convergeant uniformément vers f sur U.*

Si les fonctions f_n sont bornées sur U, alors f est bornée sur U.

Démonstration. On sait que (f_n) est bornée et que $f - f_n$ est bornée à partir d'un certain rang.
$\sup_{x\in U} |f(x)| = \sup_{x\in U} |(f-f_n)(x) + f_n(x)| \leq \underbrace{\sup_{x\in U} |f_n(x)|}_{\leq M} + \underbrace{\sup_{x\in U} |(f-f_n)(x)|}_{\leq M'}$ donc f est bornée sur U. □

2.2 Interversion de limites

Théorème 7.17. *Soit U une partie du plan complexe telle que U est un disque dans $\mathbb{C}$, ou un intervalle de $\mathbb{R}$. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions à valeurs dans E, et convergeant uniformément vers f sur U. On suppose que $x_0 \in \overline{U}$ et que $\lim\limits_{x\to x_0} f_n(x) = a_n$. Alors la suite $(a_n)_{n\in\mathbb{N}}$ converge et*

$$\lim_{x\to x_0} f(x) = \lim_{x\to x_0} (\lim_{n\in\mathbb{N}} f_n(x)) = \lim_{n\in\mathbb{N}} (\lim_{x\to x_0} f_n(x)) = \lim_{n\to+\infty} a_n.$$

Démonstration.
— Soit $\varepsilon > 0$, $\exists N$ tel que si $p, q \geq N$, $||f_p - f_q||_\infty^U < \frac{\varepsilon}{3}$
— Fixons $p \geq N$ et $q \geq N$. Puisque $\lim_{x\to x_0} f_p(x) = a_p$ et $\lim_{x\to x_0} f_q(x) = a_q$, il existe un voisinage $V_{p,q}$ de x_0 tel que si $x \in U \cap V_{p,q}$, alors $|f_p(x) - a_p| < \frac{\varepsilon}{3}$ et $|f_q(x) - a_q| < \frac{\varepsilon}{3}$.
— On a donc, pour tout $\varepsilon > 0$, $\exists N$ tel que si $p, q \geq N$, alors, en utilisant $x \in U \cap V_{p,q}$, on a
$$|a_p - a_q| \leq |a_p - f_p(x)| + |f_p(x) - f_q(x)| + |f_q(x) - a_q| < \varepsilon.$$
— Ainsi, la suite (a_n) est de Cauchy, donc convergente dans E, vers une limite que l'on notera a.

— On fixe $N' \geq N$ tel que si $p \geq N'$, alors $|a - a_p| < \frac{\varepsilon}{3}$. Par ailleurs, pour tous $p, q \geq N' \geq N$, on a $||f_p - f_q||_\infty^U < \frac{\varepsilon}{3}$, donc on obtient $||f_p - f||_\infty^U \leq \frac{\varepsilon}{3}$. Enfinn, on sait qu'il existe un voisinage V_p de x_0 tel que pour tout $x \in V_p \cap U$, on ait $|a_p - f_p(x)| < \frac{\varepsilon}{3}$. On obtient alors, pour $p \geq N'$ et $x \in V_p \cap U$:

$$|f(x) - a| \leq |a - a_p| + |a_p - f_p(x)| + |f_p(x) - f(x)| < \varepsilon.$$

Ceci montre bien que $\lim\limits_{x \to x_0} f(x) = \lim\limits_{n \to +\infty} a_n$.

□

2.3 Limite uniforme de fonctions continues

On en déduit le théorème suivant .

Théorème 7.18. *Soit U une partie du plan complexe telle que U est un disque dans $\mathbb{C}$, ou un intervalle de $\mathbb{R}$. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions à valeurs dans E, et convergeant uniformément vers f sur U. Si f_n est continue en $x_0 \in U$, $\forall n \in \mathbb{N}$, alors f est continue au point x_0.*

Corollaire 7.19. *Si f_n est continue sur U , pour tout $n \in \mathbb{N}$, alors f est continue sur U.*

Remarque 7.20. Pour démontrer la continuité en x_0, il suffit de connaître la convergence uniforme de $(f_n)_{n\in\mathbb{N}}$ vers f sur un intervalle du type $]x_0 - \alpha,\ x_0 + \alpha[$ ou un disque ouvert de la forme $D(x_0, \varepsilon) = \{z \in \mathbb{C}, |z - x_0| < \varepsilon\}$.

Corollaire 7.21. *Si (f_n) converge simplement vers f sur U et* $\left|\begin{array}{l} f_n \text{ continue en } x_0 \in U,\ \forall n \in \mathbb{N} \\ f \text{ n'est pas continue en } x_0 \in U,\ \forall n \in \mathbb{N} \end{array}\right.$
Alors $(f_n)_{n\in\mathbb{N}}$ ne converge pas uniformément sur U vers f.

2.4 Intégrabilité d'une limite uniforme

Théorème 7.22. *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions à valeurs dans $\mathbb{R}$ convergeant uniformément vers f sur un intervalle $I = [a,\ b]$ avec f_n continue par morceau pour tout $n \in \mathbb{N}$, alors*

$$\int_a^b f_n(x)dx \underset{n \to +\infty}{\longrightarrow} \int_a^b f(x)dx$$

Attention :

C'est en général faux pour les intégrales généralisées.

Démonstration.

$$\left| \int_a^b f_n(x)dx - \int_a^b f(x)dx \right| = \left| \int_a^b (f_n(x) - f(x))dx \right|$$

(linéarité)

$$\left| \int_a^b f_n(x)dx - \int_a^b f(x)dx \right| \leq \int_a^b |f_n(x) - f(x)|dx \leq \int_a^b \|f - f_n\|_\infty^{[a,b]} dx$$

(positivité de l'intégrale)

$$\left|\int_a^b f_n(x)dx - \int_a^b f(x)dx\right| \leq (b-a)\left\|f - f_n\right\|_\infty^{[a,b]} \rightarrow_{n\to+\infty} 0$$

□

2.5 Dérivabilité de la limite d'une suite de fonctions

En utilisant le théorème précédent sur l'intégrabilité d'une suite de fonctions, on peut montrer que :

Théorème 7.23. *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de I dans E telle que :*

— *$\forall n \in \mathbb{N}$, f_n est de classe C^1.*
— *$\exists x_0 \in I$ tel que $(f_n)_{n\in\mathbb{N}}$ converge simplement au point (x_0).*
— *$(f_n')_{n\in\mathbb{N}}$ converge uniformément vers une fonction g sur I.*

Alors

1. *$(f_n)_{n\in\mathbb{N}}$ converge uniformément sur tout segment $[a,b] \subset I$ vers une fonction f.*
2. *La fonction f est de classe C^1 et $f' = g$.*

Attention :

Si f_n est de classe C^1 sur tout $n \in \mathbb{N}$ et $(f_n)_{n\in\mathbb{N}}$ converge vers f sur $I \not\Rightarrow f$ est de classe C^1 et $f' = \lim f_n'$

Exemple 7.24. $f_n(x) = \frac{x^{n+1}}{n+1}$ avec $I = [0,\ 1]$
On a vu que $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers 0 sur $[0,\ 1]$ mais on sait que $(f_n')_{n\in\mathbb{N}}$ ne converge pas uniformément sur $[0,\ 1]$.

Démonstration. Dans le cas où $E = \mathbb{R}$. Il suffit d'écrire $f_n(x) = f_n(x_0) + \int_{x_0}^x f_n'(t)dt$ et d'appliquer le théorème sur les intégrales, on en déduit que $(f_n(x)) \rightarrow_{n\to+\infty} f(x) = f(x_0) + \int_{x_0}^x g(t)dt$, la fonction $x \mapsto f(x_0) + \int_{x_0}^x g(t)dt$ est de classe C^1 et de dérivée $g(x)$.

$$|f_n(x) - f(x)| \leq \left|\int_{x_0}^x f_n'(t)dt - \int_{x_0}^x g(t)dt\right| + |f_n(x_0) - f(x_0)|$$

$$|f_n(x) - f(x)| \leq \left|\int_{x_0}^x |f_n'(t) - g(t)|dt\right| + |f_n(x_0) - f(x_0)|$$

$$|f_n(x) - f(x)| \leq \underbrace{\left|\int_{x_0}^x \left\|f_n' - g\right\|_\infty^I dt\right|}_{|x-x_0|\left\|f_n'-g\right\|_\infty^I} + |f_n(x_0) - f(x_0)|$$

$$\sup_{x\in[a,b]} |f_n(x) - f(x)| \leq \underbrace{\sup_{x\in[a,b]} |x - x_0| \left\|f_n' - g\right\|_\infty^I}_{\leq (b-a)\left\|f_n'-g\right\|_\infty^I \rightarrow_{n\to+\infty} 0} + \underbrace{|f_n(x_0) - f(x_0)|}_{\rightarrow_{n\to+\infty} 0}$$

□

Dans le cas complexe, on a

Définition 7.25. Soit U un disque ouvert dans $\mathbb{C}$ et f une fonction définie sur U. Soit $z_0 \in U$, on dit que f est dérivable au sens complexe (ou holomorphe) en z_0 si le taux d'accroissement $\dfrac{f(z)-f(z_0)}{z-z_0}$ admet une limite finie dans $\mathbb{C}$ lorsque z tend vers z_0.

Plus généralement, en utilisant la définition précédente on peut montrer que :

Théorème 7.26. *Soit U une partie du plan complexe telle que U est un disque ouvert dans $\mathbb{C}$, ou un intervalle ouvert de $\mathbb{R}$. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de U dans E telle que :*
- *— $\forall n \in \mathbb{N}$, f_n est dérivable sur U.*
- *— $(f_n)_{n\in\mathbb{N}}$ converge simplement vers une fonction f sur U.*
- *— $(f_n')_{n\in\mathbb{N}}$ converge uniformément vers une fonction g sur U.*

Alors
1. *$(f_n)_{n\in\mathbb{N}}$ converge uniformément sur toute partie bornée dans U vers f.*
2. *La fonction f est dérivable sur U et $f' = g$.*

3 Séries de fonctions

3.1 Convergence normale

Définition 7.27. Soit U une partie du plan complexe telle que U est un disque dans $\mathbb{C}$, ou un intervalle de $\mathbb{R}$. On appelle série de fonctions de U à valeurs dans $E = \mathbb{R}$ ou $\mathbb{C}$ (ou un espace vectoriel normé complet) une suite de fonctions $(S_n)_{n\in\mathbb{N}}$ définie par : $S_n = \sum_{k=0}^{n} f_k$ où les $f_n : U \to E$ sont des fonctions de U à valeurs dans E.
On appelle f_n *le terme général* de la série de fonctions, et S_n *la somme partielle*. Lorsque la série est convergente au point $t \in U$, on note $\sum_{k=0}^{+\infty} f_k(t)$ sa limite, qu'on appelle *somme* de la série de fonctions.

Remarque 7.28. Les notions de convergence simple et uniforme sur un intervalle ou un disque de la série sont celles des suites de fonctions appliquées à la suite des sommes partielles.

Définition 7.29. On dit qu'une série de fonctions $\sum f_n$ sur U à valeurs dans E converge normalement sur U lorsque la série numérique $\sum ||f_n||_\infty^U$ est convergente.

Proposition 7.30. *Une série de fonctions $\sum f_n$ sur U à valeurs dans E converge normalement sur U si et seulement si il existe une suite numérique $\alpha_n \geq 0$ telle que la série $\sum \alpha_n$ converge, et telle que, $\forall z \in U, \left|f_n(z)\right| \leq \alpha_n$.*

Théorème 7.31. *Si une série de fonctions $\sum f_n$ sur U à valeurs dans E converge normalement sur l'intervalle U, alors*

1. *la série numérique $\sum f_n(z)$ est absolument convergente pour tout $z \in U$;*
2. *la série de fonctions $\sum f_n$ converge uniformément sur U ;*
3. *la somme de la série de fonctions est bornée sur U.*

Remarque 7.32. La réciproque est fausse en général : si on pose $f_n(t) = \frac{\sin(nt)}{n}$, on peut montrer que $\sum f_n$ converge uniformément, mais pas normalement sur l'intervalle $[\alpha, \pi]$ pour $\alpha \in]0, \pi[$.

Démonstration. On montre que la série est une suite uniformément de Cauchy sur U. Posons $s_n = \sum_{k=0}^{n} ||f_k||_\infty^U$. Puisque cette suite est convergente, c'est une suite de Cauchy dans $\mathbb{R}$. Donc pour $\varepsilon > 0$, il existe $N \in \mathbb{N}$ tel que pour tous $p \geq q \geq N$, on ait : $\sum_{k=q+1}^{p} ||f_k||_\infty^{I'} = s_p - s_q = |s_p - s_q| < \varepsilon$. Mais alors, on a :

$$\forall t \in U, |S_p(t) - S_q(t)| = \left| \sum_{k=q+1}^{p} f_k(t) \right| \leq \sum_{k=q+1}^{p} \left| f_k(t) \right| \leq \sum_{k=q+1}^{p} ||f_k||_\infty^U < \varepsilon$$

ce qui montre que la suite de fonctions S_n est uniformément convergente sur U. □

3.2 Propriétés de la somme d'une série de fonctions

En appliquant les théorèmes sur la limite d'une suite de fonctions aux séries, on obtient les théorèmes suivants. Soit U une partie du plan complexe telle que U est un disque dans $\mathbb{C}$, ou un intervalle de $\mathbb{R}$. On considère des fonctions f_n définies sur U et à valeurs dans E.

Théorème 7.33. *On suppose que la série de fonctions $\sum f_n$ converge uniformément sur U. Alors*

1. *Si les fonctions f_n sont bornées sur U, alors $\sum_{n=0}^{+\infty} f_n$ est bornée sur U.*
2. *Si $x_0 \in \overline{U}$, et $\lim_{x \to x_0} f_n(x) = a_n$, alors la série $\sum a_n$ converge et $\lim_{x \to x_0} \sum_{n=0}^{+\infty} f_n(x) = \sum_{n=0}^{+\infty} a_n$*
3. *Si les fonctions f_n sont continues en $x_0 \in U$, alors $\sum_{n=0}^{+\infty} f_n$ est continue en $x_0 \in U$.*
4. *Si les fonctions f_n sont continues sur U, alors $\sum_{n=0}^{+\infty} f_n$ est continue sur U.*
5. *Si les fonctions f_n sont à valeurs réelles et continues par morceaux sur un intervalle réel I, alors pour tous $a, b \in I$, on a*

$$\int_a^b \sum_{n=0}^{+\infty} f_n(t)\, dt = \sum_{n=0}^{+\infty} \int_a^b f_n(t)\, dt.$$

Théorème 7.34. *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions d'un intervalle réel I dans E telle que :*
- *f_n est de classe C^1 sur I, pour tout $n \in \mathbb{N}$*
- *$\exists x_0 \in I$ tel que la série $\sum f_n(x_0)$ converge.*
- *la série de fonctions $\sum f_n'$ converge uniformément sur I.*

Alors

1. *la série $\sum f_n$ converge uniformément sur tout segment inclus dans I*
2. *La fonction somme $f = \sum_{n=0}^{+\infty} f_n$ est de classe C^1 sur I et $f' = \sum_{n=0}^{+\infty} f_n'$.*

Théorème 7.35. *Soit U une partie du plan complexe telle que U est un disque ouvert dans $\mathbb{C}$, ou un intervalle ouvert de $\mathbb{R}$. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de U dans E telle que :*
- *$\forall n \in \mathbb{N}$, f_n est dérivable sur U.*
- *$\sum f_n$ converge simplement vers une fonction f sur U.*
- *$\sum f_n'$ converge uniformément vers une fonction g sur U.*

Alors

1. *$\sum f_n$ converge uniformément sur toute partie bornée dans U vers f.*
2. *La fonction somme $f = \sum_{n=0}^{+\infty} f_n$ est dérivable sur U et $f' = g$.*

3.3 Lemme d'Abel

Proposition 7.36. *Soit U une partie du plan complexe telle que U est un disque ouvert dans $\mathbb{C}$, ou un intervalle ouvert de $\mathbb{R}$. Soient $(u_n)_{n\in\mathbb{N}}$ une suite numérique réelle et $(v_n)_{n\in\mathbb{N}}$ une suite de fonctions de U dans E.*

On pose $U_n = \sum_{k=0}^{n} u_k$ et $V_n = \sum_{k=0}^{n} v_k$.

Alors on a pour tous $p \geq q$ entiers

$$\sum_{n=q+1}^{p} u_n v_n = \sum_{n=q+1}^{p-1} (u_n - u_{n+1})V_n + u_p V_p - u_{q+1}V_q.$$

Théorème 7.37. *Soit U une partie du plan complexe telle que U est un disque ouvert dans $\mathbb{C}$, ou un intervalle ouvert de $\mathbb{R}$. Soit u_n une suite numérique réelle qui est positive, décroissante, et qui tend vers 0. Soit v_n une suite de fonctionsde U dans E telle que $\exists M$, $\forall n \in \mathbb{N}, \forall t \in U$, $\left|\sum_{k=0}^{n} v_k(t)\right| < M$. Alors la série de fonctions $\sum u_n v_n$ de terme général $u_n v_n$ est uniformément convergente sur U.*

Démonstration. On montre que la suite des sommes partielles de cette série est uniformément de Cauchy sur U : $\forall \varepsilon > 0$, $\exists N$ tel que si $p, q \geq N$, $t \in U$ alors $\left|\sum_{n=q+1}^{p} u_n v_n(t)\right| < \varepsilon$.

Or on sait que

$$\begin{aligned}
\left|\sum_{n=q+1}^{p} u_n v_n(t)\right| &= \left|\sum_{n=q+1}^{p} (u_{n+1}-u_n)V_n(t) + u_p V_p(t) - u_{q+1}V_q(t)\right| \\
&\le \sum_{n=q+1}^{p-1} |u_n - u_{n+1}|\,|V_n(t)| + |u_p|\,|V_p(t)| + |u_{q+1}|\,|V_{q+1}(t)| \\
&\le M\left(\sum_{n=q+1}^{p-1} |u_n - u_{n+1}| + u_p + u_{q+1}\right) \\
&\le M\left(\sum_{n=q+1}^{p-1} (u_n - u_{n+1}) + u_p + u_{q+1}\right) = M(u_{q+1} - u_p + u_p + u_{q+1}) = 2Mu_{q+1}.
\end{aligned}$$

Soit $\varepsilon > 0$. Soit N tel que si $n \geq N, u_n < \dfrac{\varepsilon}{2M}$. Alors si $p \geq q > N$ on a $\sup_{t\in I}\left|\sum_{n=q+1}^{p} u_n v_n(t)\right| < \varepsilon$. Cette série est uniformément de Cauchy sur I donc uniformément convergente sur I. □

Exemple 7.38. Application du théorème d'Abel uniforme. $\sum u_n v_n$ avec $(u_n)_{n\in\mathbb{N}}$ une suite positive, décroissante, qui tend vers 0,
et $v_n(\theta) = \sin(n\theta)$ avec $\theta \in]0, \pi[$.
Alors si $\alpha \in]0, \pi[$ la série de fonctions $\sum u_n \sin(n\theta)$ converge uniformément sur l'intervalle $[\alpha, 2\pi - \alpha]$.
Il suffit de montrer que $V_n(\theta)$ est uniformément bornée sur l'intervalle $[\alpha, 2\pi - \alpha]$

$$\begin{aligned}
V_n = \sum_{k=0}^{n} \sin(k\theta) &= \mathrm{Im}\left(\sum_{k=0}^{n} e^{ik\theta}\right) = \mathrm{Im}\left(\sum_{k=0}^{n} (e^{i\theta})^k\right) \\
&= \mathrm{Im}\left(\frac{1 - e^{i(n+1)\theta}}{1 - e^{i\theta}}\right) = \mathrm{Im}\left(\frac{e^{\frac{i(n+1)\theta}{2}}}{e^{\frac{i\theta}{2}}} \times \frac{e^{-\frac{i(n+1)\theta}{2}} - e^{\frac{i(n+1)\theta}{2}}}{e^{-\frac{i\theta}{2}} - e^{\frac{i\theta}{2}}}\right) \\
&= \mathrm{Im}\left(\frac{e^{\frac{in\theta}{2}} \times \sin\left(\frac{(n+1)\theta}{2}\right)}{\sin\left(\frac{\theta}{2}\right)}\right) = \mathrm{Im}\left(e^{\frac{in\theta}{2}}\right) \times \frac{\sin\left(\frac{(n+1)\theta}{2}\right)}{\sin\left(\frac{\theta}{2}\right)}.
\end{aligned}$$

Donc $|V_n(\theta)| = \dfrac{\left|\sin\left(\frac{n\theta}{2}\right) \times \sin\left(\frac{(n+1)\theta}{2}\right)\right|}{\left|\sin\left(\frac{\theta}{2}\right)\right|} \leq \dfrac{1}{\left|\sin\left(\frac{\theta}{2}\right)\right|} \leq \dfrac{1}{\left|\sin\left(\frac{\alpha}{2}\right)\right|}$

Chapitre 8

Séries entières

1 Définition, rayon de convergence

1.1 Définition

Définition 8.1. Une *série entière* est une série de fonctions de la variable complexe z, et de la forme $\sum a_n z^n$ où $(a_n)_{n\geq 0}$ est une suite de nombres réels ou complexes.

Quelques exemples :

$$\sum_{n=0}^{\infty} z^n \qquad \sum_{n=1}^{\infty} \frac{z^n}{n} \qquad \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

De plus, tout polynôme est un cas particulier de série entière, avec $a_n = 0$ pour n assez grand. On peut donc voir les séries entières comme une généralisation des polynômes.

1.2 Rayon de convergence.

Le domaine de convergence d'une série entière n'est pas quelconque, il satisfait nécessairement certaines contraintes que nous allons voir ci-dessous. Commençons par le lemme fondamental suivant.

Lemme 8.2. *Soit $\sum a_n z^n$ une série entière, et soit $z_0 \in \mathbb{C}$.*

1. *Si $\sum a_n z_0^n$ converge absolument, alors la série entière $\sum a_n z^n$ converge normalement dans le disque fermé $\overline{D}(0, |z_0|)$.*
2. *Si la suite $(a_n z_0^n)_{n\geq 0}$ est bornée, alors la série entière $\sum a_n z^n$ converge normalement dans tout disque fermé $\overline{D}(0, r)$ avec $r < |z_0|$. En particulier, la série entière $\sum a_n z^n$ converge simplement dans le disque ouvert $D(0, |z_0|)$.*

Démonstration. Pour le point 1, soit $z_0 \in \mathbb{C}$ tel que $\sum |a_n z_0^n|$ converge. Alors pour tout $z \in \overline{D}(0, |z_0|)$ on a :

$$|a_n z^n| = |a_n| \cdot |z|^n \leq |a_n| \cdot |z_0|^n.$$

Comme le membre de droite est le terme général d'une série convergente et qui ne dépend pas de z, on a bien la convergence normale de la série entière sur $\overline{D}(0, |z_0|)$.

Pour le point 2, il suffit d'après le point précédent de montrer que $\sum a_n z^n$ est absolument convergente dès que $|z| < |z_0|$. Le résultat est trivial si $z_0 = 0$, on suppose donc sans perte de généralité que $|z_0| > 0$, et on fixe $z \in \mathbb{C}$ tel que $|z| < |z_0|$. Alors :

$$|a_n z^n| = \left|\frac{z}{z_0}\right|^n \cdot |a_n z_0|^n.$$

Or la suite $(a_n z_0^n)_{n\geq 0}$ est supposée bornée, disons par M. Posons $\theta = |z/z_0|$, on obtient donc :

$$|a_n z^n| \leq M\theta^n$$

qui est le terme général d'une série convergente puisque $0 \leq \theta < 1$, d'où le résultat.

Le raisonnement ci-dessus a montré que la série $\sum a_n z^n$ converge absolument dès que $|z| < |z_0|$, ce qui entraîne en particulier la convergence simple de la série entière dans le disque ouvert $D(0, |z_0|)$. □

Définition 8.3. Soit $\sum a_n z^n$ une série entière. Le *rayon de convergence* de la série est le nombre positif ou égal à $+\infty$ défini par :

$$\begin{aligned} R &= \sup\{r \in [0, +\infty[\,:\, \sum |a_n| r^n \text{ converge}\} \\ &= \sup\{r \in [0, +\infty[\,:\, (|a_n| r^n)_{n\geq 0} \text{ est bornée}\}, \end{aligned}$$

Le *disque ouvert de convergence* est le disque ouvert $D(0, R)$.

Remarque 8.4.

1. Le rayon de convergence est bien défini car $0 \in A = \{r \in [0, +\infty[\,:\, \sum |a_n| r^n \text{ converge}\}$ donc A est non vide. Le rayon de convergence est fini si A est majoré et infini sinon.
2. L'égalité entre les deux bornes supérieures dans la définition provient du lemme 8.2.
3. Si $R = 0$, le disque ouvert $D(0, R)$ est vide : $D(0, R) = \emptyset$.
4. Si $R = +\infty$, le disque ouvert $D(0, R)$ est $D(0, R) = \mathbb{C}$.

Théorème 8.5. *Une série entière converge absolument en tout point de son disque ouvert de convergence. Elle converge normalement sur tout disque fermé inclus dans le disque ouvert de convergence. En particulier, la fonction somme est continue dans le disque ouvert de convergence.*

Démonstration. Soit $f(z) = \sum a_n z^n$ une série entière de rayon de convergence R. Il suffit de montrer que $f(z)$ converge normalement dans tout disque fermé $\overline{D}(0, r)$ avec $r < R$, et les deux autres points en découlent.

Soit donc $r < R$. Alors il existe r' tel que $r < r' < R$ et tel que $\sum |a_n|(r')^n$ converge. D'après le lemme 8.2, il s'ensuit que $f(z)$ est normalement convergente sur le disque fermé $\overline{D}(0, r')$, et donc en particulier sur $\overline{D}(0, r)$, ce qu'on voulait montrer. □

Remarque 8.6.

1. La preuve du théorème précédent montre aussi que l'ensemble $\{r \in [0, +\infty[\,:\, \sum |a_n| r^n \text{ converge}\}$ est un intervalle, de la forme $[0, R]$ ou de la forme $[0, R[$, et dans les deux cas R est le rayon de convergence.
2. Soit $R < \infty$ le rayon de convergence d'une série entière. *Il n'y a pas* de résultat général concernant la convergence de la série entière sur le cercle de centre 0 et de rayon R. De même, *il n'y a pas* de résultat général concernant la convergence normale sur le disque ouvert, ni sur le disque fermé de convergence. C'est ce que montrent les exemples suivants.

Exemple 8.7. On considère les séries entières suivantes :

$$U(z) = \sum_{n=0}^{\infty} z^n \qquad V(z) = \sum_{n=1}^{\infty} \frac{z^n}{n} \qquad W(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^2},$$

et on note R_U, R_V et R_W leurs rayons de convergence respectifs. Nous allons montrer que $R_U = R_V = R_W = 1$, et donc que les trois séries entières ont le même disque ouvert de convergence, qui est $D(0,1)$. Mais nous allons aussi montrer que les trois séries ne se comportent pas de la même façon sur le bord du disque.

Pour les trois séries, la règle de Cauchy s'applique en tout point z tel que $|z| \neq 1$:

$$\lim_{n\to\infty} \frac{|z|^{n+1}}{|z|^n} = |z| \qquad \lim_{n\to\infty} \frac{|z|^{n+1}}{n+1} \cdot \frac{n}{|z|^n} = |z| \qquad \lim_{n\to\infty} \frac{|z|^{n+1}}{(n+1)^2} \cdot \frac{n^2}{|z|^n} = |z|.$$

Donc les trois séries $U(z)$, $V(z)$ et $W(z)$ convergent absolument pour $|z| < 1$ et divergent pour $|z| > 1$. Il s'ensuit que $R_U = R_V = R_W = 1$.

Il est clair que la série $U(z)$ ne converge en aucun point du cercle $C(0,1)$ puisque son terme général ne tend pas vers 0. Il est clair aussi que la série $W(z)$ converge normalement sur le cercle $C(0,1)$ et donc sur le disque fermé $\overline{D}(0,1)$. Tandis que la série $V(z)$ diverge en $z = 1$ mais converge en -1 d'après le critère des séries alternées. On voit donc que toutes les situations sont possibles sur le cercle $C(0,1)$.

1.3 Calcul pratique du rayon de convergence.

Soit R le rayon de convergence d'une série entière $\sum a_n z^n$. Pour déterminer R, on peut déjà utiliser le lemme élémentaire suivant, qui découle immédiatement de la définition 8.3.

Lemme 8.8. *Soit R le rayon de convergence d'une série entière $\sum a_n z^n$, et soit $r \geq 0$ un réel.*

1. *Si la suite $(|a_n| r^n)_{n\geq 0}$ est bornée, alors $R \geq r$.*
2. *Si la suite $(|a_n| r^n)_{n\geq 0}$ est non bornée, alors $R \leq r$.*

Certaines règles permettent de calculer directement le rayon de convergence, cependant elles ne permettent pas de conclure dans tous les cas. On adopte, dans le reste du chapitre, la convention $\frac{1}{0} = +\infty$ et $\frac{1}{+\infty} = 0$.

Proposition 8.9.
Soit R le rayon de convergence d'une série entière $\sum a_n z^n$.

1. *Si $|a_n|^{1/n} \longrightarrow \ell$ pour $n \to \infty$, alors $R = 1/\ell$.*
2. *Si $a_n \neq 0$ à partir d'un certain rang, et si $|a_{n+1}|/|a_n| \longrightarrow \ell$ pour $n \to \infty$, alors $R = 1/\ell$.*

Démonstration.

