

Notes de cours

MATH-105(a)
Analyse avancée II – analyse vectorielle
(Section MA)

Boris Buffoni, Fabio Nobile

2024-2025

Dernière mise à jour : 13 février 2025



Table des matières

1	L'espace \mathbb{R}^n et sa topologie	7
1.1	Espaces vectoriels normés	7
1.2	L'espace \mathbb{R}^n	9
1.3	Suites dans \mathbb{R}^n	10
1.4	Topologie de \mathbb{R}^n	11
2	Fonctions de plusieurs variables réelles	19
2.1	Notions de limite	19
2.2	Fonctions continues	24
2.3	Prolongement de fonctions par continuité	26
2.4	Fonctions continues sur un compact	27
3	Dérivabilité	31
3.1	Dérivées partielles et directionnelles; différentielle	31
3.2	Dérivation de fonctions composées	39
3.3	Théorème des accroissements finis	42
4	Dérivées d'ordres supérieurs	45
4.1	Dérivées secondes	45
4.2	Dérivées partielles d'ordres supérieurs à 2	48
4.3	Développement limité et Formule de Taylor	49
5	Intégrales qui dépendent de paramètres	53
5.1	Intégrales sur un intervalle fermé et borné	54
5.2	Intégrales avec des bornes variables	55
5.3	Intégrales généralisées dépendant de paramètres	56
6	Difféomorphismes locaux et fonctions implicites	59
6.1	Fonctions bijectives et difféomorphismes locaux	59
6.2	Théorème d'inversion locale	61
6.3	Hypersurfaces et fonctions implicites	66
6.4	Théorème des fonctions implicites – cas scalaire	67
6.5	Théorème des fonctions implicites – cas vectoriel	72

7	Extrema de fonctions réelles	77
7.1	Extrema libres	77
7.2	Extrema liés	83
7.3	Méthode des multiplicateurs de Lagrange	87
7.4	Extrema sous contraintes multiples	87
7.5	Conditions suffisantes	89
8	Intégrale multiple au sens de Riemann	91
8.1	Pavés de \mathbb{R}^n	91
8.2	Fonctions intégrables sur un pavé	93
8.3	Intégrales itérées sur un pavé et théorème de Fubini	98
8.4	Intégrale de Riemann sur un ensemble quelconque	101
8.5	Ensembles mesurables au sens de Jordan	103
8.6	Caractérisation des fonctions intégrables	107
8.7	Généralisation de la formule des intégrales itérées	110
8.8	Changement de variables	113
8.9	Quelques changements de variables usuels	115
8.10	Intégrale de Riemann généralisée	119
9	Équations différentielles ordinaires	123
9.1	Problème de Cauchy	124
9.2	Quelques méthodes de résolution d'EDO scalaires	126
9.3	EDO scalaires linéaires du premier ordre	131
9.4	Existence et unicité de solutions	135
9.5	EDO scalaires linéaires du second ordre	147

Notations

$$\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$$

$$\mathbb{R}_+^* = \{x \in \mathbb{R} : x > 0\}$$

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$$

$$\mathbb{N} = \{0, 1, 2, \dots\}$$

$$\mathbb{N}^* = \{1, 2, \dots\}$$

Chapitre 1

L'espace \mathbb{R}^n et sa topologie

1.1 Espaces vectoriels normés

On rappelle ici les notions générales d'espace vectoriel, norme, distance et produit scalaire.

Définition 1.1 (Espace vectoriel réel). *Un ensemble V est un espace vectoriel réel si les opérations de somme et multiplication par un scalaire (réel) sont définies sur V avec les propriétés suivantes :*

1. *somme : $V \times V \rightarrow V$, $(x, y) \in V \times V \mapsto z = x + y \in V$,*
 - $\forall x, y \in V, \quad x + y = y + x$,
 - $\forall x, y, z \in V, \quad (x + y) + z = x + (y + z)$,
 - \exists *élément nul* $0 : \quad x + 0 = x$,
 - $\forall x \in V, \exists$ *élément opposé* $-x : \quad x + (-x) = 0$,
2. *multiplication par un scalaire : $\mathbb{R} \times V \rightarrow V$, $(\lambda, x) \in \mathbb{R} \times V \mapsto z = \lambda x \in V$,*
 - $\forall \lambda, \mu \in \mathbb{R}, \forall x \in V, \quad \lambda(\mu x) = (\lambda\mu)x$,
 - $\forall x \in V, \quad 1 \cdot x = x$,
 - $\forall \lambda, \mu \in \mathbb{R}, \forall x \in V, \quad (\lambda + \mu)x = \lambda x + \mu x$,
 - $\forall \lambda \in \mathbb{R}, \forall x, y \in V, \quad \lambda(x + y) = \lambda x + \lambda y$.

Définition 1.2 (Norme). *Soit V un espace vectoriel réel. Une norme sur V est une application $N : V \rightarrow \mathbb{R}_+$ qui satisfait les propriétés suivantes :*

1. $\forall x \in V, \quad N(x) \geq 0$, *et* $N(x) = 0 \Leftrightarrow x = 0$,
2. $\forall x \in V, \lambda \in \mathbb{R}, \quad N(\lambda x) = |\lambda|N(x)$,
3. $\forall x, y \in V, \quad N(x + y) \leq N(x) + N(y)$ *(inégalité triangulaire).*

On note souvent une norme par $\|\cdot\|$ ($N(x) = \|x\|$).

Un espace vectoriel muni d'une norme est appelé **espace vectoriel normé** et souvent noté $(V, \|\cdot\|)$. On dit que deux normes N_1 et N_2 sur un espace vectoriel V sont équivalentes s'ils existent deux constants $\underline{c}, \bar{c} > 0$ telles que $\underline{c}N_1(x) \leq N_2(x) \leq \bar{c}N_1(x)$ pour tout $x \in V$.

Définition 1.3 (distance). Soit X un ensemble. Une distance ou métrique sur X est une application $d : X \times X \rightarrow \mathbb{R}_+$ qui satisfait les propriétés suivantes :

1. $\forall x, y \in X, \quad d(x, y) \geq 0, \quad \text{et} \quad d(x, y) = 0 \Leftrightarrow x = y,$
2. $\forall x, y \in X, \quad d(x, y) = d(y, x),$
3. $\forall x, y, z \in X, \quad d(x, y) \leq d(x, z) + d(z, y) \quad (\text{inégalité triangulaire}).$

Un ensemble X muni d'une distance (X, d) est appelé **espace métrique**. Si $(V, \|\cdot\|)$ est un espace vectoriel normé, alors l'application

$$d : V \times V \rightarrow \mathbb{R}, \quad d(x, y) = \|x - y\|, \quad \forall x, y \in V,$$

est une distance (vérifiez-le) appelée la distance induite par la norme $\|\cdot\|$. Donc $(V, d(x, y) = \|x - y\|)$ est un espace métrique.

Exercice 1.4. Soit V un espace vectoriel, $d(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}_+$ une distance sur V et $h : [0, \infty) \rightarrow \mathbb{R}_+$ une fonction différentiable telle que $h(0) = 0$, $h'(x) > 0$ pour $x > 0$ et $h'(x)$ décroissante sur $[0, \infty)$. Montrer que $\tilde{d} = h \circ d$ est aussi une distance sur V . Vérifier que les hypothèses sont satisfaites par la fonction $h(x) = x/(1+x)$, mais que la distance $\tilde{d} = h \circ d$ n'est pas induite par une norme lorsque $V \neq \{0\}$, même si ceci est vrai pour d .

Définition 1.5 (Produit scalaire). Un produit scalaire sur un espace vectoriel réel V est une application $b : V \times V \rightarrow \mathbb{R}$ qui satisfait les propriétés suivantes :

1. (symétrie) $\forall x, y \in V, \quad b(x, y) = b(y, x),$
2. (bi-linéarité) $\forall x, y \in V, \quad \forall \alpha, \beta \in \mathbb{R}, \quad b(\alpha x + \beta y, z) = \alpha b(x, z) + \beta b(y, z),$
3. (positivité) $\forall x \in V, \quad b(x, x) \geq 0, \quad \text{et} \quad b(x, x) = 0 \Leftrightarrow x = 0.$

Un produit scalaire satisfait l'importante inégalité suivante :

Lemme 1.6 (Inégalité de Cauchy–Schwarz). Soit V un espace vectoriel réel et $b : V \times V \rightarrow \mathbb{R}$ un produit scalaire sur V . Alors

$$\forall x, y \in V, \quad |b(x, y)| \leq b(x, x)^{\frac{1}{2}} b(y, y)^{\frac{1}{2}}.$$

Démonstration. $\forall \alpha \in \mathbb{R}, \forall x, y \in V$

$$0 \leq b(\alpha x + y, \alpha x + y) = \alpha^2 b(x, x) + 2\alpha b(x, y) + b(y, y) = p_2(\alpha)$$

où $p_2(\alpha)$ est un polynôme de degré 2 en α . Par la positivité de b , on obtient la condition suivante pour le discriminant : $\Delta \leq 0$ ce qui implique $b^2(x, y) - b(x, x)b(y, y) \leq 0$, d'où la thèse. \square

Grâce à cette propriété, un espace vectoriel réel muni d'un produit scalaire est toujours un espace normé.

Théorème 1.7. Soit $b : V \times V \rightarrow \mathbb{R}$ un produit scalaire sur un espace vectoriel réel V . Alors $\|x\|_b = b(x, x)^{\frac{1}{2}} : V \rightarrow \mathbb{R}_+$ est une norme et $(V, \|\cdot\|_b)$ un espace vectoriel normé.

Démonstration. Il faut vérifier que $\|\cdot\|_b$ satisfait toutes les propriétés d'une norme selon la Définition 1.2. Les propriétés 1. et 2. suivent directement de la positivité du produit scalaire (propriété 3. de 1.5) et de sa bi-linéarité (propriété 2. de 1.5).

Quant à l'inégalité triangulaire (propriété 3.) elle est une conséquence de l'inégalité de Cauchy-Schwarz :

$$\begin{aligned}\|x + y\|_b^2 &= b(x + y, x + y) = b(x, x) + 2b(x, y) + b(y, y) \\ &\leq b(x, x) + b(y, y) + 2b(x, x)^{\frac{1}{2}}b(y, y)^{\frac{1}{2}} \\ &= (\|x\|_b + \|y\|_b)^2\end{aligned}$$

□

1.2 L'espace \mathbb{R}^n

On note $\mathbb{R}^n = \overbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}^{n \text{ fois}}$ l'ensemble des n -uples $\mathbf{x} = (x_1, \dots, x_n)$, avec $x_i \in \mathbb{R}$, pour $i = 1, \dots, n$, muni des opérations de

- somme : pour tout $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{y} = (y_1, \dots, y_n)$, $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \dots, x_n + y_n)$,
- multiplication par un scalaire : pour tout $\mathbf{x} = (x_1, \dots, x_n)$, $\lambda \in \mathbb{R}$, $\lambda \mathbf{x} = (\lambda x_1, \dots, \lambda x_n)$.

Ainsi, \mathbb{R}^n a une structure d'espace vectoriel réel. Sur \mathbb{R}^n , on peut introduire plusieurs normes. Voici les plus communes :

- **Norme euclidienne :**

$$\|\mathbf{x}\|_E = \sqrt{\sum_{i=1}^n x_i^2},$$

- **Norme p :**

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}, \quad p \geq 1,$$

- **Norme ∞ :**

$$\|\mathbf{x}\|_\infty = \max_{1 \leq j \leq n} |x_j|.$$

La norme euclidienne correspond à la norme p avec $p = 2$, i.e. $\|\mathbf{x}\|_E = \|\mathbf{x}\|_2$. On vérifie (exercice) que toutes les applications $\|\cdot\|_p : \mathbb{R}^n \rightarrow \mathbb{R}_+$, pour tout $p \geq 1$ et $p = \infty$, sont des normes. Par contre, $\|\mathbf{x}\|_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$ avec $0 < p < 1$ n'est pas une norme lorsque $n \geq 2$ (elle ne vérifie pas l'inégalité triangulaire).

Toutes les normes $\|\cdot\|_p$, $p \geq 1$, sont équivalentes, c'est-à-dire, $\forall p, q \geq 1$, il existe $0 < c_1(p, q) < c_2(p, q)$:

$$c_1(p, q)\|\mathbf{x}\|_p \leq \|\mathbf{x}\|_q \leq c_2(p, q)\|\mathbf{x}\|_p, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Plus généralement, sur \mathbb{R}^n , **toutes les normes sont équivalentes**, c'est-à-dire, si $\|\cdot\|$ et $\|\cdot\|'$ sont deux normes sur \mathbb{R}^n , alors $\exists 0 < c_1 < c_2$ tels que.

$$c_1\|\mathbf{x}\| \leq \|\mathbf{x}\|' \leq c_2\|\mathbf{x}\| \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

La preuve de ce résultat sera proposée plus loin dans le cours.

Seule la norme euclidienne $\|\mathbf{x}\|_2 = (\sum_{i=1}^n x_i^2)^{\frac{1}{2}}$ parmi toutes les normes $\|\cdot\|_p$, $p \geq 1$ est une norme induite par un produit scalaire, nommément le **produit scalaire euclidien** :

$$(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n x_i y_i = \mathbf{y}^\top \cdot \mathbf{x} = \mathbf{y}^\top \mathbf{x}.$$

Dans ces deux dernières expressions, \mathbf{x} et \mathbf{y} sont des vecteurs colonnes, \mathbf{y}^\top est la transposée de \mathbf{y} et le produit est la multiplication matricielle. En effet, $\|\mathbf{x}\|_2 = (\mathbf{x}, \mathbf{x})^{\frac{1}{2}}$ et l'inégalité de Cauchy Schwarz sur \mathbb{R}^n devient :

$$|(\mathbf{x}, \mathbf{y})| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2.$$

1.3 Suites dans \mathbb{R}^n

Définition 1.8 (Suite convergente). Soit $\{\mathbf{x}^{(k)}\}_{k=0}^\infty \subset \mathbb{R}^n$ une suite d'éléments de \mathbb{R}^n , $\mathbf{x}^{(k)} = (x_1^{(k)}, \dots, x_n^{(k)}) \in \mathbb{R}^n$. On dit que $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}}$ converge s'il existe $\mathbf{x} \in \mathbb{R}^n$ tel que $\lim_{k \rightarrow \infty} \|\mathbf{x} - \mathbf{x}^{(k)}\| = 0$, c.-à-d. :

$$\forall \epsilon > 0, \quad \exists N > 0 : \quad \forall k \geq N, \quad \|\mathbf{x} - \mathbf{x}^{(k)}\| \leq \epsilon.$$

Dans ce cas, on note $\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}$.

Puisque toutes les normes sont équivalentes sur \mathbb{R}^n , la convergence de la suite $\mathbf{x}^{(k)}$ **ne dépend pas** de la norme choisie. De même que la valeur limite \mathbf{x} (si elle existe) ne dépend pas de la norme. En particulier, si on prend la norme $\|\cdot\|_\infty$ dans la définition de $\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}$, on en tire la propriété suivante.

Lemme 1.9. Une suite $\{\mathbf{x}^{(k)}\}_{k=0}^\infty \subset \mathbb{R}^n$ converge vers $\mathbf{x} \in \mathbb{R}^n$ si et seulement si $\{x_j^{(k)}\}_{k=0}^\infty \subset \mathbb{R}$ converge vers $x_j \in \mathbb{R}$ pour toute composante $j = 1, \dots, n$.

Démonstration. Soit $\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}$, et considérons la norme $\|\cdot\|_\infty$ dans la définition de convergence d'une suite de \mathbb{R}^n . Alors,

$$\forall \epsilon > 0, \quad \exists N > 0 : \quad \forall k \geq N, \quad \|\mathbf{x} - \mathbf{x}^{(k)}\|_\infty = \max_{j=1, \dots, n} |x_j - x_j^{(k)}| \leq \epsilon,$$

ce qui implique que $\lim_{k \rightarrow \infty} x_j^{(k)} = x_j$, $\forall j = 1, \dots, n$.

Réciproquement, supposons qu'il existe $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ tel que $\lim_{k \rightarrow \infty} x_j^{(k)} = x_j$, $\forall j = 1, \dots, n$. Alors,

$$\forall \epsilon > 0, \quad \exists N_j > 0 : \quad \forall k \geq N_j, \quad |x_j - x_j^{(k)}| \leq \epsilon.$$

En prenant $\bar{N} = \max\{N_1, \dots, N_n\}$ on a $\max_{j=1, \dots, n} |x_j - x_j^{(k)}| \leq \epsilon$ pour tout $k \geq \bar{N}$, ce qui implique $\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}$. \square

Définition 1.10 (Suite de Cauchy). Une suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$ est dite de **Cauchy** si

$$\forall \epsilon > 0, \exists N > 0 : \quad \forall k, j \geq N, \quad \|\mathbf{x}^{(k)} - \mathbf{x}^{(j)}\| \leq \epsilon.$$

En suivant la même démonstration du lemme 1.9, on peut montrer qu'une suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$ est de Cauchy si et seulement si chaque suite $\{x_j^{(k)}\}_{k \in \mathbb{N}} \subset \mathbb{R}$, $j = 1, \dots, n$ est de Cauchy.

Théorème 1.11. Une suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$ est convergente si et seulement si elle est de Cauchy.

Démonstration. Ceci est vrai pour $n = 1$ (voir cours d'Analyse I). En utilisant la norme $\|\cdot\|_\infty$ et le lemme 1.9, on a : $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}}$ converge $\iff \{x_i^{(k)}\}_{k \in \mathbb{N}}$ converge pour tout i $\iff \{x_i^{(k)}\}_{k \in \mathbb{N}}$ est de Cauchy pour tout i $\iff \{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}}$ est de Cauchy. \square

Théorème 1.12 (Bolzano–Weierstrass sur \mathbb{R}^n). Soit $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$ une suite **bornée**, c.-à-d., $\exists M \in]0, +\infty[$ tel que $\|\mathbf{x}^{(k)}\| \leq M$, $\forall k \geq 0$. Alors il existe une sous-suite $\{\mathbf{x}^{(k_j)}\}_{j \in \mathbb{N}} \subset \{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}}$ qui est convergente.

Démonstration. On utilise le théorème de Bolzano–Weierstrass sur \mathbb{R} : puisque $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}}$ est bornée, en particulier $\{x_1^{(k)}\}_{k \in \mathbb{N}}$ est bornée où $\mathbf{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$. Donc on peut extraire une sous-suite $x_1^{(k_j)}$ qui converge vers $x_1 \in \mathbb{R}$. Prenons maintenant la suite $y_2^{(j)} = x_2^{(k_j)}$. Puisqu'elle est bornée, on peut extraire une sous-suite $y_2^{(j_\ell)}$ qui converge vers $x_2 \in \mathbb{R}$. Ainsi $\lim_{\ell \rightarrow \infty} x_2^{(k_{j_\ell})} = x_2$ et $\lim_{\ell \rightarrow \infty} x_1^{(k_{j_\ell})} = x_1$. En itérant ce raisonnement n fois, on peut extraire une sous-suite de $\mathbf{x}^{(k)}$ dont chaque composante converge vers (x_1, x_2, \dots, x_n) . \square

Ce qui est important dans la preuve de ce théorème est que l'on fait un nombre **fini** d'itérations (n est fini). Si n était ∞ la preuve ne porterait pas à conclusion.

1.4 Topologie de \mathbb{R}^n

1.4.1 Concepts de base

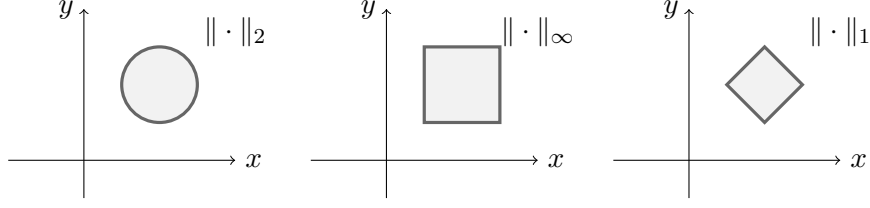
On s'intéresse ici à l'étude et classification des sous-ensembles de \mathbb{R}^n . On commence par définir les *boules*.

Définition 1.13 (Boule de \mathbb{R}^n). Pour tout $\mathbf{x} \in \mathbb{R}^n$ et $\delta > 0$, on appelle

- $B(\mathbf{x}, \delta) = \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\| < \delta\}$: la boule ouverte centrée en \mathbf{x} et de rayon δ ,
- $S(\mathbf{x}, \delta) = \partial B(\mathbf{x}, \delta) = \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\| = \delta\}$: la sphère centrée en \mathbf{x} et de rayon δ ,
- $\overline{B}(\mathbf{x}, \delta) = B(\mathbf{x}, \delta) \cup S(\mathbf{x}, \delta) = \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\| \leq \delta\}$: la boule fermée centrée en \mathbf{x} et de rayon δ .

La Figure 1.1 montre la forme des boules de \mathbb{R}^2 selon la norme qu'on choisit. On travaille par la suite avec la norme euclidienne $\|\mathbf{x}\| = \|\mathbf{x}\|_2$ mais toutes les définitions ci après s'appliquent à n'importe quelle norme puisque toutes les normes sont équivalentes sur \mathbb{R}^n .

On considère maintenant un sous-ensemble quelconque E de \mathbb{R}^n .

FIGURE 1.1 – Forme des boules de \mathbb{R}^2 pour des normes différentes

Définition 1.14 (sous-ensembles ouverts, fermés, bornés). Soit $E \subset \mathbb{R}^n$. On dit que

- E est **ouvert** si $\forall \mathbf{x} \in E, \exists \delta > 0$ tel que $B(\mathbf{x}, \delta) \subset E$. En particulier E est ouvert s'il est vide.
- E est **fermé** si son complémentaire $E^c = \mathbb{R}^n \setminus E = \{\mathbf{x} \in \mathbb{R}^n, \mathbf{x} \notin E\}$ est ouvert.
- E est **borné** s'il existe $M > 0$ tel que $\|\mathbf{x}\| \leq M, \forall \mathbf{x} \in E$.

On vérifie facilement que si un sous-ensemble E est ouvert par rapport à une norme, il est aussi ouvert par rapport à n'importe quelle autre norme puisque toutes les normes sont équivalentes sur \mathbb{R}^n . La même conclusion est vraie pour les ensembles fermés ou bornés. L'ensemble des ouverts de \mathbb{R}^n est appelé la **topologie** de \mathbb{R}^n (induite par une norme).

Remarque 1.15. On vérifie facilement que $B(\mathbf{x}, \delta)$ est ouvert et $\overline{B}(\mathbf{x}, \delta)$ est fermé. En effet, $\forall \mathbf{z} \in B(\mathbf{x}, \delta), B(\mathbf{z}, \delta - \|\mathbf{z} - \mathbf{x}\|) \subset B(\mathbf{x}, \delta)$, donc $B(\mathbf{x}, \delta)$ est ouvert. De même, $\forall \mathbf{z} \in \mathbb{R}^n \setminus \overline{B}(\mathbf{x}, \delta), B(\mathbf{z}, \|\mathbf{z} - \mathbf{x}\| - \delta) \subset \mathbb{R}^n \setminus \overline{B}(\mathbf{x}, \delta)$, donc $\overline{B}(\mathbf{x}, \delta)$ est fermé.

Étant donné un sous-ensemble $E \subset \mathbb{R}^n$, on peut classifier les points $\mathbf{x} \in \mathbb{R}^n$ par rapport à E de la façon suivante :

Définition 1.16. Soit $E \subset \mathbb{R}^n$ et $\mathbf{x} \in \mathbb{R}^n$. On dit que

- \mathbf{x} est un **point intérieur** de E si

$$\exists \delta > 0 : B(\mathbf{x}, \delta) \subset E.$$

L'ensemble des points intérieurs de E est noté \mathring{E} ou $\text{int}(E)$ et appelé l'**intérieur** de E .

- \mathbf{x} est un **point frontière** si

$$\forall \delta > 0, B(\mathbf{x}, \delta) \cap E \neq \emptyset \quad \text{et} \quad B(\mathbf{x}, \delta) \cap E^c \neq \emptyset.$$

L'ensemble des points frontières de E est noté ∂E et appelé la **frontière** ou le **bord** de E .

- \mathbf{x} est un **point adhérent** à E si

$$\forall \delta > 0, B(\mathbf{x}, \delta) \cap E \neq \emptyset.$$

Un point adhérent est soit un point intérieur, soit un point frontière. L'ensemble des points adhérents à E est noté \overline{E} , et appelé l'**adhérence** ou la **fermeture** de E , et coïncide avec $\overline{E} = E \cup \partial E = \mathring{E} \cup \partial E$.

— \mathbf{x} est un **point isolé** de E si

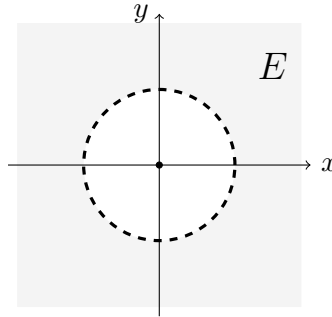
$$\exists \delta > 0 : B(\mathbf{x}, \delta) \cap E = \{\mathbf{x}\}$$

— \mathbf{x} est un **point d'accumulation** de E si $\forall \delta > 0$, $B(\mathbf{x}, \delta)$ contient au moins un point de E autre que \mathbf{x} , c.-à-d. $B(\mathbf{x}, \delta) \cap (E \setminus \{\mathbf{x}\}) \neq \emptyset$. (Rappel : $E \setminus \{\mathbf{x}\} = E$ si $\mathbf{x} \notin E$.)

Il s'ensuit que si \mathbf{x} est un point d'accumulation de E , alors $\forall \delta > 0$, $B(\mathbf{x}, \delta)$ contient une infinité de points de E . Les points d'accumulation de E sont tous les points de $\overline{E} = E \cup \partial E$ qui ne sont pas isolés.

On remarque que pour tout $\mathbf{y} \in \mathbb{R}^n$ et $\delta > 0$, on a $B(\mathbf{y}, \delta) \cap E \neq \emptyset \Leftrightarrow B(\mathbf{y}, \delta) \cap \overline{E} \neq \emptyset$. En effet, l'implication \Rightarrow est claire car $E \subset \overline{E}$. Pour montrer l'implication \Leftarrow , soit $\mathbf{z} \in B(\mathbf{y}, \delta) \cap \overline{E}$. Puisque \mathbf{z} est point adhérent à E , il existe $\mathbf{w} \in B(\mathbf{z}, \delta - \|\mathbf{z} - \mathbf{y}\|) \cap E \subset B(\mathbf{y}, \delta) \cap E$. Donc $B(\mathbf{y}, \delta) \cap E \neq \emptyset$.

Exercice 1.17. Soit $E = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 > 1\} \cup \{(0, 0)\}$ le sous-ensemble montré en figure. Déterminer son intérieur $\overset{\circ}{E}$, sa frontière ∂E , son complémentaire E^c , son adhérence \overline{E} , ainsi que tous ses points isolés et d'accumulation.



Définition 1.18. Soit $\mathbf{x} \in \mathbb{R}^n$. On dit qu'un ensemble $V \subset \mathbb{R}^n$ est un voisinage de \mathbf{x} s'il existe $\delta > 0$ tel que $B(\mathbf{x}, \delta) \subset V$. Autrement dit, tout ensemble V qui a \mathbf{x} comme point intérieur est un voisinage de \mathbf{x} .

Quelques remarques sur les ensemble ouverts

— $\overset{\circ}{E}$ est ouvert.

Démonstration. En effet, si $\mathbf{x} \in \overset{\circ}{E}$, il existe $\delta > 0$ tel que $B(\mathbf{x}, \delta) \subset E$. Vérifions que $B(\mathbf{x}, \delta) \subset \overset{\circ}{E}$, ce qui prouvera que $\overset{\circ}{E}$ est ouvert. Pour tout $\mathbf{z} \in B(\mathbf{x}, \delta)$, on a $B(\mathbf{z}, \delta - \|\mathbf{z} - \mathbf{x}\|) \subset B(\mathbf{x}, \delta) \subset E$, et donc $\mathbf{z} \in \overset{\circ}{E}$ et $B(\mathbf{x}, \delta) \subset \overset{\circ}{E}$. \square

— E est ouvert si et seulement si $E = \overset{\circ}{E}$.

- Toute réunion quelconque (même *infinie*, dénombrable ou non) de sous-ensembles ouverts de \mathbb{R}^n est un sous-ensemble ouvert.

Démonstration. Soit $E = \bigcup_{\alpha} E_{\alpha}$ avec E_{α} ouvert. Pour tout $\mathbf{x} \in E$, il existe $\alpha : \mathbf{x} \in E_{\alpha}$. Mais, E_{α} étant ouvert, $\exists \delta > 0 : B(\mathbf{x}, \delta) \subset E_{\alpha} \subset E$. Donc E est ouvert. \square

- Toute intersection *finie* de sous-ensembles ouverts de \mathbb{R}^n est un sous-ensemble ouvert.

Démonstration. Soit $E = \bigcap_{i=1}^m E_i$, avec E_i ouvert. Si $\mathbf{x} \in E$, alors $\mathbf{x} \in E_i \quad \forall i = 1, \dots, m$ et, puisque chaque E_i est ouvert, $\exists \delta_i : B(\mathbf{x}, \delta_i) \subset E_i$. Soit $\delta = \min\{\delta_1, \dots, \delta_m\}$ alors $B(\mathbf{x}, \delta) \subset B(\mathbf{x}, \delta_i) \subset E_i, \forall i$ et donc $B(\mathbf{x}, \delta) \subset \bigcap_{i=1}^m E_i = E$. \square

- \mathbb{R}^n est ouvert.

Quelques remarques sur les ensembles fermés

- $\mathbb{R}^n \setminus \overline{E} = \text{int}(\mathbb{R}^n \setminus E)$ et $\overline{\mathbb{R}^n \setminus E} = \mathbb{R}^n \setminus \overset{\circ}{E}$.
- L'adhérence \overline{E} d'un ensemble $E \subset \mathbb{R}^n$ est toujours fermé.

Démonstration. La preuve est par “passage au complémentaire” : son complémentaire $\mathbb{R}^n \setminus \overline{E} = \text{int}(\mathbb{R}^n \setminus E)$ est en effet ouvert. \square

- E est fermé si et seulement si $E = \overline{E}$. (Preuve : par passage aux complémentaires.)
- Toute intersection quelconque (même *infinie*, dénombrable ou non) de sous-ensembles fermés est fermée. (On passe aux complémentaires pour montrer cette propriété et la suivante.)
- Toute union *finie* de sous-ensembles fermés est fermée.
- \emptyset et \mathbb{R}^n sont fermés.

Une caractérisation importante des ensembles fermés est la suivante.

Lemme 1.19. *Un ensemble $E \subset \mathbb{R}^n$ non vide est fermé si et seulement si toute suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \subset E$ convergente, converge vers un élément de E .*

Démonstration. Soit E fermé et $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \subset E$ une suite convergente vers $\mathbf{x} \in \mathbb{R}^n$. Alors \mathbf{x} adhère à E et, puisque E est fermé, $\mathbf{x} \in E$.

Réciproquement, supposons que E n'est pas fermé, autrement dit, que $\mathbb{R}^n \setminus E$ n'est pas ouvert. Il existe donc $\overline{\mathbf{x}} \in \mathbb{R}^n \setminus E$ tel que $\forall \delta > 0 \quad B(\overline{\mathbf{x}}, \delta) \not\subset (\mathbb{R}^n \setminus E)$. En choisissant $\delta = 1/k$ avec $k \in \mathbb{N}^*$, on obtient $\mathbf{x}^{(k)} \in B(\overline{\mathbf{x}}, 1/k) \cap E$. La suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}^*}$ dans E converge alors vers $\overline{\mathbf{x}} \notin E$. \square

On montre de même que $\mathbf{x} \in \overline{E}$ ssi il existe une suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \subset E$ qui converge vers \mathbf{x} .

1.4.2 Ensembles compacts

On peut donner plusieurs définitions équivalentes d'un ensemble compact en \mathbb{R}^n . On présente ici la définition la plus “facile”, mais non pas celle qui caractérise le mieux la notion de compacité.

Définition 1.20 (Compacité). *Un ensemble $E \subset \mathbb{R}^n$ est **compact** s'il est à la fois borné et fermé. L'ensemble vide sera considéré comme compact.*

Les deux autres caractérisations (équivalentes en \mathbb{R}^n) sont montrées dans les théorèmes suivants.

Théorème 1.21 (Caractérisation de la compacité par sous-suites convergentes). *Un sous-ensemble non vide $E \subset \mathbb{R}^n$ est compact (fermé et borné) si et seulement si de toute suite d'éléments de E on peut extraire une sous-suite qui converge vers un élément de E .*

Démonstration.

1. Soit E compact (fermé et borné). Par le théorème de Bolzano–Weierstrass, de toute suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \subset E$ bornée (car E est borné), on peut extraire une sous-suite $\{\mathbf{x}^{(k_j)}\}_{j \in \mathbb{N}} \subset E$ convergente telle que $\lim_{j \rightarrow \infty} \mathbf{x}^{(k_j)} = \mathbf{x} \in \mathbb{R}^n$. Puisque E est fermé, $\mathbf{x} \in E$.
2. Supposons que E n'est pas compact, autrement dit, qu'il n'est pas fermé ou qu'il n'est pas borné (ou ni l'un ni l'autre). Si E n'est pas fermé, il existe $\bar{\mathbf{x}} \in \mathbb{R}^n \setminus E$ et une suite $\{\mathbf{y}^{(k)}\}_{k \in \mathbb{N}} \subset E$ telle que $\lim_{k \rightarrow \infty} \mathbf{y}^{(k)} = \bar{\mathbf{x}}$. Une telle suite n'a aucune sous-suite qui converge vers un élément de E (car $\bar{\mathbf{x}} \notin E$). Si E n'est pas borné, il existe une suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \subset E$ telle que $\|\mathbf{x}^{(k)}\| > k$ pour tout $k \in \mathbb{N}$. Toute sous-suite $\{\mathbf{x}^{(k_j)}\}_{j \in \mathbb{N}}$ satisfait $\|\mathbf{x}^{(k_j)}\| > k_j \geq j$ et $\{\mathbf{x}^{(k_j)}\}_{j \in \mathbb{N}}$ ne converge pas.

□

Théorème 1.22 (de Heine-Borel-Lebesgue – Caractérisation de la compacité par recouvrements finis). *Un sous-ensemble non vide $E \subset \mathbb{R}^n$ est compact (fermé et borné) si et seulement si de toute famille de sous-ensembles ouverts de \mathbb{R}^n constituant un recouvrement de E , c.-à-d. $E \subset \bigcup_{\alpha} U_{\alpha}$, avec U_{α} ouvert, on peut extraire une famille **finie** qui est encore un recouvrement de E .*

On dit qu'un ensemble E satisfait la *propriété de Heine–Borel* si de toute famille de sous-ensembles ouverts de \mathbb{R}^n constituant un recouvrement de E on peut extraire une famille finie qui est encore un recouvrement de E . Le théorème précédent affirme donc que un ensemble E est compact si et seulement si il satisfait la propriété de Heine–Borel. On montre deux exemples d'application de la propriété de Heine–Borel pour montrer qu'un ensemble n'est pas compact.

Exemple 1.23. \mathbb{R}^2 n'est pas compact. En fait, on peut écrire $\mathbb{R}^2 = \bigcup_{k \in \mathbb{N}^*} B(0, k)$ mais de ce recouvrement, on ne peut pas extraire de sous-recouvrement fini.

Exemple 1.24. $E = \overline{B}(0, 1) \setminus \{0\}$ n'est pas compact. En fait, on peut écrire $E \subset \bigcup_{k \in \mathbb{N}^*} \left(\overline{B}(0, \frac{1}{k}) \right)^c$ qui est un recouvrement de E mais duquel on ne peut pas extraire un sous-recouvrement fini.

Quelques remarques sur les ensembles compacts

- La caractérisation du théorème 1.21, qui peut être prise comme définition alternative de compacité, dit qu'un ensemble E est compact si et seulement si toute suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}}$ de E admet une sous-suite qui converge vers un élément de E , c'est-à-dire qu'il existe (au moins) un point $\mathbf{x} \in E$ (point d'accumulation de la suite) tel que, pour toute boule $B(\mathbf{x}, \delta)$, $\delta > 0$, l'ensemble d'indices $\{k \in \mathbb{N} : \mathbf{x}^{(k)} \in B(\mathbf{x}, \delta)\}$ est infini. Donc E est suffisamment contraignant (compact) pour que toute suite de E s'accumule quelque part dans E .
- La caractérisation du théorème 1.22 est la définition la plus générale de compacité, mais aussi la plus abstraite. Elle exprime le fait qu'on puisse décrire un ensemble compact par un nombre *fini* de termes et est à la base de toute étape d'approximation. Soit $E \subset \mathbb{R}^n$ un sous-ensemble quelconque. Clairement, pour tout $\varepsilon > 0$, $\bigcup_{\mathbf{x} \in E} B(\mathbf{x}, \varepsilon)$ est un recouvrement de E . Si E est compact, on peut extraire un sous-recouvrement fini, c.-à-d. il existe $s = s(\varepsilon) \in \mathbb{N}^*$ et $\{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(s)}\} \in E$ tels que $E \subset \bigcup_{i=1}^s B(\mathbf{x}^{(i)}, \varepsilon)$. Donc, E est bien approché par l'ensemble fini $\hat{E} = \{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(s)}\}$ au sens que pour tout $\mathbf{x} \in E$, $\text{dist}(\mathbf{x}, \hat{E}) = \inf_{i=1, \dots, s} \|\mathbf{x} - \mathbf{x}^{(i)}\| < \varepsilon$. Le nombre $s = s(\varepsilon)$ est appelé *nombre de recouvrement* de E et est un indicateur de la difficulté d'approcher E par un ensemble fini.