1. Supposons $\ell > 0$. Fixons $\varepsilon > 0$ avec $\varepsilon < \ell$. Alors pour n assez grand, disons $n \geq N$, on a :

$$\ell - \varepsilon < |a_n|^{1/n} < \ell + \varepsilon.$$

donc

$$(\ell - \varepsilon)^n \leq |a_n| \leq (\ell + \varepsilon)^n.$$

Ainsi, pour tout réel r :

$$\big(r(\ell-\varepsilon)\big)^n \leq |a_n| r^n \leq \big(r(\ell+\varepsilon)\big)^n.$$

On en déduit que $\sum |a_n| r^n$ converge si $r < (\ell+\varepsilon)^{-1}$ et diverge si $r > (\ell-\varepsilon)^{-1}$, et donc :

$$\frac{1}{\ell+\varepsilon} \leq R \leq \frac{1}{\ell-\varepsilon}.$$

ceci étant vrai pour tout $\varepsilon > 0$ assez petit, il s'ensuit que $R = 1/\ell$. Un raisonnement analogue permet de conclure si $\ell = 0$ ou $\ell = +\infty$.

2. Supposons $\ell > 0$. Fixons $\varepsilon > 0$ avec $\varepsilon < \ell$. Alors pour n assez grand, disons $n \geq N$, on a :

$$\ell - \varepsilon < \frac{|a_{n+1}|}{|a_n|} < \ell + \varepsilon.$$

On obtient donc, par récurrence, pour tout $n \geq N$:

$$|a_N| \cdot (\ell-\varepsilon)^{n-N} \leq |a_n| \leq |a_N| \cdot (\ell+\varepsilon)^{n-N}.$$

Donc, pour tout réel r :

$$\frac{|a_N|}{(\ell-\varepsilon)^N} \cdot \big(r(\ell-\varepsilon)\big)^n \leq |a_n| r^n \leq \frac{|a_N|}{(\ell+\varepsilon)^N} \cdot \big(r(\ell+\varepsilon)\big)^n.$$

On en déduit que $\sum |a_n| r^n$ converge si $r < (\ell+\varepsilon)^{-1}$ et diverge si $r > (\ell-\varepsilon)^{-1}$, et donc :

$$\frac{1}{\ell+\varepsilon} \leq R \leq \frac{1}{\ell-\varepsilon}.$$

ceci étant vrai pour tout $\varepsilon > 0$ assez petit, il s'ensuit que $R = 1/\ell$. Un raisonnement analogue permet de conclure si $\ell = 0$ ou $\ell = +\infty$.

□

Remarque 8.10. Le cas 2 est en fait un cas particulier du cas 1 : le théorème de Césaro appliqué à la suite $\ln(a_n)$ permet de montrer que si $a_n \neq 0$ à partir d'un certain rang, et si $|a_{n+1}|/|a_n| \longrightarrow \ell$, alors $|a_n|^{1/n} \longrightarrow \ell$.
En effet, si la suite $\left|\frac{a_{n+1}}{a_n}\right|$ admet une limite ℓ, alors $\ln\left|\frac{a_{n+1}}{a_n}\right|$ tend vers $\ln(\ell)$ par continuité de la fonction ln (avec les conventions $\ln(0) = -\infty$ et $\ln(+\infty) = +\infty$. Le théorème de Césaro nous donne alors la convergence vers $\ln(\ell)$ de :

$$\frac{1}{n}[\sum_{k=0}^{n-1} \ln\left|\frac{a_{k+1}}{ak}\right| = \frac{1}{n}[\sum_{k=0}^{n-1} (\ln|a_{k+1}| - \ln|a_k|) = \frac{1}{n}\ln|a_n| - \frac{1}{n}\ln|a_0|.$$

On en déduit que $\lim \frac{1}{n}\ln|a_n| = \ln(\ell)$ et donc, par continuité de l'exponentielle, que

$$\lim \sqrt[n]{|a_n|} = \lim \exp\left(\frac{1}{n}\ln|a_n|\right) = \ell.$$

On peut donner une formule "explicite" du rayon de convergence en utilisant la notion de $\limsup$, qui est définie comme la borne supérieure de l'ensemble des valeurs d'adhérence d'une suite (avec la convention qu'on prend $+\infty$ si la suite n'est pas majorée).

Proposition 8.11. **(Hors programme)**
Soit R le rayon de convergence d'une série entière $\sum a_n z^n$. On a $R = \dfrac{1}{\limsup \sqrt[n]{|a_n|}}$.

Démonstration.

1. On suppose $\ell = \limsup \sqrt[n]{|a_n|} < +\infty$. Si $r < \dfrac{1}{\limsup \sqrt[n]{|a_n|}}$, alors $\sqrt[n]{|a_n|}r$ est une suite majorée à partir d'un certain rang par $r\ell < 1$. Par comparaison avec une série géométrique, on en conclut que la série $\sum |a_n| r^n$ est convergente.
2. On suppose $\ell = \limsup \sqrt[n]{|a_n|} > 0$. Soit $r > \frac{1}{\ell}$ alors $\sqrt[n]{|a_n|}r$ est une suite admettant une valeur d'adhérence $m > 1$. On en déduit que la suite $a_n r^n$ n'est pas bornée donc la série $\sum a_n r^n$ diverge.

□

Exemple 8.12. Soit la série entière $f(z) = \displaystyle\sum_{n=0}^{+\infty} z^n$ avec $a_n = 1$ pour tout entier $n \geq 0$. Alors $\frac{a_{n+1}}{a_n} = 1 \longrightarrow +\infty$, donc $R = 1$ (on retrouve bien le résultat connu pour les séries géométriques).

Exemple 8.13. On définit la fonction exponentielle sur $\mathbb{C}$ en posant :

$$\exp z = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

Alors la règle de Cauchy s'applique avec $a_n = \frac{1}{n!}$. On a $\frac{a_{n+1}}{a_n} = \frac{1}{n+1} \longrightarrow 0$, donc $R = +\infty$.

Exemple 8.14. Soit la série entière $f(z) = \displaystyle\sum_{n=0}^{+\infty} a_n z^n$ avec $a_n = n!$ pour tout entier $n \geq 0$. Alors la règle de Cauchy s'applique : $\frac{a_{n+1}}{a_n} = n + 1 \longrightarrow +\infty$, donc $R = 0$.

Exemple 8.15 (Séries entières lacunaires)**.** Soit la série entière $f(z) = \displaystyle\sum_{n=0}^{+\infty} 2^n z^{2n}$. Attention $a_n \neq 2^n$, mais on a : $a_{2n} = 2^n$ et $a_{2n+1} = 0$!
Cette série est géométrique, de raison $r = 2z^2$, et obtient ainsi que le rayon de convergence de f est $R = \frac{1}{\sqrt{2}}$.

On utilise aussi souvent le résultat suivant.

Proposition 8.16. *Si $\frac{P}{Q}$ est une fraction rationnelle, alors le rayon de convergence de la série entière $g(z)$ définie pour n assez grand par $\sum \frac{P(n)}{Q(n)} a_n z^n$ est égal au rayon de convergence de la série entière $f(z) = \displaystyle\sum_{n=0}^{+\infty} a_n z^n$.*
En particulier $\sum n a_n z^{n-1}$ et $\sum \frac{1}{n+1} a_n z^{n+1}$ ont même rayon de convergence que $\sum a_n z^n$.

Démonstration. Soit R le rayon de convergence de $f(z)$, et soit ρ le rayon de convergence de $g(z)$. On pose $b_n = \frac{P(n)}{Q(n)} a_n$.

1. Montrons que $\rho \geq R$, et pour cela supposons sans perte de généralité que $R > 0$. Soit $r \in [0, R[$ et choisissons $\theta \in]r, R[$. Alors :
$$|b_n r^n| = \frac{|P(n)|}{|Q(n)|} \cdot |a_n| r^n = \left(|a_n| \theta^n\right) \cdot \left(\frac{|P(n)|}{|Q(n)|} \cdot \left(\frac{r}{\theta}\right)^n\right).$$
Cette suite est bornée puisque c'est le produit d'une suite bornée et d'une suite tendant vers 0, car il existe $m \in \mathbb{Z}$ et $c > 0$ tel que $\dfrac{|P(n)|}{|Q(n)|} \sim cn^m$ et on a bien $n^m \left(\frac{r}{\theta}\right)^n \rightarrow 0$ et on a donc bien $\rho \geq r$. Ceci étant vrai quel que soit $r < R$, il s'ensuit que $\rho \geq R$.
2. Montrons l'inégalité $\rho \leq R$.
On remarque que (pour n assez grand), on a $a_n = \frac{Q(n)}{P(n)} b_n$, donc en appliquant le raisonnement précédent, on obtient $R \geq \rho$.

□

2 Somme et produit

Proposition 8.17. *Si $\sum a_n z^n$ et $\sum b_n z^n$ sont deux séries entières de rayon de convergence respectifs R_a et R_b alors la série entière $\sum (a_n + b_n) z^n$ est de rayon de convergence $R \geq R_m =$* $\min(R_a; R_b)$ *et sur le disque ouvert de centre* 0 *et de rayon R_m, on a :* $\displaystyle\sum_{n=0}^{+\infty} (a_n + b_n) z^n = \sum_{n=0}^{+\infty} a_n z^n + \sum_{n=0}^{+\infty} b_n z^n$.

On rappelle la définition de la série produit de deux séries absolument convergentes :

Définition 8.18. Soit $\sum a_n$ et $\sum b_n$ deux séries. On pose $c_n = \displaystyle\sum_{k=0}^{n} a_k b_{n-k}$. Alors on définit la série produit par $\sum c_n$.

Proposition 8.19. *Soit $\sum a_n$ et $\sum b_n$ deux séries absolument convergentes, alors la série produit $\sum c_n$ est absolument convergente et vérifie :*
$$\sum_{n=0}^{+\infty} c_n = \left(\sum_{n=0}^{+\infty} a_n\right) \times \left(\sum_{n=0}^{+\infty} b_n\right).$$

Démonstration. Posons $A_n = \sum_{k=0}^{n} a_k$, $B'_n = \sum_{k=0}^{n} b_k$, $C'_n = \sum_{k=0}^{n} c_k$, et $A'_n = \sum_{k=0}^{n} |a_k|$, $B'_n = \sum_{k=0}^{n} |b_k|$, $C'_n = \sum_{k=0}^{n} |c_k|$ et $C''_n = \sum_{k=0}^{n} c'_k$, avec $c'_n = \sum_{k=0}^{n} |a_k b_{n-k}|$. . On a $C'_n \leq C''_n$ et $C''_n \leq A'_n B'_n \leq C''_{2n}$. Puisque A'_n et B'_n convergent, elles sont majorées, et donc C''_n est croissante et majorée, et C'_n également, et C_n est donc absolument convergente.
De plus, on a $|A_n B_n - C_n| \leq A'_n B'_n - C''_n \rightarrow 0$. □

Proposition 8.20. *Si $\sum a_n z^n$ et $\sum b_n z^n$ sont deux séries entières de rayon de convergence respectifs R_a et R_b alors la série entière $\sum c_n z^n$ est de rayon de convergence $R \geq R_m = \min(R_a; R_b)$ et sur le disque ouvert de centre* 0 *et de rayon R_m, on a* $\displaystyle\sum_{n=0}^{+\infty} c_n z^n = \left(\sum_{n=0}^{+\infty} a_n z^n\right) \times \left(\sum_{n=0}^{+\infty} b_n z^n\right)$

3 Dérivabilité

Théorème 8.21. *Soit* $\sum a_n z^n$ *une série entière de rayon de convergence* $R > 0$*. Alors la somme de la série entière* $f(z) = \sum_{n=0}^{+\infty} a_n z^n$ *est continue et dérivable (holomorphe) sur le disque ouvert de centre* 0 *et de rayon* R*, de dérivée* $f(z) = \sum_{n=0}^{+\infty} n a_n z^{n-1}$*.*

Corollaire 8.22. *Soit* $f(z) = \sum_{n=0}^{+\infty} a_n z^n$ *une série entière de rayon de convergence* R*, et considérons la restriction de* f *à l'intervalle réel* $I =]-R, R[$*. Alors* f *est dérivable sur* I*, et sa dérivée en un point* $t \in I$ *est donnée par la somme de la série dérivée* $f'(t)$*.*

Corollaire 8.23. *La dérivée* $f'(z) = \sum_{n=0}^{+\infty} (n+1) a_{n+1} z^n$ *est une série entière de même rayon de convergence* R*.*
La somme d'une série entière est donc indéfiniment dérivable sur son disque ouvert de convergence.

Démonstration. On applique le théorème 7.35. On peut également donner une démonstration directe, en utilisant le théorème d'interversion de limites.
On fixe z tel que $|z| \leq r < R$ et on considère $h \in \mathbb{C}$ tel que $|h| \leq r - |z|$. On étudie alors $\dfrac{f(z+h) - f(z)}{h} = \sum_{n=0}^{+\infty} a_n u_n(h)$, avec $u_n(h) = \dfrac{(z+h)^n - z^n}{h}$. En utilisant la formule $a^n - b^n = (a-b)\sum_{k=0}^{n-1} a^k b^{n-1-k}$, on obtient $u_n(h) = \sum_{k=0}^{n-1} (z+h)^k z^{n-1-k}$. On observe que $\lim_{h \to 0} u_n(h) = \sum_{k=0}^{n-1} (z)^k z^{n-1-k} = n z^{n-1}$ et par ailleurs, on a $|z+h| \leq r$ et $|z| \leq r$ donc $|a_n u_n(h)| \leq |a_n| r^n$ est le terme général d'une série de fonctions de la variable h normalement convergente sur le disque fermé de centre 0 et de rayon $r - |z|$. On a donc, par le théorème d'interversion de limites :

$$\lim_{h \to 0} \frac{f(z+h) - f(z)}{h} = \lim_{h \to 0} \sum_{n=0}^{+\infty} a_n u_n(h) = \sum_{n=0}^{+\infty} a_n \lim_{h \to 0} u_n(h) = \sum_{n=0}^{+\infty} n a_n z^{n-1}.$$

□

Proposition 8.24. *Soit* f *et* g *deux séries entières, et soit* $\lambda \in \mathbb{C}$*. Alors :*

$$(f + \lambda g)' = f' + \lambda g' \qquad\qquad (fg)' = f'g + fg'.$$

Démonstration. L'égalité $(f + \lambda g)' = f' + \lambda g'$ est évidente. La deuxième égalité découle aussi d'un calcul direct. □

Dérivées successives.

Définition 8.25. Les dérivée successives $f^{(0)}, f^{(1)}, f^{(2)}, \ldots, f^{(p)}, \ldots$, d'une série entière $f(z) = \sum_{n=0}^{+\infty} a_n z^n$ sont définies récursivement par :

$$f^{(0)} = f \qquad\qquad f^{(p+1)} = \left(f^{(p)}\right)'.$$

En particulier, $f' = f^{(1)}$ et on note pour les premières dérivées : $f'' = f^{(2)}$ et $f''' = f^{(3)}$.

On vérifie facilement par récurrence sur p que $f^{(p)}$ a l'expression suivante, pour $f(z) = \sum_{n=0}^{+\infty} a_n z^n$:

$$f^{(p)}(z) = \sum_{n=0}^{+\infty} (n+1) \cdot \ldots \cdot (n+p) a_{n+p} z^n.$$

En utilisant un raisonnement par récurrence et le corollaire 8.22, on obtient aussi le résultat suivant.

Corollaire 8.26. *Soit $f(z) = \sum a_n z^n$ une série entière de rayon de convergence R, et considérons la restriction de f à l'intervalle réel $I =]-R, R[$. Alors f est de classe C^∞ sur I, et pour tout entier $p \geq 0$, la dérivée p^e de f sur I est donnée par :*

$$f^{(p)}(t) = \sum_{n \geq 0} (n+1) \cdot \ldots \cdot (n+p) a_{n+p} t^n.$$

Relation entre les coefficients d'une série entière et la fonction somme de la série.

Soit $f(z) = \sum a_n z^n$ une série entière de rayon de convergence $R > 0$. La première relation, élémentaire mais fondamentale, est la suivante :

$$a_0 = f(0).$$

Appliquons la même relation aux séries dérivées successives de f, on obtient :

$$a_1 = f'(0) \qquad 2a_2 = f''(0) \qquad \ldots \qquad p!a_p = f^{(p)}(0).$$

On obtient donc la formule suivante, valide pour tout entier $p \geq 0$:

$$a_p = \frac{f^{(p)}(0)}{p!}.$$

4 Fonctions analytiques

Définition 8.27. Soit $f : U \to \mathbb{C}$ une fonction définie sur un ouvert de $\mathbb{C}$. On dit que f est analytique au point $z_0 \in U$ s'il existe $r > 0$ tel que sur le disque ouvert de centre 0 et de rayon r, il existe une série entière $\sum a_n w^n$ de rayon de convergence $R \geq r$ telle que si $|z - z_0| < r$, alors $f(z) = \sum_{n=0}^{+\infty} a_n (z - z_0)^n$. On dit aussi que f est développable en série entière au point z_0.

Corollaire 8.28. *Une fonction analytique en z_0 est dérivable (holomorphe) dans un disque ouvert de centre z_0.*

Démonstration. Il suffit d'écrire $f(z) = g(z - z_0)$ avec $g(z) = \sum_{n=0}^{+\infty} a_n z^n$ qui est une fonction dérivable en 0. □

Théorème 8.29. Théorème des zéros isolés *Soit $\sum a_n z^n$ une série entière à coefficients non tous nuls, de rayon de convergence R et de somme f. Si f s'annule en 0, alors il existe un disque ouvert D de rayon $r > 0$ centré en 0 tel que si $z \in D \setminus \{0\}$, alors $f(z) \neq 0$.*

Démonstration. Soit m le plus petit entier non nul pour lequel $a_m \neq 0$. Un tel m existe, par hypothèse (avec $n \geq 1$). On a alors

$$f(z) = \sum_{n=0}^{+\infty} a_n z^n = \sum_{n=m}^{+\infty} a_n z^n = z^m g(z) \text{ avec } g(z) = \sum_{n=m}^{+\infty} a_n z^{n-m}.$$

La fonction g est la somme d'une série entière de rayon de convergence R, elle donc en particulier continue, et on a $g(0) = a_m \neq 0$. Il existe donc un disque ouvert de rayon $r > 0$ centré en 0 sur lequel g ne s'annule pas. On en déduit que dans ce disque D, on a $f(z) \neq 0$ si $z \neq 0$. □

Corollaire 8.30. *Si f est une fonction sur un ouvert U qui est analytique en $z_0 \in U$ et qui s'annule en z_0, alors z_0 est un zéro isolé ou f est identiquement nulle au voisinage de z_0.*

On en déduit

Théorème 8.31. Unicité du développement en série entière *Si $f : U \to \mathbb{C}$ est une fonction définie sur un ouvert de $\mathbb{C}$ développable en série entière au point z_0, alors le développement en série entière est unique.*

Démonstration. Si on a deux développements analytiques de f en un point z_0, on a l'égalité $\displaystyle\sum_{n=0}^{+\infty}(a_n - b_n)(z - z_0)^n = 0$ sur un disque ouvert de centre z_0, et donc pour tout $n \in \mathbb{N}$, on a $a_n = b_n$. □

Théorème 8.32. Prolongement analytique *Si f et g sont deux fonctions analytiques sur un ouvert* **connexe** *U de $\mathbb{C}$ qui coïncident sur une partie de U qui contient un point d'accumulation de U, alors elle sont égales.*

Démonstration. Soit $z_0 \in U$ qui soit un point d'accumulation de U. D'après le théorème précédent, f et g coïncident sur un disque ouvert de centre z_0. L'ensemble $F = \{z \in U, f(z) = g(z)\}$ est fermé, par définition, et ouvert car si $z_1 \in F$, z_1 est un point d'accumulation de U car U est connexe et n'admet donc pas de point isolé. □

Dans le cas particulier $z_0 = 0$, une fonction définie dans un disque ouvert centrée en 0 est développable en série entière en 0 si elle est égale, sur un disque ouvert centré en o et non vide, à la somme d'une série entière.

D'après le corollaire 8.26, les fonctions développables en série entière en 0 sont nécessairement de classe C^∞ sur $]-\eta, \eta[$, pour $\eta \leq r$ avec r le rayon d'un disque centré en 0 sur lequel la fonction coïncide avec la some dune série entière (de rayon de convergence plus grand que r). Cependant toutes les fonctions de classe C^∞ ne sont pas développables en série entière.

Exemple 8.33. La fonction définie par $f(0)=0$ et $f(x)=e^{-\frac{1}{x^2}}$ si $x\neq 0$ est de classe C^∞ sur $\mathbb{R}$ mais n'est pas développable en série entière en 0.

Remarque 8.34. On peut aussi construire des fonctions de classe C^∞ mais dont la série de Taylor en 0, $\sum \frac{f^{(n)}(0)}{n!}z^n$ est de rayon de convergence nul.

Les théorèmes précédents permettent notamment de démontrer :

Théorème 8.35. *Soit $f :]-r,r[\to \mathbb{C}$ avec $r>0$ une fonction développable en série entière sur $]-r,r[$. Alors :*

1. *Le développement en série entière est unique, donné par $\sum a_n t^n$ avec :*
$$\forall n\geq 0 \quad a_n=\frac{f^{(n)}(0)}{n!}.$$
2. *Si f est paire, tous les termes impairs de son développement sont nuls. Si f est impaire, tous les termes pairs de son développement sont nuls.*
3. *Si f est à valeurs réelles, alors tous les termes de son développement sont réels.*
4. *Les dérivées successives à tout ordre de f sont développables en série entière sur $]-r,r[$.*
5. *La primitive de f sur $]-r,r[$ qui s'annule en 0 est développable en série entière sur $]-r,r[$, et est donnée par :*
$$F(t)=\sum_{n\geq 1}\frac{a_{n-1}}{n}t^n.$$
6. *Il existe une unique fonction $\tilde{f} : D(0,r)\longrightarrow \mathbb{C}$, appelée prolongement analytique de f, qui est développable en série entière en 0 et dont la restriction à $\mathbb{R}$ est égale à f.*
7. *L'ensemble des fonctions développables en série entière sur $]-r,r[$ forme une $\mathbb{C}$-algèbre (c'est-à-dire, est stable par combinaisons linéaires à coefficients complexes et par produit).*

Théorème 8.36. Analycité de la somme d'une série entière *Si $\sum a_n z^n$ est une série entière de rayon de convergence $R>0$, alors sa somme est analytique sur le disque ouvert de convergence. et on a $f(z)=\displaystyle\sum_{n=0}^{+\infty}\frac{f^{(n)}(z_0)}{n!}(z-z_0)^n$.*

Démonstration. Montrons que la série $\sum \frac{f^{(p)}(z_0)}{p!}(z-z_0)^p$ converge pour tout z tel que $|z-z_0|<R-|z_0|$.
On a
$$|f^{(p)}(z_0)|=\left|\sum_{n=p}^{+\infty}\frac{n!}{(n-p)!}a_n z_0^{n-p}\right|\leq \sum_{n=p}^{+\infty}\frac{(n)!}{(n-p)!}|a_n||z_0|^{n-p}.$$

Donc pour tout $N \in \mathbb{N}$

$$\begin{aligned}
\sum_{p=0}^{N} \frac{|f^{(p)}(z_0)|}{p!}|z-z_0|^p &= \sum_{p=0}^{N}\sum_{n=p}^{+\infty} \frac{(n)!}{(n-p)!p!}|a_n||z_0|^{n-p}|z-z_0|^p \\
&= \sum_{n=0}^{+\infty} |a_n| \sum_{p=0}^{\min(n,N)} \frac{(n)!}{(n-p)!p!}|z_0|^{n-p}|z-z_0|^p \\
&\leq \sum_{n=0}^{+\infty} |a_n| \underbrace{\sum_{p=0}^{n} \frac{(n)!}{(n-p)!p!}|z_0|^{n-p}|z-z_0|^p} \\
&\leq \sum_{n=0}^{+\infty} |a_n| \quad (|z_0|+|z-z_0|)^n \\
&\leq \sum_{n=0}^{+\infty} |a_n| R^n < +\infty
\end{aligned}$$

D'après ce qui précède, la série double $\displaystyle\sum_{q=0,p=0}^{+\infty} \frac{(p+q)!}{p!q!} a_{p+q} z_0^q (z-z_0)^p$ est absolument convergente, et

$$\begin{aligned}
\sum_{q=0,p=0}^{+\infty} \frac{(p+q)!}{p!q!} a_{p+q} z_0^q (z-z_0)^p &= \sum_{p=0}^{+\infty} \left(\sum_{q=0}^{+\infty} \frac{(p+q)!}{q!} a_{p+q} z_0^q \right) \frac{(z-z_0)^p}{p!} \\
&= \sum_{p=0}^{+\infty} \frac{f^{(p)}(z_0)}{p!}(z-z_0)^p \\
&= \sum_{n=0}^{+\infty} \left(\sum_{p=0}^{n} \frac{(n)!}{(n-p)!p!} a_n z_0^{n-p} (z-z_0)^p \right) \\
&= \sum_{n=0}^{+\infty} a_n z^n = f(z)
\end{aligned}$$

□

On admet le théorème suivant

Théorème 8.37. Fonctions holomorphes
Soit U un ouvert convexe de $\mathbb{C}$. Toute fonction analytique sur U est holomorphe sur U. Réciproquement, toute fonction holomorphe sur U est analytique sur U.

5 La fonction exponentielle et et les fonctions trigonométriques

Définition 8.38. On considère la série entière $\sum \frac{z^n}{n!}$ et on appelle $\exp(z)$ sa somme. On définit les fonctions suivantes, pour z dans le disque de convergence.

$$\cos(z) = \frac{\exp(iz) + \exp(-iz)}{2}\ ,\ \sin(z) = \frac{\exp(iz) - \exp(-iz)}{2i}\ ,$$

et

$$ch(z) = \frac{\exp(z) + \exp(-z)}{2} = \cos(iz)\ ,\ \ sh(z) = \frac{\exp(z) - \exp(-z)}{2} = -i\sin(iz).$$

Proposition 8.39. *A l'intérieur du disque de convergence, on a immédiatement les propriétés suivantes :*

1. $\exp(iz) = \cos(z) + i\sin(z)$ *et* $\exp(z) = ch(z) + sh(z)$.
2. $\cos' = -\sin$ *et* $\sin' = \cos$.
3. $ch' = sh$ *et* $sh' = ch$.
4. $\cos(z) = \displaystyle\sum_{n=0}^{+\infty}(-1)^n \frac{z^{2n}}{(2n)!}$ *et* $\sin(z) = \displaystyle\sum_{n=0}^{+\infty}(-1)^n \frac{z^{2n+1}}{(2n+1)!}$.
5. $ch(z) = \displaystyle\sum_{n=0}^{+\infty} \frac{z^{2n}}{(2n)!}$ *et* $sh(z) = \displaystyle\sum_{n=0}^{+\infty} \frac{z^{2n+1}}{(2n+1)!}$.
6. $\cos(-z) = \cos(z)$ *et* $\sin(-z) = -\sin(z)$.
7. $ch(-z) = ch(z)$ *et* $sh(-z) = -sh(z)$.
8. $\cos(z + z') = \cos(z)\cos(z') - \sin(z)\sin(z')$.
9. $\sin(z + z') = \sin(z)\cos(z') + \cos(z)\sin(z')$.
10. $ch(z + z') = ch(z)\,ch(z') + sh(z)\,sh(z')$.
11. $sh(z + z') = sh(z)\,ch(z') + ch(z)\,sh(z')$.
12. $\cos^2(z) + \sin^2(z) = 1$ *et* $ch^2(z) - sh^2(z) = 1$.

Théorème 8.40.

1. *Le rayon de convergence de* $\exp(z)$ *est* $+\infty$.
2. *Pour tous* $a, b \in \mathbb{C}$, *on a* $\exp(a + b) = \exp(a)\exp(b)$.
3. *Pour tout* $z \in \mathbb{C}$, *on a* $\exp(z) \neq 0$ *et* $\exp(-z) = \frac{1}{\exp(z)}$.
4. *La fonction* $\exp$ *est égale à sa dérivée.*
5. *Pour tout* $z \in \mathbb{C}$, *on a* $\exp(\overline{z}) = \overline{\exp(z)}$. *La restriction de la fonction* $\exp$ *à l'axe réel est une fonction positive strictement croissante, dont la limite en* $-\infty$ *est* 0 *et la limite en* $+\infty$ *est* $+\infty$.
6. *Sur l'intervalle* $]0, 2[$, *la fonction* $x \mapsto \sin(x)$ *est positive, et la fonction* $x \mapsto \cos(x)$ *est de classe* C^∞ *et strictement décroissante.*
7. *Il existe un nombre réel strictement positif* $\pi \in]0, 4[$ *tel que* $\exp(i\frac{\pi}{2}) = i$.
8. *La fonction* $\exp$ *est périodique, de période* $2\pi i$ *et on a* $\exp(z) = \exp(z')$ *si et seulement si* $\frac{z-z'}{2\pi i}$ *est un entier relatif.*

9. *L'application* exp *est surjective de* $\mathbb{C}$ *dans* $\mathbb{C}^*$, *en particulier, l'application* $\mathbb{R} \longrightarrow \mathbb{U}$ *donnée par* $t \mapsto \exp(it)$ *est une fonction surjective sur le cercle et* 2π*-périodique.*

Démonstration.

1. On utilise que $R = \dfrac{1}{\lim \frac{a_{n+1}}{a_n}}$, avec $a_n = \frac{1}{n!}$.
2. On a, par absolue convergence,
$$\sum_{k=0}^{+\infty} \frac{a^k}{k!} \sum_{m=0}^{+\infty} \frac{b^m}{m!} = \sum_{n=0}^{+\infty} \frac{1}{n!} \sum_{k=0}^{+\infty} \frac{n!}{k!(n-k)!} a^k b^{n-k} = \sum_{n=0}^{+\infty} \frac{(a+b)^n}{n!}.$$
3. On a $\exp(z)\exp(-z) = \exp(z-z) = \exp(0) = 1$.
4. On a $\exp'(z) = \sum_{n=0}^{+\infty} n \frac{z^{n-1}}{n!} = \sum_{n=1}^{+\infty} \frac{z^{n-1}}{(n-1)!} = \exp(z)$.
5. Puisque tous les coefficients de la série entière définissant exp sont positifs (et non tous nuls) la restriction de exp à $\mathbb{R}$ est strictement croissante sur $\mathbb{R}_+$. De plus, pour $x \geq 0$, on a $\exp(x) \geq 1 + x$, ce qui montre que sur R_+ la fonction est positive et tend vers $+\infty$. En utilisant $\exp(-x) = \frac{1}{\exp(x)}$, on en déduit le résultat pour $x \in \mathbb{R}_-$.
6. On utilise la fonction $\cos(x) = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n}}{(2n)!}$ et la fonction $\sin(x) = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$. Les fonctions cos et sin sont holomorphes sur $\mathbb{C}$, donc leur restriction à $\mathbb{R}$ sont de classe C^∞, et on a $\cos' = -\sin$. En utilisant la série alternée à terme général décroissant en valeur absolue qui définit $\sin(x)$, on a que $\sin x$ est minoré par la somme des deux premiers termes de la série : $\sin x \geq g(x) = x - \frac{x^3}{6}$. En étudiant g sur $]0, 2[$, on voit que g est strictement croissante sur cet intervalle, et en particulier, pour tout $x \in]0, 2[$, on a $\sin(x) \geq g(x) > 0$. On en déduit donc que que la fonction cos est strictement décroissante sur $]0, 2[$.
7. Montrons qu'il existe un unique élément $x_0 \in]0, 2[$ tel que $\cos(x_0) = 0$ et que l'on a alors $\sin(x_0) = 1$. On définira alors $\pi := 2x_0$ On a $\cos(0) = 1 > 0$ et $\cos(2) - 1 = \sum_{n=1}^{+\infty} (-1)^n \frac{2^{2n}}{(2n)!}$. Cette série est une série alternée dont le terme général $u_n = (-1)^n \frac{2^{2n}}{(2n)!}$ est décroissant en valeur absolue et dont le premier terme est négatif. On a donc $\cos(2) - 1 \leq u_1 + u_2 = -\frac{4}{3}$. Ainsi, par le théorème des valeurs intermédiaires, il existe un élément $x_0 \in]0, 2[$ tel que $\cos(x_0) = 0$, et puisque que la fonction cos est strictement décroissante sur $]0, 2[$, et donc injective, on a unicité de ce point $x_0 \in]0, 2[$ tel que $\cos(x_0) = 0$.
   En utilisant la relation $\cos^2(z) + \sin^2(z) = 1$, on en déduit que $\sin(x_0) = \pm 1$, mais puisque sin est positive sur $]0, 2[$ alors $\sin(x) = 1$ pour tout point $x \in]0, 2[$ tel que $\cos(x) = 0$. En particulier, on a $\sin(x_0) = 1$
8. On a $\exp(2\pi i) = (\exp(i\frac{\pi}{2}))^4 = i^4 = (-1)^2 = 1$, donc $\exp(z + 2n\pi i) = \exp(z)$ pour tout $z \in \mathbb{C}$ et tout $n \in \mathbb{Z}$.
   Supposons à présent que $\exp(z) = 1$. Alors, en écrivant $z = x + iy$, avec $x, y \in \mathbb{R}$, on obtient $\exp(z) = \exp(x)\exp(iy)$. en particulier, on a $1 = |\exp(z)| = |\exp(x)||\exp(iy)| = |\exp(x)| = \exp(x)$. Puisque la restriction de l'exponentielle à la droite réelle est injective, cela implique $x = 0$.
   On cherche donc $y \in \mathbb{R}$ tel que $\exp(iy) = 1$. Supposons à présent que y n'est pas un multiple de 2π. Puisque $\exp(z + 2n\pi i) = \exp(z)$ pour tout entier n, on peut supposer que $y \in]0, 2\pi[$, et on a donc $\frac{y}{4} \in]0, \frac{\pi}{2}[$. D'après la question précédente,

on sait alors que $\exp(i\frac{y}{4}) = u + iv$ avec $u = \cos(\frac{y}{4}) > 0$ et $v = \sin(\frac{y}{4}) > 0$. On a alors $\exp(iy) = (u + iv)^4 = (u^4 - 6u^2v^2 + v^4) + i(4uv(u^2 - v^2))$. La partie imaginaire est nulle si et seulement si $u^2 = v^2$, et puisque $u^2 + v^2$, on a alors $u^2 = v^2 = \frac{1}{2}$ et donc $\exp(iy) = (u^4 - 6u^2v^2 + v^4) = -1 \neq 1$. On vient donc de montrer que pour tout $y \in]0, 2\pi[$, on a $\exp(iy) \neq 1$.

9. Soit $w \in \mathbb{C}^*$. On sait qu'il existe $x \in \mathbb{R}$ tel que $\exp(x) = |w|$. On cherche à présent $y \in \mathbb{R}$ tel que $\exp(iy) = z = \frac{w}{|w|}$. Posons $z = u + iv$. On a alors $u^2 + v^2 = 1$, et on cherche $y \in \mathbb{R}$ tel que $\cos(y) = u$ et $\sin(y) = v$.
   (a) On suppose $u \geq 0$ et $v \geq 0$. Par la propriété de décroissance de la fonction continue $\cos(t)$ sur $[0, \frac{\pi}{2}]$, on sait qu'il existe un unique $y \in [0, \frac{\pi}{2}]$ tel que $\cos(y) = u$. Dans ce cas, puisque $\sin(y) \geq 0$, on a $\sin(y) = \sqrt{1 - \cos^2(y)} = \sqrt{1 - u^2} = \sqrt{v^2} = v$. On a donc bien trouvé dans ce cas $y \in [0, \frac{\pi}{2}]$ tel que $\exp(iy) = z$.
   (b) On suppose $u < 0$ et $v \geq 0$. Alors $-iz = -i(u + iv) = v - iu$. D'après ce qui précède, il existe $y' \in [0, \frac{\pi}{2}]$ tel que $\exp(iy') = -iz$. Alors $\exp(i(y' + \frac{\pi}{2})) = -iz \times i = z$, et on pose dans ce cas $y = y' + \frac{\pi}{2}$.
   (c) On suppose à présent $v < 0$. Alors $-z$ est dans l'un des deux cas précédents, donc il existe $y' \in [0, \frac{\pi}{2}]$ tel que $\exp(iy') = -z$ Alors $\exp(i(y' + \pi)) = -z \times (-1) = z$, et on pose dans ce cas $y = y' + \pi$.

   On a donc bien montré que pour tout $z \in \mathbb{U}$ il existe $y \in [0, 2\pi]$ tel que $\exp(iy) = z$.

□

6 Logarithme complexe

Définition 8.41. On appelle *logarithme néperien* la fonction $\ln :]0, +\infty[\to \mathbb{R}$ qui est la bijection réciproque de la restriction de la fonction exp à $\mathbb{R}$.

Définition 8.42. Soit $z \in \mathbb{C}^*$, on dit que $\theta \in \mathbb{R}$ est un argument de z si $\exp(i\theta) = \frac{z}{|z|}$

Proposition 8.43.

1. *Si θ est un argument de z, l'ensemble des arguments de z est $\{\theta + 2n\pi, n \in \mathbb{Z}\}$.*
2. *Si θ est un argument de $z \in \mathbb{C}^*$, alors $\exp(\ln|z| + i\theta) = z$.*

On se pose la question de savoir si on peut choisir un argument a priori parmi les autres pour tout nombre complexe z. Nous allons voir que cela ne peut se faire continûment sur $\mathbb{C}^*$, mais qu'il est possible de déterminer une fonction : $\mathrm{Arg}_p : \mathbb{C} \setminus \mathbb{R}_- \to]-\pi, \pi[$ qui à un nombre complexe associe continûment un argument, appelé argument principal, dans $]-\pi, \pi[$

Définition 8.44. Soit U un ouvert connexe de $\mathbb{C}^*$, on appelle détermination du logarithme sur U toute fonction continue $f : U \to \mathbb{C}$ telle que $\exp(f(z)) = z$.

Proposition 8.45.

1. *Si on considère le disque ouvert de centre* 1 *et de rayon* 1*, il existe une détermination du logarithme, donnée par :* $f(1+z) = \sum_{n=0}^{+\infty}(-1)^n \frac{z^{n+1}}{n+1}$ *pour* $|z| < 1$*.*
2. *Si on considère* $U = \mathbb{C}^*$ *(ou le disque ouvert de centre* 0 *et de rayon* 1 *et privé de* 0*), il n'existe pas de détermination du logarithme sur cet ouvert.*
3. *Si* f_1 *et* f_2 *sont deux déterminations du logarithme sur un ouvert connexe* U*, alors il existe* $n \in \mathbb{Z}$ *tel que* $f_1 - f_2 = 2n\pi i$*.*

Démonstration.

1. Cette série entière est convergente sur le disque ouvert de centre 1 et de rayon 1, donc continue, et de dérivée $f'(1+z) = \sum_{n=0}^{+\infty}(-z)^n = \frac{1}{1+z}$ pour $|z| < 1$. Comme $f(1) = 0$, et que leur dérivées sont égales, la fonction $f(1+x)$ coïncide avec la fonction $\ln(1+x)$ sur $]0,2[$. Puisque exp et f sont analytiques, leur composée $\exp \circ f$ est encore analytique (car holomorphe), donc par le principe du prolongement analytique, on a $\exp(f(1+z)) = 1+z$ pour $|z| < 1$.
2. Soit $\theta \in \mathbb{R}$. On pose $v(\theta) = f(\exp(i\theta))$. On a alors v continue et $\exp(v(\theta)) = \exp(i\theta)$, donc on a $\theta \mapsto \frac{v(\theta)-i\theta}{2\pi i}$ qui est une fonction continue de $\mathbb{R}$ dans $\mathbb{Z}$, donc elle est constante. Ainsi il existe $n \in \mathbb{Z}$ tel que pour tout $\theta \in \mathbb{R}$, on a $v(\theta) = i\theta + 2n\pi i$. Mais alors $v(\theta + 2\pi) = i\theta + 2\pi i + 2n\pi i = v(\theta) = i\theta + 2n\pi i$ ce qui aboutit à une contradiction.
3. On a $\exp(f_1(z) - f_2(z)) = 1$ donc pour tout z il existe $n(z) \in \mathbb{Z}$, tel que $f_1(z) - f_2(z) = 2n(z)\pi i$ et la fonction $z \mapsto n(z)$ est ainsi continue d'un ouvert connexe de $\mathbb{C}^*$ dans $\mathbb{Z}$, donc constante.

□

Proposition 8.46. *La fonction* $z \mapsto Arg_p(z) = 2\arctan(\frac{y}{|z|+x})$ *si* $z = x+iy$ *est continue sur l'ouvert* $\mathbb{C} \setminus \mathbb{R}_-$ *et* $Arg_p(\exp(i\theta)) = \theta$ *pour tout* $\theta \in]-\pi,\pi[$

Démonstration. La fonction arctan est définie sur $\mathbb{R}$ comme bijection réciproque de la fonction continue impaire et strictement croissante $\tan = \frac{\sin}{\cos} :]-\frac{\pi}{2}, +\frac{\pi}{2}[\to \mathbb{R}$, elle donc continue sur $\mathbb{R}$ et $z \mapsto \frac{\mathrm{Im}(z)}{|z|+\mathrm{re}(z)}$ est bien définie et continue de $\mathbb{C} \setminus \mathbb{R}_-$ dans $\mathbb{R}$.

On utilise les formules $\cos(\theta) = \frac{1-\tan^2(\frac{\theta}{2})}{1+\tan^2(\frac{\theta}{2})}$ et $\sin(\theta) = \frac{2\tan(\frac{\theta}{2})}{1+\tan^2(\frac{\theta}{2})}$ d'où l'on tire

$$\frac{\sin(\theta)}{1+\cos(\theta)} = \frac{2\tan(\frac{\theta}{2})}{1-\tan^2(\frac{\theta}{2})+1+\tan^2(\frac{\theta}{2})} = \tan(\frac{\theta}{2}).$$

□

Définition 8.47. On appelle *détermination principale du logarithme* sur l'ouvert $\mathbb{C} \setminus \mathbb{R}_-$ l'application $f : \mathbb{C} \setminus \mathbb{R}_- \to \mathbb{C}$ définie par $\log(z) = \ln(|z|) + i\mathrm{Arg}_p(z)$, avec $\mathrm{Arg}_p(z) = 2\arctan(\frac{y}{|z|+x})$ si $z = x+iy$.

Cette définition donne pour tout nombre complexe dans $\mathbb{C} \setminus \mathbb{R}_-$ un argument dans $]-\pi, \pi[$, et on définit alors par π l'argument d'un nombre sur la demi-droite $\mathbb{R}_-^*$.

Définition 8.48. Soit A, B, C, D quatre points dans le plan complexe. On appelle angle (orienté) entre le vecteur $\vec{AB}$ et le vecteur $\vec{CD}$ un argument du nombre complexe $\frac{z_D - z_C}{z_B - z_A}$ où z_M désigne l'affixe complexe du point M

7 Exemples et applications

7.1 La fonction $(1+x)^\alpha$.

Partons de la série géométrique, qui a un rayon de convergence égal à 1 :

$$\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n.$$

Puis dérivons l'égalité précédente par rapport à une variable t réelle, on obtient :

$$\frac{1}{(1-t)^2} = \sum_{n=0}^{\infty} (n+1)t^n. \tag{8.1}$$

Or la fonction $1/(1-z)^2$ est développable en série entière avec un rayon de convergence au moins égal à 1 comme produit de fonctions développables en séries entières. Pour étendre la formule précédente de la variable réelle à la variable complexe, on utilise le point 1 du théorème 8.35. Puisque les deux fonctions, de part et d'autre du signe égal, sont développables en série entière et égales sur $]-1, 1[$, leurs coefficients sont égaux. Donc les deux fonctions sont égales dans le disque de convergence. On obtient donc la formule :

$$|z| < 1 \qquad \frac{1}{(1-z)^2} = \sum_{n=0}^{\infty} (n+1)z^n.$$

On en déduit aussi une expression pour la somme de la série entière $\sum_{n=1}^{\infty} nz^n$:

$$\sum_{n=1}^{\infty} nz^n = \frac{1}{(1-z)^2} - \frac{1}{1-z} = \frac{z}{(1-z)^2}.$$

Remarque 8.49. Cette formule est utile en probabilités. En effet, si X est une variable aléatoire entière distribuée suivant la loi géométrique de paramètre p, c'est-à-dire $\mathbb{P}(X = n) = (1-p)p^n$ pour tout entier n, alors son espérance est donnée par :

$$\mathbb{E}(X) = \sum_{n=1}^{\infty} n(1-p)p^n = \frac{p}{1-p}.$$

En posant $z' = -z$ dans le développement de $1/(1-z)^2$, on obtient :

$$|z| < 1 \qquad \frac{1}{(1+z)^2} = \sum_{n=0}^{\infty} (-1)^n (n+1) z^n.$$

Par dérivations successives, il est facile d'obtenir par récurrence la formule suivante, pour tout valeur entière et positive de α :

$$\begin{aligned} |z| < 1 \qquad (1+z)^{-\alpha} &= \sum_{n=0}^{\infty} \frac{(-1)^n (n+1) \cdot \ldots \cdot (n+\alpha-1)}{(\alpha-1)!} z^n \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n \cdot \alpha \cdot (\alpha+1) \cdot \ldots \cdot (\alpha+n-1)}{n!} z^n. \end{aligned}$$

On va généraliser cette formule à toute valeur réelle de α. Définissons le binôme généralisé comme suit :

$$\text{Pour } \alpha \in \mathbb{R} \text{ et } k \in \mathbb{N} : \quad \binom{\alpha}{k} = \frac{\alpha \cdot (\alpha-1) \cdot \ldots \cdot (\alpha-k+1)}{k!}.$$

Remarquons qu'on retrouve le binôme habituel si α et k sont entiers avec $k \le \alpha$. Si α est entier et $k > \alpha$ alors $\binom{\alpha}{k} = 0$, tandis que $\binom{\alpha}{k}$ ne s'annule pour aucune valeur de k si α n'est pas entier ou si α est négatif.

Proposition 8.50. *La formule suivante est valide pour toute valeur réelle de α :*

$$t \in \mathbb{R}, \quad |t| < 1 \qquad (1+t)^{\alpha} = \sum_{k=0}^{\infty} \binom{\alpha}{k} t^k. \tag{8.2}$$

Si α est entier et positif, alors tous les termes de la somme pour $k > \alpha$ dans (8.2) sont nuls, et on retrouve la formule du binôme de Newton. Pour un exposant négatif, on a :

$$\binom{-\alpha}{k} = \frac{-\alpha \cdot (-\alpha-1) \cdot \ldots \cdot (-\alpha-k+1)}{k!} = \frac{(-1)^k \cdot \alpha \cdot (\alpha+1) \cdot \ldots \cdot (\alpha+k-1)}{k!}$$

donc on obtient bien la formule vue précédemment pour $(1+z)^{-\alpha}$ avec $\alpha \in \mathbb{N}$.

Démonstration de la proposition 8.50. Supposons $\alpha \notin \mathbb{N}$, sinon le résultat découle directement de la formule de binôme de Newton. Soit :

$$g(t) = \sum_{k=0}^{\infty} \binom{\alpha}{k} t^k.$$

Cette série entière a pour rayon de convergence 1, car :

$$\left| \binom{\alpha}{k+1} \Big/ \binom{\alpha}{k} \right| = \frac{|\alpha - k|}{k+1} \longrightarrow_{k\to\infty} 1.$$

La dérivée de $g(t)$ s'obtient en dérivant la série terme à terme :

$$g'(t) = \sum_{k=0}^{\infty} (k+1) \binom{\alpha}{k+1} t^k.$$

Calculons la quantité $(1+t)g'(t)$:

$$(1+t)g'(t) = \sum_{k=0}^{\infty}\left(\binom{\alpha}{k+1} + k\binom{\alpha+1}{k+1}\right)t^k = \sum_{k=0}^{\infty}\alpha\binom{\alpha}{k}t^k = \alpha g(t).$$

Les deux fonctions $g(t)$ et $(1+t)^\alpha$ vérifient la même équation différentielle $(1+t)g'(t) = \alpha g(t)$ sur $]-1,1[$ et valent toutes les deux 1 en 0, donc elles coïncident sur $]-1,1[$. □

7.2 Développement en séries entières d'autres fonctions usuelles.

On a vu comment l'opération de dérivation permet d'obtenir de nouveaux développements en série entière. On va voir maintenant comment utiliser l'opération d'intégration pour obtenir de nouveaux développements en série entière.

Partons à nouveau du développement de la série géométrique :

$$|z|<1 \qquad \frac{1}{1+z} = \sum_{n=0}^{\infty}(-1)^n z^n. \tag{8.3}$$

Appliquons le théorème d'intégration à la fonction obtenue avec une variable réelle $t \in]-1,1[$:

$$|t|<1 \qquad \log(1+t) = \sum_{n=1}^{\infty}(-1)^{n+1}\frac{t^n}{n}.$$

Et de même en substituant la variable réelle t^2 à z dans (8.3), puis en intégrant :

$$|t|<1 \qquad \arctan t = \sum_{n=0}^{\infty}(-1)^n\frac{t^{2n+1}}{2n+1}.$$

7.3 Développement en série entière et développements limités.

On a reconnu que les développements en série entière des fonctions usuelles coïncident avec leur DL en 0. Ce n'est pas un hasard, puisque les coefficients $(a_n)_{n\geq 0}$ du développement en série entière d'une fonction f sont donnés par $a_n = f^{(n)}(0)/n!$ comme on l'a vu précédemment, et ce sont les mêmes coefficients qui interviennent dans le DL de f, si elle admet un DL.

Comme toute fonction développable en série entière est nécessairement de classe C^∞ sur un intervalle ouvert centré en 0, on sait d'avance que tout toute fonction développable en série entière admet toujours un DL à tout ordre en 0. Mais l'écriture du développement en série entière permet de retrouver directement ce résultat de manière simple.

Proposition 8.51. *Soit $f(z) = \sum a_n z^n$ une fonction développable en série entière, de rayon de convergence > 0. Alors f admet un DL en 0 à tout ordre, donné par :*

$$f(z) = a_0 + a_1 z + \cdots + a_n z^n + o(z^n).$$

Démonstration. En effet, on a :

$$f(z) = a_0 + \cdots + a_n z^n + z^{n+1} p(z), \qquad \text{avec} \quad p(z) = \sum_{k=0}^{\infty} a_{n+1+k} z^k.$$

La série entière définissant $p(z)$ a même rayon de convergence R que $f(z)$, donc $p(z)$ est bornée dans l'intervalle $[-R/2, R/2]$. Donc $z^{n+1}p(z) = o(z^n)$, ce qu'on voulait montrer. □

7.4 Développement en série entière et équations différentielles linéaires.

On peut utiliser les séries entières pour résoudre des équations différentielles linéaires à coefficient polynomiaux, de la forme

$$P_n(x)y^{(n)}(x) + P_{n-1}(x)y^{(n-1)}(x) + \cdots + P_1(x)y'(x) + P_0 y(x) = Q(x). \tag{8.4}$$

où $Q, P_0, P_1, \ldots P_k$ sont des polynômes

Méthode de résolution

1. On suppose qu'il existe une fonction f développable en série entière en0 et solution de (8.4) dans un voisinage de 0. On écrit $f(x)$ sour la forme $f(x) = \sum_{n=0}^{+\infty} a_n x^n$ où les a_n sont des coefficients inconnus.
2. On remplace y par $f(x) = \sum_{n=0}^{+\infty} a_n x^n$ dans l'équation (8.4) en utilisant la dérivation terme à terme.
3. On obtient, en faisant des changements d'indice, une équation de la forme $\sum_{n=0}^{+\infty} b_n x^n = 0$ où les b_n s'expriment linéairement en fonctions des coefficients $a_{n-l}, \ldots, a_n, a_{n+1}, \ldots a_{n+m}$.
4. Par unicité du développement en série entière, on obtient ainsi des relations linéaires sur les coefficients a_n.
5. On calcule les coefficients a_n en utilisant les relations précédentes.
6. On vérifie que la série entière $\sum a_n x^n$ a bien un rayon de convergence strictement positif.
7. Eventuellement, on peut donner une expression explicite de cette solution f en calculant la somme de la série entière.

Exemple 8.52. Soit $\omega \neq 0$ un réel fixé. On considère l'équation différentielle

$$(E) \quad xy''(x) + 2y'(x) + \omega^2 xy(x) = 0\,.$$

1. On cherche une solution y de l'équation différentielle (E) développable en série entière en $x = 0$. Si y est une série entière de rayon de convergence $R > 0$, on sait

que la fonction y est de classe C^∞ sur $]-R,R[$ et indéfiniment dérivable terme à terme. On a donc pour tout $x\in]-R,R[$

$$\begin{aligned} y(x) &= \sum_{n=0}^{+\infty} a_n x^n, \\ y'(x) &= \sum_{n=1}^{+\infty} n a_n x^{n-1}, \\ y''(x) &= \sum_{n=2}^{+\infty} n(n-1) a_n x^{n-2}. \end{aligned}$$

On a donc, pour tout $x\in]-R,R[$

$$\begin{aligned} xy''(x)+2y'(x)+\omega^2 xy(x) &= \sum_{n=2}^{+\infty} n(n-1)a_n x^{n-1} + 2\sum_{n=1}^{+\infty} n a_n x^{n-1} + \sum_{n=0}^{+\infty} \omega^2 a_n x^{n+1} \\ &= a_1 + \sum_{n=1}^{+\infty} [(n+1)n+2(n+1)]a_{n+1}x^n + \sum_{n=1}^{+\infty} \omega^2 a_{n-1}x^n \\ &= a_1 + \sum_{n=1}^{+\infty} [(n+1)(n+2)a_{n+1} + \omega^2 a_{n-1}]x^n. \end{aligned}$$

On a donc que y est solution de (E) sur $]-R,R[$ si et seulement si

$$\forall x\in\mathbb{R}, a_1 + \sum_{n=1}^{+\infty}[(n+1)(n+2)a_{n+1}+\omega^2 a_{n-1}]x^n = 0.$$

Par unicité du développement en série entière, on en déduit qu'une telle fonction y est solution de (E) si et seulement si les coefficients a_n vérifient

$$a_1 = 0 \text{ et } \forall n>0,\ (n+2)(n+1)a_{n+1} + \omega^2 a_{n-1} = 0.$$

2. On a $a_1 = 0$ et $\forall n\in\mathbb{N}, a_{n+2} = -\dfrac{\omega^2}{(n+2)(n+3)}a_n$. On en déduit donc, par récurrence, que pour tout n entier, on a

$$a_{2n} = a_0 \prod_{k=1}^{n} \frac{-\omega^2}{(2k)(2k+1)} = a_0 \frac{(-1)^n\omega^{2n}}{(2n+1)!}$$

et

$$a_{2n+1} = a_1 \prod_{k=2}^{n+1} \frac{-\omega^2}{(2k)(2k-1)} = 0.$$

La série entière $y(x) = \sum_{n=0}^{+\infty} \frac{(-1)^n x^{2n}\omega^{2n}}{(2n+1)!}$ est donc l'unique solution possible de (E), qui vérifie $y(0)=a_0=1$. Le rayon de convergence de cette série entière est $R=+\infty$. En effet, pour tout $x\in\mathbb{R}$, la série $\sum \dfrac{(-1)^n x^{2n}\omega^{2n}}{(2n+1)!}$ est absolument convergente, par le critère de d'Alembert : en posant $u_n = \left|\dfrac{(-1)^n x^{2n}\omega^{2n}}{(2n+1)!}\right|$ on a (pour $x \neq 0$), $\dfrac{u_{n+1}}{u_n} = \dfrac{\omega^2 x^2}{2n+3} \longrightarrow_{n\to\infty} 0 < 1$.

3. La fonction $f(x) = \displaystyle\sum_{n=0}^{+\infty} \frac{(-1)^n x^{2n}\omega^{2n}}{(2n+1)!}$ est donc une fonction de classe C^∞ sur $\mathbb{R}$, solution de l'équation (E) telle que $f(0) = 1$ et développable en série entière en 0. D'après la question précédente, c'est l'unique solution de (E) vérifiant ces propriétés.
4. Pour tout $x \neq 0$, on a

$$f(x) = \frac{1}{\omega x}\sum_{n=0}^{+\infty}(-1)^n \frac{(\omega x)^{2n+1}}{(2n+1)!} = \frac{\sin(\omega x)}{\omega x}.$$

La fonction $g(x) = \dfrac{\sin(\omega x)}{\omega x}$ est prolongeable par continuité en 0 par la valeur $1 = f(0)$. On a donc bien $f = g$, avec g la fonction définie sur $\mathbb{R}$ par $g(0) = 1$ et $g(x) = \dfrac{\sin(\omega x)}{\omega x}$ si $x \neq 0$.
5. On peut résoudre l'équation (E) sur un intervalle ouvert I.
L'équation E est résolue sur $]-\infty, 0[$ et $]0, +\infty[$
   (a) Supposons que I est un intervalle ouvert inclus dans l'un de ces deux intervalles, et cherchons une solution de (E) sous la forme $y = uf$. On a $y' = u'f + uf'$ et $y'' = u''f + uf'' + 2(u'f')$. Une telle fonction y est donc solution de (E) si et seulement si

$$xu''f + xuf'' + 2xu'f' + 2u'f + 2uf' + \omega^2 xf = xfu'' + (2xf' + 2f)u' = 0.$$

   En utilisant que $g(x) = \omega x f(x) = \sin(\omega x)$, et en posant $u' = v$ on a à résoudre l'équation suivante en v : $gv' + 2g'v = 0$ dont les solutions sont de la forme $v = \frac{\lambda}{g^2}$, avec $\lambda \in \mathbb{R}$, sur un intervalle ouvert sur lequel g ne s'annule pas. La fonction u est donc une primitive de $\frac{\lambda}{g^2} = \frac{\lambda}{\sin^2(\omega x)}$. En utilisant le fait que $(\cot x)' = \frac{1}{\sin^2 x}$, on en déduit que u s'écrit sous la forme : $u(x) = A\cot(\omega x) + B$, avec A et B des réels. Les solutions de (E) sur un intervalle ouvert I ne contenant pas 0 forment un espace vectoriel de dimension 2 donné par :

$$\mathcal{S}_{(E)}(I) = \{A\frac{\cos(\omega x)}{\omega x} + B\frac{\sin(\omega x)}{\omega x}, (A, B) \in \mathbb{R}^2\}.$$

   (b) Si I est à présent un intervalle ouvert contenant 0, on a $I =]-a, b[=]-a, 0[\cup \{0\} \cup]0, b[$. Les solutions de (E) sur I sont donc en particulier des solutions de (E) sur $I_1 =]-a, 0[$ et $I_2 =]0, b[$. Une fonction y est solution de E sur I si et seulement si
   — la fonction y est de classe C^2 sur I ;
   — il existe des réels A_1 et B_1 tels que sur I_1, $y(x) = A_1\dfrac{\cos(\omega x)}{\omega x} + B_1\dfrac{\sin(\omega x)}{\omega x}$;
   — il existe des réels A_2 et B_2 tels que sur I_2, $y(x) = A_2\dfrac{\cos(\omega x)}{\omega x} + B_2\dfrac{\sin(\omega x)}{\omega x}$.