1.4.3 Ensembles connexes et connexes par arcs

Intuitivement, un ensemble $E \subset \mathbb{R}^n$ est connexe s'il est fait "d'un seul morceau". Plus rigoureusement, on dit qu'un ensemble E ouvert est connexe si on ne peut pas le séparer en deux parties *ouvertes* non vides et *disjointes*. La définition générale pour un ensemble quelconque est la suivante :

Définition 1.25 (Connexité). *Soit $E \subset \mathbb{R}^n$. On dit que E est **connexe** s'il n'existe pas deux ouverts $A, B \subset \mathbb{R}^n$ disjoints ($A \cap B = \emptyset$) tels que $A \cap E \neq \emptyset$, $B \cap E \neq \emptyset$ et $E \subset A \cup B$.*

En particulier \emptyset est connexe. Les ensembles connexes de \mathbb{R} sont les intervalles, par exemple \emptyset , \mathbb{R} , $[0, 1]$, $]0, 1[$, $[0, 1[$, $] - \infty, 0[$, $]0, \infty[$, etc. Une notion un peu plus forte de connexité est celle de *connexité par arcs*.

Définition 1.26. *Soit $E \subset \mathbb{R}^n$ un ensemble non vide. On appelle chemin de E une application $\gamma : [0, 1] \rightarrow E$, $t \mapsto \gamma(t) = (\gamma_1(t), \dots, \gamma_n(t)) \in E$, dont les fonctions $\gamma_i : [0, 1] \rightarrow \mathbb{R}$ sont continues.*

Définition 1.27 (Connexité par arcs). *Un ensemble non vide $E \subset \mathbb{R}^n$ est connexe par arcs si pour tout $\mathbf{x}, \mathbf{y} \in E$, il existe un chemin $\gamma : [0, 1] \rightarrow E$ tel que $\gamma(0) = \mathbf{x}$, $\gamma(1) = \mathbf{y}$ (et $\gamma(t) \in E, \forall t \in [0, 1]$). Nous considérerons \emptyset comme connexe par arcs.*

On peut montrer que tout ensemble $E \subset \mathbb{R}^n$ connexe par arcs est aussi connexe. Le réciproque n'est toutefois pas vraie. On verra dans le chapitre suivant que les propriétés de compacité, connexité et connexité par arcs sont des propriétés topologiques, préservées par les applications continues. Autrement dit, si $E \subset \mathbb{R}^n$ est un ensemble compact (resp. connexe ou connexe par arcs) et $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ une fonction continue, alors $\mathbf{f}(E) \subset \mathbb{R}^m$ est compact (resp. connexe ou connexe par arcs).

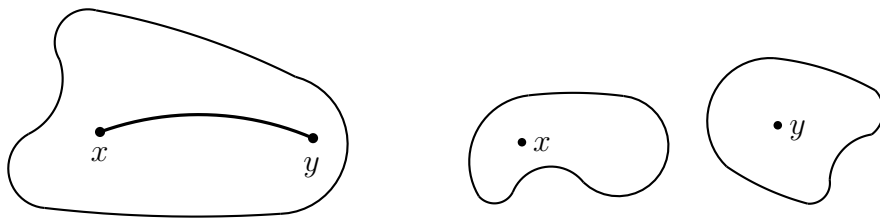


FIGURE 1.2 – Gauche : ensemble connexe par arcs. Droite : ensemble non connexe

Chapitre 2

Fonctions de plusieurs variables réelles ; limites et continuité

Soit $E \subset \mathbb{R}^n$ un ensemble non vide. On appelle *fonction sur E à valeurs réelles* une application $f : E \rightarrow \mathbb{R}$. C'est à dire, $\forall \mathbf{x} = (x_1, \dots, x_n) \in E$, $f(\mathbf{x}) = f(x_1, \dots, x_n) \in \mathbb{R}$ est l'image de \mathbf{x} par f . La fonction f est donc une fonction de n variables réelles. On note :

- E ou $D(f)$ le domaine de f ;
- $\text{Im}(f) = \{f(\mathbf{x}) \in \mathbb{R} : \mathbf{x} \in E\}$ l'image de f (notée aussi $f(E)$);
- $\mathcal{G}(f) = \{(\mathbf{x}, f(\mathbf{x})) \in \mathbb{R}^{n+1} : \mathbf{x} \in E\}$ le graphe de f .

Une fonction de 2 variables réelles à valeurs réelles, $(x, y) \mapsto f(x, y) \in \mathbb{R}$, peut être visualisée par son graphe (surface de \mathbb{R}^3), ou par ces lignes de niveau $N_f(c) = \{(x, y) \in \mathbb{R}^2 : f(x, y) = c\}$. La figure 2.1 montre le graphe et les lignes de niveau de la fonction $f(x, y) = e^{-x^2-y^2}$, $(x, y) \in [-1, 1]^2$.

2.1 Notions de limite

Définition 2.1 (limite). Soit $f : E \rightarrow \mathbb{R}$ et $\mathbf{x}_0 \in \mathbb{R}^n$ un point d'accumulation de E . On dit que $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})$ existe et est égale à $l \in \mathbb{R}$ si

$$\forall \epsilon > 0 \exists \delta > 0 : \forall \mathbf{x} \in E \left(0 < \|\mathbf{x} - \mathbf{x}_0\| \leq \delta \Rightarrow |f(\mathbf{x}) - l| \leq \epsilon \right).$$

On écrit alors $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = l$

La propriété $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = l$ ne dépend pas du choix de la norme $\|\cdot\|$ sur \mathbb{R}^n car les normes sur \mathbb{R}^n sont deux à deux équivalentes. On note que dans la définition de limite ci dessus, on exclut le point \mathbf{x}_0 de l'ensemble $\{\mathbf{x} \in E, 0 < \|\mathbf{x} - \mathbf{x}_0\| \leq \delta\}$. Cette limite est parfois appelée *limite épointée* et notée aussi $\lim_{\substack{\mathbf{x} \rightarrow \mathbf{x}_0 \\ \neq}} f(\mathbf{x})$ ou $\lim_{\substack{\mathbf{x} \rightarrow \mathbf{x}_0 \\ \mathbf{x} \neq \mathbf{x}_0}} f(\mathbf{x})$. Dans ces notes on entendra toujours par $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})$ la limite épointée. Le fait que $\mathbf{x}_0 \in E$ ou $\mathbf{x}_0 \notin E$ n'intervient pas dans cette définition.

Le théorème suivant donne une caractérisation équivalente de limite par les suites.

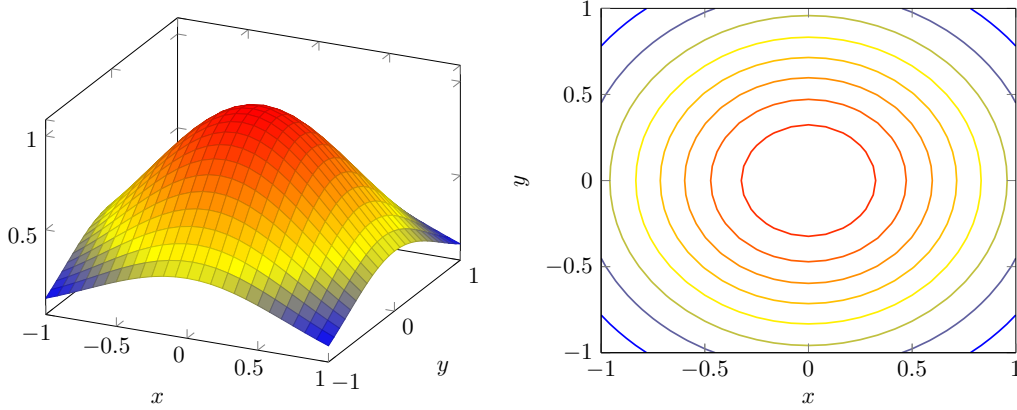


FIGURE 2.1 – Graphe (gauche) et lignes de niveau (droite) de la fonction $f(x, y) = e^{-x^2-y^2}$, $(x, y) \in [-1, 1]^2$.

Théorème 2.2. Soit $f : E \rightarrow \mathbb{R}$ et $\mathbf{x}_0 \in \mathbb{R}^n$ un point d'accumulation de E . Alors f admet pour limite l lorsque \mathbf{x} tend vers \mathbf{x}_0 si et seulement si, pour toute suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \subset E \setminus \{\mathbf{x}_0\}$ telle que $\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}_0$, on a $\lim_{k \rightarrow \infty} f(\mathbf{x}^{(k)}) = l$. De plus la limite l est unique (si elle existe).

Démonstration. Identique au cas d'une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ en remplaçant la boule 1D $]x_0 - \delta, x_0 + \delta[$ par la boule de \mathbb{R}^n , $B(\mathbf{x}_0, \delta)$. Voici la démonstration complète.

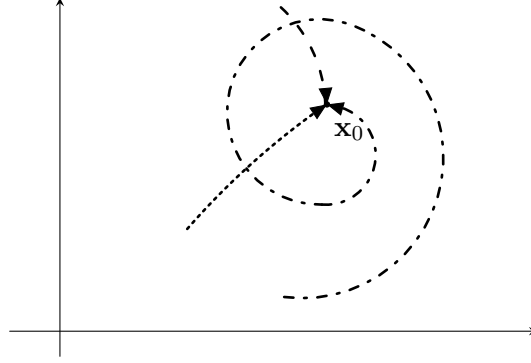
1. Supposons que f admet pour limite l lorsque \mathbf{x} tend vers \mathbf{x}_0 et donc, pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que, pour tout $\mathbf{x} \in E$ vérifiant $0 < \|\mathbf{x} - \mathbf{x}_0\| \leq \delta$, on a $|f(\mathbf{x}) - l| \leq \epsilon$. Soit, maintenant, $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \subset E \setminus \{\mathbf{x}_0\}$ une suite telle que $\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}_0$. Alors il existe $N \in \mathbb{N}$ tel que $\|\mathbf{x}^{(k)} - \mathbf{x}_0\| \leq \delta$ pour tout $k \geq N$ et donc $|f(\mathbf{x}^{(k)}) - l| \leq \epsilon$, ce qui montre que $\lim_{k \rightarrow \infty} f(\mathbf{x}^{(k)}) = l$.

2. Supposons que $\lim_{k \rightarrow \infty} f(\mathbf{x}^{(k)}) = l$ pour toute suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}}$ qui converge vers \mathbf{x}_0 . En raisonnant par l'absurde, supposons que f n'admet pas pour limite l lorsque \mathbf{x} tend vers \mathbf{x}_0 . Alors il existe $\epsilon > 0$ tel que pour tout $\delta > 0$ on a l'existence de $\mathbf{x}_0 \neq \mathbf{x} \in E$, $\|\mathbf{x} - \mathbf{x}_0\| \leq \delta$ (puisque \mathbf{x}_0 est un point d'accumulation de E) tel que $|f(\mathbf{x}) - l| > \epsilon$. Prenons $\delta = \frac{1}{k}$, $k \in \mathbb{N}^*$. Alors il existe une suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}^*}$ telle que $\|\mathbf{x}^{(k)} - \mathbf{x}_0\| \leq \frac{1}{k}$ (et donc $\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}_0$) et $|f(\mathbf{x}^{(k)}) - l| > \epsilon$, ce qui est contradictoire. □

Bien que la définition de limite soit la même pour des fonctions d'une seule ou de plusieurs variables réelles, le calcul des limites pour des fonctions de plusieurs variables réelles est bien plus compliqué. Prenons la caractérisation de limite par les suites :

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = l \iff \begin{aligned} & \forall \{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \subset E \setminus \{\mathbf{x}_0\}, \lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}_0 \\ & \text{on a } \lim_{k \rightarrow \infty} f(\mathbf{x}^{(k)}) = l \end{aligned}$$

Pour affirmer que $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = l$ existe, il faut s'assurer que $f(\mathbf{x}^{(k)}) \xrightarrow{k \rightarrow \infty} l$ pour n'importe quelle suite s'approchant de \mathbf{x}_0 .

FIGURE 2.2 – Exemples de chemins possibles qu'on peut suivre pour atteindre \mathbf{x}_0 .

Exemple 2.3. Considérons la fonction $f : E = \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}$, $f(x, y) = \frac{x}{\sqrt{x^2 + y^2}}$ et $\mathbf{x}_0 = (0, 0)$, qui est un point d'accumulation de E . Est-ce que $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})$ existe ?

Prenons la suite

$$\mathbf{x}^{(k)} = \left(\frac{1}{k}, 0\right) \quad k \in \mathbb{N}^* \quad \mathbf{x}^{(k)} \xrightarrow{k \rightarrow \infty} (0, 0).$$

On a que

$$f(\mathbf{x}^{(k)}) = \frac{\frac{1}{k}}{\sqrt{\frac{1}{k^2}}} = 1 \quad \forall k \quad \Rightarrow \quad \lim_{k \rightarrow \infty} f(\mathbf{x}^{(k)}) = 1.$$

Prenons maintenant la suite

$$\mathbf{y}^{(k)} = \left(0, \frac{1}{k}\right) \quad k \in \mathbb{N}^* \quad \mathbf{y}^{(k)} \xrightarrow{k \rightarrow \infty} (0, 0).$$

On a que

$$f(\mathbf{y}^{(k)}) = 0 \quad \forall k \quad \Rightarrow \quad \lim_{k \rightarrow \infty} f(\mathbf{y}^{(k)}) = 0.$$

On a donc trouvé deux suites différentes $\{(\frac{1}{k}, 0)\}_{k \in \mathbb{N}^*}$ et $\{(0, \frac{1}{k})\}_{k \in \mathbb{N}^*}$ qui donnent des limites différentes. On conclut donc que $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})$ n'existe pas.

Exemple 2.4. Considérons la fonction $f : E = \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}$, $f(x, y) = \frac{xy^2}{x^2 + y^4}$, et $\mathbf{x}_0 = (0, 0)$, qui est un point d'accumulation de E . Est-ce que $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})$ existe ?

Prenons la suite $\mathbf{x}^{(k)} = (\frac{\alpha}{k}, \frac{\beta}{k})$, $k \geq 1$, avec $\alpha, \beta \in \mathbb{R}$ non nuls en même temps. On a que

$$f(\mathbf{x}^{(k)}) = \frac{\frac{\alpha\beta^2}{k^3}}{\frac{\alpha^2}{k^2} + \frac{\beta^4}{k^4}} = \frac{\alpha\beta^2 k}{\alpha^2 k^2 + \beta^2} \xrightarrow{k \rightarrow \infty} 0, \quad \forall (\alpha, \beta) \neq (0, 0).$$

Donc si on se rapproche de \mathbf{x}_0 par un chemin "droit" la limite est 0. Toutefois, si on prend la suite $\mathbf{x}^{(k)} = (\frac{1}{k^2}, \frac{1}{k}) \xrightarrow{k \rightarrow \infty} (0, 0)$, on a que

$$f(\mathbf{x}^{(k)}) = \frac{\frac{1}{k^4}}{\frac{1}{k^4} + \frac{1}{k^4}} = \frac{1}{2} \quad \forall k \geq 1 \quad \Rightarrow \quad \lim_{k \rightarrow \infty} f(\mathbf{x}^{(k)}) = \frac{1}{2}.$$

Donc $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})$ n'existe pas !

2.1.1 Propriétés de l'opération de limite

L'opération de limite $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})$ pour des fonctions $f : E \subset \mathbb{R}^n \rightarrow \mathbb{R}$ de plusieurs variables réelles a les mêmes propriétés que pour des fonctions $f : \mathbb{R} \rightarrow \mathbb{R}$ d'une seule variable réelle.

Théorème 2.5. Soit $E \subset \mathbb{R}^n$, \mathbf{x}_0 un point d'accumulation de E et $f, g : E \rightarrow \mathbb{R}$ tels que $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = l_1$ et $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x}) = l_2$. Alors

- $\forall \alpha, \beta \in \mathbb{R}, \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} (\alpha f(\mathbf{x}) + \beta g(\mathbf{x})) = \alpha l_1 + \beta l_2$
- $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})g(\mathbf{x}) = l_1 l_2$
- Si $l_2 \neq 0$, $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{f(\mathbf{x})}{g(\mathbf{x})} = \frac{l_1}{l_2}$

Comme pour les fonctions $f : \mathbb{R} \rightarrow \mathbb{R}$ d'une seule variable réelle, on a un critère de comparaison qui peut être très utile pour établir l'existence d'une limite.

Théorème 2.6 (des deux gendarmes). Soient $f, g, h : E \subset \mathbb{R}^n \rightarrow \mathbb{R}$, \mathbf{x}_0 un point d'accumulation de E et $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} h(\mathbf{x}) = \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x}) = l$. S'il existe $\alpha > 0$ tel que

$$h(\mathbf{x}) \leq f(\mathbf{x}) \leq g(\mathbf{x}) \quad \forall \mathbf{x} \in E, 0 < \|\mathbf{x} - \mathbf{x}_0\| \leq \alpha$$

alors $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = l$.

Remarque 2.7. Dans le théorème des deux gendarmes, on utilise souvent des fonctions h et g qui dépendent uniquement de la distance $\|\mathbf{x} - \mathbf{x}_0\|$. Soit par exemple $g : E \rightarrow \mathbb{R}$ avec $E = \mathbb{R}^n$ ou $E = \mathbb{R}^n \setminus \{\mathbf{x}_0\}$, et supposons que

$$\forall \mathbf{x} \in E \setminus \{\mathbf{x}_0\} \quad g(\mathbf{x}) = \tilde{g}(\|\mathbf{x} - \mathbf{x}_0\|),$$

où $\tilde{g} :]0, \infty[\rightarrow \mathbb{R}$. Alors $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x}) = l$ si et seulement si $\lim_{r \rightarrow 0^+} \tilde{g}(r) = l$. En effet

$$\forall \epsilon > 0 \exists \delta > 0 : \forall \mathbf{x} \in \overline{B}(\mathbf{x}_0, \delta) \setminus \{\mathbf{x}_0\} \quad |g(\mathbf{x}) - l| = |\tilde{g}(\|\mathbf{x} - \mathbf{x}_0\|) - l| \leq \epsilon$$

si et seulement si

$$\forall \epsilon > 0 \exists \delta > 0 : \forall r \in]0, \delta] \quad |\tilde{g}(r) - l| \leq \epsilon.$$

Exemple 2.8. Soit $f : E = \mathbb{R}^2 \setminus \{(1, 0)\} \rightarrow \mathbb{R}$, $f(x, y) = y \ln((x-1)^2 + y^2)$. Calculer si elle existe $\lim_{(x,y) \rightarrow (1,0)} f(x, y)$.

Prenons la suite $\mathbf{x}^{(k)} = (1, \frac{1}{k})$. Alors $\lim_{k \rightarrow \infty} f(\mathbf{x}^{(k)}) = \lim_{k \rightarrow \infty} \frac{1}{k} \ln \frac{1}{k^2} = 0$. Donc si la limite existe elle doit être égale à $l = 0$. On a de plus

$$\begin{aligned} 0 \leq |f(x, y)| &= |y| |\ln((x-1)^2 + y^2)| \\ &\leq \sqrt{(x-1)^2 + y^2} |\ln((x-1)^2 + y^2)|. \end{aligned}$$

Notons $\rho = \sqrt{(x-1)^2 + y^2} = \|(x, y) - (1, 0)\|$. Puisque

$$\lim_{(x,y) \rightarrow (1,0)} \sqrt{(x-1)^2 + y^2} |\ln((x-1)^2 + y^2)| = \lim_{\rho \rightarrow 0^+} \rho |\ln \rho^2| = \lim_{\rho \rightarrow 0^+} \rho (-\ln \rho^2) = 0$$

(par la remarque appliquée à la norme euclidienne), on conclut par le théorème des deux gendarmes que $\lim_{(x,y) \rightarrow (1,0)} |f(x, y)| = 0$ et $\lim_{(x,y) \rightarrow (1,0)} f(x, y) = 0$.

Théorème 2.9 (Critère de Cauchy). *Soit $f : E \subset \mathbb{R}^n \rightarrow \mathbb{R}$ et \mathbf{x}_0 un point d'accumulation de E . Alors $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})$ existe (dans \mathbb{R}) si et seulement si*

$$\forall \epsilon > 0 \exists \delta > 0 : \forall \mathbf{x}, \mathbf{y} \in \overline{B}(\mathbf{x}_0, \delta) \cap (E \setminus \{\mathbf{x}_0\}) \quad |f(\mathbf{x}) - f(\mathbf{y})| \leq \epsilon.$$

Démonstration. Le sens \Rightarrow est clair. Montrons le sens \Leftarrow . Il existe $\delta_1 > 0$ tel que

$$\forall \mathbf{x}, \mathbf{y} \in \overline{B}(\mathbf{x}_0, \delta_1) \cap (E \setminus \{\mathbf{x}_0\}) \quad |f(\mathbf{x}) - f(\mathbf{y})| \leq 1.$$

Pour $\delta \in]0, \delta_1]$, soit les nombres réels

$$\alpha(\delta) = \inf\{f(\mathbf{x}) : \mathbf{x} \in \overline{B}(\mathbf{x}_0, \delta) \cap (E \setminus \{\mathbf{x}_0\})\}, \quad \beta(\delta) = \sup\{f(\mathbf{x}) : \mathbf{x} \in \overline{B}(\mathbf{x}_0, \delta) \cap (E \setminus \{\mathbf{x}_0\})\}$$

(deux fonctions monotones en δ). Il en résulte que

$$\forall \epsilon > 0 \exists \delta \in]0, \delta_1] \quad 0 \leq \beta(\delta) - \alpha(\delta) \leq \epsilon$$

et donc, comme $\beta - \alpha$ est croissante en δ , $\lim_{r \rightarrow 0^+} (\beta(r) - \alpha(r)) = 0$. Posons $\ell = \lim_{r \rightarrow 0^+} \beta(r) = \lim_{r \rightarrow 0^+} \alpha(r)$. Comme

$$\forall \mathbf{x} \in \overline{B}(\mathbf{x}_0, \delta_1) \cap (E \setminus \{\mathbf{x}_0\}) \quad \alpha(\|\mathbf{x} - \mathbf{x}_0\|) \leq f(\mathbf{x}) \leq \beta(\|\mathbf{x} - \mathbf{x}_0\|),$$

le théorème des deux gendarmes assure que $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = \ell$. □

2.1.2 Limite de fonctions à valeurs dans \mathbb{R}^m

La définition de limite s'étend sans difficultés aux fonctions à valeurs dans \mathbb{R}^m . Soit $\mathbf{f} : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, c'est-à-dire,

$$\forall \mathbf{x} \in E \quad \mathbf{f}(\mathbf{x}) = \mathbf{f}(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))$$

où $f_i : E \rightarrow \mathbb{R}$. Donc une fonction $\mathbf{f} : E \rightarrow \mathbb{R}^m$ est une collection de m fonctions $f_i : E \rightarrow \mathbb{R}$. Lorsque $n = m$, la norme dans l'espace de départ n'est pas nécessairement la même que celle dans l'espace d'arrivée.

Dans ce qui suit, nous choisirons la norme euclidienne, sauf mention du contraire.

Définition 2.10. *Soit $\mathbf{f} : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ et \mathbf{x}_0 un point d'accumulation de E . On dit que $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{l} \in \mathbb{R}^m$ si*

$$\forall \epsilon > 0, \exists \delta > 0 : \quad \forall \mathbf{x} \in E, \quad 0 < \|\mathbf{x} - \mathbf{x}_0\| \leq \delta \quad \|\mathbf{f}(\mathbf{x}) - \mathbf{l}\| \leq \epsilon.$$

Comme pour les fonctions à valeurs dans \mathbb{R} , $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{l}$ existe si et seulement si pour toute suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \subset E \setminus \{\mathbf{x}_0\}$ telle que $\mathbf{x}^{(k)} \xrightarrow{k \rightarrow \infty} \mathbf{x}_0$ on a $\lim_{k \rightarrow \infty} \mathbf{f}(\mathbf{x}^{(k)}) = \mathbf{l}$ (limite dans \mathbb{R}^m). Il est aussi facile de montrer que $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{l}$ si et seulement si toutes les limites $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f_i(\mathbf{x}) = l_i$, $i = 1, \dots, m$ existent.

2.2 Fonctions continues

Soit $E \subset \mathbb{R}^n$ un ensemble non vide et $f : E \rightarrow \mathbb{R}$.

Définition 2.11 (fonction continue en un point). Soit $\mathbf{x}_0 \in E$.

- Si \mathbf{x}_0 est un point isolé, on admettra (par définition) que f est continue en \mathbf{x}_0 .
- Si \mathbf{x}_0 n'est pas isolé (il est donc un point d'accumulation de E) on dit que f est continue en \mathbf{x}_0 si $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})$ existe et $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0)$.

De la définition 2.1 de limite d'une fonction ainsi que du théorème 2.2, caractérisant les limites par les suites, il s'ensuit que, pour tout $\mathbf{x}_0 \in E$, les trois affirmations suivantes sont équivalentes :

- i. $f : E \rightarrow \mathbb{R}$ est continue en \mathbf{x}_0 ;
- ii. $\forall \epsilon > 0 \exists \delta = \delta(\epsilon, \mathbf{x}_0) > 0 : \forall \mathbf{y} \in E \left(\|\mathbf{y} - \mathbf{x}_0\| \leq \delta \Rightarrow |f(\mathbf{y}) - f(\mathbf{x}_0)| \leq \epsilon \right)$;
- iii. $\lim_{k \rightarrow \infty} f(\mathbf{x}^{(k)}) = f(\mathbf{x}_0)$ pour toute suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \subset E$ qui converge vers \mathbf{x}_0 .

Définition 2.12 (fonction continue sur un ensemble). On dit que $f : E \rightarrow \mathbb{R}$ est continue sur E si elle est continue en tout point $\mathbf{x} \in E$. Dans ce cas, on note $f \in C^0(E)$ (ou $f \in C^0(E, \mathbb{R})$).

Il en résulte que f est continue en tout $\mathbf{x} \in E$ si et seulement si

$$\forall \epsilon > 0 \forall \mathbf{x} \in E \exists \delta = \delta(\epsilon, \mathbf{x}) > 0 : \forall \mathbf{y} \in E \left(\|\mathbf{x} - \mathbf{y}\| \leq \delta \Rightarrow |f(\mathbf{x}) - f(\mathbf{y})| \leq \epsilon \right).$$

Définition 2.13 (fonction uniformément continue). On dit que $f : E \rightarrow \mathbb{R}$ est uniformément continue sur E si

$$\forall \epsilon > 0 \exists \delta = \delta(\epsilon) > 0 : \forall \mathbf{x}, \mathbf{y} \in E \left(\|\mathbf{x} - \mathbf{y}\| \leq \delta \Rightarrow |f(\mathbf{x}) - f(\mathbf{y})| \leq \epsilon \right).$$

On note que ici δ peut être choisi de manière qui ne dépend pas de \mathbf{x} , contrairement à la caractérisation précédente de continuité sur E .

On remarque qu'une fonction $f : E \rightarrow \mathbb{R}$ uniformément continue sur E est aussi continue sur E . Le contraire n'est pas nécessairement vrai.

Exemple 2.14. La fonction $\mathbf{x} \mapsto \|\mathbf{x}\| \in \mathbb{R}_+$ est uniformément continue sur \mathbb{R}^n (et donc continue) car, pour tous $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\left| \|\mathbf{x}\| - \|\mathbf{y}\| \right| \leq \|\mathbf{x} - \mathbf{y}\|$$

(inégalité triangulaire inverse, qui découle de l'inégalité triangulaire).

Exemple 2.15. Toute fonction constante $\mathbf{x} \mapsto \mathbf{C} \in \mathbb{R}^m$ et, pour $i \in \{1, \dots, n\}$ fixé, la i -ème projection $\mathbf{x} \mapsto x_i \in \mathbb{R}$ sont des fonctions uniformément continues sur \mathbb{R}^n . En effet $|x_i - y_i| \leq \|\mathbf{x} - \mathbf{y}\|$ (norme euclidienne) pour tous $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Les définitions de continuité et continuité uniforme s'étendent sans difficultés à des fonctions $\mathbf{f} : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Définition 2.16. Soit $\mathbf{x}_0 \in E$.

- Si \mathbf{x}_0 est un point isolé, on admettra (par définition) que \mathbf{f} est continue en \mathbf{x}_0 .
- Si \mathbf{x}_0 n'est pas isolé, on dit que \mathbf{f} est continue en \mathbf{x}_0 si $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{f}(\mathbf{x})$ existe et $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0)$.

Il s'ensuit que, pour tout $\mathbf{x}_0 \in E$, les quatre affirmations suivantes sont équivalentes :

- i. $\mathbf{f} = (f_1, \dots, f_m) : E \rightarrow \mathbb{R}^m$ est continue en \mathbf{x}_0 ;
- ii. $\forall \epsilon > 0 \exists \delta = \delta(\epsilon, \mathbf{x}_0) > 0 : \forall \mathbf{y} \in E \left(\|\mathbf{y} - \mathbf{x}_0\| \leq \delta \Rightarrow \|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}_0)\| \leq \epsilon \right)$;
- iii. $\lim_{k \rightarrow \infty} \mathbf{f}(\mathbf{x}^{(k)}) = \mathbf{f}(\mathbf{x}_0)$ pour toute suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \subset E$ qui converge vers \mathbf{x}_0 ;
- iv. pour chaque $i \in \{1, \dots, m\}$ la fonction $f_i : E \rightarrow \mathbb{R}$ est continue en \mathbf{x}_0 .

Il en résulte aussi que \mathbf{f} est continue en tout $\mathbf{x} \in E$ si et seulement si

$$\forall \epsilon > 0 \forall \mathbf{x} \in E \exists \delta = \delta(\epsilon, \mathbf{x}) > 0 : \forall \mathbf{y} \in E \left(\|\mathbf{x} - \mathbf{y}\| \leq \delta \Rightarrow \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq \epsilon \right).$$

Dans ce cas, on note $f \in C^0(E, \mathbb{R}^m)$ (ou simplement $f \in C^0(E)$).

Définition 2.17. On dira que $\mathbf{f} : E \rightarrow \mathbb{R}^m$ est uniformément continue sur E si

$$\forall \epsilon > 0 \exists \delta = \delta(\epsilon) > 0 : \forall \mathbf{x}, \mathbf{y} \in E \left(\|\mathbf{x} - \mathbf{y}\| \leq \delta \Rightarrow \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq \epsilon \right). \quad (2.1)$$

Remarque 2.18. Soit $\emptyset \neq A \subset E$ et la restriction $\mathbf{f} : A \rightarrow \mathbb{R}^m$ de \mathbf{f} à A (notée aussi $\mathbf{f}|_A$). Si $\mathbf{f} : E \rightarrow \mathbb{R}^m$ est continue sur E , alors $\mathbf{f} : A \rightarrow \mathbb{R}^m$ est continue sur A .

2.2.1 Propriétés des fonctions continues

Théorème 2.19. Soient f et g deux fonctions de $E \subset \mathbb{R}^n$ dans \mathbb{R} continues en $\mathbf{x}_0 \in E$.

- $\forall \alpha, \beta \in \mathbb{R}, \alpha f + \beta g$ est continue en \mathbf{x}_0 ;
- $f \cdot g, |f|, |g|$ sont continues en \mathbf{x}_0 ;
- si $g(\mathbf{x}_0) \neq 0$, alors $\frac{f}{g}$ est continue en \mathbf{x}_0 .

Théorème 2.20 (Composition de fonctions continues). Soit $\mathbf{f} : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ continue en $\mathbf{x}_0 \in E$ et $\mathbf{g} : A \subset \mathbb{R}^p \rightarrow \mathbb{R}^n$ continue en $\mathbf{y}_0 \in A$ et telle que $\mathbf{x}_0 = \mathbf{g}(\mathbf{y}_0)$. Alors $\mathbf{h} = \mathbf{f} \circ \mathbf{g} : B = \{\mathbf{y} \in A : \mathbf{g}(\mathbf{y}) \in E\} \rightarrow \mathbb{R}^m$ est continue en \mathbf{y}_0 .

Démonstration. Pour toute suite $\{\mathbf{y}^{(k)}\}_{k \in \mathbb{N}} \subset B$ telle que $\lim_{k \rightarrow \infty} \mathbf{y}^{(k)} = \mathbf{y}_0$, on a par la continuité de \mathbf{g} que $\mathbf{g}(\mathbf{y}^{(k)}) \xrightarrow{k \rightarrow \infty} \mathbf{g}(\mathbf{y}_0) = \mathbf{x}_0$ et par la continuité de \mathbf{f} en \mathbf{x}_0 , $\mathbf{f}(\mathbf{g}(\mathbf{y}^{(k)})) \xrightarrow{k \rightarrow \infty} \mathbf{f}(\mathbf{x}_0)$. Donc, $\mathbf{h}(\mathbf{y}^{(k)}) \xrightarrow{k \rightarrow \infty} \mathbf{f}(\mathbf{x}_0) = \mathbf{f}(\mathbf{g}(\mathbf{y}_0)) = \mathbf{h}(\mathbf{y}_0)$ ce qui montre la continuité de \mathbf{h} en \mathbf{y}_0 . \square

2.3 Prolongement de fonctions par continuité

Définition 2.21. Soit $\mathbf{f} : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ et $\mathbf{x}_0 \in \bar{E} \setminus E$. Une fonction $\tilde{\mathbf{f}} : E \cup \{\mathbf{x}_0\} \rightarrow \mathbb{R}^m$ est appelée un prolongement par continuité de \mathbf{f} en \mathbf{x}_0 si $\tilde{\mathbf{f}} = \mathbf{f}$ sur E et $\tilde{\mathbf{f}}$ est continue en \mathbf{x}_0 .

Il est facile de vérifier qu'un prolongement $\tilde{\mathbf{f}}$ existe si et seulement si $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{f}(\mathbf{x})$ existe (dans \mathbb{R}^m), auquel cas $\tilde{\mathbf{f}}$ est uniquement déterminée par $\tilde{\mathbf{f}}(\mathbf{x}) = \mathbf{f}(\mathbf{x})$, $\forall \mathbf{x} \in E$ et $\tilde{\mathbf{f}}(\mathbf{x}_0) = \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{f}(\mathbf{x})$.

Théorème 2.22. Soit $E \subset \mathbb{R}^n$ un ensemble non vide et $\mathbf{f} : E \rightarrow \mathbb{R}^m$ une fonction continue sur E . Supposons que, pour tout $\mathbf{x} \in \bar{E} \setminus E$, la limite $\lim_{\mathbf{y} \rightarrow \mathbf{x}} \mathbf{f}(\mathbf{y})$ existe. Alors la fonction $\tilde{\mathbf{f}} : \bar{E} \rightarrow \mathbb{R}^m$ définie par $\tilde{\mathbf{f}}(\mathbf{x}) = \mathbf{f}(\mathbf{x})$ si $\mathbf{x} \in E$ et $\tilde{\mathbf{f}}(\mathbf{x}) = \lim_{\mathbf{y} \rightarrow \mathbf{x}} \mathbf{f}(\mathbf{y})$ si $\mathbf{x} \in \bar{E} \setminus E$ est continue et appelée le prolongement de \mathbf{f} par continuité sur \bar{E} .

Démonstration. Soit $\mathbf{x} \in \bar{E}$ et prouvons la continuité de $\tilde{\mathbf{f}}$ en \mathbf{x} à l'aide de suites. Soit donc une suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \subset \bar{E}$ qui converge vers \mathbf{x} . Remplaçons-la par une suite $\{\mathbf{y}^{(k)}\}_{k \in \mathbb{N}} \subset E$ qui converge aussi vers \mathbf{x} et telle que $\lim_{k \rightarrow +\infty} \|\mathbf{f}(\mathbf{y}^{(k)}) - \tilde{\mathbf{f}}(\mathbf{x}^{(k)})\| = 0$: si $\mathbf{x}^{(k)} \in E$, on pose $\mathbf{y}^{(k)} = \mathbf{x}^{(k)}$ et, si $\mathbf{x}^{(k)} \in \bar{E} \setminus E$, on choisit $\mathbf{y}^{(k)} \in E$ tel que

$$\|\mathbf{y}^{(k)} - \mathbf{x}^{(k)}\| \leq 2^{-k} \quad \text{et} \quad \|\mathbf{f}(\mathbf{y}^{(k)}) - \tilde{\mathbf{f}}(\mathbf{x}^{(k)})\| \leq 2^{-k}.$$

Par définition de $\tilde{\mathbf{f}}(\mathbf{x})$, on a $\lim_{k \rightarrow +\infty} \mathbf{f}(\mathbf{y}^{(k)}) = \tilde{\mathbf{f}}(\mathbf{x})$; en effet, si $\mathbf{x} \in E$, alors $\tilde{\mathbf{f}}(\mathbf{x}) = \mathbf{f}(\mathbf{x})$ et ceci découle de la continuité de \mathbf{f} en \mathbf{x} et, si $\mathbf{x} \in \bar{E} \setminus E$, ceci découle de la définition de $\tilde{\mathbf{f}}(\mathbf{x})$. Donc $\lim_{k \rightarrow +\infty} \tilde{\mathbf{f}}(\mathbf{x}^{(k)}) = \lim_{k \rightarrow +\infty} (\tilde{\mathbf{f}}(\mathbf{x}^{(k)}) - \mathbf{f}(\mathbf{y}^{(k)})) + \lim_{k \rightarrow +\infty} \mathbf{f}(\mathbf{y}^{(k)}) = \tilde{\mathbf{f}}(\mathbf{x})$. \square

Le prochain théorème est un résultat important. Il montre que sous l'hypothèse que la fonction soit *uniformément continue* sur E , on n'a pas besoin de vérifier l'existence des limites au bord pour pouvoir prolonger la fonction par continuité. L'hypothèse d'uniforme continuité est bien plus facile à vérifier que l'existence de la limite en chaque point du bord. On verra, par exemple, dans le chapitre suivant, que si la fonction est dérivable sur un ensemble E convexe avec dérivées partielles bornées, alors elle est uniformément continue sur E .

Théorème 2.23 (Prolongement de fonctions uniformément continues). Soit $E \subset \mathbb{R}^n$ un ensemble non vide et $\mathbf{f} : E \rightarrow \mathbb{R}^m$ une fonction uniformément continue sur E . Alors \mathbf{f} peut être prolongée par continuité sur \bar{E} et son prolongement continu $\tilde{\mathbf{f}} : \bar{E} \rightarrow \mathbb{R}^m$ est uniformément continu.