En 0, on a $\cos x \sim 1$, on en déduit donc que y admet une limite finie à gauche et à droite en 0 si et seulement si $A_1 = A_2 = 0$. Dans ce cas, on a $\lim_{x \to 0^-} y(x) = B_1$ et $\lim_{x \to 0^+} y(x) = B_2$, et donc y est prolongeable par continuité en 0 si et seulement si $B_1 = B_2$. On en conclut que y est une solution de E sur I, un intervalle ouvert contenant 0 si et seulement si $y = \lambda f$ avec $\lambda \in \mathbb{R}$ et f l'unique solution développable en série entière trouvée précédemment.

Chapitre 9

Fonctions de plusieurs variables

On va s'intéresser aux généralisations des notions de dérivée et d'intégration pour les fonctions de plusieurs variables. La plupart des exemples traités concernent des fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$ ou de $\mathbb{R}^3$ dans $\mathbb{R}$, mais la généralisation à $\mathbb{R}^n$ se fait sans difficulté.

1 Voisinages, limites, continuité

1.1 Norme, distance, voisinages

On se place pour commencer dans $\mathbb{R}^2$, muni de son produit scalaire canonique et on note $\|\cdot\|$ la norme du produit scalaire.

Définition 9.1.

1. Soit $X = (x, y) \in \mathbb{R}^2$, et $r > 0$, on appelle *boule ouverte de centre X et de rayon r pour la norme euclidienne* le disque ouvert
$$D(X, r) = \{X' \in \mathbb{R}^2, \|X - X'\| < r\}$$
$$= \{(x', y') \in \mathbb{R}^2, \sqrt{(x'-x)^2 + (y'-y)^2} < r\}.$$
2. On dit qu'une partie $A \subset \mathbb{R}^2$ est un voisinage de $X = (x, y) \in \mathbb{R}^2$ s'il existe ε tel que A contient la boule ouverte de centre X et de rayon ε.
3. On dit qu'une partie $O \subset \mathbb{R}^2$ est un ouvert de $\mathbb{R}^2$ si pour tout point $X \in O$, l'ensemble O est est un voisinage de $X = (x, y) \in \mathbb{R}^2$: il existe ε tel que O contient la boule ouverte de centre X et de rayon ε.

Exemple 9.2. Si on se donne $]a, b[$ et $]c, d[$ des intervalles bornés ouverts dans $\mathbb{R}$, alors l'ensemble $O =]a, b[\times]c, d[$ est un ouvert de $\mathbb{R}^2$. En effet, pour tout $(x, y) \in]a, b[\times]c, d[$, en notant $\varepsilon = \frac{1}{2}\min(x - a, b - x, y - c, d - y) > 0$, on voit immédiatement que la boule ouverte de centre (x, y) et de rayon ε est contenue dans $]a, b[\times]c, d[$ qui est donc bien un voisinage de chacun de ses points.

On peut généraliser cette définition pour n'importe quel espace vectoriel muni d'une norme. Rappelons qu'une norme sur un $\mathbb{R}$-e.v.est définie par :

Définition 9.3. Soit E un $\mathbb{R}$-e.v.Une fonction $N : E \to \mathbb{R}$, qu'on note souvent $N(x) = \|x\|$, est appelée une *norme sur E* si elle vérifie les trois propriétés suivantes :

1. (Positivité) $\|x\| \geq 0$ et $\|x\| = 0 \iff x = 0$ pour tout $x \in E$.
2. (Homogénéité) $\|\lambda x\| = |\lambda| \cdot \|x\|$ pour tout $x \in E$ et tout $\lambda \in \mathbb{R}$.
3. (Inégalité triangulaire) $\|x + y\| \leq \|x\| + \|y\|$ pour tous $x, y \in E$.

La définition précédente se généralise alors immédiatement à :

Définition 9.4.

1. Soit $X \in E$ muni de la norme $\|\cdot\|_E$ et $r > 0$, on appelle *boule ouverte de centre X et de rayon r pour la norme* $\|\cdot\|_E$ le disque ouvert
$$B(X,r) = \{X' \in E, \|X - X'\|_E < r\}$$
2. On dit qu'une partie $A \subset E$ est un voisinage de $X \in E$ s'il existe ε tel que A contient la boule ouverte de centre X et de rayon ε.
3. On dit qu'une partie $O \subset E$ est un ouvert de E si pour tout point $X \in O$, l'ensemble O est est un voisinage de X : il existe ε tel que O contient la boule ouverte de centre X et de rayon ε.

Dans le cas de $\mathbb{R}^2$, on a vu que le produit scalaire canonique fournit une norme, notée $\| \|_2$. Mais on peut en construire bien d'autres. Par exemple :

Proposition 9.5. *Les applications suivantes, définies de $\mathbb{R}^2$ dans $\mathbb{R}$ sont des normes :*

1. $\|(x,y)\|_1 = |x| + |y|$;
2. $\|(x,y)\|_\infty = \max(|x|, |y|)$;
3. *pour tout* $p \in \mathbb{N}^*$, $\|(x,y)\|_p = \sqrt[p]{|x|^p + |y|^p})$.

Remarque 9.6.

1. Les notations sont cohérentes : dans l'item 3 ci-dessus, si on prend $p = 1$, on retrouve le premier item, et le cas $p = 2$ correspond bien à la norme euclidienne. Si p tend vers $+\infty$ on peut montrer que $\|(x,y)\|_p \longrightarrow \|(x,y)\|_\infty$.
2. Ces définitions s'étendent de manière évidente à $\mathbb{R}^n$

Décrivons dans $\mathbb{R}^2$ les boules ouvertes de centre r en 0 pour les normes $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$.

1. Pour la norme $\|\cdot\|_1$, on vérifie que la boule ouverte $B_1(0,r)$ est la portion du plan contenue à l'intérieur du carré dont les sommets sont les points $(r,0), (0,r), (-r,0), (0,-r)$.
2. Pour la norme $\|\cdot\|_2$, on a vu précédemment que la boule ouverte $B_2(0,r)$ est l'intérieur du disque de centre 0 et de rayon r.
3. Pour la norme $\|\cdot\|_\infty$, on vérifie que la boule ouverte $B_\infty(0,r)$ est est la portion du plan contenue à l'intérieur du carré dont les sommets sont les points $(r,r), (-r,r), (-r,-r), (r,-r)$.

On remarque que, dans les exemples ci-dessus dans $\mathbb{R}^2$, si la notion de boule ouverte dépend évidemment de la norme choisie, les notions de voisinage d'un point ou d'ouvert, n'en dépendent pas :

En effet, on a , pour $r > 0$
$$B_\infty(0, \frac{r}{2}) \subset B_1(0,r) \subset B_2(0,r) \subset B_\infty(0,r)$$

ce qui montre qu'un ensemble qui contient une boule ouverte en 0 pour l'une de ces normes, contient également une boule ouverte en 0 pour les autres (quitte à changer de rayon).

Ce n'est pas un fait isolé, mais un fait général et vous montrerez l'an prochain :

Théorème 9.7 (Equivalence des normes)**.**
Si E est un espace vectoriel de dimension finie, alors toutes les normes sur E sont équivalentes. En particulier, les notions de voisinage d'un point ou d'ouvert ne dépendent pas de la norme choisie.

Dans la suite, on utilisera (par exemple) la norme euclidienne canonique sur $\mathbb{R}^n$, en étant conscients du fait que les notions que nous allons voir (continuité, dérivabilité, ...) ne dépendent en fait pas de la norme choisie.

1.2 Limites

La notion de voisinage dans un espace vectoriel permet de définir la notion de limite pour une fonction ou une suite (indépendamment de la norme). Si E est un espace vectoriel normé on dit que y tend vers x lorsque la quantité réelle $\|x-y\|$ tend vers 0.

Détaillons la signification de cette limite dans le cas où $E = \mathbb{R}^2$.

Définition 9.8. Soit $f : \mathbb{R}^2 \backslash \{a\} \to \mathbb{R}$ une fonction définie au voisinage d'un point $a = (u_0, v_0)$. Alors f *tend vers* $\ell \in \mathbb{R}$ *en* a si :

$$\forall \varepsilon > 0 \quad \exists \alpha > 0 \quad \forall x \in \mathbb{R}^2 \quad \|x-a\| < \alpha \implies |f(x) - f(a)| < \varepsilon.$$

On écrit alors :

$$\lim_{(u,v)\to(u_0,v_0)} f(u,v) = \ell.$$

Les propriétés de la norme impliquent l'unicité de la limite

Proposition 9.9. *Si f tend vers ℓ et vers ℓ' en a, alors $\ell = \ell'$.*

Proposition 9.10 (composition des limites, 1)**.** *Soir deux fonctions :*

$$f : \begin{cases} X \subseteq \mathbb{R}^2 \to \mathbb{R} \\ (x_1, x_2) \mapsto f(x_1, x_2) \end{cases} \qquad \varphi : \begin{cases} D \subseteq \mathbb{R} \\ t \mapsto \varphi(t) \end{cases}$$

et soit $a \in X$. Si :

$$\lim_{x\to a} f(x) = \ell \qquad \text{et} \qquad \lim_{t\to\ell} \varphi(t) = \ell',$$

alors :

$$\lim_{x\to a} \varphi \circ f(x) = \ell'.$$

Proposition 9.11 (composition des limites, 2). *Soient deux fonctions :*

$$\gamma : \begin{cases} D \subseteq \mathbb{R} \to \mathbb{R}^2 \\ t \mapsto \gamma(t) = \big(u(t), v(t)\big) \end{cases} \qquad f : \begin{cases} X \subseteq \mathbb{R}^2 \to \mathbb{R} \\ (u, v) \mapsto f(u, v) \end{cases}$$

telles que $\gamma(D) \subseteq X$. *Si :*

$$\lim_{t \to t_0} \gamma(t) = (u_0, v_0) \qquad \text{et} \qquad \lim_{(u,v) \to (u_0,v_0)} f(u, v) = \ell$$

alors :

$$\lim_{t \to t_0} f \circ \gamma(t) = \ell.$$

On peut aussi définir dans ce cadre la notion de limite d'une suite dans $\mathbb{R}^2$

Définition 9.12. Soit $(u_n)_{n \geq 0}$ une suite de points de $\mathbb{R}^2$, et soit $\ell = (x, y) \in \mathbb{R}^2$. La suite $(u_n)_{n \geq 0}$ tend vers ℓ si :

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geq N \implies \|u_n - \ell\| < \varepsilon).$$

Remarque 9.13.

1. La limite d'une suite, si elle existe, est unique.
2. Dans la définition ci-dessus, on peut remplacer la norme euclidienne $\| \cdot \|$ par n'importe quelle norme, par exemple la norme infinie $\|(u, v)\|_\infty = \max\big(|u|, |v|\big)$.

1.3 Continuité

On introduit dans $\mathbb{R}^2$ la notion de continuité par rapport *aux deux variables simultanées*. On verra que cette nouvelle notion est plus forte que la continuité par rapport à chaque variable prise séparément.

On munit $\mathbb{R}^2$ de son produit scalaire canonique et on note $\| \cdot \|$ la norme du produit scalaire.

Définition 9.14 (continuité). Soit $O \subset \mathbb{R}^2$ un ouvert. Une fonction $F : O \to \mathbb{R}$ est *continue* en un point $x_0 = (u_0, v_0) \in O$ si :

$$\forall \varepsilon > 0 \quad \exists \eta > 0 \quad \forall x \in O \quad \big(\|x - x_0\| < \eta \implies |f(x) - f(x_0)| < \varepsilon\big).$$

La fonction F est *continue sur* O si elle est continue en tout point de O.

On peut comparer cette notion de continuité par rapport aux deux variables simultanément avec la notion, plus faible, de continuité par rapport à chacune des variables, c'est à dire de la continuité des applications partielles

Définition 9.15. Soit O un ouvert de $\mathbb{R}^2$ et soit $F : O \to \mathbb{R}$ une fonction. On définit les applications partielles comme les applications suivantes

1. Pour chaque $(u_0, v_0) \in O$, on note $F_{u_0,\cdot} : V \to \mathbb{R}$ la fonction définie par $F_{u_0,\cdot}(v) = F(u_0, v)$ sur un intervalle ouvert V contenant v_0.
2. Pour chaque $(u_0, v_0) \in O$, on note $F_{\cdot,v_0} : V \to \mathbb{R}$ la fonction définie par $F_{\cdot,v_0}(x) = F(u, v_0)$ sur un intervalle ouvert U contenant u_0.

On dit que F est *continue par rapport à la première variable* en (u_0, v_0) si la fonction $F_{\cdot,v_0}$ est continue en u_0 et, de même que F est *continue par rapport à la seconde variable* en (u_0, v_0) si la fonction $F_{u_0,\cdot}$ est continue en v_0.

Remarque 9.16.

1. Il revient au même de changer la norme $\|\cdot\|$ par la *norme infinie* $\|(u, v)\|_\infty = \max\big(|u|, |v|\big)$. La continuité en (u_0, v_0) s'écrit alors :

$$\forall \varepsilon > 0 \quad \exists \eta > 0 \quad \forall (u, v) \in O \qquad \big((|u - u_0| < \eta \wedge |v - v_0| < \eta) \implies |f(u, v) - f(u_0, v_0)| < \varepsilon\big).$$

2. Il est clair que la continuité implique la continuité par rapport à chacune des variables.
3. En revanche, la continuité par rapport à chacune des variables n'implique pas la continuité globale, comme le montre l'exemple suivant. Soit la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par :

$$f(u, v) = \begin{cases} \frac{uv}{u^2+v^2}, & \text{si } (u, v) \neq (0, 0), \\ 0, & \text{si } (u, v) = (0, 0). \end{cases}$$

Alors, quel que soit $u \in \mathbb{R}$, f est continue par rapport à la deuxième variable en tout point $v \in \mathbb{R}$, et *idem* en échangeant les rôles de u et v. Cependant, f n'est pas continue au sens de la définition 9.14 au point $(0, 0)$. En effet, considérons $r > 0$ et $\theta \in [0, 2\pi[$, et posons $x(r) = (r \cos \theta, r \sin \theta)$. Supposons que f est continue en 0. Nous avons $\|x(r)\| = r$, donc on devrait avoir :

$$\lim_{r \to 0} f\big(x(r)\big) = f(0, 0), \tag{9.1}$$

et ceci indépendemment de la valeur choisie pour θ. Or, on a :

$$f\big(x(r)\big) = (\cos \theta) \cdot (\sin \theta),$$

donc (9.1) n'a pas lieu pour tout θ. Cette contradiction montre que f n'est pas continue en $(0, 0)$.

Proposition 9.17 (limite et continuité, plusieurs variables)**.**
Soit $f : O \to \mathbb{R}$ une fonction définie sur un ouvert de $\mathbb{R}^2$ et soit $X_0 \in O$. Les propositions suivantes sont équivalentes :

1. *f est continue en X_0.*

2. *Pour toute suite* $(u_n)_{n\geq 0}$ *de* O *de limite* X_0*, la suite* $(f(u_n))_{n\geq 0}$ *tend vers* $f(X_0)$*.*
3. $\lim\limits_{X\to X_0} f(X) = f(X_0)$.

Démonstration. $1 \implies 2$. Soit $(u_n)_{n\geq 0}$ une suite de O qui tend vers x_0. On a donc :

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad \|u_n - x_0\| < \varepsilon. \tag{9.2}$$

Choisissons $\varepsilon > 0$ pour montrer la convergence de la suite $(f(u_n))_{n\geq 0}$ vers $f(x_0)$. D'après la continuité de f en x_0, choisissons $\eta > 0$ tel que :

$$\forall x \in I \times J \quad \|x - x_0\| < \eta \implies \left|f(x) - f(x_0)\right| < \varepsilon. \tag{9.3}$$

Choisissons N comme dans (9.2) associé à η. Alors pour tout entier $n \geq N$, on a :

$$\|u_n - x_0\| < \eta\,, \qquad \text{et donc } \left|f(u_n) - f(x_0)\right| < \varepsilon \text{ d'après (9.3),}$$

ce qui montre que $f(u_n) \to f(x_0)$.

$2 \implies 3$. Sous l'hypothèse 2, démontrons 3 par l'absurde. Posons $\ell = f(x_0)$ et écrivons la négation de 3 :

$$\exists \varepsilon > 0 \forall \eta > 0 \quad \exists x \in I \times J \quad (\|x - x_0\| < \eta \wedge |f(x) - f(x_0)| \geq \varepsilon).$$

Choisissons $\varepsilon > 0$ comme ci-dessus, et posons $\eta_n = 1/n$ pour tout entier $n \geq 1$. D'après la propriété ci-dessus, il existe donc pour chaque entier $n \geq 1$ un élément $x_n \in I \times J$ tel que $\|x_n - x_0\| < \eta_n$ et $|f(x_n) - f(x_0)| \geq \varepsilon$. La suite $(x_n)_n$ converge vers x_0. Donc la suite $(f(x_n))_n$ converge vers $f(x_0)$, ce qui contredit $|f(x_n) - f(x_0)| \geq \varepsilon$.

$3 \implies 1$. C'est évident sur les définitions. □

Comme pour les fonctions d'une seule variable, on a, en combinant les résultats des propositions 9.10, 9.11, 9.17

Théorème 9.18. *Une composée de fonctions continues est continue.*

1.4 Exemples d'applications continues

1. Les applications $(x, y) \mapsto x$, $(x, y) \mapsto y$, $(x, y) \mapsto x + y$ sont continues.

   En effet $|x - x_0| \leq \|(x, y) - (x_0, y_0)\| = \sqrt{(x - x_0)^2 + (y - y_0)^2}$, donc $\|(x, y) - (x_0, y_0)\| < \varepsilon \Rightarrow |x - x_0| < \varepsilon$.
   De même, $|y - y_0| \leq \|(x, y) - (x_0, y_0)\|$, donc $\|(x, y) - (x_0, y_0)\| < \varepsilon \Rightarrow |x - x_0| < \varepsilon$.
   Enfin $|(x + y) - (x_0 + y_0)| \leq |x - x_0| + |y - y_0| \leq 2\|(x, y) - (x_0, y_0)\|$, et donc

   $$\|(x, y) - (x_0, y_0)\| < \varepsilon/2 \Rightarrow |(x + y) - (x_0 + y_0)| < \varepsilon.$$

2. L' application $(x, y) \mapsto xy$ est continue.

   En effet $xy - x_0y_0 = x(y - y_0) + (x - x_0)y_0$. Donc $|xy - x_0y_0| \leq (|x| + |y_0|)\|(x, y) - (x_0, y_0)\|$. Soit $\varepsilon > 0$. Posons $\alpha = \min\left(1, \frac{\varepsilon}{|x_0|+|y_0|+1}\right)$. Si $\|(x, y) - (x_0, y_0)\| < \alpha$, alors $|x| \leq |x_0| + 1$, donc $|xy - x_0y_0| \leq (|x| + |y_0|)\|(x, y) - (x_0, y_0)\| < \varepsilon$.

3. L' application $(x, y) \mapsto e^{xy}$ est continue.

   C'est la composée des applications continues $(x, y) \mapsto xy$ et $\exp : \mathbb{R} \to \mathbb{R}$.

4. Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ une application. Écrivons $f(x, y) = \big(u(x, y), v(x, y)\big)$. Alors f est continue si et seulement si les applications $u, v : \mathbb{R}^2 \to \mathbb{R}$ sont continues.

   Notons $p : (x, y) \mapsto x$ et $q : (x, y) \mapsto y$. On a vu que les applications p et q sont continues et l'on a $u = p \circ f$ et $v = q \circ f$. Donc, si f est continue, alors u et v sont continues.

   Inversement, si u et v sont continues en (x_0, y_0), alors pour tout $\varepsilon > 0$, il existe $\alpha_1 > 0$ et α_2 tels que $\|(x, y) - (x_0, y_0)\| < \alpha_1 \Rightarrow |u(x, y) - u(x_0, y_0)| < \varepsilon$ et $\|(x, y) - (x_0, y_0)\| < \alpha_2 \Rightarrow |v(x, y) - v(x_0, y_0)| < \varepsilon$.

   Posons alors $\alpha = \min(\alpha_1, \alpha_2)$. Si $\|(x, y) - (x_0, y_0)\| < \alpha$, alors $|u(x, y) - u(x_0, y_0)| < \varepsilon$ et $|v(x, y) - v(x_0, y_0)| < \varepsilon$, donc

$$\begin{aligned}\Big\|\big(u(x, y), v(x, y)\big) - \big(u(x, y), v(x, y)\big)\Big\| &= \sqrt{\big(u(x, y) - u(x_0, y_0)\big)^2 + \big(v(x, y) - v(x_0, y_0)\big)^2} \\ &< \sqrt{2}\,\varepsilon.\end{aligned}$$

5. L' application $(x, y) \mapsto (x - y)\sin(x + y)$ est continue.

   L'application $(x, y) \mapsto (x + y)$ est continue ; comme sin est continue, l'application composée $(x, y) \mapsto \sin(x + y)$ est continue ; l'application $(x, y) \mapsto (x - y)$ est continue, et donc l'application $F : (x, y) \mapsto \Big(\sin(x + y), x - y\Big)$ (de $\mathbb{R}^2$ dans $\mathbb{R}^2$) est continue. Enfin, l'application $(x, y) \mapsto (x - y)\sin(x + y)$ est la composée $m \circ F$ de F avec l'application continue $m : (x, y) \mapsto xy$.

1.5 Continuité uniforme

Comme dans le cas unidimensionnel, on a la notion plus forte de continuité uniforme.

Définition 9.19 (continuité uniforme, module de continuité uniforme). Une fonction $F : O \to \mathbb{R}$ est *uniformément continue sur* O si :

$$\forall \varepsilon > 0 \quad \exists \eta > 0 \quad \forall x, y \in O \quad \|x - y\| < \eta \implies \big|f(x) - f(y)\big| < \varepsilon.$$

Un nombre $\eta > 0$ qui convient pour un nombre $\varepsilon > 0$ donné est appelé *module de continuité uniforme* associé à ε.

On admettra le résultat suivant.

Théorème 9.20 (Heine, plusieurs variables). *Soit $U = [a, b]$ et $V = [c, d]$ deux segments de $\mathbb{R}$, et soit $F : U \times V \to \mathbb{R}$ une fonction continue sur $U \times V$. Alors F est uniformément continue sur $U \times V$.*

2 Dérivées partielles, fonctions de classe C^1

2.1 Dérivées partielles

Définition 9.21. Soit O un ouvert de $\mathbb{R}^2$ et soit $f : O \to \mathbb{R}$ une fonction. Soit $X_0 = (x_0, y_0) \in O$. On dit que f admet des *dérivées partielles en* (x_0, y_0) si l'application $x \mapsto f(x, y_0)$ admet une dérivée notée $\frac{\partial f}{\partial x}(x_0, y_0)$ en x_0 et si l'application $y \mapsto f(x_0, y)$ admet une dérivée notée $\frac{\partial f}{\partial y}(x_0, y_0)$ en y_0.

On a donc (lorsque ces limites existent)

$$\frac{\partial f}{\partial x}(x_0, y_0) = \lim_{x \to x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0} \quad \text{et} \quad \frac{\partial f}{\partial y}(x_0, y_0) = \lim_{y \to y_0} \frac{f(x_0, y) - f(x_0, y_0)}{y - y_0}.$$

Remarque 9.22. Les dérivées partielles au point (x_0, y_0) sont les dérivées (au sens ordinaire) des applications partielles $x \mapsto f_{\cdot, y_0}(x) = f(x, y_0)$ en x_0 et $y \mapsto f_{x_0, \cdot}(y) = f(x_0, y)$ en y_0.

2.2 Exemples de dérivées partielles

1. — Si $f : (x, y) \mapsto x$, alors $\frac{\partial f}{\partial x}(x, y) = 1$ et $\frac{\partial f}{\partial y}(x, y) = 0$;
   — si $f : (x, y) \mapsto y$, alors $\frac{\partial f}{\partial x}(x, y) = 0$ et $\frac{\partial f}{\partial y}(x, y) = 1$;
   — si $f : (x, y) \mapsto x + y$, alors $\frac{\partial f}{\partial x}(x, y) = 1$ et $\frac{\partial f}{\partial y}(x, y) = 1$.
2. Pour $m : (x, y) \mapsto xy$ on a $\frac{\partial m}{\partial x}(x, y) = y$ et $\frac{\partial m}{\partial y}(x, y) = x$.
3. Posons $f : (x, y) \mapsto e^{xy}$; on a $\frac{\partial f}{\partial x}(x, y) = ye^{xy}$ et $\frac{\partial f}{\partial y}(x, y) = xe^{xy}$.
4. Posons $f : (x, y) \mapsto (x - y)\sin(x + y)$;
   on a $\frac{\partial f}{\partial x}(x, y) = \sin(x + y) + (x - y)\cos(x + y)$ et
   $\frac{\partial f}{\partial y}(x, y) = -\sin(x + y) + (x - y)\cos(x + y)$.

2.3 Fonctions de classe C^1

Définition 9.23. Soit O un ouvert de $\mathbb{R}^2$ et soit $f : O \to \mathbb{R}$ une fonction. On dit que f est *de classe* C^1 si elle admet des dérivées partielles en (x, y) pour tout $(x, y) \in O$ et si les deux applications $\frac{\partial f}{\partial x} : (x, y) \mapsto \frac{\partial f}{\partial x}(x, y)$ et $\frac{\partial f}{\partial y} : (x, y) \mapsto \frac{\partial f}{\partial y}(x, y)$ de O dans $\mathbb{R}$ sont continues (au sens des fonctions de plusieurs variables).

Gradient

Définition 9.24. Soit O un ouvert de $\mathbb{R}^2$ et soit $f : O \to \mathbb{R}$ une fonction de classe C^1. On appelle *gradient de f en (x, y)* le vecteur $\nabla f_{(x,y)} = \left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y)\right)$.

Formule de Taylor à l'ordre 1, différentielle

Théorème 9.25. *Si une fonction f est de classe C^1 sur un ouvert $O \subset \mathbb{R}^2$, on a, pour $(x, y) \in O$ et h et k suffisamment petits,*

$$f(x+h, y+k) = f(x, y) + \frac{\partial f}{\partial x}(x, y)\,h + \frac{\partial f}{\partial y}(x, y)\,k + o(\|(h, k)\|).$$

Démonstration. Écrivons $f(x+h, y+k) - f(x, y) - \frac{\partial f}{\partial x}(x, y)\,h - \frac{\partial f}{\partial y}(x, y)\,k = U + V$ avec

$$\begin{aligned} U &= f(x+h, y) - f(x, y) - \frac{\partial f}{\partial x}(x, y)h \\ V &= f(x+h, y+k) - f(x+h, y) - \frac{\partial f}{\partial y}(x, y)k. \end{aligned}$$

Démontrons que U et V sont des $o(\|(h, k)\|)$.

Posons

$$\begin{aligned} \varphi(s) &= f(x+s, y) - f(x, y) - \frac{\partial f}{\partial x}(x, y)s \\ \psi(t) &= f(x+h, y+t) - f(x+h, y) - \frac{\partial f}{\partial y}(x, y)t. \end{aligned}$$

On a $\varphi'(s) = \frac{\partial f}{\partial x}(x+s, y) - \frac{\partial f}{\partial x}(x, y)$ et $\psi'(t) = \frac{\partial f}{\partial y}(x+h, y+t) - \frac{\partial f}{\partial x}(x, y)$. Par continuité de $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$, pour $\varepsilon > 0$, il existe α tel que si $(s, t) \in \mathbb{R}^2$ sont tels que $\sqrt{s^2 + t^2} < \alpha$ alors $(x+s, y+t) \in U$ et $\left|\frac{\partial f}{\partial x}(x+s, y+t) - \frac{\partial f}{\partial x}(x, y)\right| < \varepsilon$ et $\left|\frac{\partial f}{\partial y}(x+s, y+t) - \frac{\partial f}{\partial y}(x, y)\right| < \varepsilon$.

Si $\|(h, k)\| < \alpha$, alors pour tout s avec $|s| < |h|$ et pour tout $|t| < |k|$ on a $|\varphi'(s)| < \varepsilon$ et $|\psi'(t)| < \varepsilon$ donc $|U| = |\varphi(h)| < \varepsilon|h|$ et $V = |\psi(k)| < \varepsilon|k|$. □

Dérivée suivant un vecteur

La notion de dérivée partielle peut-être généralisée à la notion de dérivée par par rapport à un vecteur donné.

Définition 9.26. Soit O un ouvert de $\mathbb{R}^2$ et soit $f : O \to \mathbb{R}$ une fonction, $X = (x_0, y_0) \in O$ un point de O et $V = (h, k)$ un vecteur non nul. On dit que f admet une dérivée dans la direction du vecteur V au point (x_0, y_0) si l'application $t \mapsto f(x_0 + th, y_0 + tk)$ admet une dérivée notée $D_V f(x_0, y_0)$ en $t = 0$.

Corollaire 9.27. *Si f est de classe C^1 sur un ouvert O, alors f est dérivable en tout point (x,y) de O dans la direction du vecteur $V=(h,k)$ pour tout vecteur V non nul de $\mathbb{R}^2$. Sa dérivée dans la direction de V est donnée par*

$$D_V f(x,y) = \frac{\partial f}{\partial x}(x,y)\,h + \frac{\partial f}{\partial y}(x,y)\,k = \left\langle \nabla f_{(x,y)} | (h,k) \right\rangle$$

Remarque 9.28.

1. L'existence de la dérivée partielle par rapport à la première variable équivaut à l'existence de la dérivée dans la direction du vecteur $(1,0)$. Celle de de la dérivée partielle par rapport à la deuxième variable équivaut à l'existence de la dérivée dans la direction du vecteur $(0,1)$.
2. L'existence d'une dérivée dans toutes les directions n'implique pas que la fonction est de classe $C1$ (ou différentiable) en ce point (voir contre-exemple en TD).

Différentielle

Définition 9.29. Supposons $f : O \to \mathbb{R}$ de classe C^1. L'application linéaire $(h,k) \mapsto \frac{\partial f}{\partial x}(x,y)\,h + \frac{\partial f}{\partial y}(x,y)\,k$ s'appelle la *différentielle* de f en (x,y) et se note $df_{(x,y)}$.

Remarque 9.30.

1. On a donc $df_{(x,y)}(h,k) = \left\langle \nabla f_{(x,y)} | (h,k) \right\rangle$.
2. On écrit $df = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy$ où $dx : (h,k) \mapsto h$ et $dy : (h,k) \mapsto k$.

2.4 Fonctions de classe C^1 sur $\mathbb{R}^2$ et fonctions holomorphes

On a vu, dans les chapitres précédents, la notion de fonction holomorphe, c'est à dire de fonction de $\mathbb{C}$ dans $\mathbb{C}$ dérivable au sens complexe. En utilisant l'isomorphisme naturel $\mathbb{R}^2 \to \mathbb{C}, (x,y) \mapsto x+iy$, on peut se demander quelle relation existe entre la notion de fonction holomorphe et celle de fonction de classe C^1 que nous venons de définir.

On peut montrer facilement

Proposition 9.31. *Soit O un ouvert de $\mathbb{C}$. Une fonction holomorphe $f : O \to \mathbb{C}$ est de classe C^1 (et même C^∞) sur O vue comme fonction de deux variables, de $O \subset \mathbb{R}^2$ à valeurs dans $\mathbb{R}^2$*

Démonstration. Ecrivons $f = u + iv$, et notons f' la dérivée au sens complexe de f. Alors f est de classe C^1 sur O si et seulement si $u : O \to \mathbb{R}$ et $v : O \to \mathbb{R}$ sont de classe C^1. Mais on a

$$\frac{u(x_0+h,y_0)-u(x_0,y_0)}{h} = \mathrm{Re}\left(\frac{f(x_0+h+iy_0)-f(x_0+iy_0)}{h}\right) \to \mathrm{Re}\, f'(x_0+iy_0) = \frac{\partial u}{\partial x}(x_0,y_0)$$

car f est holomorphe. On a, de même,

$$\frac{u(x_0, y_0 + k) - u(x_0, y_0)}{k} = \operatorname{Re}\left(\frac{f(x_0 + iy_0 + ik) - f(x_0 + iy_0)}{k}\right) \to \operatorname{Re}\left(if'(x_0 + iy_0)\right)$$
$$\frac{v(x_0 + h, y_0) - v(x_0, y_0)}{h} = \operatorname{Im}\left(\frac{f(x_0 + h + iy_0) - f(x_0 + iy_0)}{h}\right) \to \operatorname{Im} f'(x_0 + iy_0)$$
$$\frac{v(x_0, y_0 + k) - v(x_0, y_0)}{k} = \operatorname{Im}\left(\frac{f(x_0 + iy_0 + ik) - f(x_0 + iy_0)}{k}\right) \to \operatorname{Im}\left(if'(x_0 + iy_0)\right)$$

On obtient ainsi :

$$\begin{aligned}
\frac{\partial u}{\partial x}(x_0, y_0) &= \operatorname{Re}\left(f'(x_0 + iy_0)\right) \\
\frac{\partial u}{\partial y}(x_0, y_0) &= \operatorname{Re}\left(if'(x_0 + iy_0)\right) = -\operatorname{Im} f'(x_0 + iy_0) \\
\frac{\partial v}{\partial x}(x_0, y_0) &= \operatorname{Im}\left(f'(x_0 + iy_0)\right) \\
\frac{\partial v}{\partial y}(x_0, y_0) &= \operatorname{Im}\left(if'(x_0 + iy_0)\right) = \operatorname{Re} f'(x_0 + iy_0)
\end{aligned}$$

Enfin, la fonction f' est clairement continue, et donc les dérivées partielles ci-dessus également (on admet que toute fonction holomorphe est analytique). □

En revanche la réciproque est loin d'être vérifiée : être holomorphe est une condition beaucoup plus forte qu'être de classe C^1.

Exemple 9.32. La fonction $z \mapsto \overline{z}$ est de classe C^1 mais pas holomorphe.
En effet la fonction s'exprime dans $\mathbb{R}^2$ sous la forme $(x, y) \mapsto (x, -y)$ et les fonctions $(x, y) \mapsto x$ et $(x, y) \mapsto -y$ sont bien de classe C^1 (et même C^∞) de $\mathbb{R}^2$ dans $\mathbb{R}$.
En revanche, le quotient $\frac{\overline{z - z_0}}{z - z_0}$ n'a pas de limite lorsque $z - z_0$ tend vers 0. Pour le voir, il suffit d'écrire $z - z_0$ sous la forme $z - z_0 = re^{i\theta}$. On obtient alors $\frac{\overline{z - z_0}}{z - z_0} = \frac{e^{-i\theta}}{e^{i\theta}} = e^{-2i\theta}$ qui n'a pas de limite lorsque r tend vers 0. Par exemple, si $\theta = 0$, $\frac{\overline{z - z_0}}{z - z_0} = 1$ tandis que si $\theta = \pi/2$, $\frac{\overline{z - z_0}}{z - z_0} = -1$.

On peut montrer en fait la caractérisation suivante.

Proposition 9.33. *Soit O un ouvert de $\mathbb{R}^2$. Une fonction de classe C^1 $f : O \to \mathbb{R}^2$ est holomorphe sur O vue comme fonction d'une variable complexe $O \subset \mathbb{C}$ à valeurs dans $\mathbb{C}$ si et seulement si elle vérifie*

$$\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y} = 0$$

soit encore :

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \text{ et } \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x},$$

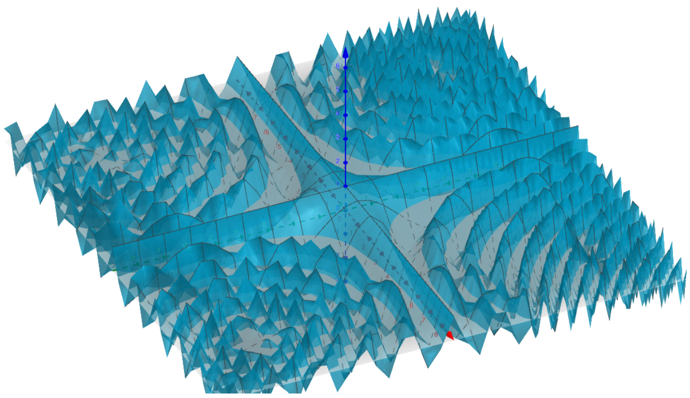
où $u(x, y)$ et $v(x, y)$ désignent les coordonnées dans $\mathbb{R}^2$ de $f(x, y) = \left(u(x, y), v(x, y)\right)$.

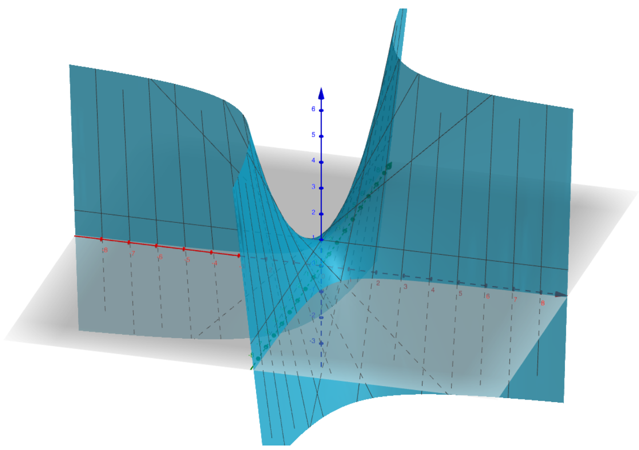
3 Représentation graphique des fonctions de deux variables à valeurs réelles

3.1 Définition

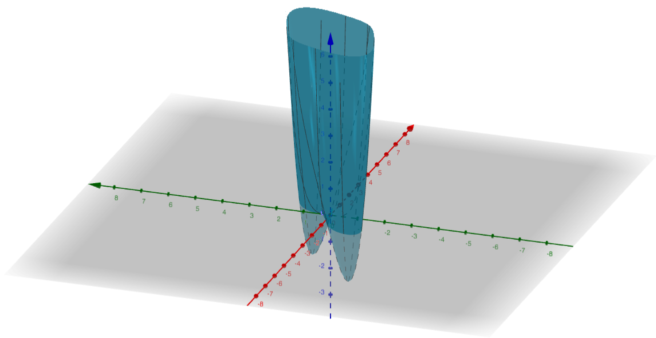
De même que les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ peuvent être représentés graphiquement dans le plan en étudiant les points du plan dans le graphe de l'application f, c'est à dire les points du plan de la forme $(x, f(x))$ lorsque x parcourt le domaine de définition de f, on peut représenter graphiquement une fonction f de deux variables réelles en étudiant dans $\mathbb{R}^3$ l'ensemble des points $(x, y, f(x, y))$ lorsque (x, y) parcourt le domaine de définition de f.
Une fonction de deux variables est donc représentée par une surface $z = f(x, y)$ dans $\mathbb{R}^3$

Définition 9.34. Soit f une fonction de deux variables et Dom_f son domaine de définition. On appelle *graphe de* f l'ensemble des points $(x, y, f(x, y))$ de $\mathbb{R}^3$ lorsque (x, y) parcourt le domaine de définition de f. Le graphe de f est représenté graphiquement par la surface $z = f(x, y)$ dans $\mathbb{R}^3$.



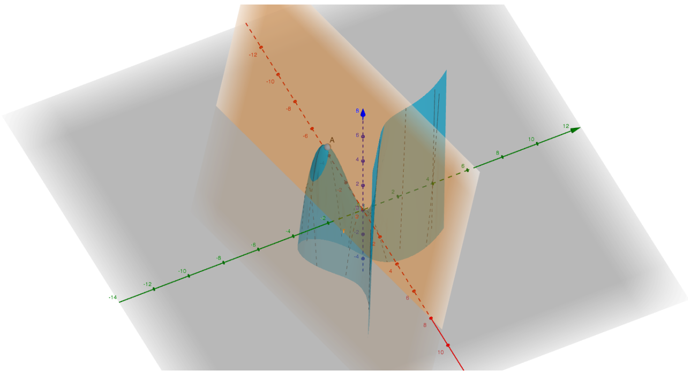


Remarquons que si f est de classe C^1 en un point (x_0, y_0), la formule de Taylor à l'ordre 1 permet de définir un plan affine qui est la meilleure approximation par un plan de la surface au point (x_0, y_0).



Définition 9.35. Soit f une fonction de deux variables de classe C^1 au point (x_0, y_0). On appelle *plan tangent graphe de f* le plan d'équation :

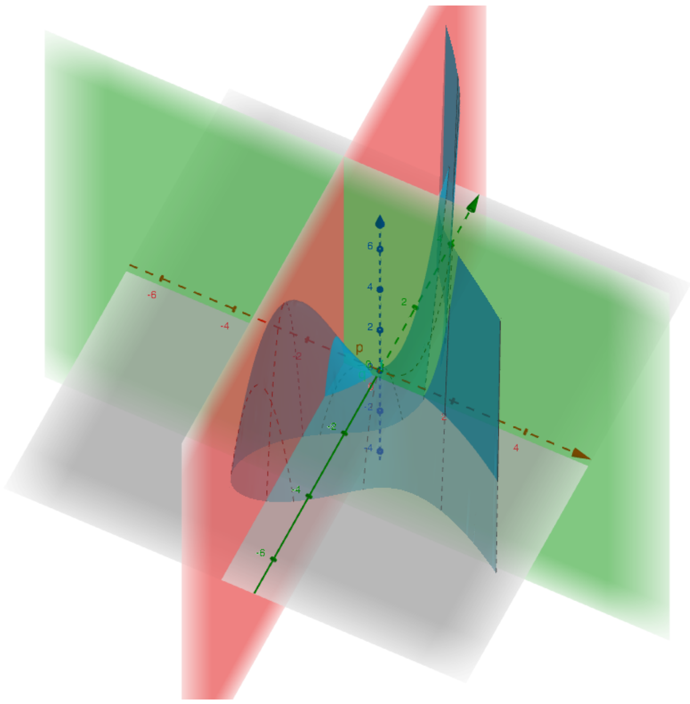
$$z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)\,(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)\,(y - y_0).$$

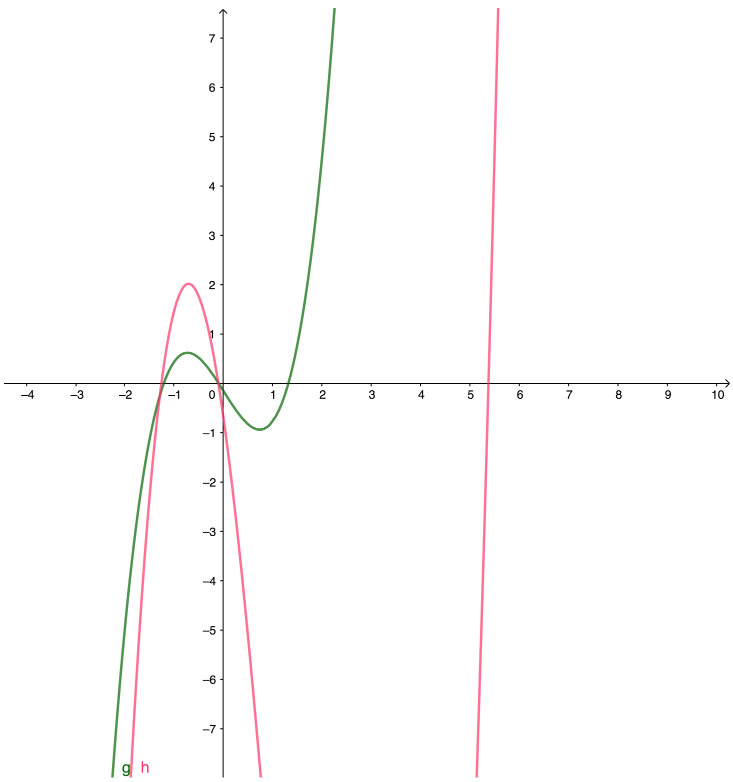


Remarque 9.36. Si on "bloque" l'une des deux variables x ou y on considère alors seulement les fonctions partielles $f_{\cdot, y_0} : x \mapsto f(x, y_0)$ et $f_{x_0, \cdot} : y \mapsto f(x_0, y)$. Les graphes de ces fonctions partielles sont donnés par l'intersection du graphe de f avec le plan $y = y_0$ ou $x = x_0$ respectivement.
Plus généralement, considérer le graphe de l'application $t \mapsto f(x_0 + th, y_0 + tk)$ revient à considérer l'intersection du graphe de f avec le plan $-k(x - x_0) + h(y - y_0) = 0$.

3.2 Lignes de niveau

Le graphe d'une fonction de deux variables est beaucoup plus difficile à tracer que le graphe d'une fonction d'une variable : difficulté du dessin, et surtout, pas de notion équivalente au tableau de variation des fonctions d'une variable. C'est pourquoi on utilise souvent d'autres modes de représentation graphique. Pour la méteo, par exemple, la pression ou la température sont classiquement représentées graphiquement par des courbes planes (isobares ou isothermes). De même, pour les cartes




g
h

géographiques représentant le relief, à la place d'une carte en trois dimensions (pas pratique à transporter), on trace les courbes d'altitude constante.

Définition 9.37. Soit f une fonction de deux variables et $\lambda \in \mathbb{R}$, on appelle *ligne de niveau* λ de la fonction f l'intersection du plan $z = \lambda$ avec le graphe de f : c'est la courbe formée par l'ensemble des points (x, y) du plan tels que $f(x, y) = \lambda$.

Proposition 9.38. *Supposons $f : O \longrightarrow \mathbb{R}$ de classe C^1.*

1. *Le gradient est orthogonal aux lignes de niveau.*
2. *Il est porté par la plus grande pente : sa direction et son sens sont donnés par le vecteur unité V tel que la dérivée de f dans la direction de V est maximale.*
3. *Sa norme est cette pente, c'est à dire cette dérivée de f dans la direction de V.*

Démonstration. 1. La formule de Taylor au point (x, y) s'écrit :

$$f(x+h, y+k) = f(x, y) + \left\langle \nabla f_{(x,y)} | (h, k) \right\rangle + o(\|(h, k\|).$$

Posons $\lambda = f(x, y)$ et considérons la ligne de niveau λ. Par définition, il s'agit des points (u, v) dans O tels que $(f(u, v) = \lambda$.
Si l'on se place dans le plan $z = \lambda = f(x, y)$, alors on a, pour tout vecteur (h, k) de norme suffisamment petite et tel que $f(x+h, y+k) = \lambda$

$$\lambda = \lambda + \left\langle \nabla f_{(x,y)} | (h, k) \right\rangle + o(\|(h, k\|).$$

Et donc on a bien, pour tout vecteur (h, k) de norme suffisamment petite et tel que $f(x+h, y+k) = \lambda$,

$$\left\langle \nabla f_{(x,y)} | (h, k) \right\rangle = 0$$

ce qui montre qu'au point (x, y), le gradient est orthogonal à la ligne de niveau λ.

2. On a vu que si V est un vecteur, la dérivée dans la direction de V est donnée par

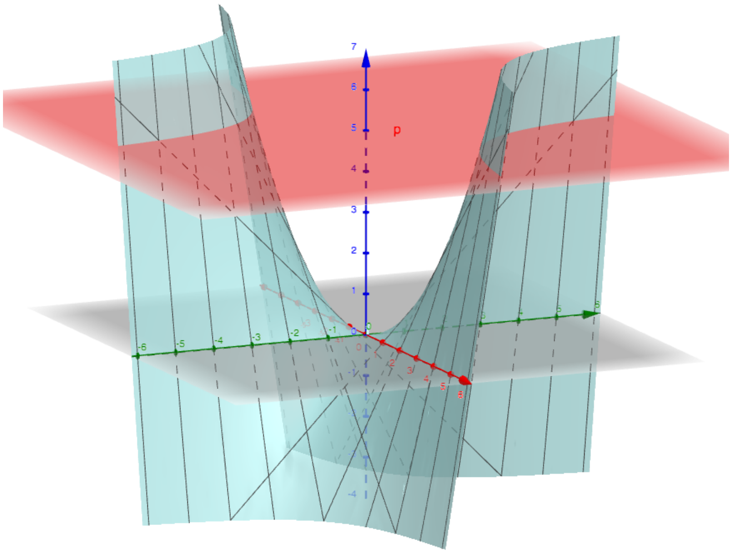
$$D_V(f)(x, y) = \left\langle \nabla f_{(x,y)} | V \right\rangle.$$

Ce nombre réel est maximal si V est colinéaire (avec un coefficient positif) au vecteur $\nabla f_{(x,y)}$

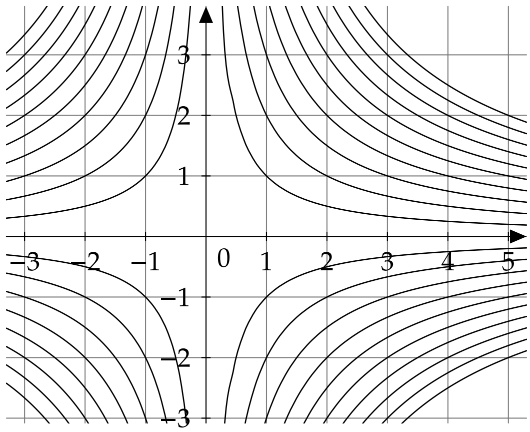
3. Si $V = \dfrac{\nabla f_{(x,y)}}{\left\|\nabla f_{(x,y)}\right\|}$, alors

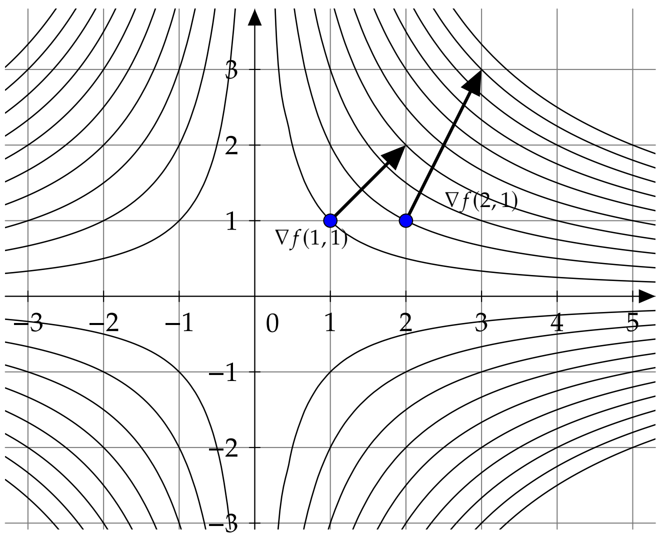
$$D_V(f)(x, y) = \left\langle \nabla f_{(x,y)} | \frac{\nabla f_{(x,y)}}{\left\|\nabla f_{(x,y)}\right\|} \right\rangle = \left\|\nabla f_{(x,y)}\right\|.$$

□



Courbes de niveau de la fonction $f : (x, y) \mapsto xy$





4 Dérivées partielles d'ordre supérieur, fonctions de classe C^2

4.1 Dérivées partielles d'ordre supérieur

Soit O un ouvert de $\mathbb{R}^2$ et soit $f : O \to \mathbb{R}$ une fonction de classe C^1. On dispose alors de deux nouvelles fonctions, les dérivées partielles $(x,y) \mapsto \frac{\partial f}{\partial x}(x,y)$ et $(x,y) \mapsto \frac{\partial f}{\partial y}(x,y)$.

Définition 9.39. On dit que $f : O \to \mathbb{R}$ est de classe C^2 si elle est de classe C^1 et si ses dérivées partielles sont elle-mêmes de classe C^1 sur O.

Remarque 9.40. On peut, en itérant ce procédé, définir également les fonctions de classes C^n pour tout entier $n > 2$

Si f est de classe C^2, on pose $\frac{\partial^2 f}{(\partial x)^2} = \frac{\partial g}{\partial x}$ et $\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial g}{\partial y}$ où $g(x,y) = \frac{\partial f}{\partial x}(x,y)$ et $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial h}{\partial x}$ et $\frac{\partial^2 f}{(\partial y)^2} = \frac{\partial h}{\partial y}$ où $h(x,y) = \frac{\partial f}{\partial y}(x,y)$.

Théorème 9.41 (Théorème de Schwarz).
Si $f : O \to \mathbb{R}$ est de classe C^2, on a

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

Remarque 9.42. L'hypothèse de classe C^2 est nécessaire. En effet, la fonction donnée par

$$f(x,y) = \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) \neq (0,0) \end{cases}$$

est deux fois dérivable en $(0,0)$ mais on a

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(0,0) \neq \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(0,0).$$

4.2 Hessienne, Formule de Taylor à l'ordre 2

Définition 9.43. Si f est de classe C^2, on appelle *hessienne de f en (x,y)* l'application $Hf_{(x,y)} : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$Hf_{(x,y)}(u,v) = \frac{\partial^2 f}{(\partial x)^2}(x,y)\,u^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x,y)\,uv + \frac{\partial^2 f}{(\partial y)^2}(x,y)v^2.$$

On appelle *matrice hessienne* de f au point (x,y) la matrice définie par :

$$\mathrm{Hess}(f)_{(x,y)} = \begin{pmatrix} \frac{\partial^2 f}{(\partial x)^2}(x,y) & \frac{\partial^2 f}{\partial x \partial y}(x,y) \\ \frac{\partial^2 f}{\partial x \partial y}(x,y) & \frac{\partial^2 f}{(\partial y)^2}(x,y) \end{pmatrix}$$

Remarque 9.44. L'application $Hf_{(x,y)}$ est une forme quadratique sur $\mathbb{R}^2$ (voir le chapitre suivant), et la matrice hessienne est la matrice associée à cette forme quadratique dans la base canonique de $\mathbb{R}^2$.

On démontre, de manière analogue au théorème 9.25.

Théorème 9.45 (Formule de Taylor à l'ordre 2)**.** *Si f est de classe C^2, on a*

$$f(x+h,y+k) = f(x,y) + df_{(x,y)}(h,k) + \frac{1}{2}Hf_{(x,y)}(h,k) + o(\|(h,k)\|^2).$$

5 Extrema locaux

5.1 Définitions

Définition 9.46. Soit O un ouvert de $\mathbb{R}^n$, et soient $f : O \longrightarrow \mathbb{R}$ une application et $a \in O$.

1. On dit que f présente un maximum (respectivement un minimum) absolu en a ou que a est un *maximum (respectivement un* minimum) *absolu de f, si pour tout $x \in O$, on a $f(x) \leq f(a)$ $\big($respectivement $f(x) \geq f(a)\big)$.*
2. *On dit que f présente un maximum absolu (respectivement un minimum absolu)* strict en a si pour tout $x \in O$, $x \neq a$ on a $f(x) < f(a)$ $\big($respectivement $f(x) > f(a)\big)$.
3. On dit que f présente un maximum (respectivement un minimum) *local (éventuellement strict) en a s'il existe $r > 0$ tel que la restriction de f à $V = \{x \in O;\ \|x-a\| < r\}$ présente un maximum (respectivement un minimum) (éventuellement strict) en a.*
4. *On dit que f présente un* extremum (absolu ou local, strict ou non) en a si f présente en a un maximum ou un minimum (absolu ou local, strict ou non).

D'après les formules de Taylor, on a :

Proposition 9.47 (Extrema locaux)**.** *Soit O un ouvert de $\mathbb{R}^2$ et soient a un point de O et $f : O \longrightarrow \mathbb{R}$ une application.*

1. *Si f est de classe C^1 et présente un extremum local en a, alors ∇f_a est nul.*
2. *Si f est de classe C^2 et présente un minimum (respectivement un maximum) local en a, alors $Hf_a(h) \geq 0$ (respectivement $Hf_a(h) \leq 0$) pour tout $h \in \mathbb{R}^n$.*
3. *Si f est de classe C^2, si $df_a = 0$ et si $Hf_a(h) > 0$ (respectivement $Hf_a(h) < 0$) pour tout $h \in \mathbb{R}^n$ non nul, alors f présente un minimum (respectivement maximum) local en a.*

Démonstration. En effet,

1. si $df_a(v) \neq 0$, alors d'après la formule de Taylor à l'ordre 1, on a $\lim_{t\to 0} \frac{f(a+tv)-f(a)}{t} = df_a(v)$; donc $f(a+tv) - f(a)$ est du signe de $df_a(v)$ pour $t > 0$ assez petit et du signe opposé pour $t < 0$ avec $|t|$ petit.
2. De même, si $df_a = 0$ et $Hf_a(v) > 0$ et $Hf_a(w) < 0$, d'après la formule de Taylor à l'ordre 2, il vient $\lim_{t\to 0} \frac{f(a+tv)-f(a)}{t^2} = Hf_a(v) > 0$ – donc f n'a pas de maximum local en a, et $\lim_{t\to 0} \frac{f(a+tw)-f(a)}{t^2} = Hf_a(w) < 0$ – donc f n'a pas de minimum local en a.

3. On utilise la formule de Taylor à l'ordre 2, pour v un vecteur quelconque et $t \in \mathbb{R}$ petit : $f(a+tv) = f(a) + t^2 Hf_a(v) + o(t^2)$. Dans ce développement limité, le signe de $f(a+tv) - f(a)$ est donné, pour t proche de 0, par le signe de $Hf_a(v)$. Ainsi, si Hf_a est définie positive, f admet un minimum local (strict) en a. On raisonne de même lorsque Hf_a est définie négative.

□

Définition 9.48. Soit O un ouvert de $\mathbb{R}^2$ et soient a un point de O et $f : O \to \mathbb{R}$ une application de classe C^1.
On dit que a est un *point critique pour* f *si* $\nabla f_a = 0$.

5.2 Determination de la signature en un point critique

Définition 9.49. Soit $A \in M_n(\mathbb{R})$ une matrice symétrique réelle . On appelle signature de A le couple d'entiers (p_+, p_-) avec p_+ le nombre de valeurs propres de A strictement positives, et p_- le nombre de valeurs propres de A strictement négatives.

On montrera, dans le chapitre sur les formes quadratiques, que l'application $x \mapsto \langle Ax, x\rangle$ de $\mathbb{R}^n$ dans $\mathbb{R}$:
— est strictement positive (si $x \neq 0$) si la signature de A est $(n, 0)$;
— est strictement négative (si $x \neq 0$) si la signature de A est $(0, n)$;
— prend des valeurs strictement positives et des valeurs strictement négatives si la signature de A est (p_+, p_-) avec $p_+ p_- \neq 0$.
On se sert de cette propriété pour étudier le signe de $Hf_a(h,k) = \langle \mathrm{Hess}(f)_a(h,k), (h,k)\rangle$.