Démonstration. Vérification que \mathbf{f} peut être prolongée par continuité sur \bar{E} . Il faut vérifier que $\lim_{\mathbf{y} \rightarrow \mathbf{x}} \mathbf{f}(\mathbf{y})$ existe en tout $\mathbf{x} \in \bar{E} \setminus E$ (la limite en $\mathbf{x} \in E$ existe car \mathbf{f} est continue sur E). Par hypothèse, la fonction \mathbf{f} est uniformément continue sur E . Pour tout $\epsilon > 0$, soit $\delta > 0$ la valeur correspondante dans la définition (2.1) de continuité uniforme.

Pour chaque $\mathbf{a} \in \bar{E} \setminus E$, choisissons une suite $\{\mathbf{a}^{(k)}\}_{k \in \mathbb{N}} \subset E$ qui converge vers \mathbf{a} . Pour $\mathbf{a} \in \bar{E} \setminus E$ fixé, la suite choisie $\{\mathbf{a}^{(k)}\}_{k \in \mathbb{N}}$ est une suite de Cauchy. Il existe donc $N = N(\delta)$ tel que $\forall j, k \geq N$ $\|\mathbf{a}^{(j)} - \mathbf{a}^{(k)}\| \leq \delta$. Il s'ensuit que, pour tous $j, k \geq N$, $\|\mathbf{f}(\mathbf{a}^{(j)}) - \mathbf{f}(\mathbf{a}^{(k)})\| \leq \epsilon$. D'où $\{\mathbf{f}(\mathbf{a}^{(k)})\}_{k \in \mathbb{N}}$ est une suite de Cauchy dans \mathbb{R}^m , qui admet pour limite un certain

$\mathbf{l} = \lim_{k \rightarrow \infty} \mathbf{f}(\mathbf{a}^{(k)}) \in \mathbb{R}^m$. Cette limite ne dépend pas de la suite choisie. En effet, soit $\{\mathbf{b}^{(k)}\}_{k \in \mathbb{N}} \subset E$ une autre suite telle que $\lim_{k \rightarrow \infty} \mathbf{b}^{(k)} = \mathbf{a}$ et $\mathbf{m} = \lim_{k \rightarrow \infty} \mathbf{f}(\mathbf{b}^{(k)})$. Alors il existe $\tilde{N} > N : \forall k \geq \tilde{N}, \|\mathbf{m} - \mathbf{f}(\mathbf{b}^{(k)})\| \leq \epsilon, \|\mathbf{l} - \mathbf{f}(\mathbf{a}^{(k)})\| \leq \epsilon, \|\mathbf{a}^{(k)} - \mathbf{b}^{(k)}\| \leq \delta$. Par conséquent,

$$\|\mathbf{l} - \mathbf{m}\| \leq \|\mathbf{l} - \mathbf{f}(\mathbf{a}^{(k)})\| + \|\mathbf{f}(\mathbf{a}^{(k)}) - \mathbf{f}(\mathbf{b}^{(k)})\| + \|\mathbf{m} - \mathbf{f}(\mathbf{b}^{(k)})\| \leq 3\epsilon$$

ce qui implique $\mathbf{l} = \mathbf{m}$ par l'arbitrarité de ϵ et donc la limite $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{f}(\mathbf{x})$ existe en tout $\mathbf{a} \in E$ et \mathbf{f} peut être prolongée par continuité. On dénote $\tilde{\mathbf{f}}$ le prolongement par continuité de \mathbf{f} .

Vérification que $\tilde{\mathbf{f}}$ est uniformément continue sur \overline{E} . Soit $\epsilon > 0$ et $\mathbf{x}, \mathbf{y} \in \overline{E}$ tels que $\|\mathbf{x} - \mathbf{y}\| \leq \frac{\delta}{3}$, où $\delta = \delta(\epsilon)$ est comme ci-dessus. Introduisons deux suites $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}}, \{\mathbf{y}^{(k)}\}_{k \in \mathbb{N}} \subset E$ qui convergent à \mathbf{x} et \mathbf{y} , respectivement. Alors, il existe $M > 0$ tel que, pour tout $k \geq M$, $\|\mathbf{x} - \mathbf{x}^{(k)}\| \leq \frac{\delta}{3}$ et $\|\mathbf{y} - \mathbf{y}^{(k)}\| \leq \frac{\delta}{3}$, d'où $\|\mathbf{x}^{(k)} - \mathbf{y}^{(k)}\| \leq \|\mathbf{x}^{(k)} - \mathbf{x}\| + \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y} - \mathbf{y}^{(k)}\| \leq \delta$.

Puisque $\tilde{\mathbf{f}}$ est continue sur \overline{E} on a $\lim_{k \rightarrow \infty} \mathbf{f}(\mathbf{x}^{(k)}) = \tilde{\mathbf{f}}(\mathbf{x})$, $\lim_{k \rightarrow \infty} \mathbf{f}(\mathbf{y}^{(k)}) = \tilde{\mathbf{f}}(\mathbf{y})$ et il existe $\tilde{M} > M$ tel que, pour tout $k \geq \tilde{M}$, $\|\tilde{\mathbf{f}}(\mathbf{x}) - \mathbf{f}(\mathbf{x}^{(k)})\| \leq \epsilon$ et $\|\tilde{\mathbf{f}}(\mathbf{y}) - \mathbf{f}(\mathbf{y}^{(k)})\| \leq \epsilon$. On a alors

$$\|\tilde{\mathbf{f}}(\mathbf{x}) - \tilde{\mathbf{f}}(\mathbf{y})\| \leq \|\tilde{\mathbf{f}}(\mathbf{x}) - \mathbf{f}(\mathbf{x}^{(k)})\| + \|\mathbf{f}(\mathbf{x}^{(k)}) - \mathbf{f}(\mathbf{y}^{(k)})\| + \|\mathbf{f}(\mathbf{y}^{(k)}) - \tilde{\mathbf{f}}(\mathbf{y})\| \leq 3\epsilon.$$

Ainsi, si $\mathbf{x}, \mathbf{y} \in \overline{E}$ satisfont $\|\mathbf{x} - \mathbf{y}\| \leq \frac{\delta}{3}$, on a $\|\tilde{\mathbf{f}}(\mathbf{x}) - \tilde{\mathbf{f}}(\mathbf{y})\| \leq 3\epsilon$, ce qui prouve que $\tilde{\mathbf{f}}$ est uniformément continue sur \overline{E} . \square

2.4 Fonctions continues sur un compact

On commence par introduire la notion de fonction bornée et de borne supérieure et inférieure.

Définition 2.24 (fonction bornée). *On dit que $f : E \subset \mathbb{R}^n \rightarrow \mathbb{R}$ est bornée s'il existe $C > 0$ tel que $|f(\mathbf{x})| \leq C, \forall \mathbf{x} \in E$.*

Définition 2.25 (bornes supérieure et inférieure). *Soit $E \subset \mathbb{R}^n$ non vide et $f : E \rightarrow \mathbb{R}$.*

- *Soit $M = \sup_{\mathbf{x} \in E} f(\mathbf{x})$. Si $M < +\infty$, alors on a $f(\mathbf{x}) \leq M, \forall \mathbf{x} \in E$ et il existe une suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \subset E$ telle que $\lim_{k \rightarrow \infty} f(\mathbf{x}^{(k)}) = M$. On dit que M est la borne supérieure ou le supremum de f sur E .*
- *S'il existe $\mathbf{x}_M \in E$ tel que $f(\mathbf{x}_M) = M$, alors on dit que M est le maximum de f sur E , $M = \max_{\mathbf{x} \in E} f(\mathbf{x})$ et f atteint son maximum au point \mathbf{x}_M . On dit aussi que \mathbf{x}_M (pas nécessairement unique) est un point de maximum de f .*
- *Si $M = +\infty$, on dit que f n'est pas bornée supérieurement.*
- *On a des définitions du même type pour la borne inférieure ou l'infimum $m = \inf_{\mathbf{x} \in E} f(\mathbf{x})$, le minimum $m = \min_{\mathbf{x} \in E} f(\mathbf{x})$ et un point de minimum $\mathbf{x}_m \in E$ tel que $f(\mathbf{x}_m) = m$.*

Théorème 2.26. *Soit $E \subset \mathbb{R}^n$ un ensemble non vide et compact, et $f : E \rightarrow \mathbb{R}$ une fonction continue. Alors f est bornée et atteint ses bornes, c'est-à-dire, il existe $\mathbf{x}_M, \mathbf{x}_m \in E$ tels que $f(\mathbf{x}_M) = \sup_{\mathbf{x} \in E} f(\mathbf{x}) = \max_{\mathbf{x} \in E} f(\mathbf{x})$ et $f(\mathbf{x}_m) = \inf_{\mathbf{x} \in E} f(\mathbf{x}) = \min_{\mathbf{x} \in E} f(\mathbf{x})$.*

Démonstration. Similaire au cas des fonctions $f : [a, b] \rightarrow \mathbb{R}$ sur un intervalle fermé et borné. Montrons d'abord que f est bornée. Ab absurdo, si f n'était pas bornée, alors $\forall k \in \mathbb{N}, \exists \mathbf{x}^{(k)} \in E : |f(\mathbf{x}^{(k)})| > k$. Puisque E est compact, on peut extraire une sous-suite $\{\mathbf{x}^{(k_j)}\}_{j \in \mathbb{N}}$ qui converge vers un certain $\mathbf{x} \in E$. Mais, f étant continue en \mathbf{x} , pour tout $\epsilon > 0$ il existe $K_\epsilon \in \mathbb{N}$ tel que $|f(\mathbf{x}^{(k_j)})| \leq |f(\mathbf{x})| + \epsilon$ pour tout $j > K_\epsilon$, ce qui contredit $|f(\mathbf{x}^{(k_j)})| > k_j \geq j, \forall j$.

Soit maintenant $M = \sup_{\mathbf{x} \in E} f(\mathbf{x})$. Il existe une suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}} \subset E$ telle que $\lim_{k \rightarrow \infty} f(\mathbf{x}^{(k)}) = M$. A nouveau, on peut extraire une sous-suite $\{\mathbf{x}^{(k_j)}\}_{j \in \mathbb{N}}$ qui converge vers un certain $\mathbf{x}_M \in E$ et, par continuité de f , on a $f(\mathbf{x}_M) = M$, ce qui prouve que $M = \sup_{\mathbf{x} \in E} f(\mathbf{x}) = f(\mathbf{x}_M) = \max_{\mathbf{x} \in E} f(\mathbf{x})$. Idem pour le minimum. \square

Théorème 2.27. *Soit $E \subset \mathbb{R}^n$ un ensemble non vide, compact et connexe par arcs, et $f : E \rightarrow \mathbb{R}$ une fonction continue sur E . Alors f atteint toutes les valeurs entre son minimum m et maximum M sur E , et $\text{Im}(f) = [m, M]$.*

Démonstration. Puisque E est compact et $f : E \rightarrow \mathbb{R}$ continue, il existe \mathbf{x}_m et \mathbf{x}_M t.q. $f(\mathbf{x}_m) = \min_{\mathbf{x} \in E} f(\mathbf{x})$ et $f(\mathbf{x}_M) = \max_{\mathbf{x} \in E} f(\mathbf{x})$. Puisque E est connexe par arcs, il existe un chemin $\gamma(t) = (\gamma_1(t), \dots, \gamma_n(t))$ avec $\gamma_i : [0, 1] \rightarrow \mathbb{R}$ continues et $\gamma(t) \in E, \forall t \in [0, 1], \gamma(0) = \mathbf{x}_m, \gamma(1) = \mathbf{x}_M$. Soit $g : [0, 1] \rightarrow \mathbb{R}, g(t) = f(\gamma(t)) = f(\gamma_1(t), \dots, \gamma_n(t))$, qui est continue sur $[0, 1]$ puisqu'elle est la composition de fonctions continues. D'après le théorème de la valeur intermédiaire d'Analyse I, $\text{Im}(g)$ est un intervalle, et il contient $g(0) = m$ et $g(1) = M$, et donc tout $[m, M]$. D'où

$$[m, M] \subset \text{Im}(g) \subset \text{Im}(f) \subset [m, M]$$

et $\text{Im}(f) = [m, M]$. \square

Les deux théorèmes précédents montrent deux propriétés importantes des fonctions continues, qui se généralisent comme suit aux fonctions à valeurs dans \mathbb{R}^m .

Soit $\mathbf{f} : E \rightarrow \mathbb{R}^m$ continue et $\emptyset \neq A \subset E \subset \mathbb{R}^n$.

- Si A est compact, alors $\mathbf{f}(A) \subset \mathbb{R}^m$ est aussi compact.
- Si A est connexe (resp. connexe par arcs), alors $\mathbf{f}(A)$ est connexe (resp. connexe par arcs).

On conclut par une propriété importante des fonctions continues sur un compact.

Théorème 2.28 (Cantor-Heine). *Soit $E \subset \mathbb{R}^n$ un ensemble non vide et compact, et $\mathbf{f} : E \rightarrow \mathbb{R}^m$ une fonction continue. Alors \mathbf{f} est uniformément continue sur E (théorème de Cantor-Heine), c-à-d :*

$$\forall \epsilon > 0 \exists \delta = \delta(\epsilon) > 0 : \forall \mathbf{x}, \mathbf{y} \in E \left(\|\mathbf{x} - \mathbf{y}\| \leq \delta \Rightarrow \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq \epsilon \right).$$

Plus généralement, soit $E \subset \mathbb{R}^n$ un ensemble quelconque, $K \subset E$ un sous-ensemble non vide et compact, et $\mathbf{f} : E \rightarrow \mathbb{R}^m$ une fonction continue. Alors

$$\forall \epsilon > 0 \exists \delta = \delta(\epsilon) > 0 : \forall \mathbf{x} \in K \forall \mathbf{y} \in E \left(\|\mathbf{x} - \mathbf{y}\| \leq \delta \Rightarrow \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq \epsilon \right)$$

(théorème de Cantor-Heine généralisé).

Démonstration. Donnons la preuve de la version générale, le cas particulier $K = E$ non vide et compact du théorème de Cantor-Heine en découlant. Ab absurdo supposons qu'il n'est pas vrai que

$$\forall \epsilon > 0 \exists \delta > 0 : \forall \mathbf{x} \in K \forall \mathbf{y} \in E \left(\|\mathbf{x} - \mathbf{y}\| \leq \delta \Rightarrow \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq \epsilon \right).$$

Ainsi

$$\exists \epsilon > 0 : \forall \delta > 0 \exists \mathbf{x} \in K \exists \mathbf{y} \in E \left(\|\mathbf{x} - \mathbf{y}\| \leq \delta \text{ et } \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| > \epsilon \right).$$

Pour un tel $\epsilon > 0$, considérons $\delta = 1/k$ avec $k \in \mathbb{N}^*$: il existe $\mathbf{x}^{(k)} \in K$, $\mathbf{y}^{(k)} \in E$ tels que

$$\|\mathbf{x}^{(k)} - \mathbf{y}^{(k)}\| \leq \frac{1}{k} \text{ et } \|\mathbf{f}(\mathbf{x}^{(k)}) - \mathbf{f}(\mathbf{y}^{(k)})\| > \epsilon.$$

La suite $\{\mathbf{x}^{(k)}\}_{k \in \mathbb{N}^*}$ étant bornée (puisque K est borné), il existe une sous-suite $\{\mathbf{x}^{(k_i)}\}_{i \in \mathbb{N}}$ qui converge vers un certain $\mathbf{x} \in K$ puisque K est fermé. On a alors

$$\|\mathbf{y}^{(k_i)} - \mathbf{x}\| \leq \|\mathbf{y}^{(k_i)} - \mathbf{x}^{(k_i)}\| + \|\mathbf{x}^{(k_i)} - \mathbf{x}\| \leq \frac{1}{k_i} + \|\mathbf{x}^{(k_i)} - \mathbf{x}\| \rightarrow 0, \quad \text{si } i \rightarrow \infty.$$

Ainsi $\lim_{i \rightarrow \infty} \mathbf{y}^{(k_i)} = \mathbf{x}$. Puisque \mathbf{f} est continue sur E on a

$$\lim_{i \rightarrow \infty} \mathbf{f}(\mathbf{x}^{(k_i)}) = \mathbf{f}(\mathbf{x}) = \lim_{i \rightarrow \infty} \mathbf{f}(\mathbf{y}^{(k_i)}),$$

ce qui contredit $\|\mathbf{f}(\mathbf{x}^{(k)} - \mathbf{f}(\mathbf{y}^{(k)}))\| > \epsilon$ pour tout k . □

Chapitre 3

Dérivabilité

On commence ce chapitre en rappelant la notion de dérivée d'une fonction réelle d'une seule variable réelle, $f : E \subset \mathbb{R} \rightarrow \mathbb{R}$, en un point $x_0 \in \overset{\circ}{E}$:

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} \quad \text{si la limite existe.}$$

Elle représente le taux d'accroissement de la fonction en x_0 . De la définition, il suit immédiatement qu'il existe $\delta > 0$ tel que $]x_0 - \delta, x_0 + \delta[\subset E$ et

$$\forall h \in]-\delta, \delta[\quad f(x_0 + h) = f(x_0) + f'(x_0)h + o(|h|)$$

c'est-à-dire,

$$f(x_0 + h) - f(x_0) = \underbrace{f'(x_0)h}_{\text{application linéaire en } h} + o(|h|)$$

et l'incrément $f(x_0 + h) - f(x_0)$ est bien approché par une application linéaire en h . Ici $o(|h|)$ dénote une fonction g définie sur un voisinage de 0 et telle que $\lim_{h \rightarrow 0} \frac{g(h)}{h} = 0$ (voir plus loin pour une définition plus générale). En fait il suffit de poser $g(h) = f(x_0 + h) - f(x_0) - f'(x_0)h$ pour $|h| < \delta$.

On a donc une double interprétation de la notion de dérivée : comme limite du taux d'accroissement en x_0 , ainsi que comme application linéaire approchant localement la fonction dans un voisinage de x_0 . En particulier, f dérivable en x_0 implique que f est aussi continue en x_0 . Dans ce chapitre, on va généraliser ces deux notions pour des fonctions $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ou $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ de plusieurs variables réelles.

3.1 Dérivées partielles, dérivées directionnelles, différentielle

Définition 3.1. Soit $f : E \subset \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction de n variables réelles, $\mathbf{x}_0 \in \overset{\circ}{E}$ un point intérieur de E et \mathbf{v} un vecteur arbitraire de \mathbb{R}^n . On dit que f est dérivable dans la direction \mathbf{v} au point \mathbf{x}_0 si la limite $\lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t}$ existe. Dans ce cas on pose

$$D_{\mathbf{v}}f(\mathbf{x}_0) = \lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t}$$

Si $\|\mathbf{v}\| = 1$ pour la norme euclidienne, on appelle $D_{\mathbf{v}}f(\mathbf{x}_0)$ la **dérivée directionnelle** de f dans la direction \mathbf{v} au point \mathbf{x}_0 .

Si on définit la fonction $\tilde{f}(t) = f(\mathbf{x}_0 + t\mathbf{v})$ dans un voisinage de 0 (ce qui est toujours possible car \mathbf{x}_0 est un point intérieur de E), alors $D_{\mathbf{v}}f(\mathbf{x}_0) = \tilde{f}'(0)$. On peut donc interpréter la dérivée directionnelle $D_{\mathbf{v}}f(\mathbf{x}_0)$ comme la limite du taux d'accroissement en suivant la direction \mathbf{v} . En particulier, si on prend $\mathbf{v} = \mathbf{e}_i$, le i -ème vecteur de la base

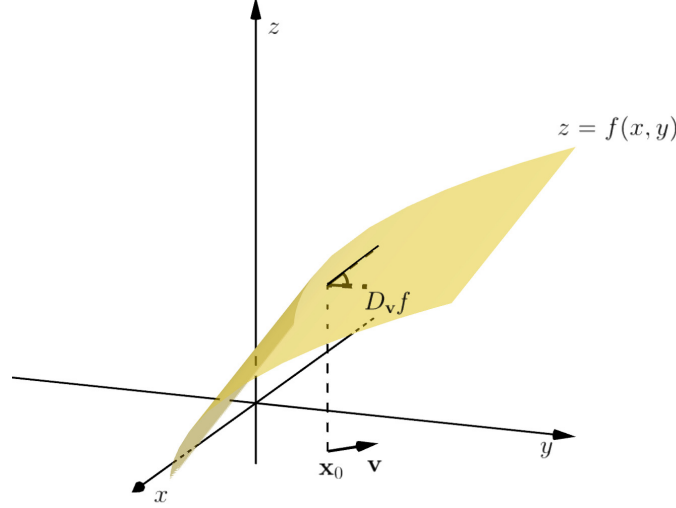


FIGURE 3.1 – Interprétation géométrique de dérivée directionnelle pour une fonction $f(x, y)$ de deux variables réelles

canonique de \mathbb{R}^n , alors la dérivée directionnelle correspondante est appelée *i -ème dérivée partielle* et est notée

$$\frac{\partial f}{\partial x_i}(\mathbf{x}_0) = D_{\mathbf{e}_i}f(\mathbf{x}_0) = \lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{e}_i) - f(\mathbf{x}_0)}{t}.$$

Autrement dit, pour $\mathbf{x} = (x_1, \dots, x_n) \in \mathring{E}$ point intérieur,

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) = \lim_{t \rightarrow 0} \frac{f(x_1, \dots, x_i + t, \dots, x_n) - f(x_1, \dots, x_i, \dots, x_n)}{t},$$

ce qui revient à calculer la dérivée de la fonction $x_i \mapsto f(\cdot, x_i, \cdot)$ pensée comme une fonction uniquement de la variable x_i , en traitant les autres variables $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ comme des paramètres. Pour cela, on peut donc utiliser les règles de dérivation des fonctions d'une seule variable réelle.

Définition 3.2 (vecteur gradient). Soit $f : E \subset \mathbb{R}^n \rightarrow \mathbb{R}$ et \mathbf{x}_0 un point intérieur de E . Si toutes les dérivées partielles de f existent en \mathbf{x}_0 , on appelle **matrice jacobienne** $Df(\mathbf{x}_0) \in \mathbb{R}^{1 \times n}$ (vecteur ligne) le vecteur

$$Df(\mathbf{x}_0) = \left[\frac{\partial f}{\partial x_1}(\mathbf{x}_0), \dots, \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \right] = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}_0) \quad \dots \quad \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \right)$$

et **vecteur gradient**, noté $\nabla f(\mathbf{x}_0) \in \mathbb{R}^{n \times 1}$, son transposé

$$\nabla f(\mathbf{x}_0) = Df(\mathbf{x}_0)^\top = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}_0) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \end{bmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}_0) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \end{pmatrix}.$$

Exemple 3.3. Soit $f(x, y) = x^2 \cos y$. Alors

$$\frac{\partial f}{\partial x}(x, y) = 2x \cos y, \quad \frac{\partial f}{\partial y}(x, y) = -x^2 \sin y, \quad \nabla f(x, y) = \begin{bmatrix} 2x \cos y \\ -x^2 \sin y \end{bmatrix}.$$

Une fonction peut avoir des dérivées partielles ou directionnelles en un point \mathbf{x}_0 sans pour autant être continue en ce point, comme les exemples suivants le montrent.

Exemple 3.4. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ donnée par

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0). \end{cases}$$

Alors,

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{t \rightarrow 0} \frac{f(t, 0) - f(0, 0)}{t} = 0, \quad \frac{\partial f}{\partial y}(0, 0) = \lim_{t \rightarrow 0} \frac{f(0, t) - f(0, 0)}{t} = 0$$

mais f n'est pas continue en $(0, 0)$. En effet, on a $\lim_{t \rightarrow 0} f(t, t) = \lim_{t \rightarrow 0} \frac{t^2}{2t^2} = \frac{1}{2} \neq 0$. Pour cette fonction, les autres dérivées directionnelles n'existent pas :

$$\lim_{t \rightarrow 0} \frac{f(tv_1, tv_2) - f(0, 0)}{t} = \lim_{t \rightarrow 0} \frac{t^2 v_1 v_2}{t^3(v_1^2 + v_2^2)} = \lim_{t \rightarrow 0} \frac{v_1 v_2}{t(v_1^2 + v_2^2)}$$

n'existe pas si $v_1 v_2 \neq 0$ (ici $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^2$, $\|\mathbf{v}\| = 1$ pour la norme euclidienne)

Exemple 3.5. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ donnée par

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2 + y^4}, & \text{si } (x, y) \neq (0, 0) \\ 0, & \text{si } (x, y) = (0, 0). \end{cases}$$

Toutes les dérivées directionnelles existent. En effet,

$$D_{\mathbf{v}}f(0, 0) = \lim_{t \rightarrow 0} \frac{v_1 v_2^2}{v_1^2 + t^2 v_2^4} = \begin{cases} \frac{v_2^2}{v_1}, & \text{si } v_1 \neq 0, \\ 0, & \text{si } v_1 = 0. \end{cases}$$

Toutefois, f n'est pas continue en $(0, 0)$ car $\lim_{t \rightarrow 0} f(t^2, t) = \frac{1}{2} \neq f(0, 0)$.

Comme l'exemple précédent le montre, la notion de dérivée partielle est trop faible et n'implique pas la continuité de la fonction. Si on veut introduire une notion de dérivabilité qui permet d'approcher une fonction au voisinage de \mathbf{x}_0 par une fonction affine, on a besoin d'une définition de dérivée plus forte. On rappelle qu'une application linéaire $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est telle que $L(\alpha \mathbf{x} + \beta \mathbf{y}) = \alpha L(\mathbf{x}) + \beta L(\mathbf{y})$, $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ et $\forall \alpha, \beta \in \mathbb{R}$. Elle peut être représentée par la matrice $A \in \mathbb{R}^{m \times n}$ (m lignes et n colonnes) telle que $L(\mathbf{e}_i)$ est la $i^{\text{ème}}$ colonne de A , où $\{\mathbf{e}_i\}_{i=1}^n$ est la base canonique de \mathbb{R}^n , de telle sorte que $L(\mathbf{x}) = A \cdot \mathbf{x} = \mathbf{Ax}$, $\forall \mathbf{x} \in \mathbb{R}^n$ (produit matriciel entre une matrice $m \times n$ et un vecteur colonne dans \mathbb{R}^n). Le produit matriciel $\mathbf{Ax} \in \mathbb{R}^m$ donne dans cette formule un vecteur colonne.

Définition 3.6 (Dérivabilité et différentielle). *Soit $f : E \subset \mathbb{R}^n \rightarrow \mathbb{R}$ et \mathbf{x}_0 un point intérieur de E . On dit que f est différentiable (ou dérivable) en \mathbf{x}_0 s'il existe une application linéaire $L : \mathbb{R}^n \rightarrow \mathbb{R}$ et une fonction $g : E \rightarrow \mathbb{R}$ tels que*

$$\forall \mathbf{x} \in E \quad f(\mathbf{x}) = f(\mathbf{x}_0) + L(\mathbf{x} - \mathbf{x}_0) + g(\mathbf{x}), \quad \text{et} \quad \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{g(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|} = 0.$$

L'application linéaire L est alors appelée la **différentielle** de f en \mathbf{x}_0 .

On remarque que la fonction g satisfait nécessairement $g(\mathbf{x}_0) = 0$ et est continue en \mathbf{x}_0 .

Notation. Pour $p \geq 0$, on utilise souvent la notation $g(\mathbf{x}) = o(\|\mathbf{x} - \mathbf{x}_0\|^p)$ pour indiquer une fonction g définie au moins sur $B(\mathbf{x}_0, \delta) \setminus \{\mathbf{x}_0\}$ pour un certain $\delta > 0$ et telle que $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{g(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|^p} = 0$.

Avec la notation $o(\cdot)$, dans la définition 3.6 on peut ainsi écrire $f(\mathbf{x}) = f(\mathbf{x}_0) + L(\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|)$ sur E (ici $p = 1$ et g est définie sur E).

Le théorème suivant met en relation l'application linéaire L de la définition précédente avec la matrice jacobienne (ou bien le gradient) de f en \mathbf{x}_0 et montre aussi que la différentielle de f en \mathbf{x}_0 , si elle existe, est unique.

Théorème 3.7. *Soit $f : E \rightarrow \mathbb{R}$ différentiable en $\mathbf{x}_0 \in \overset{\circ}{E}$. Alors, toutes les dérivées partielles de f existent en \mathbf{x}_0 et la différentielle de f en \mathbf{x}_0 est unique, donnée par*

$$L(\mathbf{x} - \mathbf{x}_0) = Df(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) = Df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$$

(produit matriciel entre un vecteur ligne et un vecteur colonne), ce qui s'écrit aussi $L(\mathbf{x} - \mathbf{x}_0) = \nabla f(\mathbf{x}_0)^\top \cdot (\mathbf{x} - \mathbf{x}_0) = \nabla f(\mathbf{x}_0)^\top (\mathbf{x} - \mathbf{x}_0)$ (produit matriciel). On peut aussi utiliser le produit scalaire usuel entre deux vecteurs colonnes : $L(\mathbf{x} - \mathbf{x}_0) = \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$ (produit scalaire). De plus, f est continue en \mathbf{x}_0 .

Démonstration. On note $a_i = L(\mathbf{e}_i)$ de sorte que $L(\mathbf{x} - \mathbf{x}_0) = \sum_{i=1}^n a_i(x_i - x_{0i})$. Par définition de la différentiabilité en \mathbf{x}_0 , on a $f(\mathbf{x}_0 + t\mathbf{e}_i) = f(\mathbf{x}_0) + tL(\mathbf{e}_i) + g(\mathbf{x}_0 + t\mathbf{e}_i)$ et, si on note $\mathbf{x}_t = \mathbf{x}_0 + t\mathbf{e}_i$,

$$\begin{aligned} \frac{\partial f}{\partial x_i}(\mathbf{x}_0) &= \lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{e}_i) - f(\mathbf{x}_0)}{t} \\ &= L(\mathbf{e}_i) + \lim_{t \rightarrow 0} \frac{g(\mathbf{x}_0 + t\mathbf{e}_i)}{t} \\ &= L(\mathbf{e}_i) + \underbrace{\lim_{t \rightarrow 0} \frac{\text{sign}(t)g(\mathbf{x}_t)}{\|\mathbf{x}_t - \mathbf{x}_0\|}}_{=0} = a_i, \end{aligned}$$

donc toutes les dérivées partielles existent et $L(\mathbf{x} - \mathbf{x}_0) = Df(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$. De plus, on a

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0) + \sum_{i=1}^n \underbrace{\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)(x_i - x_{0i})}_{=0} + \underbrace{\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x})}_{=0} = f(\mathbf{x}_0),$$

ce qui montre que f est continue en \mathbf{x}_0 . □

Une conséquence immédiate du théorème précédent est que, si $f : E \subset \mathbb{R}^n \rightarrow \mathbb{R}$ est différentiable en $\mathbf{x}_0 \in \mathring{E}$, alors toutes les dérivées directionnelles existent et $D_{\mathbf{v}}f(\mathbf{x}_0) = Df(\mathbf{x}_0) \cdot \mathbf{v} = \nabla f(\mathbf{x}_0)^\top \mathbf{v}$ (produit matriciel). En effet,

$$D_{\mathbf{v}}f(\mathbf{x}_0) = \lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t} = L(\mathbf{v}) + \underbrace{\lim_{t \rightarrow 0} \frac{g(\mathbf{x}_0 + t\mathbf{v})}{t}}_{=0} = L(\mathbf{v}) = \nabla f(\mathbf{x}_0)^\top \mathbf{v}.$$

En particulier, $D_{\mathbf{v}}f(\mathbf{x}_0)$ est dans ce cas une application linéaire en \mathbf{v} . On a de plus, pour la norme euclidienne,

$$\|\nabla f(\mathbf{x}_0)\| = \max_{\substack{\mathbf{v} \in \mathbb{R}^n \\ \|\mathbf{v}\|=1}} \nabla f(\mathbf{x}_0)^\top \mathbf{v} = \max_{\substack{\mathbf{v} \in \mathbb{R}^n \\ \|\mathbf{v}\|=1}} D_{\mathbf{v}}f(\mathbf{x}_0),$$

le maximum étant atteint en $\mathbf{v} = \frac{\nabla f(\mathbf{x}_0)}{\|\nabla f(\mathbf{x}_0)\|}$ si $\nabla f(\mathbf{x}_0)$ n'est pas nul. S'il n'est pas nul, le vecteur gradient donne donc la direction de croissance maximale au point \mathbf{x}_0 pour la fonction f . On a aussi que toute direction $\mathbf{w} \perp \nabla f(\mathbf{x}_0)$ est une direction de croissance nulle au point \mathbf{x}_0 .

Exemple 3.8. *Considérons la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = x^2 + y^2$. Son gradient est donné par $\nabla f(x, y) = (2x, 2y)^\top$. En $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$, la direction de croissance maximale au point (x, y) est $\mathbf{v} = \frac{\nabla f(x, y)}{\|\nabla f(x, y)\|} = \left(\frac{x}{\sqrt{x^2 + y^2}}, \frac{y}{\sqrt{x^2 + y^2}} \right)^\top$ (norme euclidienne normalisée à 1 ici), la direction de croissance minimale au point (x, y) est $\mathbf{v} = -\frac{\nabla f(x, y)}{\|\nabla f(x, y)\|} = -\left(\frac{x}{\sqrt{x^2 + y^2}}, \frac{y}{\sqrt{x^2 + y^2}} \right)^\top$, tandis que la direction de croissance nulle au point (x, y) est $\mathbf{w} = \pm \left(-\frac{y}{\sqrt{x^2 + y^2}}, \frac{x}{\sqrt{x^2 + y^2}} \right)^\top$.*

Enfin, si f est différentiable en \mathbf{x}_0 , on peut construire l'hyperplan tangent au graphe de f en $(\mathbf{x}_0, f(\mathbf{x}_0))$, ainsi qu'un vecteur normal au graphe de f en $(\mathbf{x}_0, f(\mathbf{x}_0))$.

Définition 3.9. *Soit $f : E \subset \mathbb{R}^n \rightarrow \mathbb{R}$ différentiable en $\mathbf{x}_0 \in \mathring{E}$ et notons $y_0 = f(\mathbf{x}_0)$ et $\mathcal{G}_f = \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} : \mathbf{x} \in E, y = f(\mathbf{x})\}$ le graphe de f . On définit l'**hyperplan tangent** à \mathcal{G}_f en (\mathbf{x}_0, y_0) comme le sous-ensemble de \mathbb{R}^{n+1}*

$$\Pi_{(\mathbf{x}_0, y_0)}(\mathcal{G}_f) = \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} : y = y_0 + \nabla f(\mathbf{x}_0)^\top (\mathbf{x} - \mathbf{x}_0)\}$$

Le vecteur

$$\mathbf{n} = \frac{1}{\sqrt{1 + \|\nabla f(\mathbf{x}_0)\|^2}} \left(-\frac{\partial f}{\partial x_1}(\mathbf{x}_0), \dots, -\frac{\partial f}{\partial x_n}(\mathbf{x}_0), 1 \right)^\top \in \mathbb{R}^{n+1}$$

est un vecteur normal au graphe de f en (\mathbf{x}_0, y_0) .

En effet, on vérifie facilement que le vecteur \mathbf{n} est normal à l'hyperplan tangent $\Pi_{(\mathbf{x}_0, y_0)}(\mathcal{G}_f)$ en (\mathbf{x}_0, y_0) , c'est-à-dire, il satisfait

$$\mathbf{n} \cdot ((\mathbf{x}, y) - (\mathbf{x}_0, y_0)) = 0, \quad \forall (\mathbf{x}, y) \in \Pi_{(\mathbf{x}_0, y_0)}(\mathcal{G}_f).$$

De plus \mathbf{n} est de norme euclidienne 1, comme d'ailleurs $-\mathbf{n}$.

Les définitions de différentiabilité et dérivées partielles/directionnelles s'étendent sans difficultés aux fonctions à valeurs dans \mathbb{R}^m .

Définition 3.10. Soit $\mathbf{f} : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$,

$$\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))^\top = \begin{bmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{bmatrix} = \begin{pmatrix} f_1(\mathbf{x}) \\ \dots \\ f_m(\mathbf{x}) \end{pmatrix},$$

noté aussi sous forme ligne (nous préciserons si nécessaire)

$$\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x})) = \begin{bmatrix} f_1(\mathbf{x}) & \dots & f_m(\mathbf{x}) \end{bmatrix} = \begin{pmatrix} f_1(\mathbf{x}) & \dots & f_m(\mathbf{x}) \end{pmatrix},$$

et soit $\mathbf{x}_0 \in \overset{\circ}{E}$. On appelle dérivée partielle de la i -ème composante de \mathbf{f} par rapport à la j -ème variable en \mathbf{x}_0 , notée $\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0)$, la quantité

$$\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0) = \lim_{t \rightarrow 0} \frac{f_i(\mathbf{x}_0 + t\mathbf{e}_j) - f_i(\mathbf{x}_0)}{t},$$

si la limite existe, et matrice jacobienne de \mathbf{f} en \mathbf{x}_0 la matrice $D\mathbf{f}(\mathbf{x}_0) \in \mathbb{R}^{m \times n}$ de composantes $(D\mathbf{f}(\mathbf{x}_0))_{ij} = \frac{\partial f_i}{\partial x_j}(\mathbf{x}_0)$, autrement dit,

$$D\mathbf{f}(\mathbf{x}_0) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}_0) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}_0) & & & \vdots \\ \vdots & & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \dots & \dots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{bmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}_0) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}_0) & & & \vdots \\ \vdots & & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \dots & \dots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix},$$

si toutes les dérivées partielles existent.

De façon similaire, on introduit la notion de différentiabilité.

Définition 3.11. On dit que $\mathbf{f} : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ est différentiable (ou dérivable) en $\mathbf{x}_0 \in \overset{\circ}{E}$ s'il existe une application linéaire $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ et une fonction $\mathbf{g} : E \rightarrow \mathbb{R}^m$ telles que

$$\forall \mathbf{x} \in E \quad \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + L(\mathbf{x} - \mathbf{x}_0) + \mathbf{g}(\mathbf{x})$$

et $\mathbf{g}(\mathbf{x}) = o(\|\mathbf{x} - \mathbf{x}_0\|)$, c.-à-d. $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\|\mathbf{g}(\mathbf{x})\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0$.

Si plus généralement $\|\mathbf{g}(\mathbf{x})\| = o(\|\mathbf{x} - \mathbf{x}_0\|^p)$ dans le sens déjà introduit, on notera simplement $\mathbf{g}(\mathbf{x}) = o(\|\mathbf{x} - \mathbf{x}_0\|^p)$. Si \mathbf{f} est différentiable en \mathbf{x}_0 , alors toutes les dérivées partielles et directionnelles existent, et \mathbf{f} est continue en \mathbf{x}_0 . De plus, l'application linéaire L est unique et donnée par

$$L(\mathbf{x} - \mathbf{x}_0) = \sum_{i=1}^m \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(\mathbf{x}_0)(x_j - x_{0j})\mathbf{e}_i$$

où \mathbf{e}_i est le $i^{\text{ème}}$ vecteur de la base canonique de \mathbb{R}^m . Ceci s'écrit aussi

$$L(\mathbf{x} - \mathbf{x}_0) = D\mathbf{f}(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) = D\mathbf{f}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0),$$

où intervient ici le produit matriciel entre une matrice $\mathbb{R}^{m \times n}$ et un vecteur colonne dans \mathbb{R}^n , ce qui donne un vecteur colonne dans \mathbb{R}^m . La dérivée de \mathbf{f} dans la direction \mathbf{v} au point \mathbf{x}_0 , $\mathbf{v} \mapsto D_{\mathbf{v}}\mathbf{f}(\mathbf{x}_0) = L(\mathbf{v}) = D\mathbf{f}(\mathbf{x}_0)\mathbf{v}$, est linéaire en $\mathbf{v} \in \mathbb{R}^n$ (si \mathbf{f} est différentiable en \mathbf{x}_0).

La condition de différentiabilité d'une fonction $f : E \subset \mathbb{R}^n \rightarrow \mathbb{R}$ (ou $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$) n'est pas immédiate à vérifier. Heureusement, on a une condition suffisante, facile à vérifier, qui nous permet de conclure si f est différentiable en $\mathbf{x}_0 \in \overset{\circ}{E}$. On donne ici la version du résultat pour une fonction à valeurs dans \mathbb{R} mais le résultat se généralise sans difficulté au cas d'une fonction à valeurs dans \mathbb{R}^m .

Théorème 3.12. *Soit $f : E \subset \mathbb{R}^n \rightarrow \mathbb{R}$ et $\mathbf{x}_0 \in \overset{\circ}{E}$. S'il existe $\delta > 0$ tel que $B(\mathbf{x}_0, \delta) \subset E$ et les dérivées partielles $\frac{\partial f}{\partial x_i}(\mathbf{x})$ existent pour tout $\mathbf{x} \in B(\mathbf{x}_0, \delta)$ et sont continues en \mathbf{x}_0 , alors f est différentiable en \mathbf{x}_0 .*

Démonstration. On utilise la norme euclidienne et, pour alléger la notation, on renomme $\mathbf{x}_0 = \mathbf{a} = (a_1, \dots, a_n)$. Pour un $\mathbf{x} = (x_1, \dots, x_n) \in B(\mathbf{a}, \delta)$ donné, on introduit la notation $\mathbf{x}^k = (x_1, \dots, x_k, a_{k+1}, \dots, a_n)$ de sorte que $\mathbf{x}^n = \mathbf{x}$ et $\mathbf{x}^0 = \mathbf{a}$. Alors, la différence $f(\mathbf{x}) - f(\mathbf{a})$ peut être écrite comme somme télescopique

$$f(\mathbf{x}) - f(\mathbf{a}) = \sum_{k=1}^n (f(\mathbf{x}^k) - f(\mathbf{x}^{k-1})).$$

Par le théorème des accroissements finis on a que $\forall k = 1, \dots, n$, il existe un $\theta_k \in]0, 1[$ tel que

$$\begin{aligned} f(\mathbf{x}^k) - f(\mathbf{x}^{k-1}) &= f(x_1, \dots, x_{k-1}, x_k, a_{k+1}, \dots, a_n) - f(x_1, \dots, x_{k-1}, a_k, a_{k+1}, \dots, a_n) \\ &= \frac{\partial f}{\partial x_k}(x_1, \dots, x_{k-1}, a_k + \theta_k(x_k - a_k), a_{k+1}, \dots, a_n)(x_k - a_k) \\ &= \frac{\partial f}{\partial x_k}(\mathbf{x}^{k-1} + \theta_k(\mathbf{x}^k - \mathbf{x}^{k-1}))(x_k - a_k). \end{aligned}$$

Puisque les dérivées partielles sont continues en \mathbf{a} , on a que

$$g_k(\mathbf{x}) := \frac{\partial f}{\partial x_k}(\mathbf{x}^{k-1} + \theta_k(\mathbf{x}^k - \mathbf{x}^{k-1})) - \frac{\partial f}{\partial x_k}(\mathbf{a}) \xrightarrow{\mathbf{x} \rightarrow \mathbf{a}} 0.$$

En effet, la continuité de $\frac{\partial f}{\partial x_k}$ en \mathbf{a} implique que

$$\forall \epsilon > 0 \exists \tilde{\delta} \in]0, \delta/2] : \forall \mathbf{y} \in B(\mathbf{a}, 2\tilde{\delta}) \quad \left| \frac{\partial f}{\partial x_k}(\mathbf{y}) - \frac{\partial f}{\partial x_k}(\mathbf{a}) \right| < \epsilon,$$

et de plus, pour $\mathbf{y}(\mathbf{x}) = \mathbf{x}^{k-1} + \theta_k(\mathbf{x}^k - \mathbf{x}^{k-1})$ avec $1 \leq k \leq n$, on a

$$\|\mathbf{y}(\mathbf{x}) - \mathbf{a}\| = \|(1 - \theta_k)(\mathbf{x}^{k-1} - \mathbf{a}) + \theta_k(\mathbf{x}^k - \mathbf{a})\| \leq \|\mathbf{x}^{k-1} - \mathbf{a}\| + \|\mathbf{x}^k - \mathbf{a}\| \leq 2\|\mathbf{x} - \mathbf{a}\|$$

et donc

$$\forall \epsilon > 0 \exists \tilde{\delta} \in]0, \delta/2] : \forall \mathbf{x} \in B(\mathbf{a}, \tilde{\delta}) \quad |g_k(\mathbf{x})| = \left| \frac{\partial f}{\partial x_k}(\mathbf{y}(\mathbf{x})) - \frac{\partial f}{\partial x_k}(\mathbf{a}) \right| < \epsilon,$$

ce qui montre que $\lim_{\mathbf{x} \rightarrow \mathbf{a}} g_k(\mathbf{x}) = 0$ pour tout $k = 1, \dots, n$. Ainsi,

$$\begin{aligned} f(\mathbf{x}) - f(\mathbf{a}) &= \sum_{k=1}^n (f(\mathbf{x}^k) - f(\mathbf{x}^{k-1})) \\ &= \sum_{k=1}^n \frac{\partial f}{\partial x_k}(\mathbf{x}^{k-1} + \theta_k(\mathbf{x}^k - \mathbf{x}^{k-1}))(x_k - a_k) \\ &= \sum_{k=1}^n \left(\frac{\partial f}{\partial x_k}(\mathbf{a})(x_k - a_k) + g_k(\mathbf{x})(x_k - a_k) \right) \\ &= Df(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) + g(\mathbf{x}) \end{aligned}$$

(produit matriciel entre un vecteur ligne et un vecteur colonne), avec $g(\mathbf{x}) = \sum_{k=1}^n g_k(\mathbf{x})(x_k - a_k)$. Par l'inégalité de Cauchy-Schwarz, on a $|g(\mathbf{x})| \leq (\sum_{k=1}^n g_k(\mathbf{x})^2)^{\frac{1}{2}} \|\mathbf{x} - \mathbf{a}\|$ et donc

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{|g(\mathbf{x})|}{\|\mathbf{x} - \mathbf{a}\|} \leq \lim_{\mathbf{x} \rightarrow \mathbf{a}} \sqrt{\sum_{k=1}^n g_k(\mathbf{x})^2} = 0$$

ce qui montre que $g(\mathbf{x}) = o(\|\mathbf{x} - \mathbf{a}\|)$. □

3.1.1 L'espace C^1

Le théorème 3.12 montre que si toutes les dérivées partielles existent dans un voisinage d'un point intérieur \mathbf{x}_0 et sont continues en \mathbf{x}_0 , alors la fonction est différentiable (et continue) en \mathbf{x}_0 et donc toutes les dérivées directionnelles existent en \mathbf{x}_0 .

Définition 3.13 (Espace C^1). Soit $E \subset \mathbb{R}^n$ ouvert non vide. On dit que $f : E \rightarrow \mathbb{R}$ est continûment différentiable sur E si toutes les dérivées partielles $\frac{\partial f}{\partial x_i}$ pour $i = 1, \dots, n$, existent et sont continues en tout point $\mathbf{x} \in E$. Dans ce cas on note $f \in C^1(E)$ (f est de classe C^1). De même, on dit que $\mathbf{f} : E \rightarrow \mathbb{R}^m$ est de classe C^1 , $\mathbf{f} \in C^1(E, \mathbb{R}^m)$, si chaque composante de $\mathbf{f} = (f_1, \dots, f_m)$ satisfait $f_k \in C^1(E)$, $k = 1, \dots, m$.

Grâce au théorème 3.12, si $f \in C^1(E)$ alors f est différentiable en tout point $\mathbf{x}_0 \in E$ (E étant ouvert) et donc elle est aussi continue, c'est-à-dire $f \in C^1(E) \Rightarrow f \in C^0(E)$; de même, on a $C^1(E, \mathbb{R}^m) \subset C^0(E, \mathbb{R}^m)$. De plus, toutes les dérivées directionnelles $D_{\mathbf{v}}f$ existent et sont continues sur E .

On vérifie facilement que $C^1(E)$ a une structure d'espace vectoriel réel. En effet, pour tous $f, g \in C^1(E)$ et toutes constantes $\lambda, \mu \in \mathbb{R}$ on a $\lambda f + \mu g \in C^1(E)$. On vérifie aussi facilement les règles de dérivation suivantes : Soit $E \subset \mathbb{R}^n$ ouvert, $f, g \in C^1(E)$ et des constantes $\lambda, \mu \in \mathbb{R}$. Alors

$$\begin{aligned} \text{i)} \quad & \lambda f + \mu g \in C^1(E), \quad D(\lambda f + \mu g)(\mathbf{x}) = \lambda Df(\mathbf{x}) + \mu Dg(\mathbf{x}), \quad \forall \mathbf{x} \in E, \\ \text{ii)} \quad & fg \in C^1(E), \quad D(fg)(\mathbf{x}) = Df(\mathbf{x})g(\mathbf{x}) + f(\mathbf{x})Dg(\mathbf{x}), \quad \forall \mathbf{x} \in E. \end{aligned}$$

3.2 Dérivation de fonctions composées

Soit $\mathbf{f} : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ et $\mathbf{g} : F \subset \mathbb{R}^m \rightarrow \mathbb{R}^p$ avec E, F ouverts non vides. Si $\text{Im}(\mathbf{f}) \subset F$, on peut définir la fonction composée

$$\varphi = \mathbf{g} \circ \mathbf{f} : E \rightarrow \mathbb{R}^p, \quad \varphi(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x})), \quad \forall \mathbf{x} \in E.$$

Par composantes :

$$\begin{aligned} \mathbf{x} &= (x_1, \dots, x_n), \quad \mathbf{f}(\mathbf{x}) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)), \\ \mathbf{y} &= (y_1, \dots, y_m), \quad \mathbf{g}(\mathbf{y}) = (g_1(y_1, \dots, y_m), \dots, g_p(y_1, \dots, y_m)), \\ \varphi(\mathbf{x}) &= (g_1(f_1(\mathbf{x}), \dots, f_m(\mathbf{x})), \dots, g_p(f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))). \end{aligned}$$

Théorème 3.14. Soient $\mathbf{f} : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbf{g} : F \subset \mathbb{R}^m \rightarrow \mathbb{R}^p$; E, F des ouverts contenant respectivement \mathbf{x}_0 et \mathbf{y}_0 ; $\text{Im}(\mathbf{f}) \subset F$ et $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$. Si \mathbf{f} est différentiable en \mathbf{x}_0 et \mathbf{g} est différentiable en \mathbf{y}_0 , alors $\varphi = \mathbf{g} \circ \mathbf{f} : E \rightarrow \mathbb{R}^p$ est différentiable en \mathbf{x}_0 et

$$D\varphi(\mathbf{x}_0) = D\mathbf{g}(\mathbf{f}(\mathbf{x}_0)) \cdot D\mathbf{f}(\mathbf{x}_0) = D\mathbf{g}(\mathbf{f}(\mathbf{x}_0))D\mathbf{f}(\mathbf{x}_0).$$

(produit matriciel des matrices jacobiniennes). Par composantes on a :

$$\forall i \in \{1, \dots, p\} \quad \forall j \in \{1, \dots, n\} \quad \frac{\partial \varphi_i}{\partial x_j}(\mathbf{x}_0) = \sum_{k=1}^m \frac{\partial g_i}{\partial y_k}(\mathbf{f}(\mathbf{x}_0)) \frac{\partial f_k}{\partial x_j}(\mathbf{x}_0).$$

Si de plus $\mathbf{f} \in C^1(E, \mathbb{R}^m)$ et $\mathbf{g} \in C^1(F, \mathbb{R}^p)$ alors $\varphi \in C^1(E, \mathbb{R}^p)$.

Démonstration. Par hypothèse on a

$$\begin{aligned} \mathbf{f}(\mathbf{x}) &= \mathbf{f}(\mathbf{x}_0) + D\mathbf{f}(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \mathbf{R}_{\mathbf{f}}(\mathbf{x}), \quad \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\mathbf{R}_{\mathbf{f}}(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|} = \mathbf{0}, \\ \mathbf{g}(\mathbf{y}) &= \mathbf{g}(\mathbf{y}_0) + D\mathbf{g}(\mathbf{y}_0) \cdot (\mathbf{y} - \mathbf{y}_0) + \mathbf{R}_{\mathbf{g}}(\mathbf{y}), \quad \lim_{\mathbf{y} \rightarrow \mathbf{y}_0} \frac{\mathbf{R}_{\mathbf{g}}(\mathbf{y})}{\|\mathbf{y} - \mathbf{y}_0\|} = \mathbf{0}. \end{aligned}$$

On écrit

$$\begin{aligned}\mathbf{R}_f(\mathbf{x}) &= \|\mathbf{x} - \mathbf{x}_0\| \mathbf{r}_f(\mathbf{x}) \quad \text{avec} \quad \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{r}_f(\mathbf{x}) = \mathbf{0} \text{ et } \mathbf{r}_f(\mathbf{x}_0) := \mathbf{0}, \\ \mathbf{R}_g(\mathbf{y}) &= \|\mathbf{y} - \mathbf{y}_0\| \mathbf{r}_g(\mathbf{y}) \quad \text{avec} \quad \lim_{\mathbf{y} \rightarrow \mathbf{y}_0} \mathbf{r}_g(\mathbf{y}) = \mathbf{0} \text{ et } \mathbf{r}_g(\mathbf{y}_0) := \mathbf{0}.\end{aligned}$$

On veut montrer qu'il existe une application linéaire $\mathbf{L}_\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^p$ et une fonction $\mathbf{R}_\varphi : E \rightarrow \mathbb{R}^p$ telles que

$$\varphi(\mathbf{x}) = \varphi(\mathbf{x}_0) + \mathbf{L}_\varphi(\mathbf{x} - \mathbf{x}_0) + \mathbf{R}_\varphi(\mathbf{x}) \quad \text{et} \quad \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\mathbf{R}_\varphi(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|} = \mathbf{0}.$$

On a :

$$\begin{aligned}\varphi(\mathbf{x}) &= \mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{g}\left(\mathbf{f}(\mathbf{x}_0) + \overbrace{D\mathbf{f}(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \|\mathbf{x} - \mathbf{x}_0\| \mathbf{r}_f(\mathbf{x})}^{\mathbf{h}(\mathbf{x})}\right) \\ &= \mathbf{g}(\mathbf{y}_0) + D\mathbf{g}(\mathbf{y}_0) \cdot \mathbf{h}(\mathbf{x}) + \|\mathbf{h}(\mathbf{x})\| \mathbf{r}_g(\mathbf{f}(\mathbf{x})) \\ &= \mathbf{g}(\mathbf{y}_0) + D\mathbf{g}(\mathbf{y}_0) \cdot D\mathbf{f}(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \underbrace{\|\mathbf{x} - \mathbf{x}_0\| D\mathbf{g}(\mathbf{y}_0) \cdot \mathbf{r}_f(\mathbf{x})}_{\mathbf{A}(\mathbf{x})} + \underbrace{\|\mathbf{h}(\mathbf{x})\| \mathbf{r}_g(\mathbf{f}(\mathbf{x}))}_{\mathbf{B}(\mathbf{x})}.\end{aligned}$$

Il faut montrer que $\mathbf{R}_\varphi(\mathbf{x}) := \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x})$ satisfait $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\mathbf{R}_\varphi(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|} = \mathbf{0}$. Pour $\mathbf{A}(\mathbf{x})$ on a

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\|\mathbf{A}(\mathbf{x})\|}{\|\mathbf{x} - \mathbf{x}_0\|} = \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \|D\mathbf{g}(\mathbf{y}_0) \cdot \mathbf{r}_f(\mathbf{x})\| \leq \|D\mathbf{g}(\mathbf{y}_0)\| \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \|\mathbf{r}_f(\mathbf{x})\| = 0,$$

où on a noté $\|C\|$ la norme matricielle $\|C\| = \sup_{\substack{\mathbf{y} \in \mathbb{R}^m \\ \mathbf{y} \neq \mathbf{0}}} \frac{\|C\mathbf{y}\|}{\|\mathbf{y}\|} < +\infty$ pour $C \in \mathbb{R}^{p \times m}$ (et on utilise finalement le théorème des deux gendarmes).

Pour $\mathbf{B}(\mathbf{x})$ on a pour tout $\mathbf{x}_0 \neq \mathbf{x} \in E$

$$\begin{aligned}\frac{\|\mathbf{B}(\mathbf{x})\|}{\|\mathbf{x} - \mathbf{x}_0\|} &= \frac{\|\mathbf{h}(\mathbf{x})\|}{\|\mathbf{x} - \mathbf{x}_0\|} \|\mathbf{r}_g(\mathbf{f}(\mathbf{x}))\| \\ &\leq \frac{\|D\mathbf{f}(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)\| + \|\mathbf{x} - \mathbf{x}_0\| \|\mathbf{r}_f(\mathbf{x})\|}{\|\mathbf{x} - \mathbf{x}_0\|} \|\mathbf{r}_g(\mathbf{f}(\mathbf{x}))\| \\ &\leq (\|D\mathbf{f}(\mathbf{x}_0)\| + \|\mathbf{r}_f(\mathbf{x})\|) \|\mathbf{r}_g(\mathbf{f}(\mathbf{x}))\|\end{aligned}$$

Puisque \mathbf{f} est continue en \mathbf{x}_0 et \mathbf{r}_g en $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$, la composée $\mathbf{r}_g \circ \mathbf{f}$ est continue en \mathbf{x}_0 . D'où $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{r}_g(\mathbf{f}(\mathbf{x})) = \mathbf{r}_g(\mathbf{f}(\mathbf{x}_0)) = \mathbf{r}_g(\mathbf{y}_0) = \mathbf{0}$. D'autre part $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{r}_f(\mathbf{x}) = \mathbf{0}$ (par hypothèse) et on conclut que

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\|\mathbf{B}(\mathbf{x})\|}{\|\mathbf{x} - \mathbf{x}_0\|} \leq (\|D\mathbf{f}(\mathbf{x}_0)\| + \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \|\mathbf{r}_f(\mathbf{x})\|) \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \|\mathbf{r}_g(\mathbf{f}(\mathbf{x}))\| = 0.$$

On a donc montré que

$$\varphi(\mathbf{x}) = \varphi(\mathbf{x}_0) + D\mathbf{g}(\mathbf{f}(\mathbf{x}_0)) \cdot D\mathbf{f}(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \mathbf{R}_\varphi(\mathbf{x}) \quad \text{où} \quad \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\mathbf{R}_\varphi(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|} = \mathbf{0}.$$

On conclut donc que φ est différentiable en \mathbf{x}_0 et

$$D\varphi(\mathbf{x}_0) = D\mathbf{g}(\mathbf{f}(\mathbf{x}_0)) \cdot D\mathbf{f}(\mathbf{x}_0).$$

Si maintenant $\mathbf{f} \in C^1(E, \mathbb{R}^m)$ et $\mathbf{g} \in C^1(F, \mathbb{R}^p)$, alors chaque composante des matrices $D\mathbf{f}(\mathbf{x})$ et $D\mathbf{g}(\mathbf{y})$ est une fonction continue de \mathbf{x} (resp. \mathbf{y}) et $D\varphi(\mathbf{x}) = D\mathbf{g}(\mathbf{f}(\mathbf{x})) \cdot D\mathbf{f}(\mathbf{x})$ est continue en \mathbf{x} pour tout $\mathbf{x} \in E$. Donc $\varphi \in C^1(E, \mathbb{R}^p)$. \square

Remarque 3.15. Dans cette preuve, on a utilisé la norme matricielle $\|C\| = \sup_{\substack{\mathbf{y} \in \mathbb{R}^m \\ \mathbf{y} \neq \mathbf{0}}} \frac{\|C\mathbf{y}\|}{\|\mathbf{y}\|} < +\infty$ pour $C \in \mathbb{R}^{p \times m}$. C'est effectivement une norme sur l'espace vectoriel des matrices $\mathbb{R}^{p \times m}$, appelée plus spécifiquement la norme "spectrale". Il y a d'autres normes matricielles possibles sur $\mathbb{R}^{p \times m}$.

Exemple 3.16. Soient $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ et $\mathbf{g} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ donnés par

$$\mathbf{f}(x, y) = \begin{bmatrix} xy \\ x + 2y \\ \sin x \end{bmatrix}, \quad \mathbf{g}(u, v, w) = \begin{bmatrix} u + w \\ v^2 \end{bmatrix},$$

avec matrices jacobiniennes

$$D\mathbf{f}(x, y) = \begin{bmatrix} y & x \\ 1 & 2 \\ \cos x & 0 \end{bmatrix} \in \mathbb{R}^{3 \times 2}, \quad D\mathbf{g}(u, v, w) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2v & 0 \end{bmatrix} \in \mathbb{R}^{2 \times 3}.$$

Alors,

$$\varphi(x, y) = \mathbf{g} \circ \mathbf{f}(x, y) = \begin{bmatrix} xy + \sin x \\ (x + 2y)^2 \end{bmatrix} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

et

$$D\varphi(x, y) = \begin{bmatrix} y + \cos x & x \\ 2(x + 2y) & 4(x + 2y) \end{bmatrix} = D\mathbf{g}(\mathbf{f}(x, y)) \cdot D\mathbf{f}(x, y) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2v & 0 \end{bmatrix} \Big|_{v=x+2y} \begin{bmatrix} y & x \\ 1 & 2 \\ \cos x & 0 \end{bmatrix}.$$

Un cas particulier du théorème 3.14 est celui de la composition d'une fonction $f : E \subset \mathbb{R}^n \rightarrow \mathbb{R}$, de classe C^1 sur l'ouvert non vide E et un chemin dans E , $\gamma : I \subset \mathbb{R} \rightarrow E$, I ouvert non vide et $t \mapsto \gamma(t) = (\gamma_1(t), \dots, \gamma_n(t))$, avec $\gamma_i \in C^1(I)$, $\forall i = 1, \dots, n$. Dans ce cas, la fonction composée $\varphi = f \circ \gamma : I \rightarrow \mathbb{R}$, $t \mapsto \varphi(t) = f(\gamma_1(t), \dots, \gamma_n(t))$ est dérivable sur I , $\varphi \in C^1(I)$ et

$$\begin{aligned} \frac{d}{dt}\varphi(t) &= Df(\gamma(t)) \cdot D\gamma(t) \\ &= \left[\frac{\partial f}{\partial x_1}(\gamma(t)), \dots, \frac{\partial f}{\partial x_n}(\gamma(t)) \right] \begin{bmatrix} \dot{\gamma}_1(t) \\ \vdots \\ \dot{\gamma}_n(t) \end{bmatrix} \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\gamma(t)) \dot{\gamma}_i(t). \end{aligned}$$

La quantité $\frac{d}{dt}\varphi(t) = \frac{d}{dt}f(\gamma_1(t), \dots, \gamma_n(t))$ est appelée la dérivée totale de f le long du chemin $\gamma(t)$.

3.3 Théorème des accroissements finis

On rappelle d'abord le théorème des accroissements finis pour une fonction réelle d'une seule variable réelle : soit $f : [a, b] \rightarrow \mathbb{R}$ continue et dérivable sur $]a, b[$, alors il existe $c \in]a, b[$ tel que $f(b) - f(a) = f'(c)(b - a)$. On vise à généraliser ce résultat pour des fonctions $f : E \subset \mathbb{R}^n \rightarrow \mathbb{R}$ ou $\mathbf{f} : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ de plusieurs variables réelles.

On introduit la notation suivante : pour $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{y}$, on note $[\mathbf{x}, \mathbf{y}]$ le segment fermé d'origine \mathbf{x} et d'extrémité \mathbf{y} : $[\mathbf{x}, \mathbf{y}] = \{\mathbf{z} \in \mathbb{R}^n : \mathbf{z} = \mathbf{x} + t(\mathbf{y} - \mathbf{x}), t \in [0, 1]\}$ et par $] \mathbf{x}, \mathbf{y} [$ le segment "ouvert" $] \mathbf{x}, \mathbf{y} [= \{\mathbf{z} \in \mathbb{R}^n : \mathbf{z} = \mathbf{x} + t(\mathbf{y} - \mathbf{x}), t \in]0, 1[\}$ (si $n \geq 2$, $] \mathbf{x}, \mathbf{y} [$ n'est ni ouvert ni fermé). Lorsque $\mathbf{x} = \mathbf{y}$, on pose $[\mathbf{x}, \mathbf{x}] = \{\mathbf{x}\}$.

Théorème 3.17. *Soit $E \subset \mathbb{R}^n$ un ouvert non vide et $f : E \rightarrow \mathbb{R}$ une fonction dérivable sur E . Si $\mathbf{x}, \mathbf{y} \in E$ distincts sont tels que $[\mathbf{x}, \mathbf{y}] \subset E$, alors il existe $\mathbf{z} \in] \mathbf{x}, \mathbf{y} [$, tel que*

$$f(\mathbf{y}) - f(\mathbf{x}) = Df(\mathbf{z})(\mathbf{y} - \mathbf{x})$$

(produit matriciel entre un vecteur ligne et un vecteur colonne).

Démonstration. On pose $g(t) = f(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))$ avec $t \in [0, 1]$. Alors g est continue sur $[0, 1]$ et dérivable sur $]0, 1[$, car composée de fonctions dérivables : $g(t) = f(\mathbf{v}(t))$, où $\mathbf{v}(t) = \mathbf{x} + t(\mathbf{y} - \mathbf{x})$. Par la règle de dérivation d'une composée, $g'(t) = Df(\mathbf{v}(t))\mathbf{v}'(t) = Df(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})$. Par le théorème des accroissements finis pour des fonctions d'une seule variable, il existe $\theta \in]0, 1[$ t.q. $g(1) - g(0) = g'(\theta)$ ce qui équivaut à

$$\begin{aligned} f(\mathbf{y}) - f(\mathbf{x}) &= Df(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x}), & \theta &\in]0, 1[\\ &= Df(\mathbf{z})(\mathbf{y} - \mathbf{x}), & \mathbf{z} &= \mathbf{x} + \theta(\mathbf{y} - \mathbf{x}) \in] \mathbf{x}, \mathbf{y} [. \end{aligned}$$

□

Dans la démonstration ci-dessus on a $g'(t) = Df(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})$. Si on suppose, maintenant, que f est de classe C^1 , on a $g \in C^1([0, 1])$ et $g(1) - g(0) = \int_0^1 g'(t)dt$. On en déduit

$$f(\mathbf{y}) - f(\mathbf{x}) = \int_0^1 Df(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})dt. \quad (3.1)$$

Considérons maintenant une fonction $\mathbf{f} : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ avec E ouvert non vide et \mathbf{f} dérivable sur E . L'analogue du théorème des accroissements finis pour une telle fonction vectorielle est *en défaut* ! En fait, on peut appliquer le théorème à chaque composante f_k , $k = 1, \dots, m$ de \mathbf{f} . Donc il existe $\mathbf{z}_k \in] \mathbf{x}, \mathbf{y} [$, $k = 1, \dots, m$ tels que

$$f_k(\mathbf{y}) - f_k(\mathbf{x}) = Df_k(\mathbf{z}_k)(\mathbf{y} - \mathbf{x}).$$

Toutefois, on n'a aucune garantie que les \mathbf{z}_k coïncident. On ne peut donc pas trouver, en général, un seul \mathbf{z} pour lequel $\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) = D\mathbf{f}(\mathbf{z})(\mathbf{y} - \mathbf{x})$ (produit matriciel entre une matrice $m \times n$ et un vecteur colonne dans \mathbb{R}^n , ce qui donne un vecteur colonne dans \mathbb{R}^m). Néanmoins, la formule (3.1) se généralise pour toute fonction vectorielle $\mathbf{f} : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ de classe C^1 .

Lemme 3.18. Soit $\mathbf{f} : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ de classe C^1 sur E ouvert non vide et $\mathbf{x}, \mathbf{y} \in E$ tels que $[\mathbf{x}, \mathbf{y}] \subset E$. Alors

$$\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) = \int_0^1 D\mathbf{f}(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})dt.$$

Ici l'intégrale d'une fonction continue $[0, 1] \rightarrow \mathbb{R}^m$ est comprise comme le vecteur dans \mathbb{R}^m obtenu en prenant l'intégrale de chaque composante (chaque composante étant une fonction continue $[0, 1] \rightarrow \mathbb{R}$).

Le Lemme précédent, demande que la fonction \mathbf{f} soit de classe C^1 sur E . On peut obtenir une *version faible* du théorème des accroissements finis qui demande uniquement que les dérivées partielles existent et soient bornées sur le segment $[\mathbf{x}, \mathbf{y}]$.

Lemme 3.19. Soit $\mathbf{f} : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ dérivable sur E ouvert non vide et $\mathbf{x}, \mathbf{y} \in E$ distincts tels que $[\mathbf{x}, \mathbf{y}] \subset E$. S'il existe $M > 0$ tel que $\|D\mathbf{f}(\mathbf{z})\| \leq M, \forall \mathbf{z} \in]\mathbf{x}, \mathbf{y}[$, alors

$$\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| \leq M\|\mathbf{y} - \mathbf{x}\|.$$

Démonstration. Le résultat est évident si $\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) = \mathbf{0}$. Supposons donc $\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) \neq \mathbf{0}$ et considérons $\mathbf{f}(\mathbf{u})$ comme un vecteur colonne, $\mathbf{u} \in E$. Soit un vecteur ligne $\mathbf{w} \in \mathbb{R}^m$ fixé et considérons la fonction $f_{\mathbf{w}} : E \rightarrow \mathbb{R}$ définie par $f_{\mathbf{w}}(\mathbf{u}) = \mathbf{w}\mathbf{f}(\mathbf{u})$ pour $\mathbf{u} \in E$, où apparaît le produit matriciel entre \mathbf{w} et $\mathbf{f}(\mathbf{u})$. On sait alors qu'il existe $\mathbf{z} \in]\mathbf{x}, \mathbf{y}[$ (qui peut dépendre de \mathbf{w}) tel que $f_{\mathbf{w}}(\mathbf{y}) - f_{\mathbf{w}}(\mathbf{x}) = Df_{\mathbf{w}}(\mathbf{z})(\mathbf{y} - \mathbf{x})$, c'est-à-dire

$$\mathbf{w}(\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})) = \mathbf{w}D\mathbf{f}(\mathbf{z})(\mathbf{y} - \mathbf{x}).$$

Choisissons $\mathbf{w} = (\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}))^\top$, ce qui donne

$$\begin{aligned} \|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\|^2 &= \mathbf{w}D\mathbf{f}(\mathbf{z})(\mathbf{y} - \mathbf{x}) \leq \|\mathbf{w}\| \|D\mathbf{f}(\mathbf{z})(\mathbf{y} - \mathbf{x})\| \\ &\leq \|\mathbf{w}\| \|D\mathbf{f}(\mathbf{z})\| \|\mathbf{y} - \mathbf{x}\| \leq \|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| M \|\mathbf{y} - \mathbf{x}\|, \end{aligned}$$

et donc $\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| \leq M\|\mathbf{y} - \mathbf{x}\|$. □

Chapitre 4

Dérivées d'ordres supérieurs

4.1 Dérivées secondes

Soit $E \subset \mathbb{R}^n$ ouvert, une fonction $f : E \rightarrow \mathbb{R}$ et un indice $j \in \{1, \dots, n\}$ fixé, tels que $\frac{\partial f}{\partial x_j}(\mathbf{x})$ existe en tout $\mathbf{x} \in E$. On considère maintenant la fonction $\frac{\partial f}{\partial x_j} : E \rightarrow \mathbb{R}$ et soit $k \in \{1, \dots, n\}$ fixé. Si cette fonction admet une dérivée partielle par rapport à x_k en tout $\mathbf{x} \in E$, on peut définir la dérivée partielle seconde $\frac{\partial \left(\frac{\partial f}{\partial x_j} \right)}{\partial x_k} : E \rightarrow \mathbb{R}$. On utilise la notation $\frac{\partial^2 f}{\partial x_k \partial x_j}(\mathbf{x}) := \frac{\partial \left(\frac{\partial f}{\partial x_j} \right)}{\partial x_k}(\mathbf{x})$. Pour $k = j$ on utilise la notation $\frac{\partial^2 f}{\partial x_j^2}(\mathbf{x}) = \frac{\partial \left(\frac{\partial f}{\partial x_j} \right)}{\partial x_j}(\mathbf{x})$.

Définition 4.1 (matrice hessienne). Soit $E \subset \mathbb{R}^n$ ouvert non vide et $f : E \rightarrow \mathbb{R}$ telle que toutes les dérivées partielles secondes $\frac{\partial^2 f}{\partial x_i \partial x_j} : E \rightarrow \mathbb{R}$, $i, j = 1, \dots, n$ existent. On appelle matrice hessienne de f en $\mathbf{x} \in E$ la matrice $H_f(\mathbf{x}) \in \mathbb{R}^{n \times n}$:

$$H_f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}) \\ \vdots & & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}) & \dots & \dots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{x}) \end{bmatrix}$$

$$= \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}) \\ \vdots & & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}) & \dots & \dots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{x}) \end{pmatrix}.$$

Définition 4.2 (espace $C^2(E)$). Soit $E \subset \mathbb{R}^n$ ouvert non vide. On dit que $f : E \rightarrow \mathbb{R}$ est de classe C^2 , $f \in C^2(E)$, si toutes les dérivées partielles secondes $\frac{\partial^2 f}{\partial x_i \partial x_j} : E \rightarrow \mathbb{R}$, $i, j = 1, \dots, n$ existent et sont continues sur E .

On peut aussi définir les dérivées directionnelles mixtes : soit $f : E \rightarrow \mathbb{R}$ différentiable sur E et $D_{\mathbf{v}}f(\mathbf{x}) = \nabla f(\mathbf{x})^\top \mathbf{v}$ la dérivée directionnelle de f dans la direction fixée $\mathbf{v} \in \mathbb{R}^n$, $\|\mathbf{v}\| = 1$ (norme euclidienne). Si $D_{\mathbf{v}}f : E \rightarrow \mathbb{R}$ admet une dérivée directionnelle dans la direction fixée \mathbf{w} en tout point $\mathbf{x} \in E$, on note

$$D_{\mathbf{w}\mathbf{v}}^2 f(\mathbf{x}) = D_{\mathbf{w}}(D_{\mathbf{v}}f)(\mathbf{x})$$

la dérivée directionnelle mixte.

Lemme 4.3. Soit $f \in C^2(E)$, $E \subset \mathbb{R}^n$ ouvert non vide et $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, $\|\mathbf{v}\| = \|\mathbf{w}\| = 1$. Alors

$$D_{\mathbf{w}\mathbf{v}}^2 f(\mathbf{x}) = \mathbf{w}^\top H_f(\mathbf{x}) \mathbf{v},$$

où il y a deux produits matriciels, et \mathbf{v} et \mathbf{w} sont des vecteurs colonnes. Ceci s'écrit aussi

$$D_{\mathbf{w}\mathbf{v}}^2 f(\mathbf{x}) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) v_j w_i.$$

Démonstration. Puisque $f \in C^2(E)$, toutes les dérivées partielles $\frac{\partial f}{\partial x_j}$, $j = 1, \dots, n$ sont continues sur E et f est différentiable sur E . Donc $D_{\mathbf{v}} f(\mathbf{x}) = Df(\mathbf{x}) \cdot \mathbf{v} = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{x}) v_j$ et $D_{\mathbf{v}} f : E \rightarrow \mathbb{R}$ est continue. De plus, ses dérivées partielles $\frac{\partial(D_{\mathbf{v}} f)}{\partial x_i}$, $i = 1, \dots, n$ existent et sont continues sur E car

$$\frac{\partial(D_{\mathbf{v}} f)}{\partial x_i} = \frac{\partial \left(\sum_{j=1}^n \frac{\partial f}{\partial x_j} v_j \right)}{\partial x_i} = \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} v_j$$

et $\frac{\partial^2 f}{\partial x_i \partial x_j} \in C^0(E)$ pour tout $i, j = 1, \dots, n$ puisque $f \in C^2(E)$. Donc $D_{\mathbf{v}} f \in C^1(E)$ et

$$D_{\mathbf{w}}(D_{\mathbf{v}} f)(\mathbf{x}) = \sum_{i=1}^n \frac{\partial(D_{\mathbf{v}} f)}{\partial x_i}(\mathbf{x}) w_i = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) v_j w_i = \mathbf{w}^\top H_f(\mathbf{x}) \mathbf{v}.$$

□

Exemple 4.4. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto f(x, y) = x^2 y$, et considérons les deux directions $\mathbf{v} = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ et $\mathbf{w} = (\frac{1}{2}, -\frac{\sqrt{3}}{2})$. On a

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= 2xy, & \frac{\partial f}{\partial y}(x, y) &= x^2, \\ \frac{\partial^2 f}{\partial x^2}(x, y) &= 2y, & \frac{\partial^2 f}{\partial y \partial x}(x, y) &= 2x, & \frac{\partial^2 f}{\partial x \partial y}(x, y) &= 2x, & \frac{\partial^2 f}{\partial y^2}(x, y) &= 0. \end{aligned}$$

Donc

$$Df(x, y) = (2xy, x^2), \quad H_f(x, y) = \begin{bmatrix} 2y & 2x \\ 2x & 0 \end{bmatrix}$$

et

$$D_{\mathbf{w}\mathbf{v}}^2 f(x, y) = \mathbf{w}^\top H_f(x, y) \mathbf{v} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 2y & 2x \\ 2x & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{\sqrt{2}}(x(1 - \sqrt{3}) + y).$$

On remarque de plus que $D_{\mathbf{w}\mathbf{v}}^2 f(x, y) = D_{\mathbf{v}\mathbf{w}}^2 f(x, y)$ puisque H_f est symétrique.