On pose $r = \frac{\partial^2 f}{(\partial x)^2}(a)$, $s = \frac{\partial^2 f}{\partial x\, \partial y}(a)$ et $t = \frac{\partial^2 f}{(\partial y)^2}(a)$.

On a alors $\mathrm{Hess}(f)_a = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$ et $Hf_a(h,k) = (rh^2 + 2s\, hk + t\, k^2)$

On sait que cette matrice $\mathrm{Hess}(f)_a = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$ est diagonalisable dans une base orthonormale de $\mathbb{R}^2$ et de déterminant égal au produit de ses valeurs propres $\delta = \lambda_1 \lambda_2$ et égal par ailleurs à $\delta = rt - s^2$. Ainsi les signes de λ_1 et λ_2 sont les mêmes si et seulement si $\delta > 0$. Ceci montre :

Proposition 9.50. *Soit O un ouvert de $\mathbb{R}^2$, et soient $f : O \to \mathbb{R}$ une application de classe C^2 et a un point critique de f.*
On pose $r = \frac{\partial^2 f}{(\partial x)^2}(a)$, $s = \frac{\partial^2 f}{\partial x\, \partial y}(a)$ et $t = \frac{\partial^2 f}{(\partial y)^2}(a)$.
On note $\delta_a = \det \mathrm{Hess}(f)_a = \det \begin{pmatrix} r & s \\ s & t \end{pmatrix}$. Alors

1. *Si $\delta_a < 0$, alors la signature de la hessienne Hf_a de f en a est $(1,1)$ et donc Hf_a ne garde pas un signe constant. On dit que a est un* point selle *ou* point col *de f.*
2. *Si $\delta_a > 0$, alors la signature de la hessienne Hf_a de f en a est*
   — *$(2,0)$ si $r > 0$ et donc, pour $(h,k) \neq (0,0)$, $Hf_a(h,k) > 0$. Ainsi f admet un minimum local strict en a.*

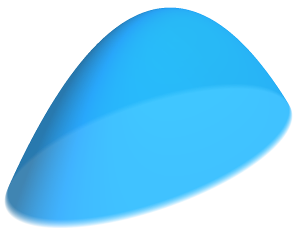
— $(0,2)$ *si* $r < 0$ *et donc, pour* $(h,k) \neq (0,0)$, $Hf_a(h,k) < 0$. *Ainsi* f *admet un maximum local strict en* a.

3. *Si* $\delta_a = 0$, *alors la forme quadratique* Hf_a *est dégénérée et on ne peut rien dire a priori sur le comportement de* f *au voisinage de* a.

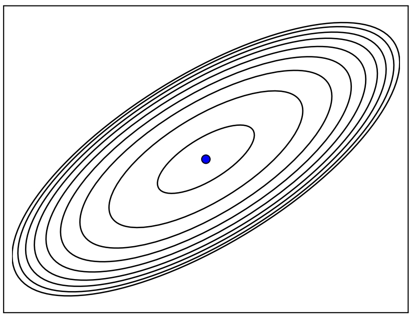
5.3 Representation graphique et lignes de niveaux au voisinage d'un point critique

Au voisinage d'un point critique a, on peut montrer que la surface et l'allure des courbes de niveaux au voisinage du point a sont de la forme suivante :

Si f admet un extremum local strict en a

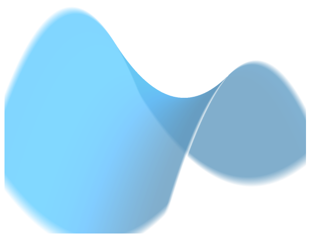


maximum local

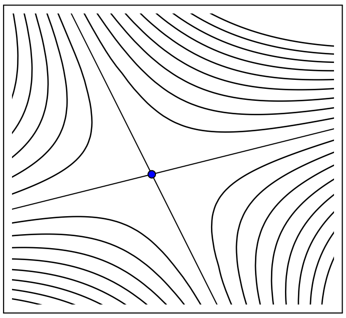


$f(x,y) = -2x^2 + 4xy - 4y^2$
$rt > s^2$: Extremum local

Si f admet un point col en a



point col



$f(x, y) = x^2 - 2xy - 2y^2$
$rt < s^2$: point col

6 Généralisation aux fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$

On considère maintenant plus généralement le cas d'applications de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$. On munit $\mathbb{R}^n$ et $\mathbb{R}^p$ de leurs normes euclidiennes canoniques respectives (ou de n'importe quelle norme). Si O est un ouvert de $\mathbb{R}^n$ et $f : O \to \mathbb{R}^p$ une application, on note $f_1, f_2, \ldots f_p$ les fonctions coordonnées de f dans la base canonique : pour tout $x \in O$, $f(x) = (f_1(x), f_2(x), \ldots, f_p(x))$ définit naturellement p fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$.

Proposition 9.51. *Si O est un ouvert de $\mathbb{R}^n$, la donnée d'une application $f : O \to \mathbb{R}^p$ équivaut à la donnée de ses p fonctions coordonnées $f_1, f_2, \ldots f_p$ de $\mathbb{R}^n$ dans $\mathbb{R}$.*

6.1 Continuité

Soit O un ouvert de $\mathbb{R}^n$, et $f : O \to \mathbb{R}^p$ une application.
On peut définir naturellement la continuité dans ce cadre.

Définition 9.52. Une fonction $f : O \to \mathbb{R}^p$ est *continue* en un point $a \in O$ si :

$$\forall \varepsilon > 0 \quad \exists \eta > 0 \quad \forall x \in O \quad \big(\|x - a\| < \eta \implies \|f(x) - f(a)\| < \varepsilon\big).$$

La fonction f est *continue sur O* si elle est continue en tout point de O.

On a, comme annoncé :

Proposition 9.53. *Soit O un ouvert de $\mathbb{R}^n$, et $f : O \to \mathbb{R}^p$ une application. La continuité de f en un point $a \in O$ équivaut à la continuité de chacune de ses p fonctions coordonnées $f_1, f_2, \ldots f_p$ en a.*

Démonstration. Soit $i \in \{1, \ldots, p\}$, si on suppose que f est continue en a, alors f_i l'est aussi car

$$\left|f_i(x) - f_i(a)\right| \leq \sqrt{\sum_{k=1}^{p} \left|f_k(x) - f_k(a)\right|^2}.$$

Réciproquement, si chacune des fonctions f_i est continue en a, pour tout $\varepsilon > 0$ il existe α_i tel que $\|x - a\| < \alpha_i$ implique $\left|f_i(x) - f_i(a)\right| < \frac{\varepsilon}{\sqrt{p}}$. Ainsi, en posant $\eta = \min\{\alpha_i, 1 \leq i \leq p\}$

, on obtient que $\|x - a\| < \eta$ implique

$$\|f(x) - f(a)\| = \sqrt{\sum_{i=1}^{p} \left|f_i(x) - f_i(a)\right|^2} < \sqrt{\sum_{i=1}^{p} \frac{\varepsilon^2}{p}} = \varepsilon.$$

□

On a également que :

Théorème 9.54. *Une composée de fonctions continues est continue.*

Démonstration. Soient U un ouvert de $\mathbb{R}^m$, V un ouvert de $\mathbb{R}^p$; soient aussi $f : U \to V$, $g : V \to \mathbb{R}^q$ des applications et $u \in U$. Supposons que f est continue en u et g est continue en $f(u)$.

Soit alors $\varepsilon > 0$.

— Comme g est continue en $f(u)$, il existe $\beta > 0$ tel que, pour $y \in V$ avec $\|y - f(u)\| < \beta$ on ait $\|g(y) - g(f(u))\| < \varepsilon$.

— Comme f est continue en u, il existe $\alpha > 0$ tel que, pour $x \in U$ avec $\|x - u\| < \alpha$ on ait $\|f(x) - f(u)\| < \beta$,
– et donc $\|g(f(x)) - g(f(u))\| < \varepsilon$.

Cela prouve que $g \circ f$ est continue en u. □

6.2 Fonctions de classe C^1

Définition 9.55. Soit O un ouvert de $\mathbb{R}^n$, et $f : O \to \mathbb{R}^p$ une application.

1. On dit que f admet des dérivées partielles si chacune de ses fonctions coordonnées admet des dérivées partielles, et qu'elle est de classe C^1 lorsque chacune de ses fonctions coordonnées est de classe C^1.
2. Lorsque f est de classe C^1, on peut assembler les gradients (ou les différentielles) de chaque application coordonnée au point $a \in O$ en une matrice à p lignes et n colonnes que l'on appelle *matrice jacobienne de* f au point a, dont la ligne i est donnée par les coordonnées du gradient de la fonction coordonnée f_i en a :

$$\mathrm{Jac(f)}_a = \begin{pmatrix} \nabla(f_1)_a \\ \nabla(f_2)_a \\ \vdots \\ \nabla(f_p)_a \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \dots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \frac{\partial f_p}{\partial x_2}(a) & \dots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix}$$

3. Lorsque f est de classe C^1, on appelle différentielle de f au point a l'application linéaire $df_a : \mathbb{R}^n$ dans $\mathbb{R}^p$ dont la matrice dans les bases canonique est la matrice jacobienne ci-dessus.

Avec ces définitions, les formules de Taylor à l'ordre 1 pour les fonctions coordonnées peuvent être assemblées en une formule globale :

Théorème 9.56. *Si une application f est de classe C^1 d'un ouvert $O \subset \mathbb{R}^n$ vers $\mathbb{R}^p$, on a, pour $a \in O$ et pour tout $h \in \mathbb{R}^n$ de norme suffisamment petite :*

$$f(a + h) = f(a) + df_a(h) + o(\|h\|).$$

6.3 Fonctions de classe C^2 dans le cas d'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$

Une application f d'un ouvert de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ et de classe C^2 si toutes les dérivées partielles de ses fonctions coordonnées sont elles-mêmes des fonctions de classe C^1. On définit ainsi par itération la notion d'application de classe C^n.
Le théorème de Schwarz se généralise dans ce cadre

Théorème 9.57 (Théorème de Schwarz)**.**
Si $f : O \longrightarrow \mathbb{R}^p$ est de classe C^2, on a pour tous i et j entre 1 et n :

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

On peut également écrire une formule de Taylor à l'ordre 2 pour une application $f : O \longrightarrow \mathbb{R}^p$ de classe C^2, cependant la formule fait intervenir la différentielle seconde de f qui sera une application bilinéaire de $\mathbb{R}^n \times \mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$. La situation se simplifie beaucoup si l'on choisit $p = 1$. Dans ce cas en effet, la différentielle seconde apparait comme une forme quadratique sur $\mathbb{R}^n$, la hessienne.

Définition 9.58. Si f est de classe C^2, on appelle *hessienne de f en a* l'application $Hf_a : \mathbb{R}^n \longrightarrow \mathbb{R}$ définie par

$$Hf_a(u_1, u_2, \ldots, u_n) = \sum_{i=1}^{n} \frac{\partial^2 f}{(\partial x_i)^2}(a)\, u_i^2 + 2 \sum_{1 \leq i < j \leq n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\, u_i u_j.$$

On appelle *matrice hessienne* de f au point a la matrice de cette forme quadratique dans la base canonique de $\mathbb{R}^n$.

$$\mathrm{Hess}(f)_{(x,y)} = \begin{pmatrix} \frac{\partial^2 f}{(\partial x_1)^2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \ldots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \frac{\partial^2 f}{(\partial x_2)^2}(a) & \ldots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(a) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \frac{\partial^2 f}{\partial x_n \partial x_2}(a) & \ldots & \frac{\partial^2 f}{(\partial x_n)^2}(a) \end{pmatrix}$$

Théorème 9.59 (Formule de Taylor à l'ordre 2)**.** *Si f est de classe C^2, on a*

$$f(a+h) = f(a) + df_a(h) + \frac{1}{2} Hf_a(h) + o(\|h\|^2).$$

Extrema

Proposition 9.60 (Extrema locaux)**.** *Soit O un ouvert de $\mathbb{R}^n$ et soient a un point de O et $f : O \longrightarrow \mathbb{R}$ une application.*

1. *Si f est de classe C^1 et présente un extremum local en a, alors ∇f_a est nul.*
2. *Si f est de classe C^2 et présente un minimum (respectivement un maximum) local en a, alors $Hf_a(h) \geq 0$ (respectivement $Hf_a(h) \leq 0$) pour tout $h \in \mathbb{R}^n$.*
3. *Si f est de classe C^2, si $df_a = 0$ et si $Hf_a(h) > 0$ (respectivement $Hf_a(h) < 0$) pour tout $h \in \mathbb{R}^n$ non nul, alors f présente un minimum (respectivement maximum) local en a.*

On a donc, pour un point critique d'une fonction de classe C^2 d'un ouvert de $\mathbb{R}^n$ dans $\mathbb{R}$.

Proposition 9.61. *Soit O un ouvert de $\mathbb{R}^n$ et soient a un point crtitque de $f : O \longrightarrow \mathbb{R}$, une application de classe C^2.*

1. *Si la forme quadratique Hf_a est définie positive, c'est à dire de signature $(n, 0)$ alors f admet un minimum local strict en a.*
2. *Si la forme quadratique Hf_a est définie négative, c'est à dire de signature $(0, n)$ alors f admet un maximum local strict en a.*
3. *Si la forme quadratique Hf_a est de signature (ℓ, k) avec $\ell k \neq 0$, alors $f - f(a)$ n'est pas de signe constant au voisinage de a. Le point a est un point col de f.*
4. *Si la forme quadratique Hf_a est dégénérée et de signe constant, c'est à dire de signature (ℓ, k) avec $k + \ell < n$ et $\ell k = 0$, alors on ne peut pas conclure sur la nature du point a pour f, en général.*

6.4 Exemple en dimension 3

Soit O un ouvert de $\mathbb{R}^3$. Une fonction $f : O \longrightarrow \mathbb{R}$ est de classe C^1 si elle a trois dérivées partielles $\frac{\partial f}{\partial x}(x, y, z)$, $\frac{\partial f}{\partial y}(x, y, z)$, $\frac{\partial f}{\partial z}(x, y, z)$ en tout point (x, y, z) de O et si les applications $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$, $\frac{\partial f}{\partial z}$ sont continues.

Le gradient de f en (x, y, z) s'écrit

$$\nabla f_{(x,y,z)} = \Big(\frac{\partial f}{\partial x}(x, y, z), \frac{\partial f}{\partial y}(x, y, z), \frac{\partial f}{\partial z}(x, y, z)\Big).$$

La différentielle de f en un point (x, y, z) s'écrit :

$$df_{(x,y,z)}(h, k, \ell) = \frac{\partial f}{\partial x}(x, y, z)h + \frac{\partial f}{\partial y}(x, y, z)k + \frac{\partial f}{\partial z}(x, y, z)\ell.$$

La formule de Taylor a l'ordre 1 s'écrit :

$$f(x + h, y + k, z + \ell) = f(x, y, z) + df_{(x,y,z)}(h, k, \ell) + o(\|(h, k, \ell)\|).$$

$$\begin{aligned} Hf_X(u, v, w) \quad = \quad & \frac{\partial^2 f}{(\partial x)^2}(X)u^2 + \frac{\partial^2 f}{(\partial y)^2}(X)v^2 + \frac{\partial^2 f}{(\partial z)^2}(X)w^2 \\ & +2\frac{\partial^2 f}{\partial x \partial y}(X)\, uv + 2\frac{\partial^2 f}{\partial x \partial z}(X)\, uw + 2\frac{\partial^2 f}{\partial y \partial z}(X)\, vw. \end{aligned}$$

La formule de Taylor a l'ordre 2 s'écrit :

$$f(x + h, y + k, z + l) = f(x, y, z) + df_{(x,y,z)}(h, k, l) + \frac{1}{2}Hf_{(x,y,z)}(h, k, l) + o(\|(h, k, l)\|^2).$$

En un point critique, la signature de la forme quadratique hessienne en ce point permet de déterminer si ce point critique est un extremum local ou un point col, à condition que la forme quadratique soit non dégénérée.

... et de même en dimension 4, en dimension n...

Chapitre 10

Intégrales dépendant d'un paramètre

1 Intégrales définies à paramètre

Soit I un intervalle de $\mathbb{R}$ et $J = [a, b]$ un segment de $\mathbb{R}$, et soit $f : I \times J \to \mathbb{R}$ une fonction. On va s'intéresser à des intégrales de la forme :

$$F(x) = \int_a^b f(x, t)dt.$$

On va énoncer des théorèmes qui donnent des conditions *suffisantes* pour garantir la continuité et la dérivabilité de la fonction F.

1.1 Continuité et limite.

Théorème 10.1 (continuité des intégrales définies à paramètres)**.** *Soit I un intervalle de $\mathbb{R}$, soit $J = [a, b]$ un segment de $\mathbb{R}$, et soit $f : I \times J \to \mathbb{R}$ une fonction continue. Alors la fonction $F : I \to \mathbb{R}$ définie par :*

$$F(x) = \int_a^b f(x, t)dt$$

est une fonction qui est bien définie et continue sur I.

Démonstration. L'intégrale $F(x)$ est bien définie pour chaque $x \in I$ car la fonction partielle $t \mapsto f(x, t)$ est continue sur J. Fixons $x_0 \in I$ et montrons que F est continue en x_0. Il existe un segment I' contenu dans I et qui contient x_0. En effet il existe $\alpha > 0$ tel que

- $I' = [x_0 - \alpha, x_0 + \alpha] \subset I$ si x_0 est un point intérieur à I,
- ou bien $I' = [x_0 - \alpha, x_0] \subset I$ si $x_0 = \max(I)$,
- ou bien $I' = [x_0, x_0 + \alpha] \subset I$ si $x_0 = \min(I)$.

Soit $\varepsilon > 0$.
D'après le théorème de Heine pour les fonctions de plusieurs variables (théorème 9.20), la fonction f est uniformément continue sur $I' \times [a, b]$. Choisissons $\eta > 0$ associé à $\varepsilon / \max(1, |b - a|)$ pour la continuité uniforme, c'est à dire tel que :

$$\forall (x, t), (x', t') \in I' \times [a, b], \|(x, t) - (x', t')\| < \eta \Rightarrow \left|f(x, t) - f(x', t')\right| < \frac{\varepsilon}{\max(1, |b - a|)}.$$

Soit $x \in I'$ tel que $|x - x_0| < \eta$. Alors :

$$\left|F(x) - F(x_0)\right| \leq \int_a^b \left|f(x, t) - f(x_0, t)\right|dt < \varepsilon,$$

ce qui montre la continuité en x_0, et donc sur I (puisque c'est vrai en tout point x_0 de I). $\square$

Théorème 10.2 (limite des intégrales définies à paramètre). *Soit I un intervalle de $\mathbb{R}$, soit $J = [a, b]$ un segment de $\mathbb{R}$, et soit $f : I \times J \longrightarrow \mathbb{R}$ une fonction continue. On suppose que $x_0 \in \overline{I}$ (c'est à dire que x_0 est à l'intérieur de I ou une borne de I, y compris infinie) et que $\lim_{x \to x_0} f(x,t) = h(t)$ en supposant que la convergence de $f(x,.)$ vers h est uniforme sur J, c'est à dire qu'il existe α une fonction définie sur I et telle que :*

$$|f(x,t) - h(t)| \leq \alpha(x), \qquad \lim_{x \to x_0} \alpha(x) = 0.$$

Alors h est continue sur J et on peut intervertir limite et intégrale :

$$\lim_{x \to x_0} \int_a^b f(x,t)dt = \int_a^b h(t)dt.$$

Démonstration. — Montrons tout d'abord que sous les hypothèses du théorème, la fonction h est continue sur J. Soit $t_0 \in J$ et $\varepsilon > 0$, et $x_1 \in I$ tel que $\alpha(x_1) < \frac{\varepsilon}{3}$. On sait que f est continue au point (x_1, t_0) et donc qu'il existe $\eta > 0$ tel que $\forall (x,t) \in I \times J, \|(x,t) - (x_1,t_0)\| < \alpha \Rightarrow \left|f(x,t) - f(x_1,t_0)\right|$. On en déduit que si $t \in J$ et $|t - t_0| < \eta$, alors

$$\begin{aligned} |h(t) - h(t_0)| &\leq \left|h(t) - f(x_1,t)\right| + \left|f(x_1,t) - f(x_1,t_0)\right| + \left|f(x_1,t_0) - h(t_0)\right| \\ &\leq \alpha(x_1) + \left|f(x_1,t) - f(x_1,t_0)\right| + \alpha(x_1) \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned}$$

On a donc bien montré la continuité de h.

— On évalue à présent la différence :

$$\left|\int_a^b f(x,t)dt - \int_a^b h(t)dt\right| \leq \int_a^b \left|f(x,t) - h(t)\right|dt \leq (b-a)\alpha(x) \to_{x \to x_0} 0.$$

□

Remarque 10.3.

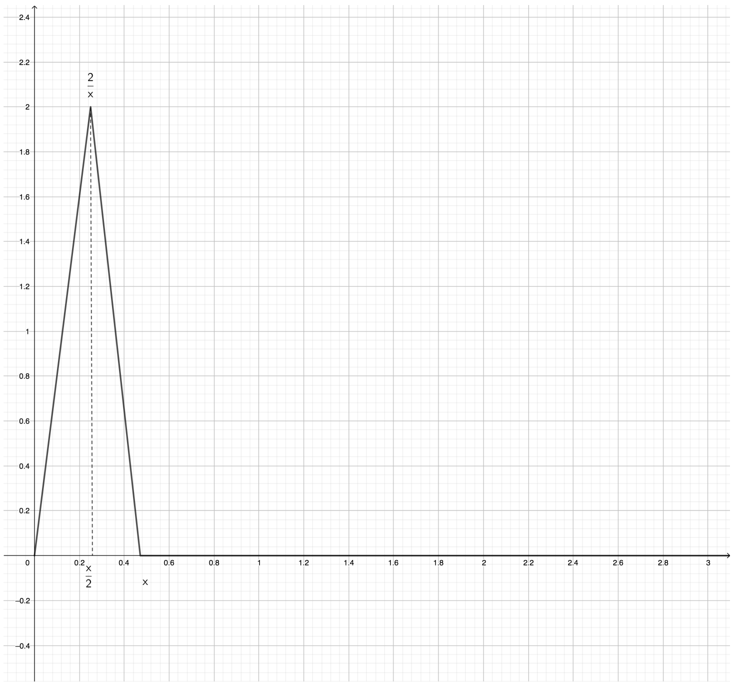
1. Ce théorème n'est vraiment intéressant qu'aux bornes de I qui ne sont pas contenues dans I. En effet si x_0 est un point intérieur de I, le théorème précédent nous assure que l'interversion a bien lieu, en posant $h(t) = f(x_0, t)$. La condition de convergence uniforme sur J de $f(x,.)$ vers h est satisfaite automatiquement grâce à l'uniforme continuité de f sur un produit de segments de la forme $I' \times J$ où I' est un segment contenu dans I et qui contient x_0.
2. L'hypothèse de l'existence de la limite $h(t)$ pour tout $t \in J$ est indispensable en général.
   Si l'on prend la fonction $f :]0,1[\times [0,1] \longrightarrow \mathbb{R}$ donnée par $f(x,t)\frac{1}{x+t}$, cette fonction est bien continue sur $]0,1[\times [0,1]$ et pour chaque $t \neq 0$, on a $f(x,t) \longrightarrow \frac{1}{t}$ mais l'intégrale $\int_0^1 \frac{dt}{t}$ est divergente.
3. L'hypothèse de convergence uniforme sur J de $f(x,.)$ vers h est indispensable en général.

Si l'on prend la fonction $f :]0,1[\times[0,1] \to \mathbb{R}$ donnée par

$$f(x,t) = \begin{cases} \frac{2t}{x^2} & \text{si } 0 \leq t \leq \frac{x}{2} \\ \frac{2(2x-t)}{x^2} & \text{si } \frac{x}{2} \leq t \leq x \\ 0 & \text{si } t \geq x \end{cases}$$

On peut montrer que cette cette fonction est bien continue sur $]0,1[\times[0,1]$ et pour chaque $t \in [0,1]$, on a $f(x,t) \underset{x\to 0}{\longrightarrow} h(t) = 0$ mais

$$\int_0^1 f(x,t)dt = 1 \underset{x\to 0}{\longrightarrow} 1 \neq 0 = \int_0^1 \lim_{x\to 0} f(x,t)dt.$$



Exemple 10.4.

1.2 Dérivabilité.

Théorème 10.5. *Soit $f : I \times J \to \mathbb{R}$ une fonction continue, où I est un intervalle de $\mathbb{R}$ et $J = [a,b]$ un segment. On suppose que $\partial f/\partial x$ existe et est continue sur $I \times J$. Alors la fonction $F : I \to \mathbb{R}$ définie par*

$$F(x) = \int_a^b f(x,t)dt$$

est de classe C^1 et vérifie :

$$F'(x) = \int_a^b \frac{\partial f}{\partial x}(x,t)dt \qquad \text{(on dérive sous le signe } \textstyle\int\text{).}$$

Démonstration. Supposons d'abord que J est un intervalle fermé et borné.

Soit $x_0 \in I$, montrons que F est dérivable en x_0 avec $F'(x_0)$ comme dans l'énoncé. On considère un segment I' qui contient x_0 et qui est contenu dans I. Comme dans la démonstration du théorème 10.1, on choisit :

— $I' = [x_0 - \alpha, x_0 + \alpha] \subset I$ si x_0 est un point intérieur à I,
— ou bien $I' = [x_0 - \alpha, x_0] \subset I$ si $x_0 = \max(I)$,
— ou bien $I' = [x_0, x_0 + \alpha] \subset I$ si $x_0 = \min(I)$.

Posons, pour tout $h \neq 0$ et tel que $x_0 + h \in I'$:

$$\begin{aligned}\delta(h) &= \frac{1}{h}\Big(F(x_0+h) - F(x_0)\Big) - \int_a^b \frac{\partial f}{\partial x}(x_0,t)dt \\ &= \int_a^b \left(\frac{1}{h}\Big(f(x_0+h,t) - f(x_0,t)\Big) - \frac{\partial f}{\partial x}(x_0,t)\right)dt\end{aligned}$$

Choisissons $\varepsilon > 0$. La fonction $\dfrac{\partial f}{\partial x}$ est continue sur $I \times J$ par hypothèse, donc est uniformément continue sur $I' \times J$ d'après le théorème de Heine 9.20. Soit donc $\eta > 0$ un module de continuité uniforme associé à ε.

Pour $t \in [a,b]$ fixé, appliquons l'inégalité des accroissements finis à la fonction

$$g : h \mapsto f(x_0+h,t) - h\frac{\partial f}{\partial x}(x_0,t).$$

La fonction g est dérivable, de dérivée

$$g'(h) = \frac{\partial f}{\partial x}(x_0+h,t) - \frac{\partial f}{\partial x}(x_0,t),$$

donc $|g'(t)| \leq \varepsilon$ pour $|h| < \eta$. Donc, pour $|h| < \eta$, on a :

$$\Big|g(h) - g(0)\Big| \leq \varepsilon|h|, \qquad \text{c'est-à-dire : } \quad \left|\frac{f(x_0+h,t) - f(x_0,t)}{h} - \frac{\partial f}{\partial x}(x_0,t)\right| \leq \varepsilon.$$

On en déduit $|\delta(h)| \leq \varepsilon(b-a)$, ce qui montre que $F'(x_0)$ existe et a bien la valeur annoncée.
Enfin, la continuité de $F'(x) = \int_a^b \frac{\partial f}{\partial x}(x,t)dt$ sur I se déduit du théorème 10.1. □

2 Intégrales généralisées à paramètre

Étant donnés un intervalle I de $\mathbb{R}$ et $f : I \times [0, +\infty[$ une fonction continue, on s'intéresse aux propriétés de la fonction $F : I \to \mathbb{R}$ donnée, sous réserve d'existence, par :

$$F(x) = \int_0^\infty f(x,t)dt.$$

On formulera tous les énoncés dans ce cadre, mais ils s'adaptent sans difficulté à des intervalles d'intégration différents.

Définition 10.6. L'intégrale

$$F(x) = \int_0^\infty f(x,t)dt$$

converge *simplement sur* I si l'intégrale $F(x)$ est convergente pour chaque $x \in I$.

L'intégrale $F(x)$ est *normalement convergente* sur I s'il existe une fonction $g : [0, +\infty[\longrightarrow \mathbb{R}_+$ continue et telle que :

$$(*) \quad \forall (x,t) \in I \times [0, +\infty[\quad |f(x,t)| \leq g(t), \qquad \int_0^\infty g(t)dt \text{ converge.}$$

On dit que g est une *fonction de domination* pour f.

Si $F(x)$ est normalement convergente sur I, en particulier elle converge absolument en tout point de I, et elle est donc simplement convergente sur I.

Lemme 10.7. *Supposons que $F(x)$ définie comme ci-dessus converge normalement sur I. Pour chaque entier n, soit $\varphi_n : I \longrightarrow \mathbb{R}$ définie par :*

$$\varphi_n(x) = \int_0^n f(x,t)dt.$$

Alors $(\varphi_n)_{n\geq 0}$ est une suite de fonctions convergeant uniformément sur I vers F.

Démonstration. Soit g une fonction de domination pour f. En effet, pour chaque $x \in I$, on a :

$$\left|\varphi_n(x) - F(x)\right| \leq \int_n^\infty |f(x,t)|dt \leq \int_n^\infty g(t)dt.$$

Le membre de droite ne dépend pas de x, on a donc la majoration suivante de la norme uniforme :

$$\|\varphi_n - F\|_\infty \leq \int_n^\infty g(t)dt,$$

qui tend vers 0 quand $n \to \infty$ puisque l'integrale $\int_0^\infty g(t)dt$ est convergente. □

Théorème 10.8 (continuité des intégrales généralisées à paramètre)**.** *Supposons que $f : I \times [0, +\infty[\longrightarrow \mathbb{R}$ est une fonction continue sur $I \times [0, +\infty[$ et que l'intégrale $F(x)$ est normalement convergente sur I. Alors F est une fonction continue sur I.*

Démonstration. Soit $(\varphi_n)_{n\geq 0}$ la suite de fonctions définies dans le lemme 10.7. Pour chaque entier n, la fonction φ_n est continue sur I d'après le théorème de continuité des intégrales définies à paramètre (théorème 10.1). Comme F est la limite uniforme sur I d'une suite de fonctions continues, elle est elle-même continue sur I. □

Théorème 10.9 (limite des intégrales généralisées à paramètres)**.** *Soit I un intervalle de $\mathbb{R}$, $x_0 \in \overline{I}$ (c'est à dire que x_0 est à l'intérieur de I ou une borne de I, y compris infinie). On considère $f : I \times [0, +\infty[\longrightarrow \mathbb{R}$ une fonction continue sur $I \times [0, +\infty[$. On suppose que :*

1. *L'intégrale* $F(x) = \int_0^\infty f(x,t)dt$ *est normalement convergente sur* I.
2. *Pour chaque* $t \in [0, +\infty[$, *la limite :*

$$h(t) = \lim_{x \to x_0} f(x,t)$$

existe.

3. *On suppose de plus, cette convergence est uniforme en* t *sur chaque intervalle de la forme* $[0, A]$ *avec* $A \in \mathbb{R}_+$, *c'est à dire que pour tout* $A > 0$, *il existe une fonction* α *telle que*

$$|f(x,t) - h(t)| \leq \alpha(x), \qquad \lim_{x \to x_0} \alpha(x) = 0.$$

Alors, la fonction h *est continue, l'intégrale :*

$$\int_0^\infty h(t)dt$$

est absolument convergente, et on a :

$$\lim_{x \to x_0} F(x) = \int_0^\infty h(t)dt.$$

Démonstration. Soit $(\varphi_n)_{n \geq 0}$ la suite de fonction définie comme dans le lemme 10.7. Soit $n \in \mathbb{N}$ fixé ; montrons que $\lim_{x \to x_0} \varphi_n(x) = \int_0^n h(t)dt$.

En effet, puisque la convergence $f(x,t) \to_{x \to x_0} h(t)$ est supposée uniforme en t sur tout intervalle de la forme $[0, A]$, en particulier sur $[0, n]$, on est dans le cas d'application du théorème d'inversion limite et intégrale (théorème 10.2).

Puis on applique le théorème sur les limites des limites uniformes de suites de fonctions puisque $(\varphi_n)_{n \geq 0}$ converge uniformément vers F d'après le lemme 10.7. □

Théorème 10.10 (dérivabilité des intégrales généralisées à paramètre). *Soit* $f : I \times [0, +\infty[\to \mathbb{R}$ *une fonction. On suppose que :*

1. $\int_0^\infty f(x,t)dt$ *converge simplement sur* I.
2. f *a une dérivée partielle par rapport à* x, *qui est continue sur* $I \times [0, +\infty[$.
3. *L'intégrale :*

$$\int_0^\infty \frac{\partial f}{\partial x}(x,t)dt$$

converge normalement sur I.

Alors

$$F(x) = \int_0^\infty f(x,t)dt$$

est dérivable sur I, *de dérivée :*

$$F'(x) = \int_0^\infty \frac{\partial f}{\partial x}(x,t)dt.$$

Démonstration. Pour chaque entier $n \geq 0$, soit $\varphi_n : I \to \mathbb{R}$ la fonction définie par :

$$\varphi_n(x) = \int_0^n f(x,t)dt.$$

Par hypothèse, la suite $(\varphi_n)_{n\geq 0}$ converge simplement sur I vers F. D'après le théorème de dérivation des intégrales définies à paramètres (théorème 10.5), chaque fonction φ_n est de classe C^1 sur I, de dérivée :

$$\varphi_n'(x) = \int_0^n \frac{\partial f}{\partial x}(x,t)dt.$$

D'après le lemme 10.7 appliqué à l'intégrale $K(x) = \int_0^\infty \partial f/\partial x(x,t)dt$, la suite de fonctions $(\varphi_n')_{n\geq 0}$ converge uniformément vers la fonction K sur I.

D'après le théorème de dérivabilité des limites uniformes, on en déduit que F est dérivable sur I, de dérivée K. □

Exemple.

Soit $F(x)$ l'intégrale :

$$F(x) = \int_0^\infty \frac{\exp(-tx)}{1+t^2}dt\,.$$

Montrer que F est solution sur $\mathbb{R}_+^*$ de limite nulle en $+\infty$ de l'équation différentielle $y'' + y = 1/x$.

Soit $f(x,t) = \exp(-xt)/(1+t^2)$ pour $(x,t) \in \mathbb{R}_+^* \times [0,+\infty[$. On a la domination : $|f(x,t)| \leq 1/(1+t^2)$, donc $F(x)$ est bien définie sur $\mathbb{R}_+^*$, et en fait sur $\mathbb{R}_+$.

Montrons que $\lim_{x\to\infty} F(x) = 0$. Soit $\varepsilon > 0$. Pour $\alpha > 0$ qu'on choisira plus tard, on a :

$$F(x) = \int_0^\alpha f(x,t)dt + \int_\alpha^\infty f(x,t)dt.$$

Examinons chacun des deux termes :

$$\int_0^\alpha f(x,t)dt \leq \alpha.$$

Pour le deuxième terme, on cherche une majoration du terme à intégrer uniforme en t. En effet, pour $t \geq \alpha$, on a :

$$\frac{\exp(-tx)}{1+t^2} \leq \exp(-\alpha x),$$

et le membre de droite tend vers 0 indépendamment de t. Donc, d'après le théorème de limite des intégrales généralisées, on a :

$$\lim_{x\to\infty} \int_\alpha^\infty f(x,t)dt < \varepsilon/2$$

pour x assez grand. Il suffit donc de prendre $\alpha < \varepsilon/2$ pour avoir $0 \leq F(x) < \varepsilon$ pour x assez grand, ce qui montre :

$$\lim_{x\to\infty} F(x) = 0.$$

Étudions la dérivabilité de F. La fonction f est continue et admet une dérivée partielle par rapport à x, donnée par :

$$\frac{\partial f}{\partial x}(x,t) = \frac{-t\exp(-tx)}{1+t^2}.$$

La majoration $\exp(-tx) \leq 1$ ne permet pas de conclure à la convergence normale sur $[0,+\infty[$, il faut donc une majoration plus fine.

Soit $\alpha > 0$. Pour $x \in [\alpha,+\infty[$, et pour tout $t \in [0,+\infty[$, on a :

$$0 \leq \frac{-t\exp(-tx)}{1+t^2} \leq \frac{t\exp(-t\alpha)}{1+t^2}$$

et le membre de droite converge et est indépendant de x. Le théorème 10.10 s'applique, et F est de classe C^1 sur $[\alpha,+\infty[$ de dérivée :

$$F'(x) = \int_0^\infty \frac{-t\exp(-tx)}{1+t^2}dt.$$

Ceci étant vrai pour tout $\alpha > 0$, on en déduit que F est dérivable sur $]0,+\infty[$. Un raisonnement similaire permet d'obtenir que F' est C^1, de dérivée :

$$F''(x) = \int_0^\infty \frac{t^2\exp(-tx)}{1+t^2}dt.$$

On en déduit :

$$F(x) + F''(x) = \int_0^\infty \exp(-tx)dt = \left[\frac{\exp(-tx)}{-x}\right]_{t=0}^{t=\infty} = \frac{1}{x}.$$

Chapitre 11

Intégrales doubles

1 Intégrale de Riemann sur un pavé rectangulaire

On peut transposer les étapes suivies pour la construction de l'intégrale des fonctions réglées sur un segment I de $\mathbb{R}$ aux fonctions réglées sur un pavé $I \times J$ de $\mathbb{R}^2$.

Définition 11.1. Soient I et J deux segments de $\mathbb{R}$. Une fonction f sur $I \times J$ est une *fonction en escalier* s'il existe des subdivisions $\sigma = \{x_0, x_1, \ldots, x_k, \ldots, x_n\}$ du segment I et $\tau = \{y_0, x_1, \ldots, x_\ell, \ldots, y_m\}$ du segment J, telle que f est constante sur chaque pavé ouvert $]x_k, x_{k+1}[\times]y_\ell, y_{\ell+1}[$ pour $0 \le k \le n-1, 0 \le \ell \le m-1$. On note $\lambda_{k,\ell}$ la valeur de f sur $]x_k, x_{k+1}[\times]y_\ell, y_{\ell+1}[$ et on note $I_k =]x_k, x_{k+1}[$ et $J_\ell =]y_\ell, y_{\ell+1}[$

L'*intégrale* d'une telle fonction en escalier est définie par :

$$\int_{I\times J} f(x,y)dxdy = \sum_{0\le k\le n-1, 0\le \ell\le m-1} \lambda_{k,\ell} A(I_k \times J_\ell)$$
$$\text{avec} \quad A(]a,b[\times]c,d[) = (b-a)(d-c),$$

et ce nombre ne dépend pas de la décomposition de f comme combinaison linéaire de fonctions caractéristiques de pavés. On remarque également, puisque l'aire d'un segment (horizontal ou vertical) est nulle que seule la valeur de f sur l'intérieur des pavés contribue à l'intégrale : on peut changer la valeur d'une fonction en escalier sur le bord d'un rectangle sans changer la valeur de l'intégrale.

Une fonction réglée sur $I \times J$ est la limite uniforme de fonctions en escalier sur $I \times J$.

On a alors les résultats standards sur l'intégrale des fonctions en escalier : c'est une forme linéaire positive. Puis on étend la définition de l'intégrale aux fonctions réglées, en procédant comme pour les fonctions réglées sur un intervalle de $\mathbb{R}$. L'intégrale ainsi définie est toujours une forme linéaire positive sur l'espace des fonctions réglées.

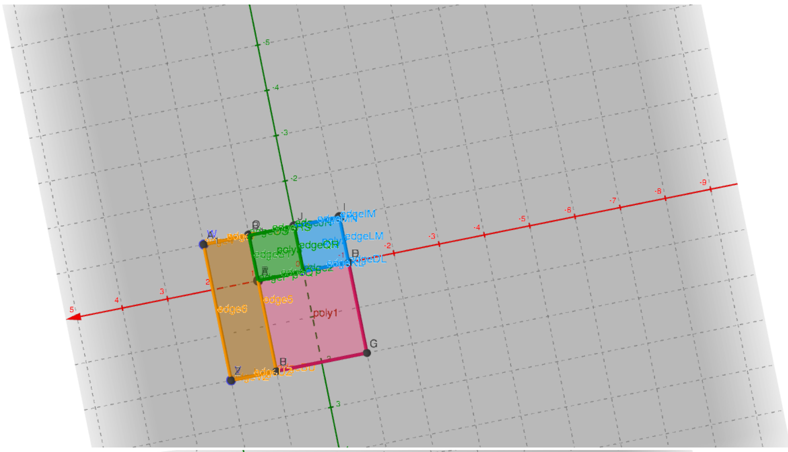
Par ailleurs, on vérifie grâce au théorème de Heine que toute fonction continue sur $I \times J$ est une fonction réglée sur $I \times J$.

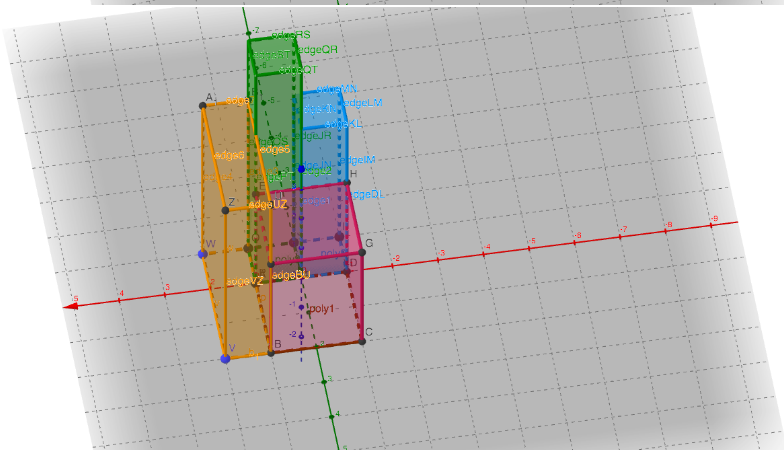
2 Aire d'un domaine, Intégrale sur un domaine

2.1 Aire d'un domaine, domaine quarrable

Étant donné un pavé fermé (un produit de segments) $I \times J$, et une partie $D \subseteq I \times J$, on aimerait définir l'aire de D par

$$a(D) = \int_D dxdy = \int_{I\times J} \mathbb{1}_D(x,y)dxdy\,.$$





où $\mathbb{1}_D$ est la fonction caractéristique de D.

Ceci n'est pas possible en général ! C'est l'étude de cette question abordée par Jordan qui a conduit Borel et Lebesgue à la définition d'une nouvelle notion de mesure que vous verrez l'an prochain.

En ce qui concerne l'intégrale de Riemann, il est évidemment possible de calculer l'aire d'un tel domaine si D s'écrit comme une réunion finie de pavés d'intérieurs disjoints. On peut généraliser cette définition aux domaines quarrables :

Définition 11.2.

1. Étant donné un pavé fermé (un produit de segments) $I \times J$, on appelle pavage (par des rectangles) de $I \times J$ une décomposition de $I \times J$ en un nombre fini de sous-pavés fermés $I'_k \times J'_\ell$ telle que l'intérieur des pavés est disjoint. Un pavage P de $I \times J$ correspond à la donnée de subdivisions $\sigma = \{x_0, x_1, \ldots, x_k, \ldots, x_n\}$ du segment I et $\tau\{y_0, x_1, \ldots, x_\ell, \ldots, y_m\}$ du segment J, et les sous-pavés $I'_k \times J'_\ell$ sont alors donnés par $[x_k, x_{k+1}] \times [y_\ell, y_{\ell+1}]$. On appelle pas du pavage et on note $\delta(P)$ la longueur maximale de la diagonale d'un sous-pavé.
2. Étant donné un pavé fermé (un produit de segments) $I \times J$, et une partie $D \subseteq I \times J$, et un pavage $P = (P_i)$ de $I \times J$, on appelle
   — *aire extérieure de D dans le pavage P* et on note $\mathrm{A}_{\mathrm{ext}}(D, P)$ la somme des aires des pavés P_i d'intersection non vide avec D
$$\mathrm{A}_{\mathrm{ext}}(D, P) = \sum_{i, P_i \cap D \neq \emptyset} \mathrm{A}(P_i).$$

— *aire intérieure de D dans le pavage P* et on note $\mathrm{A}_{\text{int}}(D,P)$ la somme des aires des pavés P_i contenus dans D

$$\mathrm{A}_{\text{int}}(D,P) = \sum_{i,P_i \subset D} \mathrm{A}(P_i).$$

3. Étant donné un pavé fermé (un produit de segments) $I\times J$, et une partie $D \subseteq I\times J$, on dit que D est *quarrable* si pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que si P est un pavage $I\times J$ dont le pas vérifie $\delta(P) < \alpha$, alors la différence des aires extérieures et intérieures vérifie :

$$\mathrm{A}_{\text{ext}}(D,P) - \mathrm{A}_{\text{int}}(D,P) < \varepsilon.$$

4. Dans ce cas on a que $\inf_P \mathrm{A}_{\text{ext}}(D,P) = \sup_P \mathrm{A}_{\text{int}}(D,P)$, lorsque P parcourt l'ensemble des pavages de $I\times J$ et on appelle aire de D la valeur commune.

Remarque 11.3.

1. On peut montrer qu'un domaine D est quarrable si est seulement si sa frontière est quarrable et d'aire nulle.
2. En particulier les domaines bornés délimités par une courbe fermée formée d'un nombre fini d'arcs paramétrés continus sont quarrables.
3. Il existe des ensembles mesurables (au sens de la mesure de Lebesgue) mais non quarrables.

On va s'intéresser plus particulièrement à des domaines quarrables particulièrement simples :

Définition 11.4.
On dit qu'une partie bornée D est un *domaine élémentaire*

1. de type "vertical" ou "en piles" s'il existe des fonctions continues c et d définies sur $[a,b]$ et telles que

$$D = \{(x,y) \in \mathbb{R}^2, a \le x \le b, c(x) \le y \le d(x)\}$$

2. de type "horizontal" ou "en tranches" s'il existe des fonctions continues a et b définies sur $[c,d]$ et telles que

$$D = \{(x,y) \in \mathbb{R}^2, a(y) \le x \le b(y), c \le y \le d\}$$

2.2 Intégration sur un domaine quarrable

Définition 11.5. On considère à présent un pavé fermé (un produit de segments) $I\times J$, un domaine $D \subseteq I\times J$ quarrable et une fonction réglée $f : I\times J \to \mathbb{R}$ alors on définit :

$$\int_D f(x,y)dxdy = \int_{I\times J} f(x,y)\mathbb{1}_D(x,y)dxdy\,,$$

où $\mathbb{1}_D$ est la fonction caractéristique de D.

une fonction réglée $f : I \times J \longrightarrow \mathbb{R}$ pose :

$$\int_D f(x,y)dxdy = \int_{I\times J} f(x,y)\mathbb{1}_D(x,y)dxdy\,,$$

où $\mathbb{1}_D$ est la fonction caractéristique de D.

Proposition 11.6. *Soit D un domaine borné quarrable dans $\mathbb{R}^2$ et $f : D \longrightarrow \mathbb{R}$ une fonction continue, alors l'intégrale $\int_D f(x,y)dxdy$ est bien définie.*

2.3 Propriétés de l'intégrale double d'une fonction continue sur un domaine quarrable

On retrouve les propriétés analogues à celle de l'intégrale d'une fonction d'une variable sur un segment.

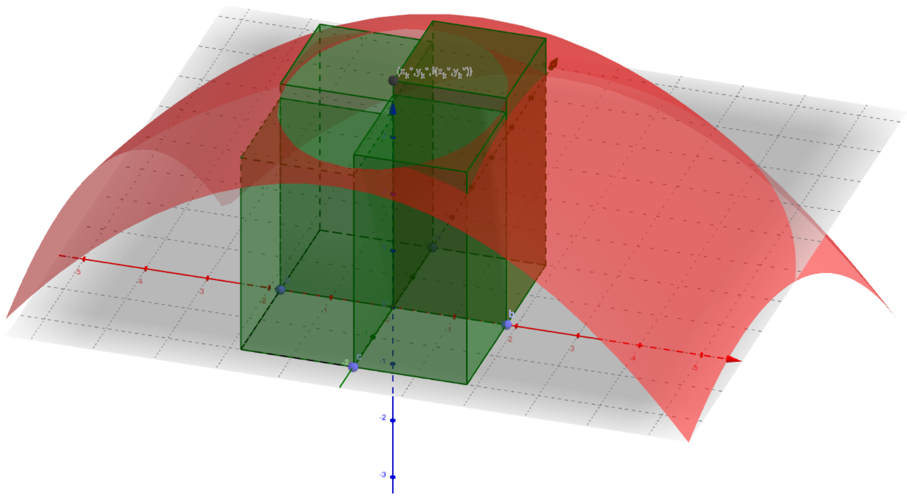
Théorème 11.7.

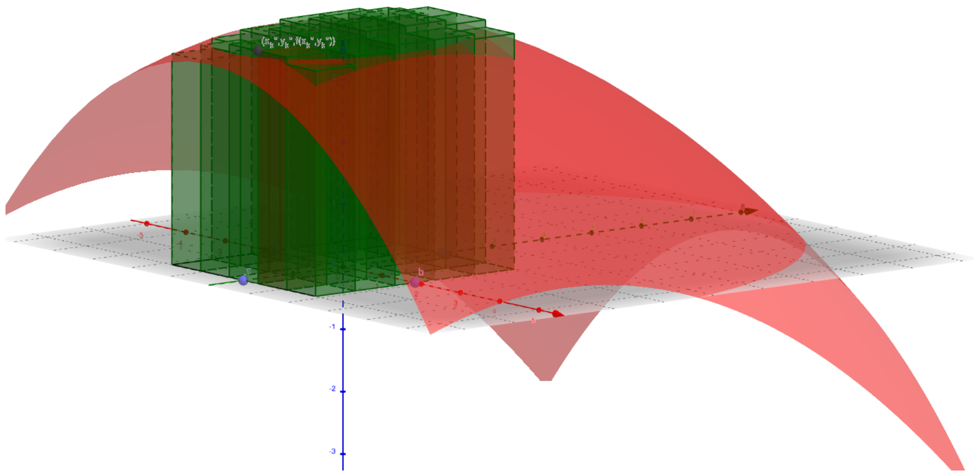
Soit D un domaine fermé borné quarrable, d'aire non nulle, dans $\mathbb{R}^2$, et $f, g : D \longrightarrow \mathbb{R}$ des fonctions continues, alors on a les propriétés suivantes :

1. **Linéarité** *Pour tout $\lambda \in \mathbb{R}$,*
$$\int_D (\lambda f(x,y) + g(x,y))dxdy = \lambda \int_D f(x,y)dxdy + \int_D g(x,y)dxdy.$$
2. **Positivité** *Si $f \geq 0$, alors*
$$\int_D f(x,y)dxdy \geq 0.$$
3. **Continuité** *Si on note m et M respectivement un minorant et un majorant de f sur D, en notant $A(D)$ l'aire du domaine D, alors*
$$mA(D) \leq \int_D f(x,y)dxdy \leq MA(D).$$
4. **Analogue de la relation de Chasles** *Si $D = \cup_{i=1}^n D_i$ où les D_i sont des sous-domaines de D fermés quarrables et dont les intérieurs sont disjoints, alors*
$$\int_D f(x,y)dxdy = \sum_{i=1}^n \int_{D_i} f(x,y)dxdy.$$
5. *Si $f \geq 0$ et $D' \subset D$ est un sous-domaine quarrable, alors*
$$\int_D f(x,y)dxdy \geq \int_{D'} f(x,y)dxdy.$$
6. *Si $f \geq 0$, et $\int_D f(x,y)dxdy = 0$, alors f est identiquement nulle sur D.*
7. **Inégalité triangulaire** *On a*
$$\left|\int_D f(x,y)dxdy\right| \leq \int_D \left|f(x,y)\right|dxdy$$
8. **Inégalité de Cauchy-Schwarz** *On a*
$$\left|\int_D f(x,y)g(x,y)dxdy\right| \leq \sqrt{\int_D f(x,y)^2dxdy}\sqrt{\int_D g(x,y)^2dxdy}$$

Interprétation géométrique

On a vu que le graphe de f définit une surface dans $\mathbb{R}^3$. De même que l'intégrale d'une fonction positive de $\mathbb{R}$ dans $\mathbb{R}$ sur un segment $[a, b]$ s'interprète naturellement comme l'aire de la surface délimitée par les droites $x = a$ et $x = b$, l'axe des abscisses et la courbe $y = f(x)$, l'intégrale double d'une fonction sur un domaine D correspond au volume de la partie de $\mathbb{R}^3$ suivante : $\Omega = \{(x, y) \in D, 0 \leq z \leq f(x, y)\}$. Comme dans le cas des fonctions d'une variable, le volume est compté négativement lorsque $f(x, y) < 0$.





3 Théorème de Fubini

Le théorème de Fubini est essentiel car il permet de calculer explicitement des intégrales doubles à l'aide d'intégrales simples.

Théorème 11.8 (théorème de Fubini sur un pavé)**.** *Foit $f : [a, b] \times [c, d] \to \mathbb{R}$ une fonction continue. On a alors l'égalité suivante :*

$$\int_a^b \Big(\int_c^d f(u, v) dv \Big) du = \int_c^d \Big(\int_a^b f(u, v) du \Big) dv = \int_{[a,b]\times[c,d]} f(u, v) du dv.$$

Démonstration. Voici un résumé de l'idée de la démonstration.

1. Si la fonction f est constante et égale à λ sur le pavé $[a,b]\times[c,d]$, l'égalité ci-dessus est évidente (toutes les quantités ont égales à $\lambda(b-a)(d-c)$.)
2. Si la fonction f est en escalier sur le pavé $[a,b]\times[c,d]$, l'identité découle de l'identité ci-dessus sur chaque sous-pavé.
3. Pour démontrer le cas général, on encadre f par des fonctions en escaliers dont la différence des intégrales sur D est aussi petite que l'on veut, ce qui permet de conclure.

□

Le théorème de Fubini se généralise aux domaines quarrables, nous allons nous limiter ici aux domaines élémentaires.

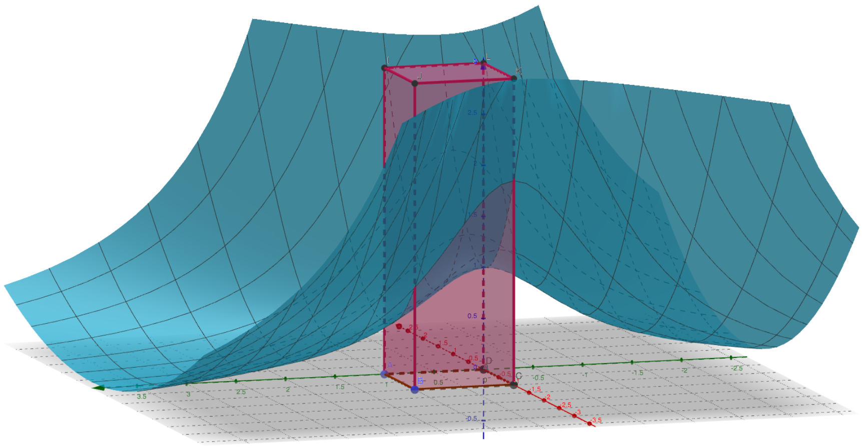
Théorème 11.9 (théorème de Fubini sur un domaine élémentaire)**.**

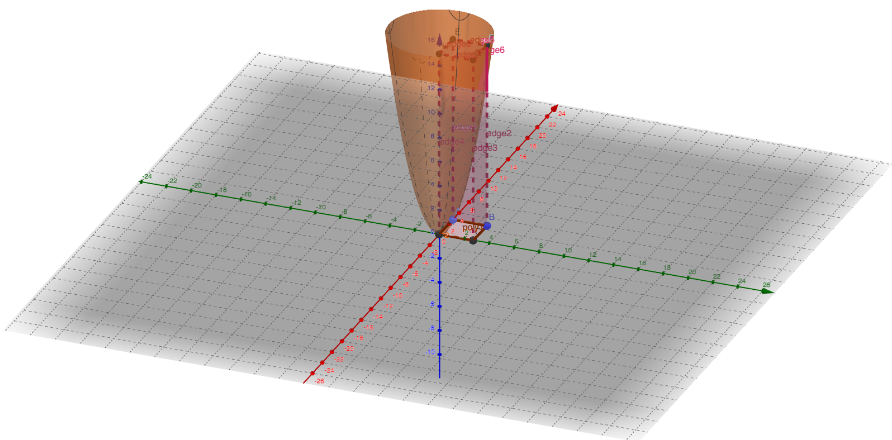
1. *Soit $f : D \longrightarrow \mathbb{R}$ une fonction continue sur un domaine élémentaire D de type "vertical" . On a alors l'égalité suivante :*
$$\int_a^b \Big(\int_{c(u)}^{d(u)} f(u,v)dv\Big)du = \int_D f(u,v)dudv.$$
2. *Soit $f : D \longrightarrow \mathbb{R}$ une fonction continue sur un domaine élémentaire D de type "horizontal" . On a alors l'égalité suivante :*
$$\int_c^d \Big(\int_{a(v)}^{b(v)} f(u,v)du\Big)dv = \int_D f(u,v)dudv.$$
3. *Si D est à la fois un domaine élémentaire horizontal et vertical, on obtient l'égalité :*
$$\int_D f(u,v)dudv = \int_a^b \Big(\int_{c(u)}^{d(u)} f(u,v)dv\Big)du = \int_c^d \Big(\int_{a(v)}^{b(v)} f(u,v)du\Big)dv.$$

Pout intégrer sur un domaine qui n'est pas élémentaire, on se sert souvent d'une décomposition du domaine d'intégration en des domaines élementaires sont les intérieurs sont disjoints.

Exemples 11.10.