Dans l'exemple précédent on a $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$. Sous des conditions assez générales, on aura toujours $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$, $i, j = 1, \dots, n$. Ceci est garanti par l'important résultat suivant :

Théorème 4.5 (de Schwarz). *Soit deux indices fixés $i, j \in \{1, \dots, n\}$, $E \subset \mathbb{R}^n$ ouvert non vide avec $n \geq 2$, $\mathbf{x} \in E$ fixé et $f : E \rightarrow \mathbb{R}$ telle que $\frac{\partial f}{\partial x_i}, \frac{\partial f}{\partial x_j}, \frac{\partial^2 f}{\partial x_i \partial x_j}, \frac{\partial^2 f}{\partial x_j \partial x_i}$ existent dans E et $\frac{\partial^2 f}{\partial x_i \partial x_j}, \frac{\partial^2 f}{\partial x_j \partial x_i}$ sont continues en cet $\mathbf{x} \in E$. Alors $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x})$.*

Démonstration. Soient $s, t > 0$ fixés suffisamment petits de sorte que le rectangle rempli fermé de sommets $\mathbf{x}, \mathbf{x} + s\mathbf{e}_i, \mathbf{x} + t\mathbf{e}_j, \mathbf{x} + s\mathbf{e}_i + t\mathbf{e}_j$ est contenu dans E . Ceci est toujours possible car \mathbf{x} est un point intérieur de E . On considère la quantité

$$\Delta(s, t) = f(\mathbf{x} + s\mathbf{e}_i + t\mathbf{e}_j) - f(\mathbf{x} + s\mathbf{e}_i) - f(\mathbf{x} + t\mathbf{e}_j) + f(\mathbf{x})$$

et les deux fonctions auxiliaires

$$g(\xi) = f(\mathbf{x} + \xi\mathbf{e}_i + t\mathbf{e}_j) - f(\mathbf{x} + \xi\mathbf{e}_i), \quad \xi \in [0, s], \quad h(\xi) = f(\mathbf{x} + s\mathbf{e}_i + \xi\mathbf{e}_j) - f(\mathbf{x} + \xi\mathbf{e}_j), \quad \xi \in [0, t].$$

La fonction g est dérivable dans $]0, s[$ par hypothèse ($\frac{\partial f}{\partial x_i}$ existe dans E) et donc, par le théorème des accroissements finis, il existe $\tilde{s} \in]0, s[$ tel que

$$g(s) - g(0) = g'(\tilde{s})s = \left(\frac{\partial f}{\partial x_i}(\mathbf{x} + \tilde{s}\mathbf{e}_i + t\mathbf{e}_j) - \frac{\partial f}{\partial x_i}(\mathbf{x} + \tilde{s}\mathbf{e}_i) \right) s.$$

Soit maintenant

$$\varphi(\eta) = \frac{\partial f}{\partial x_i}(\mathbf{x} + \tilde{s}\mathbf{e}_i + \eta\mathbf{e}_j), \quad \eta \in [0, t].$$

La fonction φ est dérivable dans $]0, t[$ car $\frac{\partial^2 f}{\partial x_j \partial x_i}$ existe dans E , et donc il existe $\tilde{t} \in]0, t[$ tel que

$$\varphi(t) - \varphi(0) = \varphi'(\tilde{t})t = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x} + \tilde{s}\mathbf{e}_i + \tilde{t}\mathbf{e}_j)t.$$

On a finalement

$$\Delta(s, t) = g(s) - g(0) = g'(\tilde{s})s = (\varphi(t) - \varphi(0))s = \varphi'(\tilde{t})st = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x} + \tilde{s}\mathbf{e}_i + \tilde{t}\mathbf{e}_j)st.$$

On répète maintenant les mêmes arguments pour la fonction h :

$$\exists \hat{t} \in]0, t[: \quad \Delta(s, t) = h(t) - h(0) = h'(\hat{t})t = \left(\frac{\partial f}{\partial x_j}(\mathbf{x} + s\mathbf{e}_i + \hat{t}\mathbf{e}_j) - \frac{\partial f}{\partial x_j}(\mathbf{x} + \hat{t}\mathbf{e}_j) \right) t$$

Soit $\psi(\eta) = \frac{\partial f}{\partial x_j}(\mathbf{x} + \eta\mathbf{e}_i + \hat{t}\mathbf{e}_j)$, $\eta \in [0, s]$, alors

$$\exists \hat{s} \in]0, s[: \quad \psi(s) - \psi(0) = \frac{\partial f}{\partial x_j}(\mathbf{x} + s\mathbf{e}_i + \hat{t}\mathbf{e}_j) - \frac{\partial f}{\partial x_j}(\mathbf{x} + \hat{t}\mathbf{e}_j) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x} + \hat{s}\mathbf{e}_i + \hat{t}\mathbf{e}_j)s$$

et donc

$$\Delta(s, t) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x} + \hat{s}\mathbf{e}_i + \hat{t}\mathbf{e}_j)st = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x} + \tilde{s}\mathbf{e}_i + \tilde{t}\mathbf{e}_j)st.$$

Si on prend maintenant $s = t \rightarrow 0^+$, on obtient $\hat{s}, \tilde{s} \rightarrow 0^+$ et $\hat{t}, \tilde{t} \rightarrow 0^+$. Puisque $\frac{\partial^2 f}{\partial x_i \partial x_j}$ et $\frac{\partial^2 f}{\partial x_j \partial x_i}$ sont continues en \mathbf{x} , on a

$$\lim_{s \rightarrow 0^+} \frac{1}{s^2} \Delta(s, s) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}).$$

□

Corollaire 4.6. *Soit $E \subset \mathbb{R}^n$ ouvert non vide avec $n \geq 2$ et $f \in C^2(E)$. Alors*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}), \quad \forall \mathbf{x} \in E, \quad \forall i, j = 1, \dots, n.$$

Le résultat du théorème de Schwarz n'est pas forcément vrai sans l'hypothèse de continuité des dérivées partielles secondes, comme l'exemple suivant le montre.

Exercice 4.7. *Considérons la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par*

$$f(x, y) = \begin{cases} \frac{xy^3}{x^2+y^2}, & (x, y) \neq (0, 0), \\ 0, & (x, y) = (0, 0). \end{cases}$$

On vérifie que $\frac{\partial^2 f}{\partial x \partial y}(0, 0)$, $\frac{\partial^2 f}{\partial y \partial x}(0, 0)$ existent mais elles ne sont pas égales.

Grâce au théorème de Schwarz on a que si $f : E \rightarrow \mathbb{R}$ admet toutes les dérivées partielles secondes $\frac{\partial^2 f}{\partial x_i \partial x_j} : E \rightarrow \mathbb{R}$, $i, j = 1, \dots, n$ et qu'elles sont continues en $\mathbf{x} \in E$, alors $H_f(\mathbf{x})$ est une matrice symétrique. Donc, en fait, il suffit de calculer seulement $\frac{n(n+1)}{2}$ dérivées partielles secondes au lieu de n^2 . Si de plus f est de classe $C^2(E)$, il s'ensuit aussi que les dérivées directionnelles secondes $D_{\mathbf{v}\mathbf{w}}f(\mathbf{x})$ existent pour tout $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ de norme 1 et que $D_{\mathbf{v}\mathbf{w}}f(\mathbf{x}) = D_{\mathbf{w}\mathbf{v}}f(\mathbf{x})$.

4.2 Dérivées partielles d'ordres supérieurs à 2

Soit $f : E \rightarrow \mathbb{R}$, avec $E \subset \mathbb{R}^n$ ouvert non vide, et $(i_1, \dots, i_p) \in \{1, \dots, n\}^p$. On généralise facilement la notion de dérivée partielle d'ordre p de f par rapport aux variables x_{i_1}, \dots, x_{i_p} :

$$\frac{\partial^p f}{\partial x_{i_p} \partial x_{i_{p-1}} \dots \partial x_{i_1}}(\mathbf{x}) = \frac{\partial \left(\frac{\partial \left(\dots \left(\frac{\partial f}{\partial x_{i_1}} \right) \dots \right)}{\partial x_{i_{p-1}}} \right)}{\partial x_{i_p}}(\mathbf{x}).$$

Définition 4.8 (espace $C^p(E)$). *Soit $E \subset \mathbb{R}^n$ ouvert non vide et $f : E \rightarrow \mathbb{R}$. On dit que f est de classe C^p , $f \in C^p(E)$, si toutes les dérivées partielles d'ordre p , $\frac{\partial^p f}{\partial x_{i_1} \dots \partial x_{i_p}} : E \rightarrow \mathbb{R}$, existent pour tout $(i_1, \dots, i_p) \in \{1, \dots, n\}^p$ et sont continues sur E .*

Grâce au théorème de Schwarz, si $f \in C^p(E)$, alors

$$\frac{\partial^p f}{\partial x_{i_1} \cdots \partial x_{i_p}} = \frac{\partial^p f}{\partial x_{i_{\sigma(1)}} \cdots \partial x_{i_{\sigma(p)}}}$$

où $\sigma(1), \dots, \sigma(p)$ est une permutation arbitraire des indices $1, \dots, p$. Autrement dit, l'ordre de dérivation n'est pas importante si $f \in C^p(E)$.

On remarque que si $f \in C^p(E)$, alors $f \in C^q(E)$, $\forall 0 \leq q \leq p$, c.-à-d., $C^p(E) \subset C^{p-1}(E) \subset \cdots \subset C^1(E) \subset C^0(E)$.

Notation par multi-entiers. Considérons une dérivée partielle d'ordre p , $\frac{\partial^p f}{\partial x_{i_1} \cdots \partial x_{i_p}}$. Soit $\alpha = (\alpha_1, \dots, \alpha_n)$ où l'entier $\alpha_j \geq 0$ est le nombre de fois que l'indice $j \in \{1, \dots, n\}$ apparaît dans la suite i_1, \dots, i_p . On note $|\alpha| = \sum_{j=1}^n \alpha_j$ (dans ce cas $|\alpha| = p$). On utilise souvent la notation

$$\frac{\partial^{|\alpha|} f}{\partial \mathbf{x}^\alpha} = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} = \frac{\partial^p f}{\partial x_{i_1} \cdots \partial x_{i_p}}$$

où $\partial x_j^{\alpha_j}$ signifie qu'on dérive α_j fois par rapport à x_j . Cette notation ne distingue pas l'ordre de dérivation.

Exemple 4.9. $f(\mathbf{x}) = x_1 x_2^2 x_3^3$. Calculons $\frac{\partial^{|\alpha|} f}{\partial \mathbf{x}^\alpha}$ pour $\alpha = (1, 2, 0)$ et $\alpha = (1, 1, 1)$.

$$\alpha = (1, 2, 0) \Rightarrow \frac{\partial^3 f}{\partial x_1 \partial x_2^2} = \frac{\partial^3 f}{\partial x_1 \partial x_2 \partial x_2} = \frac{\partial^3 f}{\partial x_2 \partial x_1 \partial x_2} = \frac{\partial^3 f}{\partial x_2 \partial x_2 \partial x_1} = 2x_3^3$$

$$\begin{aligned} \alpha = (1, 1, 1) \Rightarrow \frac{\partial^3 f}{\partial x_1 \partial x_2 \partial x_3} &= \frac{\partial^3 f}{\partial x_3 \partial x_1 \partial x_2} = \frac{\partial^3 f}{\partial x_2 \partial x_3 \partial x_1} = \frac{\partial^3 f}{\partial x_1 \partial x_3 \partial x_2} \\ &= \frac{\partial^3 f}{\partial x_2 \partial x_1 \partial x_3} = \frac{\partial^3 f}{\partial x_3 \partial x_2 \partial x_1} = 6x_2 x_3^2 \end{aligned}$$

Remarquons que le nombre de dérivées partielles d'ordre $|\alpha|$ qui correspondent à un multi-entier $\alpha \in \mathbb{N}^n$, est

$$\binom{|\alpha|}{\alpha} = \frac{|\alpha|!}{\alpha_1! \cdots \alpha_n!}.$$

4.3 Développement limité et Formule de Taylor

Rappelons d'abord la formule de Taylor pour une fonction d'une seule variable réelle définie sur un intervalle ouvert $I =]a, b[\subset \mathbb{R}$ et p fois continûment dérivable sur I , $f \in C^p(I)$: $\forall x, y \in I$

$$\begin{aligned} f(y) &= f(x) + f'(x)(y-x) + \frac{f''(x)}{2}(y-x)^2 + \cdots + \frac{f^{(p)}(x)}{p!}(y-x)^p + R_p(y) \\ &= \underbrace{\sum_{k=0}^p \frac{f^{(k)}(x)}{k!}(y-x)^k}_{=T_x^p(y)} + R_p(y) \end{aligned}$$

où $T_x^p(y) = \sum_{k=0}^p \frac{f^{(k)}(x)}{k!} (y-x)^k$ est le polynôme de Taylor de degré p et $R_p(y)$ est le reste de la formule de Taylor t.q. $\lim_{y \rightarrow x} \frac{R_p(y)}{|y-x|^p} = 0$ (ou avec notation “petit o ”, $R_p(y) = o(|y-x|^p)$). La formule de Taylor donne donc le développement limité d'ordre p de la fonction f autour de x . Si, de plus, la dérivée d'ordre $p+1$ existe sur I et est continue, alors le reste de la formule peut être caractérisé par

$$\text{Reste de Lagrange :} \quad R_p(y) = \frac{f^{(p+1)}(x + \theta(y-x))}{(p+1)!} (y-x)^{p+1} \quad \text{pour un } \theta \in]0, 1[,$$

$$\text{Reste (sous forme d') intégrale :} \quad R_p(y) = \int_0^1 (1-s)^p \frac{f^{(p+1)}(x + s(y-x))}{p!} (y-x)^{p+1} ds.$$

Considérons maintenant le **cas** $n \geq 2$. Pour un multi-entier $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, on utilise les notations suivantes :

$$\alpha! = \prod_{i=1}^n \alpha_i!, \quad \mathbf{x}^\alpha = \prod_{i=1}^n x_i^{\alpha_i}.$$

Un polynôme de degré p dans les variables $\mathbf{x} = (x_1, \dots, x_n)$ s'écrit

$$q(\mathbf{x}) = \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| \leq p}} c_\alpha \mathbf{x}^\alpha, \quad c_\alpha \in \mathbb{R}.$$

Définition 4.10. Soit $f : E \rightarrow \mathbb{R}$ et $\mathbf{x} \in \mathring{E}$. S'il existe un polynôme $\mathbf{y} = (y_1, \dots, y_n) \mapsto q(\mathbf{y}) = \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| \leq p}} c_\alpha (\mathbf{y} - \mathbf{x})^\alpha$ de degré p en \mathbf{y} et une fonction $R_p : E \rightarrow \mathbb{R}$ tels que

$$f(\mathbf{y}) = q(\mathbf{y}) + R_p(\mathbf{y}), \quad \forall \mathbf{y} \in E \quad (4.1)$$

et $\lim_{\mathbf{y} \rightarrow \mathbf{x}} \frac{R_p(\mathbf{y})}{\|\mathbf{y} - \mathbf{x}\|^p} = 0$, alors on dit que (4.1) est un développement limité d'ordre p de f autour de \mathbf{x} .

Observons que $R_p(\mathbf{y})$ est défini à partir de $f(\mathbf{y})$ et $q(\mathbf{y})$ pour tout $\mathbf{y} \in E$ par $R_p(\mathbf{y}) = f(\mathbf{y}) - q(\mathbf{y})$ mais la propriété qui le caractérise est $\lim_{\mathbf{y} \rightarrow \mathbf{x}} \frac{R_p(\mathbf{y})}{\|\mathbf{y} - \mathbf{x}\|^p} = 0$; nous verrons, sous certaines conditions, des formules plus explicites pour $R_p(\mathbf{y})$ si \mathbf{y} est suffisamment proche de \mathbf{x} . Comme pour les fonctions d'une seule variable réelle, si un développement limité d'ordre p de f autour d'un point existe, alors il est unique et peut être construit en utilisant la formule de Taylor. On va détailler sa dérivation pour une fonction $f \in C^{p+1}(E)$, avec $E \subset \mathbb{R}^n$ ouvert non vide.

Soit $\mathbf{x}, \mathbf{y} \in E$ tels que le segment $[\mathbf{x}, \mathbf{y}] = \{\mathbf{z} = \mathbf{x} + t(\mathbf{y} - \mathbf{x}), t \in [0, 1]\} \subset E$. On note $g(t) = f(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))$, qui est bien définie et de classe C^{p+1} sur un intervalle ouvert $I =]-\delta, 1 + \delta[$ qui contient $[0, 1]$, grâce au fait que $\mathbf{x}, \mathbf{y} \in E$ sont des points intérieurs. L'idée pour dériver une formule de Taylor pour f est d'utiliser la formule de Taylor pour $g : I \rightarrow \mathbb{R}$:

$$g(t) = g(0) + g'(0)t + \frac{g''(0)}{2}t^2 + \dots + \frac{g^{(p)}(0)}{p!}t^p + R_p^g(t), \quad t \in [0, 1]$$

où

$$R_p^g(t) = \int_0^1 \frac{(1-s)^p}{p!} g^{(p+1)}(st) t^{p+1} ds = \frac{g^{(p+1)}(\theta)}{(p+1)!} t^{p+1} \quad \text{pour un } \theta \in]0, t[.$$

Or, si on note $\mathbf{x}_t = \mathbf{x} + t(\mathbf{y} - \mathbf{x})$:

$$\begin{aligned} g(t) &= f(\mathbf{x}_t) \\ g'(t) &= \sum_{i_1=1}^n \frac{\partial f}{\partial x_{i_1}}(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))(y_{i_1} - x_{i_1}) = \sum_{|\alpha|=1} \frac{\partial^{|\alpha|} f}{\partial \mathbf{x}^\alpha}(\mathbf{x}_t)(\mathbf{y} - \mathbf{x})^\alpha \\ g''(t) &= \sum_{i_2=1}^n \frac{\partial}{\partial x_{i_2}} \left(\sum_{i_1=1}^n \frac{\partial f}{\partial x_{i_1}}(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))(y_{i_1} - x_{i_1}) \right) (y_{i_2} - x_{i_2}) \\ &= \sum_{i_1, i_2=1}^n \frac{\partial^2 f}{\partial x_{i_1} \partial x_{i_2}}(\mathbf{x}_t)(y_{i_1} - x_{i_1})(y_{i_2} - x_{i_2}) \\ &= \sum_{|\alpha|=2} \frac{2!}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial \mathbf{x}^\alpha}(\mathbf{x}_t)(\mathbf{y} - \mathbf{x})^\alpha \\ &\vdots \\ g^{(p)}(t) &= \sum_{i_1, \dots, i_p=1}^n \frac{\partial^p f}{\partial x_{i_1} \dots \partial x_{i_p}}(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))(y_{i_1} - x_{i_1}) \dots (y_{i_p} - x_{i_p}) \\ &= \sum_{|\alpha|=p} \frac{p!}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial \mathbf{x}^\alpha}(\mathbf{x}_t)(\mathbf{y} - \mathbf{x})^\alpha. \end{aligned}$$

Finalement, on peut écrire la formule de Taylor

$$\begin{aligned} f(\mathbf{y}) &= g(1) = \sum_{k=0}^p \frac{1}{k!} g^{(k)}(0) + R_p^g(1) \\ &= \sum_{k=0}^p \frac{1}{k!} \sum_{|\alpha|=k} \frac{k!}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial \mathbf{x}^\alpha}(\mathbf{x})(\mathbf{y} - \mathbf{x})^\alpha + R_p(\mathbf{y}) \\ &= \underbrace{\sum_{|\alpha| \leq p} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial \mathbf{x}^\alpha}(\mathbf{x})(\mathbf{y} - \mathbf{x})^\alpha}_{=T_{\mathbf{x}}^p(\mathbf{y})} + R_p(\mathbf{y}), \end{aligned}$$

où on identifie :

$$T_{\mathbf{x}}^p(\mathbf{y}) = \sum_{|\alpha| \leq p} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial \mathbf{x}^\alpha}(\mathbf{x})(\mathbf{y} - \mathbf{x})^\alpha \quad (\text{polynôme de Taylor de degré } p)$$

et

$$\begin{aligned} R_p(\mathbf{y}) &= \sum_{|\alpha|=p+1} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial \mathbf{x}^\alpha}(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})^\alpha \quad \text{pour un } \theta \in]0, 1[\quad (\text{reste de Lagrange}), \\ &= \sum_{|\alpha|=p+1} \frac{(p+1)}{\alpha!} \int_0^1 (1-s)^p \frac{\partial^{|\alpha|} f}{\partial \mathbf{x}^\alpha}(\mathbf{x} + s(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})^\alpha ds \quad (\text{reste intégrale}). \end{aligned}$$

4.3.1 Formule de Taylor pour des fonctions vectorielles

Soit $E \subset \mathbb{R}^n$ ouvert non vide, $\mathbf{f} \in C^{p+1}(E, \mathbb{R}^m)$ et $\mathbf{x}, \mathbf{y} \in E$ t.q. $[\mathbf{x}, \mathbf{y}] \subset E$. On peut appliquer la formule de Taylor à chaque composante f_k de $\mathbf{f} = (f_1, \dots, f_m)$: il existe $\theta_1, \theta_2, \dots, \theta_m \in]0, 1[$ t.q.

$$f_k(\mathbf{y}) = \sum_{|\alpha| \leq p} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f_k}{\partial \mathbf{x}^\alpha}(\mathbf{x})(\mathbf{y} - \mathbf{x})^\alpha + \sum_{|\alpha|=p+1} \frac{1}{\alpha!} \frac{\partial^{(p+1)} f_k}{\partial \mathbf{x}^\alpha}(\mathbf{x} + \theta_k(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})^\alpha.$$

Toutefois, en général, les $\theta_1, \dots, \theta_m$ ne sont pas égaux entre eux et on ne peut pas trouver un seul $\theta \in]0, 1[$ tel que $\mathbf{R}_p(\mathbf{y}) = \sum_{|\alpha|=p+1} \frac{1}{\alpha!} \frac{\partial^{(p+1)} \mathbf{f}}{\partial \mathbf{x}^\alpha}(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})^\alpha$. D'un autre côté, la formule de Taylor avec reste intégrale est valable aussi pour des fonctions vectorielles :

$$\mathbf{f}(\mathbf{y}) = \sum_{|\alpha| \leq p} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} \mathbf{f}}{\partial \mathbf{x}^\alpha}(\mathbf{x})(\mathbf{y} - \mathbf{x})^\alpha + \sum_{|\alpha|=p+1} \frac{p+1}{\alpha!} \int_0^1 (1-s)^p \frac{\partial^{(p+1)} \mathbf{f}}{\partial \mathbf{x}^\alpha}(\mathbf{x} + s(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})^\alpha ds$$

et fournit le développement limité d'ordre p de \mathbf{f} autour de \mathbf{x} .

Chapitre 5

Intégrales qui dépendent de paramètres

Soit $I \subset \mathbb{R}$ un intervalle avec une infinité de points et $E \subset \mathbb{R}$ un ensemble non vide. On considère une fonction $f : I \times E \rightarrow \mathbb{R}$, $(t, \mathbf{x}) \mapsto f(t, \mathbf{x})$ telle que, pour tout $\mathbf{x} \in E$, l'intégrale (éventuellement généralisée)

$$g(\mathbf{x}) = \int_I f(t, \mathbf{x}) dt$$

existe. On se pose la question si certaines propriétés de la fonction $g : E \rightarrow \mathbb{R}$ comme, par exemple, la continuité ou la dérivabilité, peuvent se déduire de celles de f et si on peut passer les opérations de limite et dérivation sous le signe de l'intégrale :

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x}) \stackrel{?}{=} \int_I \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(t, \mathbf{x}) dt, \quad \frac{\partial g}{\partial x_i}(\mathbf{x}) \stackrel{?}{=} \int_I \frac{\partial f}{\partial x_i}(t, \mathbf{x}) dt.$$

Que ceci ne soit pas toujours le cas est illustré par l'exemple suivant.

Exemple 5.1. Soit $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, $f(t, x) = x^2 e^{-x^2 t}$. On considère l'intégrale généralisée

$$g(x) = \int_0^\infty f(t, x) dt = \int_0^\infty x^2 e^{-x^2 t} dt, \quad \forall x \in \mathbb{R}.$$

Clairement, f est continue sur $\mathbb{R} \times \mathbb{R}$ et l'intégrale généralisée $g(x)$ existe pour tout $x \in \mathbb{R}$. Toutefois, par calcul direct, on trouve

$$g(x) = \begin{cases} 1, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

ce qui montre que la fonction g n'est pas continue sur \mathbb{R} et $\lim_{x \rightarrow 0} \int_0^\infty f(t, x) dt \neq \int_0^\infty f(t, 0) dt$.

Exemple 5.2. Soit $f :]0, 1[\times [0, 1] \rightarrow \mathbb{R}$,

$$f(t, x) = \begin{cases} \frac{1}{x} e^{-\frac{t}{x}} & t \in]0, 1[, x \in]0, 1] \\ 0 & t \in]0, 1[, x = 0. \end{cases}$$

On vérifie que f est continue sur $]0, 1[\times [0, 1]$. L'intégrale généralisée $g(x) = \int_0^1 f(t, x) dt$ existe pour tout $x \in [0, 1]$ et vaut

$$g(x) = \begin{cases} 1 - e^{-\frac{1}{x}}, & x \in]0, 1] \\ 0, & x = 0. \end{cases}$$

On a donc $\lim_{x \rightarrow 0^+} \int_0^1 f(t, x) dt = \lim_{x \rightarrow 0^+} g(x) = 1 \neq g(0) = \int_0^1 \lim_{x \rightarrow 0} f(t, x) dt$.

5.1 Intégrales sur un intervalle fermé et borné

Le fait que dans l'exemple 5.2 la continuité de la fonction f n'implique pas la continuité de la fonction g est dû à un manque d'uniformité dans la continuité de f .

On va d'abord se restreindre au cas où le domaine d'intégration est un intervalle fermé et borné (compact).

Théorème 5.3. Soient $a, b \in \mathbb{R}$, $a < b$ et $E \subset \mathbb{R}^n$ non vide. Si $f : [a, b] \times E \rightarrow \mathbb{R}$ est continue, alors la fonction $g(\mathbf{x}) = \int_a^b f(t, \mathbf{x}) dt$ est bien définie $\forall \mathbf{x} \in E$ et continue sur E . En particulier, pour tout $\mathbf{x}_0 \in E$ qui n'est pas isolé :

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x}) = g(\mathbf{x}_0) = \int_a^b f(t, \mathbf{x}_0) dt = \int_a^b \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(t, \mathbf{x}) dt.$$

Démonstration. Fixons d'abord \mathbf{x}_0 dans E et ensuite $\epsilon > 0$. D'après le théorème de Cantor-Heine généralisé 2.28 appliqué au sous-ensemble compact $K = [a, b] \times \{\mathbf{x}_0\} \subset [a, b] \times E$, il existe $\delta > 0$ tel que,

$$\forall t_0 \in [a, b] \quad \forall (t, \mathbf{x}) \in [a, b] \times E \quad \left(|t - t_0| + \|\mathbf{x} - \mathbf{x}_0\| \leq \delta \Rightarrow |f(t, \mathbf{x}) - f(t_0, \mathbf{x}_0)| \leq \frac{\epsilon}{b - a} \right).$$

En choisissant $t = t_0 \in [a, b]$, ceci donne

$$\forall (t, \mathbf{x}) \in [a, b] \times E \quad \left(\|\mathbf{x} - \mathbf{x}_0\| \leq \delta \Rightarrow |f(t, \mathbf{x}) - f(t, \mathbf{x}_0)| \leq \frac{\epsilon}{b - a} \right).$$

Pour tout $\mathbf{x} \in E$, la fonction $t \mapsto f(t, \mathbf{x})$ est continue sur l'intervalle fermé $[a, b]$, donc l'intégrale $g(\mathbf{x}) = \int_a^b f(t, \mathbf{x}) dt$ existe. De plus, pour tout $\mathbf{x} \in \overline{B}(\mathbf{x}_0, \delta) \cap E$,

$$\begin{aligned} |g(\mathbf{x}) - g(\mathbf{x}_0)| &= \left| \int_a^b (f(t, \mathbf{x}) - f(t, \mathbf{x}_0)) dt \right| \\ &\leq \int_a^b \frac{\epsilon}{b - a} dt = \epsilon, \end{aligned}$$

ce qui prouve la continuité de g en $\mathbf{x}_0 \in E$. Puisque \mathbf{x}_0 est arbitraire, g est continue sur E . \square

Exercice 5.4. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto f(x, y) = \frac{e^{x^2 y}}{x^2 + y^2 + 1}$, et $g(y) = \int_{-1}^1 f(x, y) dx$. Calculer $\lim_{y \rightarrow 0} g(y)$.

On considère maintenant la dérivabilité de la fonction $g(\mathbf{x}) = \int_a^b f(t, \mathbf{x}) dt$.

Théorème 5.5. Soient $a, b \in \mathbb{R}$, $a < b$, $E \subset \mathbb{R}^n$ ouvert non vide, $f : [a, b] \times E \rightarrow \mathbb{R}$ continue et $g(\mathbf{x}) = \int_a^b f(t, \mathbf{x}) dt$, $\forall \mathbf{x} \in E$. Soit encore $i \in \{1, \dots, n\}$ fixé. Si $\frac{\partial f}{\partial x_i} : [a, b] \times E \rightarrow \mathbb{R}$ existe et est continue, alors $\frac{\partial g}{\partial x_i}$ existe pour tout $\mathbf{x} \in E$ et $\frac{\partial g}{\partial x_i}(\mathbf{x}) = \int_a^b \frac{\partial f}{\partial x_i}(t, \mathbf{x}) dt$ est continue en tout $\mathbf{x} \in E$.

Démonstration. Comme dans la démonstration du Théorème 5.3, soit $\mathbf{x}_0 \in E$ arbitraire, $\epsilon > 0$ et $K = [a, b] \times \{\mathbf{x}_0\}$. Alors, il existe $\delta > 0$ tel que

$$\forall (t, \mathbf{x}) \in [a, b] \times E \quad \left(\|\mathbf{x} - \mathbf{x}_0\| \leq \delta \Rightarrow \left| \frac{\partial f}{\partial x_i}(t, \mathbf{x}) - \frac{\partial f}{\partial x_i}(t, \mathbf{x}_0) \right| \leq \frac{\epsilon}{b-a} \right).$$

Quitte à diminuer $\delta > 0$, on peut encore supposer que $\overline{B}(\mathbf{x}_0, \delta) \subset E$, car E est ouvert. On a alors pour tout $0 < |s| \leq \delta$,

$$\begin{aligned} \left| \frac{g(\mathbf{x}_0 + s\mathbf{e}_i) - g(\mathbf{x}_0)}{s} - \int_a^b \frac{\partial f}{\partial x_i}(t, \mathbf{x}_0) dt \right| &= \left| \int_a^b \left(\frac{f(t, \mathbf{x}_0 + s\mathbf{e}_i) - f(t, \mathbf{x}_0)}{s} - \frac{\partial f}{\partial x_i}(t, \mathbf{x}_0) \right) dt \right| \\ &= \left| \int_a^b \frac{1}{s} \left(\int_0^s \left(\frac{\partial f}{\partial x_i}(t, \mathbf{x}_0 + \sigma\mathbf{e}_i) - \frac{\partial f}{\partial x_i}(t, \mathbf{x}_0) \right) d\sigma \right) dt \right| \\ &\leq \int_a^b \frac{1}{|s|} \left| \int_0^s \left(\frac{\partial f}{\partial x_i}(t, \mathbf{x}_0 + \sigma\mathbf{e}_i) - \frac{\partial f}{\partial x_i}(t, \mathbf{x}_0) \right) d\sigma \right| dt \leq \epsilon, \end{aligned}$$

ce qui montre que $\frac{\partial g}{\partial x_i}(\mathbf{x}_0) = \lim_{s \rightarrow 0} \frac{g(\mathbf{x}_0 + s\mathbf{e}_i) - g(\mathbf{x}_0)}{s}$ existe et vaut $\int_a^b \frac{\partial f}{\partial x_i}(t, \mathbf{x}_0) dt$. Puisque $\frac{\partial f}{\partial x_i}$ est continue sur $[a, b] \times E$, $\frac{\partial g}{\partial x_i}$ est continue sur E . \square

5.2 Intégrales avec des bornes variables

Considérons maintenant le cas où les bornes d'intégration dépendent aussi de \mathbf{x} :

$$g(\mathbf{x}) = \int_{a(\mathbf{x})}^{b(\mathbf{x})} f(t, \mathbf{x}) dt$$

avec $a, b : E \subset \mathbb{R}^n \rightarrow]\alpha, \beta[\subset \mathbb{R}$ et $f :]\alpha, \beta[\times E \rightarrow \mathbb{R}$.

Théorème 5.6. Soit $-\infty \leq \alpha < \beta \leq +\infty$, $E \subset \mathbb{R}^n$ ouvert non vide, $a, b \in C^1(E)$ tels que $\text{Im}(a), \text{Im}(b) \subset]\alpha, \beta[$ et $f :]\alpha, \beta[\times E \rightarrow \mathbb{R}$ continue, avec dérivées partielles $\frac{\partial f}{\partial x_i} :]\alpha, \beta[\times E \rightarrow \mathbb{R}$, $i = 1, \dots, n$ continues. Alors $g \in C^1(E)$ et

$$\frac{\partial g}{\partial x_i}(\mathbf{x}) = \frac{\partial b}{\partial x_i}(\mathbf{x}) f(b(\mathbf{x}), \mathbf{x}) - \frac{\partial a}{\partial x_i}(\mathbf{x}) f(a(\mathbf{x}), \mathbf{x}) + \int_{a(\mathbf{x})}^{b(\mathbf{x})} \frac{\partial f}{\partial x_i}(t, \mathbf{x}) dt.$$

Démonstration. Soit $c \in]\alpha, \beta[$ et définissons la fonction $G(s, \mathbf{x}) = \int_c^s f(t, \mathbf{x}) dt$, $(s, \mathbf{x}) \in]\alpha, \beta[\times E$. Alors

$$g(\mathbf{x}) = \int_{a(\mathbf{x})}^c f(t, \mathbf{x}) dt + \int_c^{b(\mathbf{x})} f(t, \mathbf{x}) dt = G(b(\mathbf{x}), \mathbf{x}) - G(a(\mathbf{x}), \mathbf{x}).$$

Montrons que la fonction $G(s, \mathbf{x})$ a toutes les dérivées partielles en tout $(s, \mathbf{x}) \in]\alpha, \beta[\times E$. Par un résultat fondamental d'analyse I, $\frac{\partial G}{\partial s}(s, \mathbf{x})$ existe et vaut $f(s, \mathbf{x})$.

Pour $s \in]c, \beta[$, le Théorème 5.5 assure que $\frac{\partial G}{\partial x_i}(s, \mathbf{x})$ existe et vaut $\int_c^s \frac{\partial f}{\partial x_i}(t, \mathbf{x}) dt$. La même conclusion est vraie pour $s \in]\alpha, c[$ car la fonction $\mathbf{x} \mapsto \int_s^c f(t, \mathbf{x}) dt$ admet une dérivée partielle en x_i et qu'elle vaut $\int_s^c \frac{\partial f}{\partial x_i}(t, \mathbf{x}) dt$. Ainsi $\frac{\partial G}{\partial x_i}(s, \mathbf{x})$ existe et vaut $\int_c^s \frac{\partial f}{\partial x_i}(t, \mathbf{x}) dt$. Enfin, pour $s = c$ on a $G(c, \mathbf{x}) = 0$ sur E et donc $\frac{\partial G}{\partial x_i}(c, \mathbf{x})$ existe et est nul et $\frac{\partial G}{\partial x_i}(c, \mathbf{x}) = 0 = \int_c^c \frac{\partial f}{\partial x_i}(t, \mathbf{x}) dt$. En résumé, pour tout $s \in]\alpha, \beta[$ et $\mathbf{x} \in E$, $\frac{\partial G}{\partial x_i}(s, \mathbf{x})$ existe et vaut $\int_c^s \frac{\partial f}{\partial x_i}(t, \mathbf{x}) dt$.

Montrons que les dérivées partielles sont continues sur $] \alpha, \beta[\times E$. Ceci est vrai pour $\frac{\partial G}{\partial s}(s, \mathbf{x}) = f(s, \mathbf{x})$. Pour $i \in \{1, \dots, n\}$ fixé, on a

$$\frac{\partial G}{\partial x_i}(s, \mathbf{x}) = \int_c^s \frac{\partial f}{\partial x_i}(t, \mathbf{x}) dt = \int_0^1 \frac{\partial f}{\partial x_i}(c + (s - c)\tau, \mathbf{x})(s - c) d\tau.$$

Cette dernière expression est continue en (s, \mathbf{x}) en tant qu'intégrale sur τ dans le compact $[0, 1]$ d'une fonction continue en (τ, s, \mathbf{x}) . Voir le théorème 5.3.