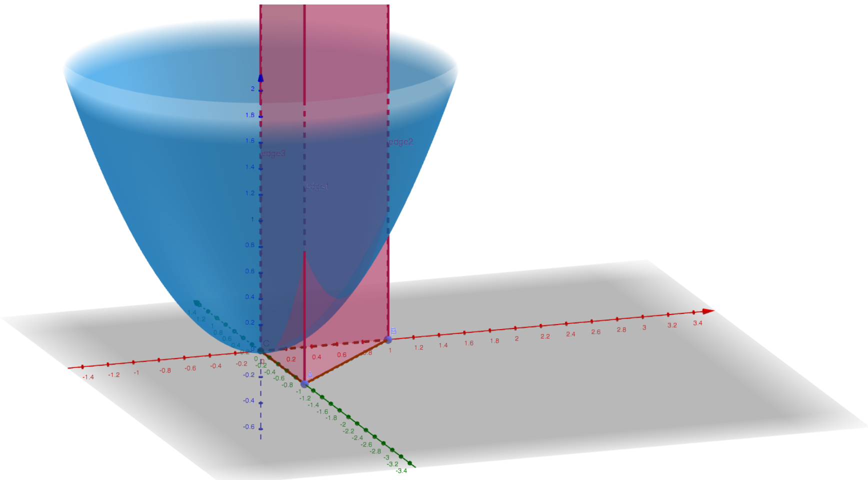
1. Calcul de $I_1 = \int_D \frac{1+x^2}{1+y^2}dxdy$ où D est le carré $[0,1]\times[0,1]$.
On a
$$I_1 = \int_{[0,1]\times[0,1]} \frac{1+x^2}{1+y^2}dxdy = \int_0^1 (1+x^2)dx \int_0^1 \frac{1}{1+y^2}dy = \frac{4}{3}\times\frac{\pi}{4} = 3\pi$$
2. Calcul de $I_2 = \int_D (x^2+y^2)dxdy$ où D est le carré $[0,1]\times[0,1]$.
On a
$$I_2 = \int_0^1 \left(\int_0^1 (x^2+y^2)dx\right)dy = \int_0^1 (\frac{1}{3}+y^2)dy = \left[\frac{y}{3}+\frac{y^3}{3}\right]_0^1 = \frac{2}{3}.$$





3. Calcul de $I_3 = \int_D (x^2+y^2)dxdy$ où D est l'intérieur du triangle de sommets $A = (1,0), B = (0,0), C = (0,1)$
L'équation de la droite passant par A et C est $y = -x+1$
On a

$$\begin{aligned} I_3 &= \int_0^1 \left(\int_0^{1-x} (x^2+y^2)dy \right) dx = \int_0^1 \left[x^2y + \frac{y^3}{3} \right]_{y=0}^{y=1-x} dx \\ &= \int_0^1 \left[x^2(1-x) + \frac{(1-x)^3}{3} \right] dx = \left[\frac{x^3}{3} - \frac{x^4}{4} - \frac{(1-x)^4}{12} \right]_{x=0}^{x=1} = \frac{1}{6}. \end{aligned}$$



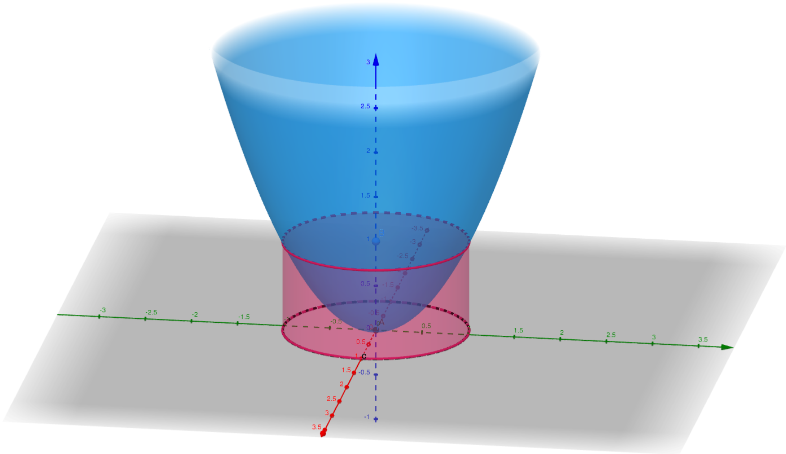
4. Calcul de $I_4 = \int_D (x^2+y^2)dxdy$ où D est le disque de centre 0 et de rayon 1
Pour $y \in [-1,1]$ fixé, x varie dans le segment $[-\sqrt{1-y^2}, \sqrt{1-y^2}]$ On a

$$\begin{aligned} I_4 &= \int_{-1}^1 \left(\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} (x^2+y^2)dx \right) dy = \int_{-1}^1 \left[\frac{x^3}{3} + xy^2 \right]_{x=-\sqrt{1-y^2}}^{x=\sqrt{1-y^2}} dy \\ &= \int_{-1}^1 2\left[\frac{\sqrt{1-y^2}^3}{3} + \sqrt{1-y^2}y^2 \right] dy = \frac{2}{3}\int_{-1}^1 \left[\sqrt{1-y^2} + 2y^2\sqrt{1-y^2} \right] dy. \end{aligned}$$

On calcule le résultat en utilisant le changement de variable $y = \sin t$ où t varie entre $-\frac{\pi}{2}$ et $\frac{\pi}{2}$

$$\begin{aligned} \int_{-1}^1 \left[\sqrt{1-y^2} + 2y^2\sqrt{1-y^2} \right] dy &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2 t \left[1 + 2(1-\cos^2 t) \right] dt \\ &= 3\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2 t dt - 2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^4 t dt \\ &= 3\frac{\pi}{2} - 2\frac{3\pi}{8} = \frac{3\pi}{4} \end{aligned}$$

et donc $I_4 = \frac{\pi}{2}$.



4 Formule du changement de variables.

Pour énoncer la formule du changement de variable, introduisons d'abord la notion de matrice jacobienne et de jacobien pour les applications $\mathbb{R}^2 \longrightarrow \mathbb{R}^2$.

Définition 11.11. Soit $\varphi : I \times J \longrightarrow \mathbb{R}^2$ une fonction de coordonnées φ_1 et φ_2. On suppose que chacune des coordonnées est dérivable par rapport à chacune de ses variables. La *matrice jacobienne* de φ en un point $(x_0, y_0) \in I \times J$ est la matrice $M\varphi$ définie par :

$$\mathrm{Jac}(\varphi) = \begin{pmatrix} \frac{\partial \varphi_1}{\partial x} & \frac{\partial \varphi_1}{\partial y} \\ \frac{\partial \varphi_2}{\partial x} & \frac{\partial \varphi_2}{\partial y} \end{pmatrix}$$

Le *jacobien* de φ est le déterminant de la matrice jacobienne :

$$J_\varphi = \begin{vmatrix} \frac{\partial \varphi_1}{\partial x} & \frac{\partial \varphi_1}{\partial y} \\ \frac{\partial \varphi_2}{\partial x} & \frac{\partial \varphi_2}{\partial y} \end{vmatrix}$$

La fonction φ est *régulière* en (x_0, y_0) si $J\varphi(x_0, y_0) \neq 0$.
On dit que φ est un C^1-difféomorphisme de Δ vers D lorsque :
- φ est de classe C^1 sur Δ ;
- φ est une bijection entre Δ et D ;
- la bijection réciproque $\varphi^{-1} : D \longrightarrow \Delta$ est également de classe C^1.

Théorème 11.12. *Soient D et Δ deux domaines d'intégration tels qu'il existe $\varphi : \Delta \longrightarrow D$ un C^1-difféomorphisme de Δ vers D. Soit $f : D \longrightarrow \mathbb{R}$ une fonction continue. Alors :*

$$\int_D f(x,y)dxdy = \int_\Delta f \circ \varphi(u,v) \big| J\varphi(u,v) \big| dudv.$$

Exemple : changement de variables en coordonnées polaires.

Soit D le disque unité. On veut poses $x = r\cos\theta$ et $y = r\sin\theta$. On considère donc le domaine $\Delta = [0,1] \times [0, 2\pi[$ et la fonction $\varphi : \Delta \longrightarrow D$ définie par

$$\varphi(r,\theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}$$

Son jacobien est :

$$J\varphi(r,\theta) = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r.$$

La fonction φ est donc régulière en tout point de coordonnée (r,θ) avec $r \neq 0$. Mais les points où $r = 0$ forment une bande du domaine d'aire nulle. Ils ne contribuent pas à l'intégrale et on peut donc remplacer Δ par $\Delta' =]0,1[\times[0,2\pi[$. On obtient la formule de changement de variables suivante :

$$\int_D f(x,y)dxdy = \int_0^1 \int_0^{2\pi} f(r\cos\theta, r\sin\theta) r dr d\theta.$$

Exemples 11.13. 1. On veut calculer $I_4 = \int_D (x^2+y^2)dxdy$ où D est l'intérieur du disque de centre 0 et de rayon 1
Le changement de coordonnées polaires donne

$$\begin{aligned} \int_D f(x,y)dxdy &= \int_0^1 \int_0^{2\pi} \left[(r\cos\theta)^2 + (r\sin\theta)^2\right] r dr d\theta \\ &= \int_0^1 \int_0^{2\pi} r^3 dr d\theta = 2\pi \left[\frac{r^4}{4}\right]_0^1 \\ &= \frac{\pi}{2} \end{aligned}$$

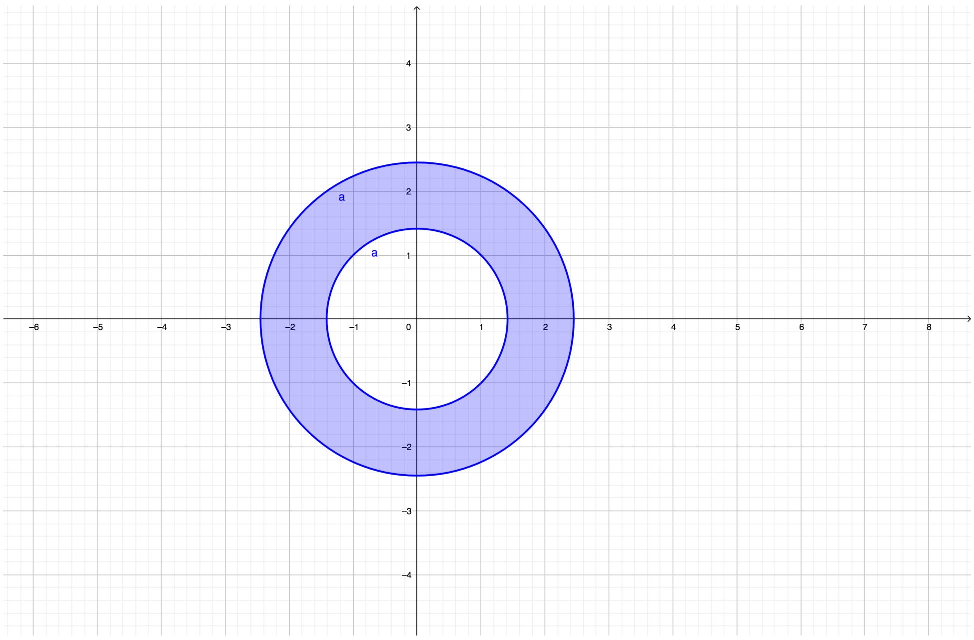
2. Surface d'une couronne.

On considère la couronne D par $D = \{(x,y) \in \mathbb{R}^2, R_1^2 \leq x^2+y^2 \leq R_2^2\}$. On a alors par changement de coordonnées polaires $\Delta = [R_1, R_2 \times [0, 2\pi]$ et donc

$$\int_D dxdy = \int_{R_1}^{R_2} \int_0^{2\pi} r dr d\theta = \pi(R_2^2 - R_1^2).$$

5 Intégrales triples, calcul de volumes

On peut itérer la construction faite en dimension deux aux dimensions supérieures. Pour $n = 3$, on remplace les pavés par des parallépipèdes rectangles. On peut alors comme précédemment construire l'intégrale d'une fonction constante sur un parallépipède rectangle, puis l'intégrale d'une fonction en escalier, et enfin celle d'une fonction réglée sur un domaine d'intégration qui est également parallépipède rectangle.



De même on peut étendre la notion de domaine de $\mathbb{R}^3$ quarrable, et la définition de l'intégrale.

On a vu précédemment comment calculer volume délimité par un "solide" S de base un domaine de $\mathbb{R}^2$ quarrable dans le plan z=0 et délimité par une surface de la forme $z = f(x, y)$ avec f une fonction positive et continue sur D.
On va se limiter ici au calcul du volume délimité par un "solide" S élémentaire dans le sens suivant : pour z variant dans le segment $[0, h]$ la surface obtenue par l'intersection du plan horizontal $z = a$ avec S définit une surface $A(a)$ élémentaire dans $\mathbb{R}^2$.

Théorème de Fubini et calculs de volume

On définit $S \subset \mathbb{R}^3$ un domaine élémentaire borné. On a alors

$$\mathrm{Vol}(S) = \int_{I\times J\times K} \mathbb{1}_S dxdydz.$$

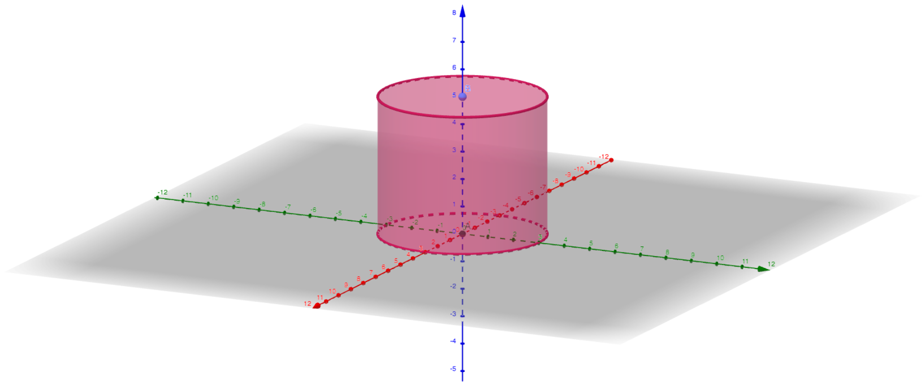
et le théorème de Fubini permet de montrer, dans ce cas :

$$\mathrm{Vol}(S) = \int_K \mathrm{A}_S(z)dz.$$

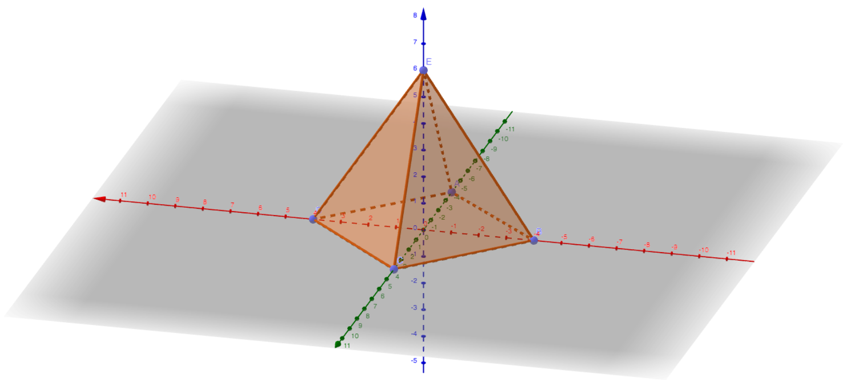
où $A_S(z)$ désigne la surface délimitée par l'intersection de S avec le plan horizontal de hauteur z.

Exemples 11.14.
1. Calcul d'un volume cylindrique de hauteur h et de base une surface d'aire A : on a $V = \int_0^h Adz = hA$.
2. Calcul du volume d'une pyramide de hauteur h et de base carrée de coté a
   D'après ce qui précède, il suffit de calculer $\int_0^h A(z)dz$, où $A(z)$ est la surface de l'intérieur du carré qui est la ligne de niveau z. Le coté du carré à la hauteur z est donné par $\ell(z) = a(1 - \frac{z}{h})$ par le théorème de Thalès, et on a donc

$$V = \int_0^h a^2(1 - \frac{z}{h})^2 dz = a^2 h \int_0^1 u^2 du = \frac{a^2 h}{3}.$$



Plus généralement, si on cherche à calculer le volume d'un objet conique de base une surface d'aire A et de hauteur h, on obtient de même $V = \frac{Ah}{3}$.



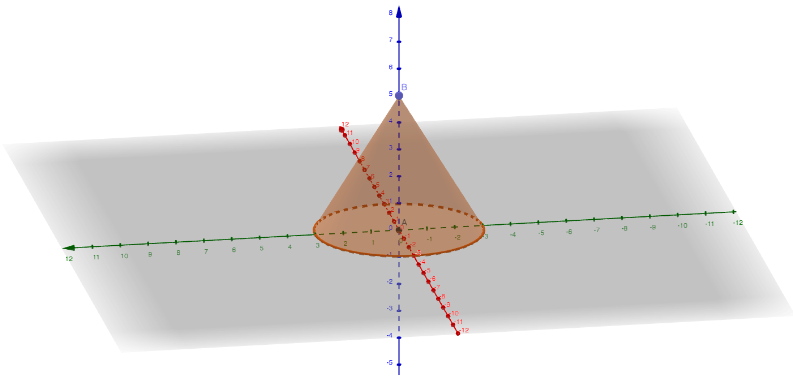
Coordonnées cylindriques

On repère un point M de coordonnées (x, y, z) dans l'espace $\mathbb{R}^3$ en considérant d'une part sa coordonnée de hauteur z et d'autre part en repérant la projection du vecteur $\overrightarrow{OM}$ dans le plan $z = 0$ par les coordonnées polaires $x = r\cos(\theta)$ et $y = r\sin(\theta)$. On considère donc la fonction$\varphi : [0, +\infty[\times[0, 2\pi[\times\mathbb{R} \to \mathbb{R}^3$ définie par

$$\varphi(r, \theta, z) = (r\cos(\theta), r\sin(\theta), z).$$

On a que le jacobien de φ est donné par

$$J_\varphi = \begin{vmatrix} \cos(\theta) & -r\sin(\theta) & 0 \\ \sin(\theta) & r\cos(\theta) & 0 \\ 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{vmatrix} = r.$$



Ainsi, on a que si $\varphi(\Delta) = D$,

$$\int_D f(x,y,z)dxdydz = \int_\Delta f(r\cos(\theta), r\sin(\theta), z)rdrd\theta dz$$

Exemples 11.15. 1. Calcul du volume d'un cylindre de base un disque de rayon R et de hauteur h.

$$\int_D dxdydz = \int_0^h \int_0^{2\pi} \int_0^2 rdrd\theta dz = 2\pi h \left[\frac{r^2}{2}\right]_0^R = \pi h R^2.$$

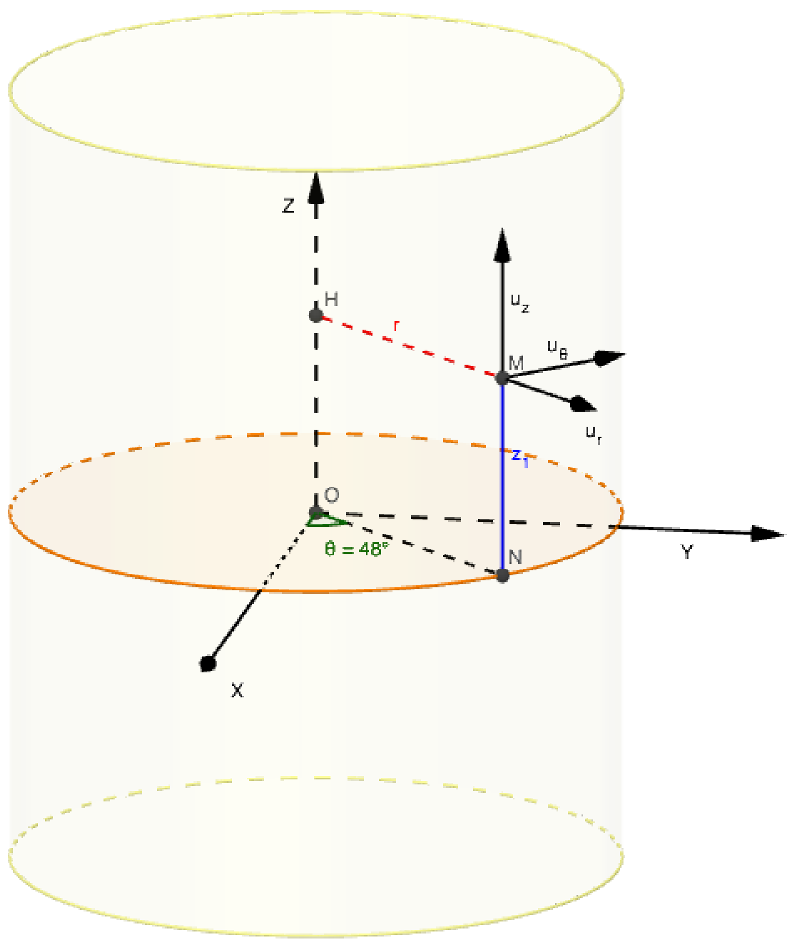
2. Calcul de $\int_D (x^2+y^2)dxdydz$ où D est l'intérieur d'un cylindre de hauteur h et de base un disque de rayon 2

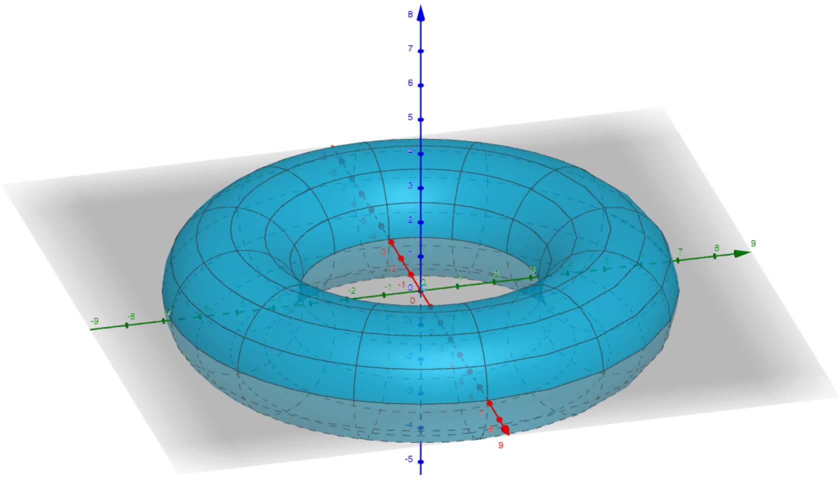
$$\int_D f(x,y,z)dxdydz = \int_0^h \int_0^{2\pi} \int_0^2 r^2 rdrd\theta dz = 2\pi h \left[\frac{r^4}{4}\right]_0^2 = 8\pi h.$$

3. Calcul du volume d'un tore Tde grand rayon R au centre et de petit rayon h

$$\begin{aligned}
\int_D \mathbb{1}_T dxdydz &= \int_{-h}^h \int_0^{2\pi} \int_{R-\sqrt{h^2-z^2}}^{R-\sqrt{h^2-z^2}} rdrd\theta dz \\
&= 2\pi \int_{-h}^h \frac{(R+\sqrt{h^2-z^2})^2 - (R+\sqrt{h^2-z^2})}{2} dz \\
&= 4\pi R \int_{-h}^h \sqrt{h^2-z^2}dz \\
&= 4\pi R \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} h^2\cos^2(t)dt \\
&= 4\pi R h^2 \frac{\pi}{2} = 2\pi^2 h^2 R
\end{aligned}$$

Coordonnées cylindriques





Coordonnées sphériques

On repère un point M de coordonnées (x, y, z) dans l'espace $\mathbb{R}^3$ en considérant

— sa distance r au point origine.

— l'angle θ formé par la demi- droite (Oz) et le vecteur $\overrightarrow{OM}$

— l'angle ϕ formé par la demi-droite (Ox) et la projection du vecteur $\overrightarrow{OM}$ dans le plan $z = 0$ par $x = r\sin(\theta)\cos(\phi)$ et $y = r\sin(\theta)\sin(\phi)$.

On considère donc la fonctionφ : $[0, +\infty[\times[0, 2\pi[\times[0, \pi[$ définie par

$$\varphi(r, \phi, \theta) = (r\cos(\phi)\sin(\theta), r\sin(\phi)\sin(\theta), r\cos(\theta)).$$

On a que le jacobien de φ est donné par

$$\begin{aligned}
J_\varphi &= \begin{vmatrix} \cos(\phi)\sin(\theta) & -r\sin(\phi)\sin(\theta) & r\cos(\phi)\cos(\theta) \\ \sin(\phi)\sin(\theta) & r\cos(\phi)\cos(\theta) & r\sin(\phi)\cos(\theta) \\ \cos(\theta) & 0 & -r\sin(\theta) \end{vmatrix} \\
&= -\cos(\theta)\begin{vmatrix} -r\sin(\phi)\sin(\theta) & r\cos(\phi)\sin(\theta) \\ r\cos(\phi)\cos(\theta) & r\sin(\phi)\cos(\theta) \end{vmatrix} + r\sin(\theta)\begin{vmatrix} \cos(\phi)\sin(\theta) & -r\sin(\phi)\sin(\theta) \\ \sin(\phi)\sin(\theta) & r\cos(\phi)\sin(\theta) \end{vmatrix} \\
&= r^2\sin(\theta)\cos^2(\theta) + r^2\sin(\theta)\sin^2(\theta) = r^2\sin(\theta).
\end{aligned}$$

Ainsi, on a que si $\varphi(\Delta) = D$,

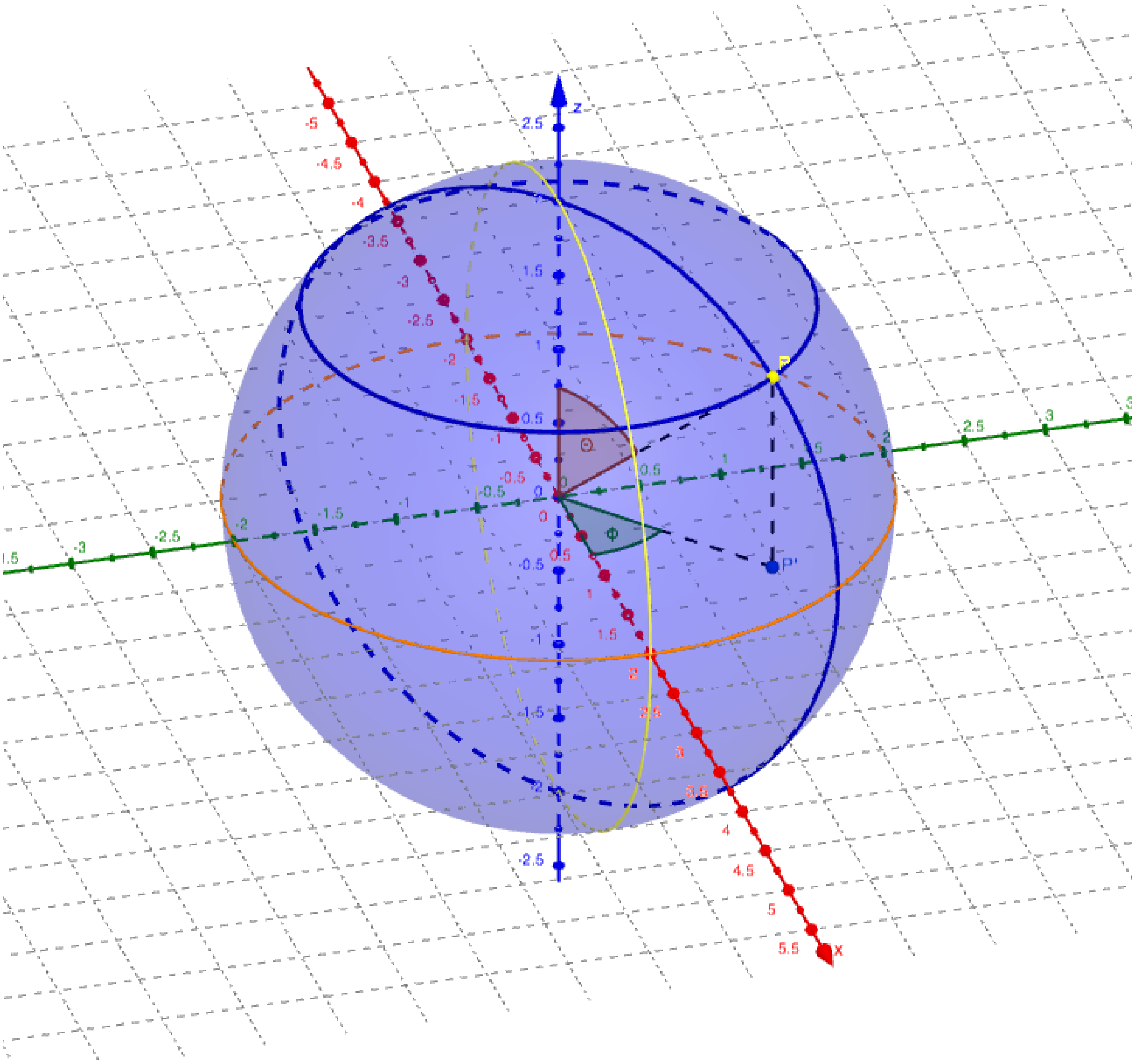
$$\int_D f(x, y, z)dxdydz = \int_\Delta f(r\cos(\phi)\sin(\theta), r\sin(\phi)\sin(\theta), r\cos(\theta))r^2\sin(\theta)drd\phi d\theta$$

Exemple 11.16. Calcul du volume délimité par une sphère de rayon R

On appelle D la partie de l'espace délimitée par la sphère de rayon R

On a que $\varphi^{-1}(D) = \Delta = [0, R] \times [0, 2\pi[\times[0, \pi]$ et donc le volume de la sphère est

$$\int_D dxdydz = \int_\Delta r^2\sin(\theta)drd\phi d\theta = \frac{R^3}{3}\cdot(2\pi)\int_0^\pi \sin(\theta)d\theta = \frac{4\pi R^3}{3}.$$


z
x
P
P'
Θ
Φ