A ce stade, nous avons donc prouvé que G est de classe C^1 sur $] \alpha, \beta[\times E$. Par les propriétés de dérivation de fonctions composées on a que $g(\mathbf{x}) = G(b(\mathbf{x}), \mathbf{x}) - G(a(\mathbf{x}), \mathbf{x})$ est de classe C^1 sur E et

$$\begin{aligned} \frac{\partial g}{\partial x_i}(\mathbf{x}) &= \frac{\partial G}{\partial s}(b(\mathbf{x}), \mathbf{x}) \frac{\partial b}{\partial x_i}(\mathbf{x}) + \frac{\partial G}{\partial x_i}(b(\mathbf{x}), \mathbf{x}) - \frac{\partial G}{\partial s}(a(\mathbf{x}), \mathbf{x}) \frac{\partial a}{\partial x_i}(\mathbf{x}) - \frac{\partial G}{\partial x_i}(a(\mathbf{x}), \mathbf{x}) \\ &= f(b(\mathbf{x}), \mathbf{x}) \frac{\partial b}{\partial x_i}(\mathbf{x}) + \int_c^{b(\mathbf{x})} \frac{\partial f}{\partial x_i}(t, \mathbf{x}) dt - f(a(\mathbf{x}), \mathbf{x}) \frac{\partial a}{\partial x_i}(\mathbf{x}) - \int_c^{a(\mathbf{x})} \frac{\partial f}{\partial x_i}(t, \mathbf{x}) dt \\ &= f(b(\mathbf{x}), \mathbf{x}) \frac{\partial b}{\partial x_i}(\mathbf{x}) - f(a(\mathbf{x}), \mathbf{x}) \frac{\partial a}{\partial x_i}(\mathbf{x}) + \int_{a(\mathbf{x})}^{b(\mathbf{x})} \frac{\partial f}{\partial x_i}(t, \mathbf{x}) dt. \end{aligned}$$

□

Exercice 5.7. Soit $g(y) = \int_1^{\sqrt{y}} \frac{e^{-xy}}{x} dx$, $y > 0$. Calculer $g'(y)$ si elle existe.

5.3 Intégrales généralisées dépendant de paramètres

On considère maintenant le cas où l'intégrale est définie sur un intervalle non compact. On se limite à discuter le cas d'un intervalle de la forme $[a, b[$ où b pourrait être $+\infty$ mais tous les autres cas se traitent de façon similaire.

Soit donc $E \subset \mathbb{R}^n$ un ensemble non vide, et $f : [a, b[\times E \rightarrow \mathbb{R}$ une fonction continue telle que l'intégrale généralisée

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

converge pour tout $\mathbf{x} \in E$, c-à-d, $\lim_{c \rightarrow b-} \int_a^c f(t, \mathbf{x}) dt$ existe pour tout $\mathbf{x} \in E$. Comme dans la section précédente, pour pouvoir établir la continuité de la fonction g , on a besoin de quelque forme d'uniformité en \mathbf{x} pour la convergence de l'intégrale.

Définition 5.8. Soit $E \subset \mathbb{R}^n$ un sous-ensemble non vide et $f : [a, b[\times E \rightarrow \mathbb{R}$ une fonction continue. On dit que l'intégrale $\int_a^b f(t, \mathbf{x}) dt$ converge uniformément sur E si elle converge pour tout $\mathbf{x} \in E$ et si

$$\forall \epsilon > 0 \exists \bar{c} \in]a, b[: \forall c \in [\bar{c}, b[\forall \mathbf{x} \in E \quad \left| \int_c^b f(t, \mathbf{x}) dt \right| \leq \epsilon.$$

Avec cette définition, on peut établir le résultat suivant.

Théorème 5.9. Soient $-\infty < a < b \leq +\infty$ et $E \subset \mathbb{R}^n$ non vide. Si $f : [a, b[\times E \rightarrow \mathbb{R}$ est continue et l'intégrale $\int_a^b f(t, \mathbf{x}) dt$ converge uniformément sur E , alors la fonction $g(\mathbf{x}) = \int_a^b f(t, \mathbf{x}) dt$ est continue sur E .

Démonstration. Soit $\mathbf{x}_0 \in E$ quelconque et $\epsilon > 0$. Alors, il existe $\bar{c} \in]a, b[$ t.q. $|\int_{\bar{c}}^b f(t, \mathbf{x}) dt| \leq \epsilon$, $\forall \mathbf{x} \in E$, grâce à la convergence uniforme de l'intégrale. En utilisant le théorème de Cantor-Heine généralisé 2.28, il existe $\delta > 0$ t.q. $|f(t, \mathbf{x}) - f(t, \mathbf{x}_0)| \leq \frac{\epsilon}{\bar{c}-a}$, $\forall t \in [a, \bar{c}]$, $\mathbf{x} \in \overline{B}(\mathbf{x}_0, \delta) \cap E$. Donc, pour tout $\mathbf{x} \in \overline{B}(\mathbf{x}_0, \delta) \cap E$,

$$|g(\mathbf{x}) - g(\mathbf{x}_0)| \leq \left| \int_a^{\bar{c}} (f(t, \mathbf{x}) - f(t, \mathbf{x}_0)) dt \right| + \left| \int_{\bar{c}}^b f(t, \mathbf{x}) dt \right| + \left| \int_{\bar{c}}^b f(t, \mathbf{x}_0) dt \right| \leq 3\epsilon,$$

ce qui montre la continuité de g en tout $\mathbf{x}_0 \in E$. \square

On a un résultat similaire pour la dérivabilité de g qu'on énonce sans démonstration.

Théorème 5.10. Soit $-\infty < a < b \leq +\infty$ et $E \subset \mathbb{R}^n$ ouvert non vide. Si $f \in C^1([a, b[\times E)$ et les intégrales $\int_a^b f(t, \mathbf{x}) dt$, $\int_a^b \frac{\partial f}{\partial x_i}(t, \mathbf{x}) dt$, $i = 1, \dots, n$, convergent uniformément sur E , alors la fonction $g(\mathbf{x}) = \int_a^b f(t, \mathbf{x}) dt$ est de classe C^1 sur E et $\frac{\partial g}{\partial x_i}(\mathbf{x}) = \int_a^b \frac{\partial f}{\partial x_i}(t, \mathbf{x}) dt$.

Vérifier la condition de convergence uniforme de l'intégrale n'est parfois pas immédiat. Voici un critère de majoration qui est souvent plus simple à vérifier.

Théorème 5.11. Soit $-\infty < a < b \leq +\infty$ et $E \subset \mathbb{R}^n$ ouvert non vide. Si $f : [a, b[\times E \rightarrow \mathbb{R}$ est continue et il existe une fonction $h : [a, b[\rightarrow \mathbb{R}_+$ telle que $\int_a^b h(t) dt$ converge et $|f(t, \mathbf{x})| \leq h(t)$, $\forall (t, \mathbf{x}) \in [a, b[\times E$, alors la fonction $g(\mathbf{x}) = \int_a^b f(t, \mathbf{x}) dt$ est continue sur E .

Si en plus f est de classe C^1 et il existe $h_1, \dots, h_n : [a, b[\rightarrow \mathbb{R}$ t.q. pour tout $i = 1, \dots, n$, $\int_a^b h_i(t) dt$ existe et $|\frac{\partial f}{\partial x_i}(t, \mathbf{x})| \leq h_i(t)$, $\forall (t, \mathbf{x}) \in [a, b[\times E$, alors $g \in C^1(E)$.

Démonstration. Il suffit de remarquer que la condition de majoration du théorème implique que l'intégrale $\int_a^b f(t, \mathbf{x}) dt$ converge uniformément. En effet, puisque $\int_a^b h(t) dt$ converge, pour tout $\epsilon > 0$ il existe $\bar{c} \in]a, b[$ t.q. $\int_{\bar{c}}^b h(t) dt \leq \epsilon$, $\forall c \in [\bar{c}, b[$, et donc $|\int_{\bar{c}}^b f(t, \mathbf{x}) dt| \leq \epsilon$, ce qui implique que l'intégrale $\int_a^b f(t, \mathbf{x}) dt$ converge uniformément. Une remarque analogue est valable pour chacune des intégrales $\int_a^b \frac{\partial f}{\partial x_i}(t, \mathbf{x}) dt$. \square

Remarque 5.12. Dans les théorèmes 5.9 et 5.10, la condition d'intégrabilité uniforme de $\int_a^b f(t, \mathbf{x})dt$ et $\int_a^b \frac{\partial f}{\partial x_i}(t, \mathbf{x})dt$, $1 \leq i \leq n$, sur tout l'ensemble E est parfois trop contraignante. Elle peut être affaiblie de la façon suivante :

Il existe un recouvrement $\{U_\alpha\}_\alpha$ de E , avec U_α ouverts et $E \subset \bigcup_\alpha U_\alpha$, tel que, pour tout α , les intégrales $\int_a^b f(t, \mathbf{x})dt$ et $\int_a^b \frac{\partial f}{\partial x_i}(t, \mathbf{x})dt$, $1 \leq i \leq n$, convergent uniformément sur $U_\alpha \cap E$.

Alors, les conclusions des théorèmes 5.9, 5.10 sont encore valables. Pour généraliser leur démonstration il suffit de remarquer que, pour tout $\mathbf{x}_0 \in E$, il existe α tel que $\mathbf{x}_0 \in U_\alpha$ (puisque $\{U_\alpha\}_\alpha$ est un recouvrement de E). Les intégrales $\int_a^b f(t, \mathbf{x})dt$ et $\int_a^b \frac{\partial f}{\partial x_i}(t, \mathbf{x})dt$ étant uniformément convergentes sur $U_\alpha \cap E$, on conclut que $g(\mathbf{x})$ est continue / différentiable sur $U_\alpha \cap E$ pour tout α et donc sur $E = \bigcup_\alpha (U_\alpha \cap E)$.

Le même raisonnement s'applique au théorème 5.11 où on peut remplacer la condition de majoration par la suivante :

Il existe un recouvrement $\{U_\alpha\}_\alpha$ de E , avec U_α ouverts et $E \subset \bigcup_\alpha U_\alpha$, et, pour chaque α , des fonctions $h_{i,\alpha} : [a, b[\rightarrow \mathbb{R}_+$, $i = 0, \dots, n$, telles que $\int_a^b h_{i,\alpha}(t)dt$ converge et

- $|f(t, \mathbf{x})| \leq h_{0,\alpha}(t)$ pour tout $(t, \mathbf{x}) \in [a, b[\times (U_\alpha \cap E)$,
- $|\frac{\partial f}{\partial x_i}(t, \mathbf{x})| \leq h_{i,\alpha}(t)$ pour tout $(t, \mathbf{x}) \in [a, b[\times (U_\alpha \cap E)$ et tout $i \in \{1, \dots, n\}$.

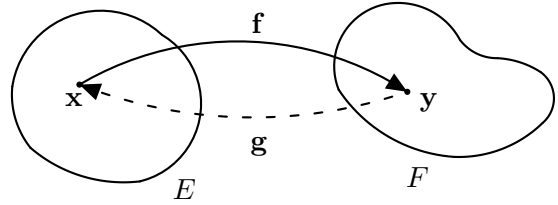
Chapitre 6

Difféomorphismes locaux et fonctions implicites

6.1 Fonctions bijectives et difféomorphismes locaux

Considérons une fonction $\mathbf{f} : E \subset \mathbb{R}^n \rightarrow F \subset \mathbb{R}^n$ bijective. Alors on peut définir l'application inverse (ou “réciproque”) $\mathbf{g} : F \rightarrow E$ telle que

$$\begin{aligned} \mathbf{g} \circ \mathbf{f} &= \text{id}_E : \quad \forall \mathbf{x} \in E \quad \mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{x} \\ \mathbf{f} \circ \mathbf{g} &= \text{id}_F : \quad \forall \mathbf{y} \in F \quad \mathbf{f}(\mathbf{g}(\mathbf{y})) = \mathbf{y}. \end{aligned}$$



Un premier intérêt à étudier la bijectivité d'une fonction est de pouvoir garantir l'existence d'une solution unique $\mathbf{x} = \mathbf{g}(\mathbf{y})$ du système d'équations non-linéaires

$$\mathbf{f}(\mathbf{x}) = \mathbf{y}, \quad \Longleftrightarrow \quad \begin{cases} f_1(x_1, \dots, x_n) = y_1 \\ \vdots \\ f_n(x_1, \dots, x_n) = y_n \end{cases}$$

pour tout $\mathbf{y} = (y_1, \dots, y_n) \in F$. Souvent, on souhaite avoir aussi de bonnes propriétés de stabilité aux petites perturbations, i.e. si $\tilde{\mathbf{x}} = \mathbf{g}(\tilde{\mathbf{y}})$ et $\tilde{\mathbf{y}} \rightarrow \mathbf{y}$ on souhaite que $\tilde{\mathbf{x}} \rightarrow \mathbf{x}$ ce qui revient à demander la continuité de la fonction inverse (en plus de la continuité de \mathbf{f}). On parle d'*homéomorphisme* lorsque \mathbf{f} est une bijection continue avec inverse continu.

Définition 6.1 (Homéomorphisme). *Soient $E, F \subset \mathbb{R}^n$ ouverts non vides. Une application $\mathbf{f} : E \rightarrow F$ est un homéomorphisme si elle est bijective et si \mathbf{f} et son inverse \mathbf{g} sont continues.*

Un autre intérêt d'étudier des bijections est pour introduire un changement de variables. Soit $\mathbf{f} : E \rightarrow F$ une bijection et $\phi : F \rightarrow \mathbb{R}$, $\mathbf{y} \mapsto \phi(\mathbf{y}) = \phi(y_1, \dots, y_n)$, une fonction réelle.

Grâce à l'application \mathbf{f} , on peut exprimer ϕ en fonction des variables $\mathbf{x} = (x_1, \dots, x_n) \in E$:

$$\tilde{\phi} = \phi \circ \mathbf{f} : E \rightarrow \mathbb{R}, \quad \tilde{\phi}(\mathbf{x}) = \phi(\mathbf{f}(\mathbf{x})), \quad \mathbf{x} \in E$$

et vice versa, étant donné $\tilde{\phi} : E \rightarrow \mathbb{R}$, on peut l'exprimer en fonction des variables \mathbf{y} :

$$\phi = \tilde{\phi} \circ \mathbf{g} : F \rightarrow \mathbb{R}, \quad \phi(\mathbf{y}) = \tilde{\phi}(\mathbf{g}(\mathbf{y})), \quad \mathbf{y} \in F,$$

où $\mathbf{g} : F \rightarrow E$ est l'inverse de \mathbf{f} . Dans ce cas, si ϕ est une fonction régulière, par exemple de classe C^k , on souhaite que la fonction transformée $\tilde{\phi}$ soit aussi de classe C^k , et vice-versa. Ceci est garanti si le changement de variables \mathbf{f} et son inverse sont de classe C^k .

Exemple 6.2. *Transformation en coordonnées polaires :*

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{f}(\rho, \theta) = \begin{pmatrix} \rho \cos \theta \\ \rho \sin \theta \end{pmatrix} \quad \mathbf{f} :]0, +\infty[\times]-\pi, \pi[\rightarrow \mathbb{R}^2 \setminus \{(x, y) \in \mathbb{R}^2 : x \leq 0, y = 0\}.$$

Cette transformation est continue, même de classe C^∞ , et inversible. De plus, l'application inverse est continue, même de classe C^∞ , sur $\mathbb{R}^2 \setminus \{(x, y) \in \mathbb{R}^2 : x \leq 0, y = 0\} \rightarrow]0, +\infty[\times]-\pi, \pi[$.

On va étudier dans ce chapitre des bijections continûment différentiables avec application inverse continûment différentiable. Ces applications sont appelées *difféomorphismes*.

Définition 6.3 (Difféomorphisme). Soient $E, F \subset \mathbb{R}^n$ ouverts non vides. Une application $\mathbf{f} : E \rightarrow F$ inversible de classe C^1 , $\mathbf{f} \in C^1(E, \mathbb{R}^n)$, est un *difféomorphisme* si l'application inverse $\mathbf{g} : F \rightarrow E$ (t.q. $\mathbf{g} \circ \mathbf{f} = \text{id}_E$ et $\mathbf{f} \circ \mathbf{g} = \text{id}_F$) est de classe C^1 sur F , $\mathbf{g} \in C^1(F, \mathbb{R}^n)$. De plus, on dit que \mathbf{f} est un *k-difféomorphisme* si $\mathbf{f} \in C^k(E, \mathbb{R}^n)$ et l'application inverse $\mathbf{g} \in C^k(F, \mathbb{R}^n)$.

Remarque 6.4. A strictement parler, $\mathbf{f} : E \rightarrow F$ est un *difféomorphisme* (E et F étant des ouverts non vides) si \mathbf{f} est différentiable en tout point de E , bijective et son inverse $\mathbf{g} : F \rightarrow E$ est différentiable en tout point de F . Dans ce cours, sauf mention contraire, nous rajouterons l'hypothèse que \mathbf{f} et \mathbf{g} sont de classe C^1 .

Établir si une application $\mathbf{f} : E \rightarrow F$ est un difféomorphisme est souvent compliqué. On introduit une définition plus faible qui est plus facile à vérifier :

Définition 6.5 (Difféomorphisme local). Soit $E \subset \mathbb{R}^n$ ouvert non vide, $\mathbf{x}_0 \in E$ et $\mathbf{f} : E \rightarrow \mathbb{R}^n$ une application de classe C^1 . On dit que \mathbf{f} est un *difféomorphisme local* en \mathbf{x}_0 s'il existe un ouvert $U \subset E$ contenant \mathbf{x}_0 et un ouvert $V \subset \mathbb{R}^n$ contenant $\mathbf{f}(\mathbf{x}_0)$ tels que $\mathbf{f} : U \rightarrow V$ est un *difféomorphisme* (fonction bijective avec inverse $\mathbf{g} : V \rightarrow U$ de classe C^1). Si \mathbf{f} et \mathbf{g} sont de classe C^k sur U et V respectivement, on dit que \mathbf{f} est un *k-difféomorphisme local*. (Nous appellerons \mathbf{g} un *inverse local* de \mathbf{f} restreinte à un voisinage de \mathbf{x}_0 .)

Soient $E, F, G \subset \mathbb{R}^n$ ouverts non vides. Il est facile de montrer (exercice!) que si $\mathbf{f} : E \rightarrow F$ est un difféomorphisme local en $\mathbf{x}_0 \in E$ et $\mathbf{h} : F \rightarrow G$ est un difféomorphisme local en $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0) \in F$, alors $\mathbf{h} \circ \mathbf{f}$ est un difféomorphisme local en \mathbf{x}_0 .

On pourrait penser que si $\mathbf{f} : E \rightarrow F$, avec $E, F \subset \mathbb{R}^n$ ouvert, est un difféomorphisme local en tout point $\mathbf{x}_0 \in E$, alors elle est un difféomorphisme global entre E et F . Ceci n'est en général pas vrai, comme l'exemple suivant le montre.

Exemple 6.6. Soit $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ donnée par

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \mathbf{f}(x_1, x_2) = \begin{pmatrix} e^{x_1} \cos x_2 \\ e^{x_1} \sin x_2 \end{pmatrix}.$$

On verra par la suite que \mathbf{f} est un difféomorphisme local en tout point $\mathbf{x}_0 \in \mathbb{R}^2$. Toutefois, elle n'est pas un difféomorphisme global car elle n'est pas une bijection : $\mathbf{f}(0, 0) = \mathbf{f}(0, 2k\pi)$, $\forall k \in \mathbb{Z}$.

Toutefois, si on ajoute l'hypothèse que \mathbf{f} est une bijection, alors le difféomorphisme est global.

Lemme 6.7. Soient $E, F \subset \mathbb{R}^n$ ouverts non vides et $\mathbf{f} : E \rightarrow F$ un difféomorphisme local en tout point $\mathbf{x} \in E$. Si \mathbf{f} est une bijection entre E et F , alors elle est un difféomorphisme global.

La démonstration de ce théorème est laissée comme exercice.

6.2 Théorème d'inversion locale

On s'intéresse à comprendre sous quelles conditions une application $\mathbf{f} : E \rightarrow \mathbb{R}^n$, avec $E \subset \mathbb{R}^n$ ouvert, est un difféomorphisme local en $\mathbf{x}_0 \in E$. Pour cela, on va d'abord étudier le cas d'une application *affine* $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$,

$$\mathbf{f}(\mathbf{x}) = A\mathbf{x} + \mathbf{b}, \quad A \in \mathbb{R}^{n \times n}, \quad \mathbf{b} \in \mathbb{R}^n$$

qui est clairement de classe C^1 (même C^∞). Cette transformation est une bijection si et seulement si la matrice A est inversible, c.-à-d. si et seulement si $\det(A) \neq 0$. Dans ce cas, l'application inverse est donnée par $\mathbf{x} = \mathbf{g}(\mathbf{y}) = A^{-1}(\mathbf{y} - \mathbf{b})$ et \mathbf{g} est de classe C^1 (même C^∞). Donc $\mathbf{f}(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ est un difféomorphisme (global) si et seulement si $\det(A) \neq 0$.

On considère maintenant le cas d'une application non linéaire $\mathbf{f} : E \rightarrow \mathbb{R}^n$ de classe C^1 sur $E \subset \mathbb{R}^n$ ouvert. La différentiabilité de \mathbf{f} en $\mathbf{x}_0 \in E$ assure que l'on peut écrire

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + D\mathbf{f}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \mathbf{R}_f(\mathbf{x}), \quad \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\|\mathbf{R}_f(\mathbf{x})\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0,$$

pour tout $\mathbf{x} \in E$. Donc, localement autour de \mathbf{x}_0 , $\mathbf{f}(\mathbf{x})$ est bien approchée par la fonction affine $\mathbf{T}_{\mathbf{x}_0}^1(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + D\mathbf{f}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$. On soupçonne alors que \mathbf{f} est un difféomorphisme local en \mathbf{x}_0 si et seulement si $\det(D\mathbf{f}(\mathbf{x}_0)) \neq 0$.

On remarque que la condition $\det(D\mathbf{f}(\mathbf{x}_0)) \neq 0$ n'est pas nécessaire pour que \mathbf{f} soit localement inversible. Par exemple, la fonction $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $(x, y) \mapsto \mathbf{f}(x, y) = (x^3, y)$ est localement inversible autour de $(x, y) = (0, 0)$ (et même globalement sur tout \mathbb{R}^2) même si $\det(D\mathbf{f}(0, 0)) = \det \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = 0$. Toutefois, cette condition est nécessaire pour que la fonction \mathbf{f} soit un difféomorphisme local, comme le lemme suivant le montre.

Lemme 6.8. Soit $\mathbf{f} : E \rightarrow \mathbb{R}^n$, avec $E \subset \mathbb{R}^n$ ouvert, un difféomorphisme local autour de $\mathbf{x}_0 \in E$. Alors $\det(D\mathbf{f}(\mathbf{x}_0)) \neq 0$.

Démonstration. Puisque $\mathbf{f} : E \rightarrow \mathbb{R}^n$ est un difféomorphisme local en $\mathbf{x}_0 \in E$, il existe un ouvert $U \subset E$ contenant \mathbf{x}_0 et un ouvert $V \subset \mathbb{R}^n$ contenant $\mathbf{f}(\mathbf{x}_0) = \mathbf{y}_0$ t.q. $\mathbf{f} : U \rightarrow V$ est bijective et la fonction inverse $\mathbf{g} : V \rightarrow U$ est de classe C^1 . Puisque $\mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{x}$, $\forall \mathbf{x} \in U$, $\mathbf{f}(\mathbf{g}(\mathbf{y})) = \mathbf{y}$, $\forall \mathbf{y} \in V$ et \mathbf{f} et \mathbf{g} sont différentiables en \mathbf{x}_0 et \mathbf{y}_0 , on a :

$$D\mathbf{f}(\mathbf{x}_0)D\mathbf{g}(\mathbf{y}_0) = D\mathbf{g}(\mathbf{y}_0)D\mathbf{f}(\mathbf{x}_0) = I$$

ce qui implique que $D\mathbf{f}(\mathbf{x}_0)$ est inversible et $\det(D\mathbf{f}(\mathbf{x}_0)) \neq 0$. \square

La condition $\det(D\mathbf{f}(\mathbf{x}_0)) \neq 0$ est aussi suffisante pour que \mathbf{f} soit un difféomorphisme local autour de \mathbf{x}_0 , comme le théorème suivant le montre.

Théorème 6.9 (d'inversion locale). *Soit $E \subset \mathbb{R}^n$ ouvert, $\mathbf{f} \in C^1(E, \mathbb{R}^n)$ et $\mathbf{x}_0 \in E$. Si $\det(D\mathbf{f}(\mathbf{x}_0)) \neq 0$, alors \mathbf{f} est un difféomorphisme local en \mathbf{x}_0 , c.-à-d. qu'il existe un ouvert $U \subset E$ contenant \mathbf{x}_0 et un ouvert $V \subset \text{Im}(f)$ contenant $\mathbf{f}(\mathbf{x}_0) = \mathbf{y}_0$ tels que $\mathbf{f} : U \rightarrow V$ est une bijection et la fonction inverse $\mathbf{g} : V \rightarrow U$ est de classe C^1 . De plus $D\mathbf{g}(\mathbf{f}(\mathbf{x})) = (D\mathbf{f}(\mathbf{x}))^{-1}$ pour tout $\mathbf{x} \in U$.*

Avant de démontrer ce théorème, nous allons prouver des résultats intermédiaires. Commençons par le théorème du point fixe de Banach.

Théorème 6.10 (du point fixe de Banach). *Soit un fermé non vide $K \subset \mathbb{R}^n$ et une fonction $\phi : K \rightarrow \mathbb{R}^n$ telle que*

- $\phi(K) \subset K$,
- *il existe $\rho \in]0, 1[$ tel que $\forall \mathbf{v}, \mathbf{w} \in K \quad \|\phi(\mathbf{v}) - \phi(\mathbf{w})\| \leq \rho \|\mathbf{v} - \mathbf{w}\|$.*

Alors il existe un unique $\mathbf{v}_ \in K$ tel que $\phi(\mathbf{v}_*) = \mathbf{v}_*$.*

Démonstration.

Existence. Fixons $\mathbf{v}^{(0)}$ dans K et définissons $\{\mathbf{v}^{(k)}\}_{k \in \mathbb{N}} \subset K$ par récurrence comme suit :

$$\mathbf{v}^{(k+1)} = \phi(\mathbf{v}^{(k)}) \in K, \quad k \geq 0.$$

Vérifions que $\{\mathbf{v}^{(k)}\}_{k \in \mathbb{N}}$ est de Cauchy. Pour $k \in \mathbb{N}$,

$$\|\mathbf{v}^{(k+1)} - \mathbf{v}^{(k)}\| = \|\phi(\mathbf{v}^{(k)}) - \phi(\mathbf{v}^{(k-1)})\| \leq \rho \|\mathbf{v}^{(k)} - \mathbf{v}^{(k-1)}\| \leq \dots \leq \rho^k \|\mathbf{v}^{(1)} - \mathbf{v}^{(0)}\|.$$

Soit $\epsilon > 0$. Alors pour $k > m \geq 0$,

$$\begin{aligned} \|\mathbf{v}^{(k)} - \mathbf{v}^{(m)}\| &\leq \|\mathbf{v}^{(k)} - \mathbf{v}^{(k-1)}\| + \dots + \|\mathbf{v}^{(m+1)} - \mathbf{v}^{(m)}\| \leq (\rho^{k-1} + \dots + \rho^m) \|\mathbf{v}^{(1)} - \mathbf{v}^{(0)}\| \\ &\leq \rho^m \left(\sum_{j=0}^{\infty} \rho^j \right) \|\mathbf{v}^{(1)} - \mathbf{v}^{(0)}\| = \rho^m (1 - \rho)^{-1} \|\mathbf{v}^{(1)} - \mathbf{v}^{(0)}\| < \epsilon \end{aligned}$$

si m (et donc k) est suffisamment grand.

Soit $\mathbf{v}_* \in \mathbb{R}^n$ la limite de $\{\mathbf{v}^{(k)}\}_{k \in \mathbb{N}}$. Alors $\mathbf{v}_* \in K$ car K est fermé, et

$$\|\mathbf{v}_* - \phi(\mathbf{v}_*)\| \leq \|\mathbf{v}_* - \mathbf{v}^{(k+1)}\| + \|\mathbf{v}^{(k+1)} - \phi(\mathbf{v}_*)\| = \|\mathbf{v}_* - \mathbf{v}^{(k+1)}\| + \|\phi(\mathbf{v}^{(k)}) - \phi(\mathbf{v}_*)\|$$

$$\leq \|\mathbf{v}_* - \mathbf{v}^{(k+1)}\| + \rho \|\mathbf{v}^{(k)} - \mathbf{v}_*\| \rightarrow 0$$

lorsque $k \rightarrow +\infty$. D'où $\phi(\mathbf{v}_*) = \mathbf{v}_*$.

Unicité. Supposons qu'il existe un autre $\mathbf{u}_* \in K$ tel que $\phi(\mathbf{u}_*) = \mathbf{u}_*$. Alors

$$\|\mathbf{u}_* - \mathbf{v}_*\| = \|\phi(\mathbf{u}_*) - \phi(\mathbf{v}_*)\| \leq \rho \|\mathbf{u}_* - \mathbf{v}_*\|$$

avec $\rho \in]0, 1[$. D'où $\|\mathbf{u}_* - \mathbf{v}_*\| = 0$ et $\mathbf{u}_* = \mathbf{v}_*$. \square

Définition 6.11. Avec les notations du théorème 6.10, puisque $\rho < 1$, on dit que l'application $\phi : K \rightarrow \mathbb{R}^n$ est contractante ou une contraction.

Pour la démonstration du théorème 6.9 nous avons encore besoin de deux lemmes suivants.

Lemme 6.12. Soit $A \in \mathbb{R}^{m \times n}$. On note par $\|A\|$ la norme spectrale de A , $\|A\| = \sup_{\substack{\xi \in \mathbb{R}^n \\ \|\xi\|=1}} \|A\xi\|$, où $\|\cdot\|$ denote la norme euclidienne, et par $\|A\|_F$ la norme de Frobenius de A , $\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n A_{ij}^2}$. Alors, on a $\|A\| \leq \|A\|_F$.

Démonstration. Grâce à l'inégalité de Cauchy-Schwarz,

$$\|A\|^2 = \sup_{\substack{\xi \in \mathbb{R}^n \\ \|\xi\|=1}} \sum_{i=1}^m \left(\sum_{j=1}^n A_{ij} \xi_j \right)^2 \leq \sup_{\substack{\xi \in \mathbb{R}^n \\ \|\xi\|=1}} \sum_{i=1}^m \left(\sum_{j=1}^n A_{ij}^2 \right) \left(\sum_{j=1}^n \xi_j^2 \right) = \|A\|_F^2.$$

\square

Lemme 6.13. Soit $-\infty < a < b < \infty$ et $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^m$, $t \mapsto \mathbf{f}(t) = (f_1(t), \dots, f_m(t))$ une fonction continue. Alors

$$\left\| \int_a^b \mathbf{f}(t) dt \right\| \leq \int_a^b \|\mathbf{f}(t)\| dt,$$

où $\|\cdot\|$ est la norme euclidienne.

Démonstration. Chaque fonction $f_i : [a, b] \rightarrow \mathbb{R}$, $i = 1, \dots, m$ est intégrale, ainsi que la fonction $g = \|\mathbf{f}\| : [a, b] \rightarrow \mathbb{R}_+$, étant une composition de fonctions continues sur $[a, b]$. Notons $\mathbf{v} \in \mathbb{R}^m$ le vecteur de composantes $v_i = \int_a^b f_i(t) dt$, $i = 1, \dots, m$. Alors

$$\left\| \int_a^b \mathbf{f}(t) dt \right\|^2 = \|\mathbf{v}\|^2 = \sum_{i=1}^m v_i v_i = \int_a^b \sum_{i=1}^m v_i f_i(t) dt \leq \int_a^b \|\mathbf{v}\| \|\mathbf{f}(t)\| dt,$$

d'où le résultat. \square

Démonstration du théorème 6.9 d'inversion locale. On va couper la preuve en trois étapes :

- (i) On montre qu'il existe $r, \tilde{r} > 0$ tel que $\forall \mathbf{y} \in B(\mathbf{y}_0, \tilde{r})$ l'équation $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ pour $\mathbf{x} \in B(\mathbf{x}_0, r)$ a une solution unique $\mathbf{x} \in B(\mathbf{x}_0, r)$ (on utilise le théorème du point fixe de Banach).

- (ii) Soit $V = B(\mathbf{y}_0, \tilde{r})$ et $U = B(\mathbf{x}_0, r) \cap \mathbf{f}^{-1}(V) = \{\mathbf{x} \in B(\mathbf{x}_0, r) : \mathbf{f}(\mathbf{x}) \in V\}$, on montre que $\mathbf{f} : U \rightarrow V$ est une bijection et la fonction inverse $\mathbf{g} : V \rightarrow U$ est continue.
- (iii) On montre que $\mathbf{g} : V \rightarrow U$ est de classe $C^1(V, \mathbb{R}^n)$ et $D\mathbf{g}(\mathbf{f}(\mathbf{x})) = (D\mathbf{f}(\mathbf{x}))^{-1}$ pour tout $\mathbf{x} \in U$.

Première étape. On étudie l'équation $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ avec $\mathbf{y} \in \mathbb{R}^n$ fixé dans un voisinage convenable de \mathbf{y}_0 . En notant par $I \in \mathbb{R}^{n \times n}$ la matrice identité, observons d'abord que l'application $\mathbf{x} \mapsto I - D\mathbf{f}(\mathbf{x}_0)^{-1}D\mathbf{f}(\mathbf{x}) \in \mathbb{R}^{n \times n}$ est continue sur E (c.-à-d. chaque composante de la matrice $I - D\mathbf{f}(\mathbf{x}_0)^{-1}D\mathbf{f}$ est une fonction continue sur E) et s'annule en \mathbf{x}_0 . De même l'application $\mathbf{x} \mapsto \det D\mathbf{f}(\mathbf{x}) \in \mathbb{R}$ est continue et, par hypothèse, ne s'annule pas en \mathbf{x}_0 . Il existe donc $r > 0$ tel que, pour tout $\mathbf{x} \in \overline{B}(\mathbf{x}_0, r)$,

- $\mathbf{x} \in E$,
- chaque composante de la matrice $I - D\mathbf{f}(\mathbf{x}_0)^{-1}D\mathbf{f}(\mathbf{x})$ est dans $\left[-\frac{1}{2n}, \frac{1}{2n}\right]$,
- $\det D\mathbf{f}(\mathbf{x}) \neq 0$ et donc la matrice jacobienne $D\mathbf{f}(\mathbf{x})$ est inversible.

L'équation $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ est équivalente à $\mathbf{x} = \phi^{\mathbf{y}}(\mathbf{x})$ où $\phi^{\mathbf{y}}(\mathbf{x}) = \mathbf{x} - D\mathbf{f}(\mathbf{x}_0)^{-1}(\mathbf{f}(\mathbf{x}) - \mathbf{y})$, grâce au fait que $D\mathbf{f}(\mathbf{x}_0)$ est inversible. On remarque, en particulier, que $\phi^{\mathbf{y}} \in C^1(E, \mathbb{R}^n)$. Alors $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ a une solution dans $\overline{B}(\mathbf{x}_0, r)$ si et seulement si $\phi^{\mathbf{y}}$ a un point fixe dans $\overline{B}(\mathbf{x}_0, r)$. Mais, pour tout $\mathbf{x} \in \overline{B}(\mathbf{x}_0, r)$,

$$D\phi^{\mathbf{y}}(\mathbf{x}) = I - D\mathbf{f}(\mathbf{x}_0)^{-1}D\mathbf{f}(\mathbf{x})$$

et donc $\left| \frac{\partial \phi_i^{\mathbf{y}}}{\partial x_j}(\mathbf{x}) \right| \leq \frac{1}{2n}$. Ceci implique

$$\|D\phi^{\mathbf{y}}(\mathbf{x})\| \leq \|D\phi^{\mathbf{y}}(\mathbf{x})\|_F = \left(\sum_{i,j=1}^n \left| \frac{\partial \phi_i^{\mathbf{y}}}{\partial x_j}(\mathbf{x}) \right|^2 \right)^{\frac{1}{2}} \leq \frac{1}{2}, \quad \forall \mathbf{x} \in \overline{B}(\mathbf{x}_0, r).$$

Par le théorème des accroissements finis (version intégrale, où apparaît le produit matricielle d'une matrice et d'un vecteur colonne, donnant un vecteur colonne) et les lemmes 6.12-6.13, on a pour $\mathbf{x}_1, \mathbf{x}_2 \in \overline{B}(\mathbf{x}_0, r)$,

$$\begin{aligned} \|\phi^{\mathbf{y}}(\mathbf{x}_2) - \phi^{\mathbf{y}}(\mathbf{x}_1)\| &= \left\| \int_0^1 D\phi^{\mathbf{y}}(\mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1)) \cdot (\mathbf{x}_2 - \mathbf{x}_1) dt \right\| \\ &\leq \int_0^1 \|D\phi^{\mathbf{y}}(\mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1)) \cdot (\mathbf{x}_2 - \mathbf{x}_1)\| dt \\ &\leq \int_0^1 \underbrace{\|D\phi^{\mathbf{y}}(\mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1))\|}_{\in \overline{B}(\mathbf{x}_0, r)} \|\mathbf{x}_2 - \mathbf{x}_1\| dt \leq \frac{1}{2} \|\mathbf{x}_2 - \mathbf{x}_1\| \end{aligned}$$

et ϕ est contractante sur $\overline{B}(\mathbf{x}_0, r)$ pour tout $\mathbf{y} \in \mathbb{R}^n$. De plus, pour tout $\mathbf{x} \in \overline{B}(\mathbf{x}_0, r)$ et $\mathbf{y} \in B(\mathbf{y}_0, \tilde{r})$, avec $\tilde{r} = \frac{r}{2\|D\mathbf{f}(\mathbf{x}_0)^{-1}\|}$, on a

$$\begin{aligned} \|\phi^{\mathbf{y}}(\mathbf{x}) - \mathbf{x}_0\| &\leq \|\phi^{\mathbf{y}}(\mathbf{x}) - \phi^{\mathbf{y}}(\mathbf{x}_0)\| + \|\phi^{\mathbf{y}}(\mathbf{x}_0) - \mathbf{x}_0\| \\ &\leq \frac{1}{2} \|\mathbf{x} - \mathbf{x}_0\| + \|D\mathbf{f}(\mathbf{x}_0)^{-1}(\mathbf{y}_0 - \mathbf{y})\| \leq \frac{r}{2} + \|D\mathbf{f}(\mathbf{x}_0)^{-1}\| \|\mathbf{y}_0 - \mathbf{y}\| < r. \end{aligned}$$

Donc, si on prend $\mathbf{y} \in B(\mathbf{y}_0, \tilde{r})$, on a $\phi^{\mathbf{y}}(\overline{B}(\mathbf{x}_0, r)) \subset B(\mathbf{x}_0, r)$. Par conséquent, $\phi^{\mathbf{y}} : \overline{B}(\mathbf{x}_0, r) \rightarrow \overline{B}(\mathbf{x}_0, r)$ a un point fixe unique $\mathbf{x} \in \overline{B}(\mathbf{x}_0, r)$. Mais puisque $\phi^{\mathbf{y}}(\overline{B}(\mathbf{x}_0, r)) \subset B(\mathbf{x}_0, r)$, ce point fixe est dans la boule ouverte $B(\mathbf{x}_0, r)$. On a donc montré que $\forall \mathbf{y} \in B(\mathbf{y}_0, \tilde{r})$, il existe un unique $\mathbf{x} \in B(\mathbf{x}_0, r)$ t.q. $\mathbf{x} = \phi^{\mathbf{y}}(\mathbf{x})$, c.-à-d. t.q. $\mathbf{f}(\mathbf{x}) = \mathbf{y}$.

Deuxième étape. Soit $V = B(\mathbf{y}_0, \tilde{r})$ et $U = B(\mathbf{x}_0, r) \cap \mathbf{f}^{-1}(V) = \{\mathbf{x} \in B(\mathbf{x}_0, r) : \mathbf{f}(\mathbf{x}) \in B(\mathbf{y}_0, \tilde{r})\}$. On remarque que $\mathbf{f}^{-1}(V)$ est ouvert (\mathbf{f} est continue sur l'ouvert E , donc la pré-image d'un ouvert est un ouvert), donc U est ouvert. Par l'étape précédente, $\mathbf{f} : U \rightarrow V$ est une bijection donc on peut définir la fonction inverse $\mathbf{g} : V \rightarrow U$. On montre que \mathbf{g} est continue. Soit $\epsilon > 0$ et $\delta = \frac{\epsilon}{2\|D\mathbf{f}(\mathbf{x}_0)^{-1}\|}$, alors, pour tous $\mathbf{y}_1, \mathbf{y}_2 \in V$ tels que $\|\mathbf{y}_1 - \mathbf{y}_2\| \leq \delta$, et en notant $\mathbf{x}_1 = \mathbf{g}(\mathbf{y}_1)$ et $\mathbf{x}_2 = \mathbf{g}(\mathbf{y}_2)$, on a

$$\begin{aligned} \|\mathbf{x}_1 - \mathbf{x}_2\| &= \|\phi^{\mathbf{y}_1}(\mathbf{x}_1) - \phi^{\mathbf{y}_2}(\mathbf{x}_2)\| \\ &\leq \|\phi^{\mathbf{y}_1}(\mathbf{x}_1) - \phi^{\mathbf{y}_1}(\mathbf{x}_2)\| + \|\phi^{\mathbf{y}_1}(\mathbf{x}_2) - \phi^{\mathbf{y}_2}(\mathbf{x}_2)\| \\ &\leq \frac{1}{2}\|\mathbf{x}_1 - \mathbf{x}_2\| + \|D\mathbf{f}(\mathbf{x}_0)^{-1}(\mathbf{y}_1 - \mathbf{y}_2)\|, \end{aligned}$$

et donc

$$\|\mathbf{x}_1 - \mathbf{x}_2\| = \|\mathbf{g}(\mathbf{y}_1) - \mathbf{g}(\mathbf{y}_2)\| \leq 2\|D\mathbf{f}(\mathbf{x}_0)^{-1}\| \|\mathbf{y}_2 - \mathbf{y}_1\| \leq \epsilon$$

ce qui montre la continuité (uniforme) de $\mathbf{g} : V \rightarrow U$. Même plus, le calcul précédent montre que \mathbf{g} est Lipschitz.

Troisième étape. Il reste à montrer que $\mathbf{g} : V \rightarrow U$ est de classe C^1 et $D\mathbf{g}(\mathbf{y}) = (D\mathbf{f}(\mathbf{x}))^{-1}$, $\forall \mathbf{y} \in V$, où $\mathbf{x} = \mathbf{g}(\mathbf{y})$. Par le choix de r , $D\mathbf{f}(\mathbf{x})$ est inversible pour tout $\mathbf{x} \in B(\mathbf{x}_0, r)$. Soit maintenant $\mathbf{y}_1 \in V$ et $\mathbf{x}_1 = \mathbf{g}(\mathbf{y}_1)$. Puisque $\mathbf{f} \in C^1(E, \mathbb{R}^n)$, on a

$$\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}) = D\mathbf{f}(\mathbf{x})(\mathbf{x}_1 - \mathbf{x}) + \mathbf{R}_{\mathbf{f}}(\mathbf{x}_1) \quad \text{avec} \quad \lim_{\mathbf{x}_1 \rightarrow \mathbf{x}} \frac{\|\mathbf{R}_{\mathbf{f}}(\mathbf{x}_1)\|}{\|\mathbf{x}_1 - \mathbf{x}\|} = 0.$$

Ceci implique

$$\mathbf{x}_1 - \mathbf{x} = \mathbf{g}(\mathbf{y}_1) - \mathbf{g}(\mathbf{y}) = D\mathbf{f}(\mathbf{x})^{-1}(\mathbf{y}_1 - \mathbf{y}) - \underbrace{D\mathbf{f}(\mathbf{x})^{-1}\mathbf{R}_{\mathbf{f}}(\mathbf{x}_1)}_{\mathbf{R}_{\mathbf{g}}(\mathbf{y}_1)}$$

On va montrer que $\lim_{\mathbf{y}_1 \rightarrow \mathbf{y}} \frac{\mathbf{R}_{\mathbf{g}}(\mathbf{y}_1)}{\|\mathbf{y}_1 - \mathbf{y}\|} = 0$. En effet,

$$\begin{aligned} \lim_{\mathbf{y}_1 \rightarrow \mathbf{y}} \frac{\|\mathbf{R}_{\mathbf{g}}(\mathbf{y}_1)\|}{\|\mathbf{y}_1 - \mathbf{y}\|} &= \lim_{\mathbf{y}_1 \rightarrow \mathbf{y}} \frac{\|D\mathbf{f}(\mathbf{x})^{-1}\mathbf{R}_{\mathbf{f}}(\mathbf{x}_1)\|}{\|\mathbf{y}_1 - \mathbf{y}\|} \leq \lim_{\mathbf{y}_1 \rightarrow \mathbf{y}} \|D\mathbf{f}(\mathbf{x})^{-1}\| \frac{\|\mathbf{R}_{\mathbf{f}}(\mathbf{x}_1)\|}{\|\mathbf{y}_1 - \mathbf{y}\|} \\ &= \|D\mathbf{f}(\mathbf{x})^{-1}\| \lim_{\mathbf{y}_1 \rightarrow \mathbf{y}} \frac{\|\mathbf{R}_{\mathbf{f}}(\mathbf{x}_1)\|}{\|\mathbf{x}_1 - \mathbf{x}\|} \frac{\|\mathbf{x}_1 - \mathbf{x}\|}{\|\mathbf{y}_1 - \mathbf{y}\|}. \end{aligned}$$

D'un autre côté, comme on l'a vu au point précédent

$$\|\mathbf{x}_1 - \mathbf{x}\| \leq 2\|D\mathbf{f}(\mathbf{x}_0)^{-1}\| \|\mathbf{y}_1 - \mathbf{y}\|,$$

et donc

$$\lim_{\mathbf{y}_1 \rightarrow \mathbf{y}} \frac{\|\mathbf{R}_{\mathbf{g}}(\mathbf{y}_1)\|}{\|\mathbf{y}_1 - \mathbf{y}\|} \leq 2\|D\mathbf{f}(\mathbf{x})^{-1}\| \|D\mathbf{f}(\mathbf{x}_0)^{-1}\| \lim_{\mathbf{x}_1 \rightarrow \mathbf{x}} \frac{\|\mathbf{R}_{\mathbf{f}}(\mathbf{x}_1)\|}{\|\mathbf{x}_1 - \mathbf{x}\|} = 0.$$

On conclut que $\mathbf{g} : V \rightarrow U$ est différentiable en tout $\mathbf{y} \in V$ et $D\mathbf{g}(\mathbf{y}) = D\mathbf{f}(\mathbf{x})^{-1}$, avec $\mathbf{x} = \mathbf{g}(\mathbf{y})$. Enfin, puisque $D\mathbf{f}(\mathbf{x})$ est continue en tout $\mathbf{x} \in B(\mathbf{x}_0, r)$, $D\mathbf{f}(\mathbf{x})^{-1}$ est continue pourvu que $\det(D\mathbf{f}(\mathbf{x})) \neq 0$, ce qui est vrai dans $B(\mathbf{x}_0, r)$. Ceci montre que $\mathbf{g} \in C^1(V, \mathbb{R}^n)$. \square

6.3 Hypersurfaces et fonctions implicites

Considérons une fonction continue $\phi : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$. Le graphe de la fonction ϕ , qu'on appellera Σ par la suite, donné par $\Sigma = \mathcal{G}(\phi) = \{\mathbf{z} = (\mathbf{x}, y) \in U \times \mathbb{R} : y = \phi(\mathbf{x})\}$ représente une *surface* de \mathbb{R}^3 . Plus généralement, si on a une fonction continue $\phi : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$, le graphe $\Sigma = \mathcal{G}(\phi) \subset \mathbb{R}^{n+1}$ sera une (*hyper*)*surface* de \mathbb{R}^{n+1} .

Le fait d'avoir une représentation explicite de l'hypersurface $\Sigma \subset \mathbb{R}^{n+1}$ comme graphe d'une fonction $\phi : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ nous permet de définir facilement certaines quantités locales comme, par exemple, l'hyperplan tangent ou un vecteur normal à l'hypersurface en un point $\mathbf{z}_0 = (\mathbf{x}_0, \phi(\mathbf{x}_0)) \in \Sigma$, si la fonction ϕ est de classe C^1 . En effet, la fonction ϕ est bien approchée, dans un voisinage de \mathbf{x}_0 , par l'application affine

$$T_\phi^1(\mathbf{x}) = \underbrace{\phi(\mathbf{x}_0)}_{=y_0} + \sum_{i=1}^n \frac{\partial \phi}{\partial x_i}(\mathbf{x}_0)(x_i - x_{0,i})$$

dont le graphe est l'hyperplan $\Pi_{\mathbf{z}_0}(\Sigma) = \{\mathbf{z} \in \mathbb{R}^{n+1} : (-D\phi(\mathbf{x}_0), 1) \cdot (\mathbf{z} - \mathbf{z}_0) = 0\}$ appelé l'**hyperplan tangent** à la surface Σ au point $\mathbf{z}_0 = (\mathbf{x}_0, \phi(\mathbf{x}_0))$. Le vecteur $\mathbf{n} = (-D\phi(\mathbf{x}_0), 1) \in \mathbb{R}^{n+1}$ est un **vecteur normal** à la surface Σ au point \mathbf{z}_0 , étant un vecteur normal à l'hyperplan tangent à Σ en \mathbf{z}_0 (voir la définition 3.9 du Chapitre 3). On parle, alors, d'une *surface différentiable* ou bien d'une *variété différentiable*. De plus, si ϕ est de classe C^2 , on peut introduire des notions de *courbure* de la surface au point $\mathbf{z}_0 \in \Sigma$, liées à la matrice hessienne de la fonction ϕ en \mathbf{x}_0 . On ne va pas détailler plus ces notions dans ce cours.

Considérons maintenant un ensemble $\Sigma \subset \mathbb{R}^{n+1}$. On peut se poser la question si, autour d'un point $\mathbf{z}_0 \in \Sigma$, l'ensemble peut être représenté localement comme le graphe d'une fonction continue $\phi : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Si ceci est le cas, on dit que Σ est une hypersurface de \mathbb{R}^{n+1} localement autour de \mathbf{z}_0 .

Définition 6.14. Soit $\Sigma \subset \mathbb{R}^{n+1}$, $\mathbf{z}_0 \in \Sigma$ et $k \in \mathbb{N}$. On dit que Σ est une hypersurface de classe C^k autour de \mathbf{z}_0 si elle est le graphe d'une fonction de classe C^k localement autour de \mathbf{z}_0 , c.-à-d., s'il existe un voisinage V de \mathbf{z}_0 , un indice $i \in \{1, \dots, n+1\}$, un ouvert $U \subset \mathbb{R}^n$ et une fonction $\phi : U \rightarrow \mathbb{R}$ de classe C^k tels que

$$\Sigma \cap V = \mathcal{G}(\phi) = \{\mathbf{x} \in \mathbb{R}^{n+1} : x_i = \phi(\mathbf{x}_{\sim i}), \mathbf{x}_{\sim i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1}) \in U\}.$$

On dit que Σ est une hypersurface de classe C^k si elle est le graphe d'une fonction de classe C^k localement autour de chacun de ses points.

On s'intéresse par la suite à des ensembles définis par

$$\Sigma = \{\mathbf{x} \in E : f(\mathbf{x}) = 0\}, \quad \text{avec } f : E \subset \mathbb{R}^{n+1} \rightarrow \mathbb{R} \text{ régulière (à préciser),}$$

c.-à-d. que Σ est la courbe de niveau zéro de la fonction f définie sur l'ouvert $E \subset \mathbb{R}^{n+1}$. Si Σ est une hypersurface, localement autour d'un point $\mathbf{z}_0 \in \Sigma$, alors elle peut être représentée comme le graphe d'une fonction $\phi : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ localement autour de \mathbf{z}_0 , c'est-à-dire il existe un ouvert $V \subset \mathbb{R}^{n+1}$ contenant \mathbf{z}_0 , un ouvert $U \subset \mathbb{R}^n$ et un indice $i \in \{1, \dots, n+1\}$ tels que

$$\begin{aligned}\Sigma \cap V &= \{\mathbf{x} \in \mathbb{R}^{n+1} : x_i = \phi(\mathbf{x}_{\sim i}), \mathbf{x}_{\sim i} \in U\} \\ &= \{\mathbf{x} \in V : f(x_1, \dots, x_{n+1}) = 0\}.\end{aligned}$$

On dit dans ce cas que l'équation $f(\mathbf{x}) = 0$ définit *implicitement* une fonction $x_i = \phi(\mathbf{x}_{\sim i})$ localement autour du point \mathbf{z}_0 . Autrement dit, la relation

$$f(x_1, \dots, x_{n+1}) = 0 \quad (6.1)$$

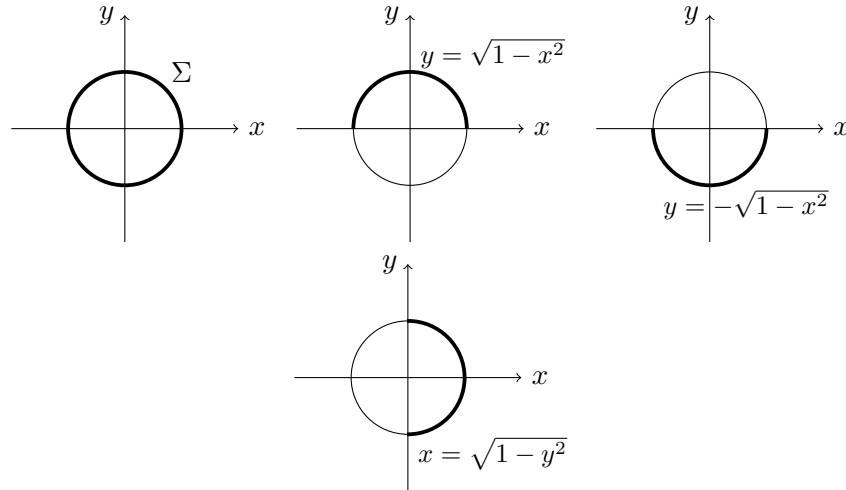
permet d'exprimer la variable x_i en fonction des autres variables, $x_i = \phi(\mathbf{x}_{\sim i})$, et le graphe de ϕ coïncide avec l'ensemble des zéros de f dans un voisinage de \mathbf{z}_0 . On se pose alors la question de savoir quand l'équation (6.1) peut être explicitée par rapport à une des variables.

6.4 Théorème des fonctions implicites – cas scalaire

Regardons plus en détail le cas d'une fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ de deux variables réelles.

Exemple 6.15. Soit $\Sigma = \{(x, y) \in \mathbb{R}^2 : f(x, y) = x^2 - y = 0\}$. L'équation $f(x, y) = 0$ définit implicitement la fonction $y = \phi(x) = x^2, \forall x \in \mathbb{R}$. C'est à dire, pour tout $x \in \mathbb{R}$, $f(x, \phi(x)) = x^2 - \phi(x) = 0$ et $\mathcal{G}(\phi) = \Sigma$.

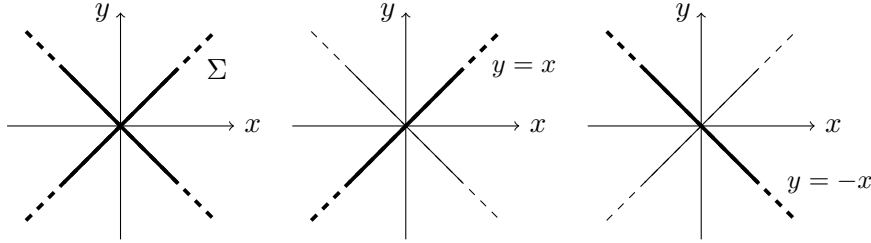
Exemple 6.16. Soit $\Sigma = \{(x, y) \in \mathbb{R}^2 : f(x, y) = x^2 + y^2 - 1 = 0\}$. Clairement, l'ensemble Σ correspond au cercle unitaire.



Les quatre fonctions $y = \pm\sqrt{1-x^2}$ et $x = \pm\sqrt{1-y^2}$ sont définies implicitement par $f(x, y) = 0$. Toutefois, on ne peut pas trouver une seule fonction $y = \phi(x)$ ou $x = \phi(y)$ qui décrit tout l'ensemble Σ . En revanche, soit $\mathbf{z}_0 \in \Sigma$ arbitraire. Alors, autour de \mathbf{z}_0 , on peut définir une fonction $y = \phi(x)$ ou $x = \phi(y)$. Par exemple, soit $\mathbf{z}_0 = (x_0, y_0) \in \Sigma$.

- Si $y_0 > 0$, alors on peut prendre $y = \phi(x) = \sqrt{1-x^2}$ et on a $f(x, \phi(x)) = 0$ dans un voisinage de x_0 et $\mathcal{G}(\phi)$ coïncide avec Σ dans un voisinage de \mathbf{z}_0 .
- Si $y_0 < 0$ alors on peut prendre $y = \phi(x) = -\sqrt{1-x^2}$ et on a $f(x, \phi(x)) = 0$ dans un voisinage de x_0 et $\mathcal{G}(\phi)$ coïncide avec Σ dans un voisinage de \mathbf{z}_0 .
- Si $y_0 = 0$, par exemple $\mathbf{z}_0 = (1, 0)$, alors on peut expliciter x en fonction de y , $x = \phi(y) = \sqrt{1-y^2}$ mais non pas y en fonction de x .

Exemple 6.17. Soit $\Sigma = \{(x, y) \in \mathbb{R}^2 : f(x, y) = x^2 - y^2 = 0\}$. Clairement, l'équation $f(x, y) = 0$ est satisfaite si et seulement si $x = y$ ou $x = -y$.



Soit $\mathbf{z}_0 = (x_0, y_0) \in \Sigma$.

- Si $x_0 y_0 > 0$ alors $f(x, y) = 0$ définit implicitement la fonction $y = \phi(x) = x$ (ou bien la fonction $x = \phi(y) = y$) et $\mathcal{G}(\phi)$ coïncide avec Σ dans un voisinage de \mathbf{z}_0 .
- Pareillement, si $x_0 y_0 < 0$ alors $f(x, y) = 0$ définit implicitement la fonction $y = \phi(x) = -x$ (ou bien la fonction $x = \phi(y) = -y$) et $\mathcal{G}(\phi)$ coïncide avec Σ dans un voisinage de \mathbf{z}_0 .
- En revanche, si $(x_0, y_0) = (0, 0)$, il n'est pas possible de trouver ni une fonction $y = \phi(x)$ ni une fonction $x = \phi(y)$ définies implicitement par l'équation $f(x, y) = 0$ qui décrivent Σ dans un voisinage de $\mathbf{z}_0 = (0, 0)$. On dit dans ce cas que $(0, 0)$ est un point singulier de Σ . On remarque que $\nabla f(0, 0) = (0, 0)$.

Exemple 6.18. Soit $\Sigma = \{(x, y) \in \mathbb{R}^2 : f(x, y) = xe^y + ye^x = 0\}$. L'équation $f(x, y) = xe^y + ye^x = 0$ ne peut pas être explicitée sous forme simple ni par rapport à x ni par rapport à y . Est-ce que l'équation $f(x, y) = 0$ définit implicitement une fonction $y = \phi(x)$ ou $x = \phi(y)$ au moins localement autour de chaque point $\mathbf{z}_0 \in \Sigma$, dont le graphe coïncide avec Σ dans un voisinage de \mathbf{z}_0 ? On verra que la réponse à cette question est positive.

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ et $\mathbf{z}_0 = (x_0, y_0)$ tel que $f(x_0, y_0) = 0$. Supposons que $f(x, y) = 0$ définit implicitement une fonction $y = \phi(x)$ autour de \mathbf{z}_0 , c'est-à-dire $\exists \delta > 0$ et $\phi :]x_0 - \delta, x_0 + \delta[\rightarrow \mathbb{R}$ tels que $y_0 = \phi(x_0)$ et $f(x, \phi(x)) = 0, \forall x \in]x_0 - \delta, x_0 + \delta[$. Supposons de plus que f et ϕ sont de classe C^1 et notons $\tilde{f}(x) = f(x, \phi(x))$. Alors, par la formule de dérivation de fonctions composées, on a :

$$0 = \frac{d}{dx} \tilde{f}(x) = \frac{d}{dx} f(x, \phi(x)) = \frac{\partial f}{\partial x}(x, \phi(x)) + \frac{\partial f}{\partial y}(x, \phi(x)) \phi'(x), \quad \forall x \in]x_0 - \delta, x_0 + \delta[.$$

On en tire que si $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$, alors

$$\phi'(x_0) = -\frac{\frac{\partial f}{\partial x}(x_0, y_0)}{\frac{\partial f}{\partial y}(x_0, y_0)}.$$

Plus généralement, si $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$, alors $\frac{\partial f}{\partial y}(x, \phi(x)) \neq 0$ pour tout x suffisamment proche de x_0 et

$$\phi'(x) = -\frac{\frac{\partial f}{\partial x}(x, \phi(x))}{\frac{\partial f}{\partial y}(x, \phi(x))}$$

Même si on ne connaît pas ϕ , on peut quand même évaluer sa dérivée $\phi'(x_0)$ si $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$. De plus, si f et ϕ sont de classe C^2 , on peut itérer le raisonnement :

$$\begin{aligned} 0 &= \frac{d^2}{dx^2} f(x, \phi(x)) \\ &= \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x}(x, \phi(x)) + \frac{\partial f}{\partial y}(x, \phi(x)) \phi'(x) \right) \\ &= \frac{\partial^2 f}{\partial x^2}(x, \phi(x)) + 2 \frac{\partial^2 f}{\partial x \partial y}(x, \phi(x)) \phi'(x) + \frac{\partial^2 f}{\partial y^2}(x, \phi(x)) (\phi'(x))^2 + \frac{\partial f}{\partial y}(x, \phi(x)) \phi''(x). \end{aligned}$$

Par conséquent,

$$\phi''(x_0) = -\frac{1}{\frac{\partial f}{\partial y}(x_0, y_0)} \left(\frac{\partial^2 f}{\partial x^2}(x_0, y_0) + 2 \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \phi'(x_0) + \frac{\partial^2 f}{\partial y^2}(x_0, y_0) (\phi'(x_0))^2 \right).$$

et, pour tout x suffisamment proche de x_0 ,

$$\phi''(x) = -\frac{1}{\frac{\partial f}{\partial y}(x, \phi(x))} \left(\frac{\partial^2 f}{\partial x^2}(x, \phi(x)) + 2 \frac{\partial^2 f}{\partial x \partial y}(x, \phi(x)) \phi'(x) + \frac{\partial^2 f}{\partial y^2}(x, \phi(x)) (\phi'(x))^2 \right).$$

Plus généralement, si f est de classe C^k , alors ϕ sera aussi de classe C^k et on peut calculer explicitement $\phi^{(k)}(x_0)$. Tous les calculs précédents sont valables sous l'hypothèse qu'une fonction implicite ϕ de classe C^k existe et que $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$. Cette dernière condition s'avère être suffisante pour l'existence d'une fonction implicite.

Théorème 6.19 (des fonctions implicites — cas $n = 2$). *Soit $f : E \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, E ouvert non vide, de classe C^1 , $\Sigma = \{(x, y) \in E : f(x, y) = 0\}$ et $\mathbf{z}_0 = (x_0, y_0) \in \Sigma$ t.q. $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$. Alors il existe un voisinage $U =]x_0 - \delta, x_0 + \delta[$ de x_0 , un ouvert $V \subset E$ contenant \mathbf{z}_0 et une unique fonction $\phi : U \rightarrow \mathbb{R}$ de classe C^1 tels que*

- $y_0 = \phi(x_0)$;
- $(x, \phi(x)) \in V$ et $f(x, \phi(x)) = 0$, $\forall x \in U$;
- $\mathcal{G}(\phi) = \Sigma \cap V$.

De plus, pour tout $x \in U$, $\frac{\partial f}{\partial y}(x, \phi(x)) \neq 0$ et

$$\phi'(x) = -\frac{\frac{\partial f}{\partial x}(x, \phi(x))}{\frac{\partial f}{\partial y}(x, \phi(x))},$$

et si $f \in C^k(E)$, alors $\phi \in C^k(U)$.

Démonstration. Supposons $\frac{\partial f}{\partial y}(x_0, y_0) > 0$ (le cas $\frac{\partial f}{\partial y}(x_0, y_0) < 0$ est identique). Puisque $\frac{\partial f}{\partial y}$ est continue sur E (f étant de classe C^1), il existe $\delta_1, \delta_2 > 0$ tels que, pour tout $(x, y) \in W := [x_0 - \delta_1, x_0 + \delta_1] \times [y_0 - \delta_2, y_0 + \delta_2]$, on a

$$(x, y) \in E \quad \text{et} \quad \frac{\partial f}{\partial y}(x, y) > 0.$$

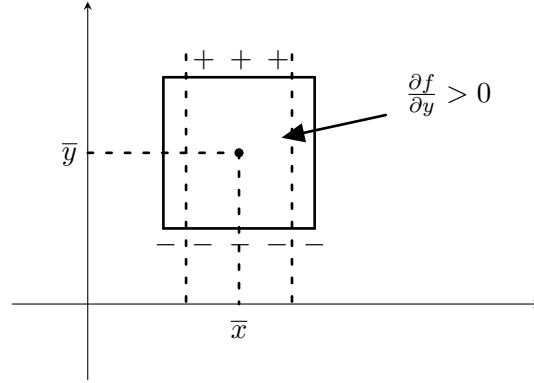
Pour tout $x \in [x_0 - \delta_1, x_0 + \delta_1]$, la fonction $y \mapsto g_x(y) = f(x, y)$ est strictement croissante dans l'intervalle $[y_0 - \delta_2, y_0 + \delta_2]$ car $g'_x(y) = \frac{\partial f}{\partial y}(x, y) > 0$. En particulier

$$g_{x_0}(y_0 - \delta_2) < g_{x_0}(y_0) = f(x_0, y_0) = 0 < g_{x_0}(y_0 + \delta_2).$$

Puisque f est continue, il existe $0 < \delta \leq \delta_1$ tel que

$$g_x(y_0 - \delta_2) = f(x, y_0 - \delta_2) < 0 \quad \text{et} \quad g_x(y_0 + \delta_2) = f(x, y_0 + \delta_2) > 0, \quad \forall x \in U :=]x_0 - \delta, x_0 + \delta[$$

comme illustré dans la figure ci dessous.



Pour $x \in U$, $g_x(y)$ est continue et strictement croissante sur l'intervalle $[y_0 - \delta_2, y_0 + \delta_2]$, et donc il existe un unique $y = \phi(x) \in]y_0 - \delta_2, y_0 + \delta_2[$ tel que $g_x(y) = f(x, y) = 0$. Cette procédure permet de définir de façon unique une fonction $\phi : U \rightarrow \mathbb{R}$ telle que

$$\phi(x_0) = y_0 \quad \text{et} \quad f(x, \phi(x)) = 0 \quad \forall x \in U.$$

De plus, si on note $V =]x_0 - \delta, x_0 + \delta[\times]y_0 - \delta_2, y_0 + \delta_2[$, on a $\mathcal{G}(\phi) = \{(x, y) \in V : f(x, y) = 0\} = \Sigma \cap V$.

Continuité de ϕ . Soit $\bar{x} \in U$ et $\bar{y} = \phi(\bar{x})$. Pour tout $\epsilon \in]0, \delta_2 - |\bar{y} - y_0|]$, considérons $W_\epsilon = [x_0 - \delta_1, x_0 + \delta_1] \times [\bar{y} - \epsilon, \bar{y} + \epsilon]$. Puisque $W_\epsilon \subset W$ on a que $\frac{\partial f}{\partial y} > 0$ sur W_ϵ . En raisonnant comme auparavant, mais sur W_ϵ au lieu de W , il existe $\delta_\epsilon > 0$ et une fonction $\bar{\phi} :]\bar{x} - \delta_\epsilon, \bar{x} + \delta_\epsilon[\cap U \rightarrow \mathbb{R}$ tels que $\bar{y} = \bar{\phi}(\bar{x})$ et $f(x, \bar{\phi}(x)) = 0, \forall x \in]\bar{x} - \delta_\epsilon, \bar{x} + \delta_\epsilon[\cap U$. Par l'unicité de la fonction implicite sur W et le choix de W_ϵ , on a $\bar{\phi}(x) = \phi(x) \in [\bar{y} - \epsilon, \bar{y} + \epsilon]$, $\forall x \in]\bar{x} - \delta_\epsilon, \bar{x} + \delta_\epsilon[\cap U$. On a donc montré que pour tout $\epsilon > 0$ il existe $\delta_\epsilon > 0$ tel que $|\phi(x) - \phi(\bar{x})| \leq \epsilon$ pour tout $x \in]\bar{x} - \delta_\epsilon, \bar{x} + \delta_\epsilon[\cap U$ ce qui montre la continuité de ϕ en tout point $\bar{x} \in U$.

Continuité de ϕ' . Soient $x_1 \neq x_2$ dans U , $y_1 = \phi(x_1)$ et $y_2 = \phi(x_2)$. Puisque $f \in C^1(E)$, on a

$$0 = f(x_2, y_2) - f(x_1, y_1) = \frac{\partial f}{\partial x}(\xi, \eta)(x_2 - x_1) + \frac{\partial f}{\partial y}(\xi, \eta)(y_2 - y_1)$$

avec $(\xi, \eta) = (x_1 + \theta(x_2 - x_1), y_1 + \theta(y_2 - y_1))$ et $\theta \in]0, 1[$. Alors

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{\phi(x_2) - \phi(x_1)}{x_2 - x_1} = -\frac{\frac{\partial f}{\partial x}(\xi, \eta)}{\frac{\partial f}{\partial y}(\xi, \eta)}$$

et $\frac{\partial f}{\partial y}(\xi, \eta) > 0$ car $(\xi, \eta) \in V \subset W$. Il s'ensuit

$$\phi'(x_1) = \lim_{x_2 \rightarrow x_1} \frac{\phi(x_2) - \phi(x_1)}{x_2 - x_1} = -\frac{\frac{\partial f}{\partial x}(x_1, y_1)}{\frac{\partial f}{\partial y}(x_1, y_1)},$$

donc ϕ est dérivable en tout $x \in U$. De plus, $\phi'(x) = -\frac{\frac{\partial f}{\partial x}(x, \phi(x))}{\frac{\partial f}{\partial y}(x, \phi(x))}$ est continue grâce au fait que $\frac{\partial f}{\partial y}$, $\frac{\partial f}{\partial x}$ et ϕ sont continues et $\frac{\partial f}{\partial y}(x, \phi(x)) \neq 0$, $\forall x \in U$. La démonstration que $\phi \in C^k(U)$ si $f \in C^k(E)$ se fait par récurrence sur $k \geq 1$. \square

On a montré le théorème des fonctions implicites pour une fonction de deux variables réelles $(x, y) \mapsto f(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$. Toutefois, la démonstration de ce théorème se généralise sans difficulté au cas d'une fonction de n variables $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

Théorème 6.20 (des fonctions implicites à plusieurs variables). *Soit $E \subset \mathbb{R}^{n+1}$ ouvert non vide, $f : E \rightarrow \mathbb{R}$ de classe C^1 , $\Sigma = \{(\mathbf{x}, y) \in E, \mathbf{x} \in \mathbb{R}^n, y \in \mathbb{R} : f(\mathbf{x}, y) = 0\}$ et $\mathbf{z}_0 = (\mathbf{x}_0, y_0) \in \Sigma$ tel que $\frac{\partial f}{\partial y}(\mathbf{x}_0, y_0) \neq 0$. Alors, il existe un ouvert $U = B(\mathbf{x}_0, \delta) \subset \mathbb{R}^n$ de \mathbf{x}_0 , un ouvert $V \subset E$ contenant \mathbf{z}_0 et une unique application $\phi : U \rightarrow \mathbb{R}$ de classe C^1 telle que :*

- $y_0 = \phi(\mathbf{x}_0)$;
- $(\mathbf{x}, \phi(\mathbf{x})) \in V$ et $f(\mathbf{x}, \phi(\mathbf{x})) = 0$, $\forall \mathbf{x} \in U$;
- $\mathcal{G}(\phi) = \Sigma \cap V$, (i.e. le graphe de ϕ décrit Σ dans un voisinage de \mathbf{z}_0).

De plus, pour tout $\mathbf{x} \in U$, $\frac{\partial f}{\partial y}(\mathbf{x}, \phi(\mathbf{x})) \neq 0$ et

$$\frac{\partial \phi}{\partial x_i}(\mathbf{x}) = -\frac{\frac{\partial f}{\partial x_i}(\mathbf{x}, \phi(\mathbf{x}))}{\frac{\partial f}{\partial y}(\mathbf{x}, \phi(\mathbf{x}))},$$

et si $f \in C^k(E)$ alors $\phi \in C^k(U)$.

Le théorème précédent montre que si $\frac{\partial f}{\partial y}(\mathbf{z}_0) \neq 0$, l'ensemble Σ coïncide avec le graphe d'une fonction $y = \phi(\mathbf{x})$ dans un voisinage de \mathbf{z}_0 . Ceci permet de définir l'hyperplan tangent à Σ en \mathbf{z}_0 , noté $\Pi_{\mathbf{z}_0}(\Sigma)$, comme l'hyperplan tangent au graphe de ϕ en \mathbf{x}_0 , qui est donné par

$$\Pi_{\mathbf{z}_0}(\Sigma) = \left\{ (\mathbf{x}, y) \in \mathbb{R}^{n+1} : y = \phi(\mathbf{x}_0) + \sum_{i=1}^n \frac{\partial \phi}{\partial x_i}(\mathbf{x}_0)(x_i - x_{0,i}) \right\}.$$

En se rappelant de l'expression des dérivées partielles de ϕ en \mathbf{x}_0 , l'hyperplan tangent peut s'écrire de façon équivalente comme

$$\frac{\partial f}{\partial y}(\mathbf{z}_0)(y - y_0) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{z}_0)(x_i - x_{0,i}) = 0,$$

qui conduit à l'expression simple

$$\Pi_{\mathbf{z}_0}(\Sigma) = \{\mathbf{z} \in \mathbb{R}^{n+1} : \nabla f(\mathbf{z}_0) \cdot (\mathbf{z} - \mathbf{z}_0) = 0\}. \quad (6.2)$$

Cette équation montre que le vecteur $\nabla f(\mathbf{z}_0)$ est un vecteur normal à Σ en \mathbf{z}_0 .

Enfin, on remarque que si $\frac{\partial f}{\partial y}(\mathbf{z}_0) = 0$ mais $\nabla f(\mathbf{z}_0) \neq \mathbf{0}$, alors il existe un $i \in \{1, \dots, n\}$ tel que $\frac{\partial f}{\partial x_i}(\mathbf{z}_0) \neq 0$ et on peut encore appliquer le théorème des fonctions implicites en exprimant la variable x_i en fonction des autres variables. L'expression de l'hyperplan tangent (6.2) et du vecteur normal sont inchangées.

6.5 Théorème des fonctions implicites – cas vectoriel

Les idées présentées dans la section précédente se généralisent au cas d'une fonction continue à valeurs vectorielles $\mathbf{f} : E \subset \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$, $(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{f}(\mathbf{x}, \mathbf{y})$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^m$, où E est ouvert non vide. Soit $\Sigma = \{(\mathbf{x}, \mathbf{y}) \in E : \mathbf{f}(\mathbf{x}, \mathbf{y}) = \mathbf{0}\}$ l'ensemble des solutions de l'équation $\mathbf{f}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$. Cette équation correspond au système sous-déterminé des m équations

$$\begin{cases} f_1(x_1, \dots, x_n, y_1, \dots, y_m) = 0 \\ \vdots \\ f_m(x_1, \dots, x_n, y_1, \dots, y_m) = 0 \end{cases} \quad (6.3)$$

des $n + m$ inconnues $(x_1, \dots, x_n, y_1, \dots, y_m)$. Du point de vue géométrique,

$$\Sigma = \bigcap_{i=1}^m \Sigma_i, \quad \text{où } \Sigma_i = \{(\mathbf{x}, \mathbf{y}) \in E : f_i(\mathbf{x}, \mathbf{y}) = 0\}.$$

Si chaque ensemble Σ_i est une surface de \mathbb{R}^{n+m} , alors Σ est l'intersection de m surfaces. Par exemple, si $f_1(x, y, z) = 0$ et $f_2(x, y, z) = 0$ définissent des surfaces en \mathbb{R}^3 , leur intersection $\Sigma = \{(x, y, z) \in \mathbb{R}^3 : f_1(x, y, z) = 0 \text{ et } f_2(x, y, z) = 0\}$ sera en générale une courbe de \mathbb{R}^3 (voir Figure 6.1). On peut se poser de nouveau la question de savoir si l'ensemble Σ peut être représenté comme le graphe d'une fonction continue, au moins localement autour de chaque point $\mathbf{z}_0 \in \Sigma$. De façon équivalente, on veut savoir si pour l'équation $\mathbf{f}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ on peut exprimer m variables, disons (y_1, \dots, y_m) , comme fonctions continues des n variables restantes (x_1, \dots, x_n) , dans un voisinage de $\mathbf{z}_0 = (\mathbf{x}_0, \mathbf{y}_0)$; autrement dit, si le système (6.3) définit implicitement une fonction continue $\phi : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ autour de \mathbf{x}_0 telle que $\mathbf{f}(\mathbf{x}, \phi(\mathbf{x})) = \mathbf{0}$, $\forall \mathbf{x} \in U$ et $\Sigma = \mathcal{G}(\phi)$ dans un voisinage de \mathbf{z}_0 .

Avant de donner le résultat d'existence générale, il est utile de considérer le cas d'une fonction \mathbf{f}_a affine :

$$\mathbf{f}_a(\mathbf{x}, \mathbf{y}) = A_1 \mathbf{x} + A_2 \mathbf{y} + \mathbf{b}, \quad A_1 \in \mathbb{R}^{m \times n}, A_2 \in \mathbb{R}^{m \times m}, \mathbf{b} \in \mathbb{R}^m.$$

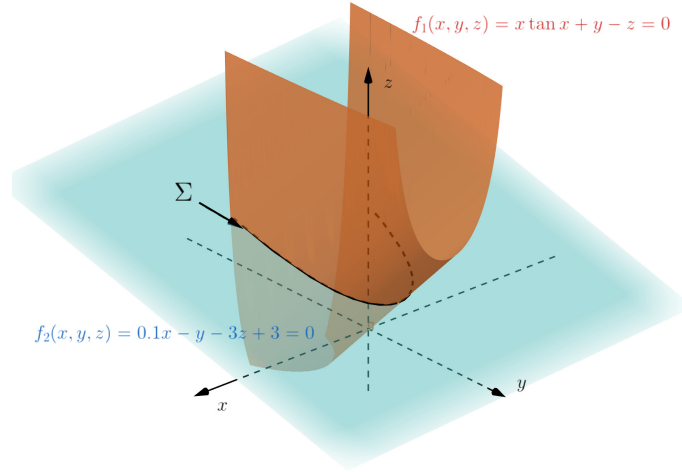


FIGURE 6.1 – Courbe $\Sigma = \{(x, y, z) \in \mathbb{R}^3 : f_1(x, y, z) = 0 \text{ et } f_2(x, y, z) = 0\}$ obtenue par l'intersection des deux surfaces $f_1(x, y, z) = x \tan x + y - z = 0$ et $f_2(x, y, z) = \frac{x}{10} - y - 3z + 3 = 0$.

L'équation $\mathbf{f}_a = \mathbf{0}$ équivaut à $A_2 \mathbf{y} = -(A_1 \mathbf{x} + \mathbf{b})$. Donc on peut écrire de façon unique \mathbf{y} en fonction de \mathbf{x} si et seulement si la matrice A_2 est inversible, c.-à-d. $\det(A_2) \neq 0$. Dans ce cas on a

$$\mathbf{y} = -A_2^{-1} A_1 \mathbf{x} - A_2^{-1} \mathbf{b}.$$

Considérons maintenant une fonction *non-linéaire* $\mathbf{f} : E \subset \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ de classe C^1 et $\mathbf{z}_0 = (\mathbf{x}_0, \mathbf{y}_0) \in \Sigma = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^m : (\mathbf{x}, \mathbf{y}) \in E, \mathbf{f}(\mathbf{x}, \mathbf{y}) = \mathbf{0}\}$. La différentiabilité de \mathbf{f} en $\mathbf{z}_0 \in E$ assure que l'on peut écrire pour tout $\mathbf{z} \in E$

$$\mathbf{f}(\mathbf{z}) = \mathbf{f}(\mathbf{z}_0) + D\mathbf{f}(\mathbf{z}_0)(\mathbf{z} - \mathbf{z}_0) + \mathbf{R}_f(\mathbf{z}), \quad \lim_{\mathbf{z} \rightarrow \mathbf{z}_0} \frac{\mathbf{R}_f(\mathbf{z})}{\|\mathbf{z} - \mathbf{z}_0\|} = \mathbf{0}.$$

Il est pratique d'écrire la matrice jacobienne $D\mathbf{f}(\mathbf{z}_0) \in \mathbb{R}^{m \times (n+m)}$ par blocs comme

$$D\mathbf{f}(\mathbf{z}_0) = \left[\begin{array}{ccc|ccc} \frac{\partial f_1}{\partial x_1}(\mathbf{z}_0) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{z}_0) & \frac{\partial f_1}{\partial y_1}(\mathbf{z}_0) & \cdots & \frac{\partial f_1}{\partial y_m}(\mathbf{z}_0) \\ \vdots & & & & & \\ \frac{\partial f_m}{\partial x_1}(\mathbf{z}_0) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{z}_0) & \frac{\partial f_m}{\partial y_1}(\mathbf{z}_0) & \cdots & \frac{\partial f_m}{\partial y_m}(\mathbf{z}_0) \end{array} \right] = \left[D_{\mathbf{x}}\mathbf{f}(\mathbf{z}_0) \mid D_{\mathbf{y}}\mathbf{f}(\mathbf{z}_0) \right].$$

Donc, dans un voisinage de \mathbf{z}_0 , la fonction \mathbf{f} est bien approchée par la fonction affine

$$\mathbf{f}_a(\mathbf{z}) = D\mathbf{f}(\mathbf{z}_0)(\mathbf{z} - \mathbf{z}_0) = D_{\mathbf{x}}\mathbf{f}(\mathbf{z}_0)(\mathbf{x} - \mathbf{x}_0) + D_{\mathbf{y}}\mathbf{f}(\mathbf{z}_0)(\mathbf{y} - \mathbf{y}_0)$$

et on s'attend à pouvoir exprimer \mathbf{y} en fonction de \mathbf{x} si $\det(D_{\mathbf{y}}\mathbf{f}(\mathbf{z}_0)) \neq 0$. Le théorème suivant rend ce raisonnement rigoureux.

Théorème 6.21. Soit $E \subset \mathbb{R}^{n+m}$ ouvert non vide, $\mathbf{f} \in C^1(E, \mathbb{R}^m)$, $\Sigma = \{(\mathbf{x}, \mathbf{y}) \in E : \mathbf{f}(\mathbf{x}, \mathbf{y}) = \mathbf{0}\}$ et $\mathbf{z}_0 = (\mathbf{x}_0, \mathbf{y}_0) \in \Sigma$ tel que $\det(D_{\mathbf{y}}\mathbf{f}(\mathbf{z}_0)) \neq 0$. Alors il existe une boule ouverte $U = B(\mathbf{x}_0, \delta) \subset \mathbb{R}^n$, un ouvert $V \subset E$ contenant \mathbf{z}_0 et une unique fonction $\phi : U \rightarrow \mathbb{R}^m$ de classe C^1 tels que

1. $\mathbf{y}_0 = \phi(\mathbf{x}_0)$;
2. pour tout $\mathbf{x} \in U$, $(\mathbf{x}, \phi(\mathbf{x})) \in V$ et $\mathbf{f}(\mathbf{x}, \phi(\mathbf{x})) = \mathbf{0}$;
3. $\Sigma \cap V = \mathcal{G}(\phi)$;
4. $\forall \mathbf{x} \in U$, $\det(D_{\mathbf{y}}\mathbf{f}(\mathbf{x}, \phi(\mathbf{x}))) \neq 0$ et $D\phi(\mathbf{x}) = -(D_{\mathbf{y}}\mathbf{f}(\mathbf{x}, \phi(\mathbf{x})))^{-1}D_{\mathbf{x}}\mathbf{f}(\mathbf{x}, \phi(\mathbf{x}))$, (dans cette dernière formule, il y a un produit matriciel).

De plus, pour tout entier $k \geq 1$, si \mathbf{f} est de classe C^k , alors ϕ l'est aussi.

Démonstration. On utilise le théorème d'inversion locale. Soit $\mathbf{F} : E \rightarrow \mathbb{R}^{n+m}$, $\mathbf{F}(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{f}(\mathbf{x}, \mathbf{y}))$. On a $\mathbf{F}(\mathbf{x}_0, \mathbf{y}_0) = (\mathbf{x}_0, \mathbf{0})$ et

$$D\mathbf{F}(\mathbf{x}_0, \mathbf{y}_0) = \begin{bmatrix} I_{n \times n} & 0 \\ D_{\mathbf{x}}\mathbf{f}(\mathbf{z}_0) & D_{\mathbf{y}}\mathbf{f}(\mathbf{z}_0) \end{bmatrix}, \quad \det(D\mathbf{F}(\mathbf{x}_0, \mathbf{y}_0)) = \det(D_{\mathbf{y}}\mathbf{f}(\mathbf{z}_0)) \neq 0.$$

Donc il existe un ouvert $V' \subset E$ contenant $(\mathbf{x}_0, \mathbf{y}_0)$ et un ouvert U' contenant $(\mathbf{x}_0, \mathbf{0})$ tels que \mathbf{F} est un difféomorphisme de V' à U' . Soit \mathbf{G} la fonction inverse. On peut toujours trouver une boule ouverte $U = B(\mathbf{x}_0, \delta) \subset \mathbb{R}^n$ et un ouvert $W \subset \mathbb{R}^m$ contenant $\mathbf{0}$ tels que $U \times W \subset U'$. On considère alors la restriction $\mathbf{F} : V \rightarrow U \times W$ où $V = \mathbf{G}(U \times W)$. La fonction inverse a la forme $\mathbf{G}(\mathbf{x}, \mathbf{w}) = (\mathbf{x}, \varphi(\mathbf{x}, \mathbf{w}))$ avec $\varphi(\mathbf{x}_0, \mathbf{0}) = \mathbf{y}_0$. On note, en particulier, que si $(\mathbf{x}, \mathbf{y}) \in \Sigma \cap V$, alors $\mathbf{F}(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{0})$ et $\mathbf{x} \in U$. En prenant la fonction inverse on a $\mathbf{G}(\mathbf{x}, \mathbf{0}) = (\mathbf{x}, \varphi(\mathbf{x}, \mathbf{0})) = (\mathbf{x}, \mathbf{y})$. La fonction implicite cherchée est alors $\phi(\mathbf{x}) = \varphi(\mathbf{x}, \mathbf{0}) : U \rightarrow \mathbb{R}^m$. En effet :

$$\forall \mathbf{x} \in U, \quad (\mathbf{x}, \mathbf{0}) = \mathbf{F} \circ \mathbf{G}(\mathbf{x}, \mathbf{0}) = \mathbf{F}(\mathbf{x}, \varphi(\mathbf{x}, \mathbf{0})) = \mathbf{F}(\mathbf{x}, \phi(\mathbf{x})) = (\mathbf{x}, \mathbf{f}(\mathbf{x}, \phi(\mathbf{x}))),$$

donc $\mathbf{f}(\mathbf{x}, \phi(\mathbf{x})) = \mathbf{0}$, $\forall \mathbf{x} \in U$, et

$$\forall (\mathbf{x}, \mathbf{y}) \in \Sigma \cap V, \quad (\mathbf{x}, \mathbf{y}) = \mathbf{G} \circ \mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{G}(\mathbf{x}, \underbrace{\mathbf{f}(\mathbf{x}, \mathbf{y})}_{=\mathbf{0}}) = (\mathbf{x}, \varphi(\mathbf{x}, \mathbf{0})) = (\mathbf{x}, \phi(\mathbf{x})),$$

donc $\Sigma \cap V = \mathcal{G}(\phi)$.

Montrons l'unicité de ϕ : si $\phi^* : U \rightarrow \mathbb{R}^m$ satisfait aussi $\Sigma \cap V = \mathcal{G}(\phi^*)$, alors $\mathcal{G}(\phi^*) = \mathcal{G}(\phi)$ et donc $\phi^* = \phi$.

Finalement, ϕ est de classe C^1 puisque \mathbf{G} est de classe C^1 , \mathbf{F} étant un difféomorphisme. De plus, $0 \neq \det(D\mathbf{F}(\mathbf{x}, \mathbf{y})) = \det(D_{\mathbf{y}}\mathbf{f}(\mathbf{x}, \mathbf{y})) \forall (\mathbf{x}, \mathbf{y}) \in V$, donc $D_{\mathbf{y}}\mathbf{f}$ est inversible sur V . Par la formule de dérivation des fonctions composées, on a

$$0 = D(\mathbf{f}(\mathbf{x}, \phi(\mathbf{x}))) = D_{\mathbf{x}}\mathbf{f}(\mathbf{x}, \phi(\mathbf{x})) + D_{\mathbf{y}}\mathbf{f}(\mathbf{x}, \phi(\mathbf{x})) \cdot D\phi(\mathbf{x}).$$

d'où

$$D\phi(\mathbf{x}) = -(D_{\mathbf{y}}\mathbf{f}(\mathbf{x}, \phi(\mathbf{x})))^{-1}D_{\mathbf{x}}\mathbf{f}(\mathbf{x}, \phi(\mathbf{x})).$$

La démonstration que $\phi \in C^k(U)$ si $\mathbf{f} \in C^k(E)$ se fait par récurrence sur $k \geq 1$. □

Remarque 6.22. Dans le théorème 6.21, la décomposition du vecteur $\mathbf{z} \in \mathbb{R}^{n+m}$ en $\mathbf{z} = (\mathbf{x}, \mathbf{y})$ est arbitraire. Le théorème reste valable sous la condition plus générale :

$$\text{Rang}(D\mathbf{f}(\mathbf{z}_0)) = m.$$

En effet, dans ce cas on sait qu'il existe m colonnes de $D\mathbf{f}(\mathbf{z}_0)$ linéairement indépendantes. Notons ces colonnes i_1, \dots, i_m . On peut alors définir $(y_1, \dots, y_m) = (z_{i_1}, \dots, z_{i_m})$ et (x_1, \dots, x_n) les variables restantes. On aura donc $\det(D_{\mathbf{y}}\mathbf{f}(\mathbf{z}_0)) \neq 0$ et on peut appliquer le théorème pour exprimer ces m variables \mathbf{y} en fonction des autres n variables \mathbf{x} .

Sous les hypothèses du théorème 6.21, grâce à l'existence d'une fonction implicite on peut définir l'hyperplan tangent à Σ en \mathbf{z}_0 :

$$\begin{aligned}\Pi_{\mathbf{z}_0}(\Sigma) &= \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{n+m} : \mathbf{y} - \mathbf{y}_0 - D\phi(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) = \mathbf{0}\} \\ &= \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{n+m} : D\mathbf{f}(\mathbf{z}_0)(\mathbf{z} - \mathbf{z}_0) = \mathbf{0}\}\end{aligned}$$

On remarque que $\Pi_{\mathbf{z}_0}(\Sigma)$ est l'ensemble des points de \mathbb{R}^{n+m} qui satisfont m équations linéaires. Puisque $D\mathbf{f}(\mathbf{z}_0)$ a rang maximal, $\Pi_{\mathbf{z}_0}(\Sigma)$ est un sous-espace affine de \mathbb{R}^{n+m} de dimension n .

Chapitre 7

Extrema de fonctions réelles

Soit $E \subset \mathbb{R}^n$ pas forcément ouvert pour le moment. On considère ici une fonction $f : E \rightarrow \mathbb{R}$ à valeurs dans \mathbb{R} (fonction scalaire) et on se pose la question de trouver le maximum et le minimum de f sur E s'ils existent. Ceci est un problème d'*optimisation*. On commence par donner la définition de minimum et maximum local ou global, ainsi que de point de minimum/maximum local/global.

Définition 7.1 (Minima et maxima locaux et globaux). *Soit $E \subset \mathbb{R}^n$ et $\mathbf{x}^* \in E$.*

— *On dit que f admet un maximum local (ou relatif) au point $\mathbf{x}^* \in E$ s'il existe $\delta > 0$:*

$$f(\mathbf{x}) \leq f(\mathbf{x}^*), \quad \forall \mathbf{x} \in B(\mathbf{x}^*, \delta) \cap E.$$

Si l'inégalité est stricte (i.e. $f(\mathbf{x}) < f(\mathbf{x}^)$, $\forall \mathbf{x} \in B(\mathbf{x}^*, \delta) \cap E$, $\mathbf{x} \neq \mathbf{x}^*$) alors le maximum local est strict. Le point \mathbf{x}^* est appelé point de maximum local (strict) pour f .*

— *On dit que f admet un minimum local (ou relatif) au point $\mathbf{x}^* \in E$ s'il existe $\delta > 0$:*

$$f(\mathbf{x}) \geq f(\mathbf{x}^*), \quad \forall \mathbf{x} \in B(\mathbf{x}^*, \delta) \cap E.$$

Le minimum est strict si l'inégalité est stricte. Le point \mathbf{x}^ est appelé point de minimum local (strict) pour f .*

— *Par extremum local (strict) on entend un minimum ou maximum local (strict).*

— *On dit que f admet un maximum (global ou absolu), resp. minimum (global ou absolu) au point $\mathbf{x}^* \in E$ si $f(\mathbf{x}) \leq f(\mathbf{x}^*)$, $\forall \mathbf{x} \in E$, resp. $f(\mathbf{x}) \geq f(\mathbf{x}^*)$, $\forall \mathbf{x} \in E$. Le maximum/minimum est strict si l'inégalité est stricte*

7.1 Extrema libres

On considère d'abord le cas où l'ensemble E est ouvert et la fonction $f : E \rightarrow \mathbb{R}$ différentiable (une ou plusieurs fois) sur E .

Rappelons ce qu'on sait dire sur la caractérisation des extrema locaux pour une fonction d'une seule variable réelle $f : I \rightarrow \mathbb{R}$, définie sur un ouvert $I \subset \mathbb{R}$. Soit $x^* \in I$.

- Si f est dérivable en x^* et x^* est un point d'extremum local de f , alors $f'(x^*) = 0$ (condition nécessaire du premier ordre).
- Si f est deux fois dérivable sur I et x^* est un point de minimum (resp. maximum) local, alors $f'(x^*) = 0$ et $f''(x^*) \geq 0$ (resp. $f''(x^*) \leq 0$) (condition nécessaire du second ordre).
- Soit f deux fois dérivable sur I et x^* un point stationnaire de f , c.-à-d. que $f'(x^*) = 0$. Si $f''(x^*) > 0$ (resp. $f''(x^*) < 0$) alors x^* est un point de minimum (resp. maximum) local strict (condition suffisante).
- Si $f'(x^*) = f''(x^*) = 0$, on ne peut rien conclure. Pour décider si x^* est un point de minimum / maximum / inflexion, il faut regarder les dérivées d'ordres supérieurs (si elles existent) en x^* , ou bien étudier le signe de la fonction $g(x) = f(x) - f(x^*)$ autour de x^* . Par exemple, si on peut montrer que g est non négative dans un voisinage de x^* , alors x^* est un point de minimum local de f . Par contre, si g change de signe en x^* , alors x^* est un point d'inflexion (sous l'hypothèse $f'(x^*) = 0$).

Considérons maintenant le cas d'une fonction de plusieurs variables $f : E \rightarrow \mathbb{R}$, avec $E \subset \mathbb{R}^n$ ouvert non vide. Soit $\mathbf{x}^* \in E$ (point intérieur car E est ouvert). On peut regarder le comportement de f le long des droites : soit $\mathbf{v} \in \mathbb{R}^n$, $\|\mathbf{v}\| = 1$ et $g_{\mathbf{v}}(t) = f(\mathbf{x}^* + t\mathbf{v})$. Si f est différentiable sur E et \mathbf{x}^* est un point d'extremum local de f , alors pour tout $\mathbf{v} \in \mathbb{R}^n$, la fonction $g_{\mathbf{v}}$ est dérivable et admet un extremum local en $t = 0$ (voir Figure 7.1). Il s'ensuit

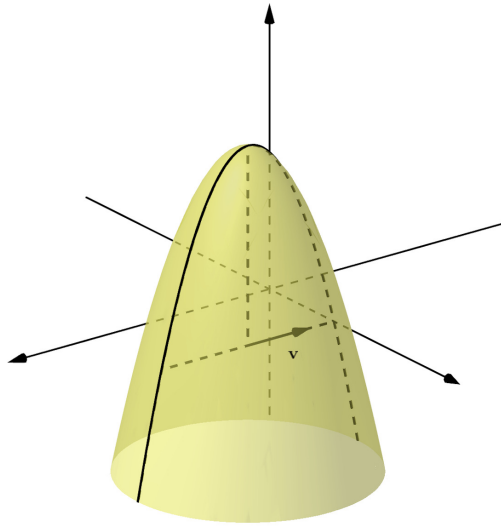


FIGURE 7.1 – Si \mathbf{x}^* est point de maximum, alors la fonction $g_{\mathbf{v}}(t) = f(\mathbf{x}^* + t\mathbf{v})$ à maximum en $t = 0$ pour tout $\mathbf{v} \in \mathbb{R}^n$.

$$\forall \mathbf{v} \in \mathbb{R}^n \quad 0 = g'_{\mathbf{v}}(0) = D_{\mathbf{v}}f(\mathbf{x}^*) = \nabla f(\mathbf{x}^*) \cdot \mathbf{v}$$

(produit scalaire entre deux vecteurs colonnes) ce qui implique $\nabla f(\mathbf{x}^*) = \mathbf{0}$. On s'attend alors à ce que $\nabla f(\mathbf{x}^*) = \mathbf{0}$ soit une condition *nécessaire* pour que f admette un extremum local en \mathbf{x}^* . Le théorème suivant formalise cette idée.

Définition 7.2 (Point stationnaire). Soit $E \subset \mathbb{R}^n$ ouvert non vide, $\mathbf{x}^* \in E$ et $f : E \rightarrow \mathbb{R}$ différentiable en \mathbf{x}^* . On dit que \mathbf{x}^* est un point stationnaire de f si $\nabla f(\mathbf{x}^*) = \mathbf{0}$ (i.e. $\frac{\partial f}{\partial x_i}(\mathbf{x}^*) = 0$, $i = 1, \dots, n$ ou, de façon équivalente, $Df(\mathbf{x}^*) = \mathbf{0}$).

Théorème 7.3 (condition nécessaire de premier ordre). Soit $E \subset \mathbb{R}^n$ un ouvert non vide, et $f : E \rightarrow \mathbb{R}$ différentiable en $\mathbf{x}^* \in E$ et admettant un extremum local en \mathbf{x}^* . Alors \mathbf{x}^* est un point stationnaire de f , c.-à-d. que $\nabla f(\mathbf{x}^*) = \mathbf{0}$.

Démonstration. Puisque E est ouvert, $\exists \delta > 0 : B(\mathbf{x}^*, \delta) \subset E$. Alors $\mathbf{x}^* + t\mathbf{e}_j \in E$ pour tout $t \in]-\delta, \delta[$ et $g_j(t) = f(\mathbf{x}^* + t\mathbf{e}_j) :]-\delta, \delta[\rightarrow \mathbb{R}$ est dérivable en $t = 0$. Puisque \mathbf{x}^* est un point d'extremum local pour f , 0 est un point d'extremum local pour g_j et $g'_j(0) = \frac{\partial f}{\partial x_j}(\mathbf{x}^*) = 0$. \square

On considère maintenant le cas $f \in C^2(E)$, qui assure que la matrice hessienne $H_f(\mathbf{x}^*)$ est symétrique. Soit \mathbf{x}^* un point de minimum local de f . Alors $t = 0$ est un point de minimum local de $g_{\mathbf{v}}$ pour tout $\mathbf{v} \in \mathbb{R}^n$ de norme 1. Il s'ensuit que

$$0 \leq g''_{\mathbf{v}}(0) = D^2_{\mathbf{v}\mathbf{v}}f(\mathbf{x}^*) = \mathbf{v}^\top H_f(\mathbf{x}^*)\mathbf{v}$$

et ainsi $\mathbf{v}^\top H_f(\mathbf{x}^*)\mathbf{v} \geq 0$ pour tout $\mathbf{v} \in \mathbb{R}^n$. Donc, la condition $\mathbf{v}^\top H_f(\mathbf{x}^*)\mathbf{v} \geq 0$, $\forall \mathbf{v} \in \mathbb{R}^n$, est aussi nécessaire pour que f admette un minimum local en \mathbf{x}^* , pourvu que f soit de classe C^2 sur E . De même, la condition $\mathbf{v}^\top H_f(\mathbf{x}^*)\mathbf{v} \leq 0$, $\forall \mathbf{v} \in \mathbb{R}^n$, est nécessaire pour que f admette un maximum local en \mathbf{x}^* , pourvu que f soit de classe C^2 sur E . On a donc démontré

Théorème 7.4 (condition nécessaire du second ordre). Soit $E \subset \mathbb{R}^n$ un ouvert non vide et $f : E \rightarrow \mathbb{R}$ de classe C^2 sur E , admettant un minimum (resp. maximum) local en $\mathbf{x}^* \in E$. Alors \mathbf{x}^* est un point stationnaire de f et $\mathbf{v}^\top H_f(\mathbf{x}^*)\mathbf{v} \geq 0$ (resp. $\mathbf{v}^\top H_f(\mathbf{x}^*)\mathbf{v} \leq 0$) pour tout $\mathbf{v} \in \mathbb{R}^n$.

Si l'inégalité dans la condition précédente est stricte, i.e. si $\mathbf{v}^\top H_f(\mathbf{x}^*)\mathbf{v} > 0$, $\forall \mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, la condition devient suffisante pour que f admette un minimum en \mathbf{x}^* , pourvu que \mathbf{x}^* soit un point stationnaire. Ceci est montré dans le théorème 7.8 ci-dessous. Avant de présenter le théorème, on rappelle quelques notions d'algèbre linéaire sur les matrices définies positives ou négatives.

Définition 7.5. On dit qu'une matrice $A \in \mathbb{R}^{n \times n}$ est

- définie positive (ou simplement « positive ») si $\mathbf{x}^\top A\mathbf{x} > 0$, $\forall \mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n$;
- semi-définie positive si $\mathbf{x}^\top A\mathbf{x} \geq 0$, $\forall \mathbf{x} \in \mathbb{R}^n$;
- définie négative (ou simplement « négative ») si $\mathbf{x}^\top A\mathbf{x} < 0$, $\forall \mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n$;
- semi-définie négative si $\mathbf{x}^\top A\mathbf{x} \leq 0$, $\forall \mathbf{x} \in \mathbb{R}^n$;
- indéfinie s'il existe $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$: $\mathbf{x}^\top A\mathbf{x} > 0$ et $\mathbf{y}^\top A\mathbf{y} < 0$.

À toute matrice $A \in \mathbb{R}^{n \times n}$ on peut associer une forme quadratique

$$Q_A(\mathbf{x}) = \mathbf{x}^\top A\mathbf{x} = \sum_{i,j=1}^n A_{ij}x_i x_j, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Lemme 7.6. Une matrice $A \in \mathbb{R}^{n \times n}$ est définie positive si et seulement si

$$\exists c > 0 : \quad \mathbf{x}^\top A \mathbf{x} \geq c \|\mathbf{x}\|^2, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

(On peut prendre n'importe quelle norme $\|\cdot\|$ sur \mathbb{R}^n ; la constante c dépendra de la norme choisie.)

Démonstration. Soit A définie positive et Q_A la forme quadratique associée à A . Clairement Q_A est une fonction continue sur \mathbb{R}^n . De plus, $\forall t \in \mathbb{R}$, $Q_A(t\mathbf{x}) = t^2 Q_A(\mathbf{x})$. Considérons l'ensemble compact $\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = 1\}$. Q_A admet un maximum et un minimum sur \mathcal{S} . Soit $c = \min_{\mathbf{x} \in \mathcal{S}} Q_A(\mathbf{x})$. Clairement, $c > 0$ car $Q_A(\mathbf{x}) > 0$, $\forall \mathbf{x} \in \mathcal{S}$, et

$$\mathbf{x}^\top A \mathbf{x} = Q_A(\mathbf{x}) = \|\mathbf{x}\|^2 Q_A\left(\frac{\mathbf{x}}{\|\mathbf{x}\|}\right) \geq c \|\mathbf{x}\|^2, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Inversement, s'il existe $c > 0$ tel que $\mathbf{x}^\top A \mathbf{x} \geq c \|\mathbf{x}\|^2$, $\forall \mathbf{x} \in \mathbb{R}^n$, on conclut immédiatement que A est définie positive. \square

Dans ce cours on appliquera toujours la Définition 7.5 et le Lemme 7.6 à des matrices *symétriques*. Le Lemme suivant donne une caractérisation alternative de matrice définie positive sous hypothèse de symétrie.

Lemme 7.7. Soit $A \in \mathbb{R}^{n \times n}$ une matrice symétrique. A est définie positive si et seulement si toutes les valeurs propres $\lambda_1, \dots, \lambda_n$ de A sont positives. De plus $\mathbf{x}^\top A \mathbf{x} \geq \lambda_{\min} \|\mathbf{x}\|_2^2$, où $\lambda_{\min} = \min\{\lambda_i : 1 \leq i \leq n\}$.

Démonstration. On sait de l'algèbre linéaire qu'une matrice $A \in \mathbb{R}^{n \times n}$ symétrique est toujours diagonalisable avec valeurs propres réelles $\lambda_1, \dots, \lambda_n$ et n vecteurs propres orthonormés $\mathbf{v}_1, \dots, \mathbf{v}_n$ (de norme euclidienne 1 et deux à deux orthogonaux ; chacun est écrit sous forme colonne dans ce qui suit) :

$$AV = VD, \quad D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}, \quad \lambda_i \in \mathbb{R}, \quad V = [\mathbf{v}_1, \dots, \mathbf{v}_n] : V^\top V = VV^\top = I.$$

Si A est définie positive, c'est-à-dire $\mathbf{x}^\top A \mathbf{x} > 0$, $\forall \mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n$, en particulier l'inégalité de l'énoncé est vraie pour $\mathbf{x} = \mathbf{v}_j$ et

$$\mathbf{v}_j^\top A \mathbf{v}_j = \lambda_j \mathbf{v}_j^\top \mathbf{v}_j = \lambda_j > 0, \quad \forall j = 1, \dots, n.$$

Inversement, supposons $\lambda_j > 0$, $\forall j = 1, \dots, n$. Puisque $\{\mathbf{v}_j\}_{j=1}^n$ forme une base orthonormée de \mathbb{R}^n , on peut écrire de façon unique tout vecteur $\mathbf{x} \in \mathbb{R}^n$ comme $\mathbf{x} = \sum_{j=1}^n \beta_j \mathbf{v}_j = V\boldsymbol{\beta}$ ($\boldsymbol{\beta}$ étant un vecteur colonne) et $\|\mathbf{x}\|_2 = \|\boldsymbol{\beta}\|_2$. Donc $\mathbf{x}^\top A \mathbf{x} = \boldsymbol{\beta}^\top V^\top A V \boldsymbol{\beta} = \boldsymbol{\beta}^\top D \boldsymbol{\beta} = \sum_{j=1}^n \lambda_j \beta_j^2 \geq \lambda_{\min} \|\boldsymbol{\beta}\|_2^2 = \lambda_{\min} \|\mathbf{x}\|_2^2$, avec égalité si $\mathbf{x} = \mathbf{v}_{\min}$, le vecteur propre correspondant à λ_{\min} . \square

Le Lemme précédent montre que la plus grande constante c possible du Lemme 7.6 est $c = \lambda_{\min}$ si la matrice A est symétrique et si on utilise la norme euclidienne. On revient maintenant à la question de caractériser les points d'extremum local d'une fonction $f : E \rightarrow \mathbb{R}$.

Théorème 7.8 (condition suffisante pour extrema locaux). *Soit $E \subset \mathbb{R}$ ouvert non vide, $f : E \rightarrow \mathbb{R}$ une fonction de classe C^2 sur E et $\mathbf{x}^* \in E$ un point stationnaire de f , c'est-à-dire $\nabla f(\mathbf{x}^*) = \mathbf{0}$. Si $H_f(\mathbf{x}^*)$ est définie positive (resp. définie négative) alors f a un minimum (resp. maximum) local strict en \mathbf{x}^* .*

Démonstration. Considérons le cas $H_f(\mathbf{x}^*)$ définie positive (la démonstration pour $H_f(\mathbf{x}^*)$ définie négative est la même). Puisque f est de classe C^2 sur E on peut écrire un développement limité d'ordre 2 de f au point \mathbf{x}^* :

$$f(\mathbf{x}) = f(\mathbf{x}^*) + \underbrace{\nabla f(\mathbf{x}^*) \cdot (\mathbf{x} - \mathbf{x}^*)}_{=0} + \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^\top H_f(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) + R_f(\mathbf{x}), \quad \forall \mathbf{x} \in E,$$

où on a utilisé que $\nabla f(\mathbf{x}^*) = \mathbf{0}$ (\mathbf{x}^* est un point stationnaire) et où $R_f(\mathbf{x})$ est tel que $\lim_{\mathbf{x} \rightarrow \mathbf{x}^*} \frac{R_f(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}^*\|^2} = 0$. Puisque $H_f(\mathbf{x}^*)$ est définie positive, $\exists c > 0 : \mathbf{v}^\top H_f(\mathbf{x}^*)\mathbf{v} \geq c\|\mathbf{v}\|^2$ pour tout $\mathbf{v} \in \mathbb{R}^n$. De plus, puisque $\lim_{\mathbf{x} \rightarrow \mathbf{x}^*} \frac{R_f(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}^*\|^2} = 0$, il existe $\delta > 0$ tel que, pour tout $\mathbf{x} \in B(\mathbf{x}^*, \delta) \setminus \{\mathbf{x}^*\}$, on a $\mathbf{x} \in E$ et $|R_f(\mathbf{x})| \leq \frac{c}{4}\|\mathbf{x} - \mathbf{x}^*\|^2$. Donc

$$\begin{aligned} \forall \mathbf{x} \in B(\mathbf{x}^*, \delta) \setminus \{\mathbf{x}^*\}, \quad f(\mathbf{x}) &= f(\mathbf{x}^*) + \frac{1}{2} \underbrace{(\mathbf{x} - \mathbf{x}^*)^\top H_f(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*)}_{\geq c\|\mathbf{x} - \mathbf{x}^*\|^2} + \underbrace{R_f(\mathbf{x})}_{\geq -\frac{c}{4}\|\mathbf{x} - \mathbf{x}^*\|^2} \\ &\geq f(\mathbf{x}^*) + \frac{c}{4}\|\mathbf{x} - \mathbf{x}^*\|^2 > f(\mathbf{x}^*), \end{aligned}$$

ce qui montre que \mathbf{x}^* est un point de minimum local strict de f . □

D'après ce théorème, si la matrice hessienne est définie (positive ou négative) en un point stationnaire, on peut conclure que ce point stationnaire est un point d'extremum local. On se pose alors la question de savoir la nature d'un point stationnaire lorsque la matrice hessienne en ce point est semi-définie ou indéfinie. On donne d'abord la définition suivante.

Définition 7.9 (Point selle). *Soit $f : E \rightarrow \mathbb{R}$, où $E \subset \mathbb{R}^n$ est ouvert non vide et f est différentiable en un certain point stationnaire $\mathbf{x}^* \in E$. On dit que \mathbf{x}^* est un point selle de f si*

$$\forall \delta > 0 \exists \mathbf{x}, \mathbf{y} \in B(\mathbf{x}^*, \delta) \cap E \quad \left(f(\mathbf{x}) > f(\mathbf{x}^*) \text{ et } f(\mathbf{y}) < f(\mathbf{x}^*) \right).$$

Avec cette définition et le résultat du théorème 7.8, on a la classification suivante des points stationnaires $\{\mathbf{x}^* \in E : \nabla f(\mathbf{x}^*) = \mathbf{0}\}$, où les $\lambda_i(H_f(\mathbf{x}^*))$ sont les valeurs propres de $H_f(\mathbf{x}^*)$ ($i = 1, \dots, n$).

- Si $H_f(\mathbf{x}^*)$ est *définie positive* (i.e. $\mathbf{x}^\top H_f(\mathbf{x}^*)\mathbf{x} > 0$, $\forall \mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n$ ou, de façon équivalente, $\lambda_i(H_f(\mathbf{x}^*)) > 0$, $\forall i = 1, \dots, n$), alors \mathbf{x}^* est un point de minimum local strict. Lorsque $n = 2$, $H_f(\mathbf{x}^*)$ est définie positive ssi à la fois la trace de $H_f(\mathbf{x}^*)$ est > 0 et $\det(H_f(\mathbf{x}^*)) > 0$.
- Si $H_f(\mathbf{x}^*)$ est *définie négative* (i.e. $\mathbf{x}^\top H_f(\mathbf{x}^*)\mathbf{x} < 0$, $\forall \mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n$ ou, de façon équivalente, $\lambda_i(H_f(\mathbf{x}^*)) < 0$, $\forall i = 1, \dots, n$), alors \mathbf{x}^* est un point de maximum local strict. Lorsque $n = 2$, $H_f(\mathbf{x}^*)$ est définie négative ssi à la fois la trace de $H_f(\mathbf{x}^*)$ est < 0 et $\det(H_f(\mathbf{x}^*)) > 0$.

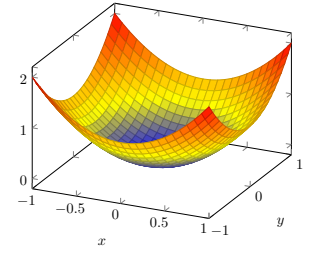
- Si $H_f(\mathbf{x}^*)$ est *indéfinie* (i.e. $\exists \mathbf{x}, \mathbf{y} \in \mathbb{R}^n : \mathbf{x}^\top H_f(\mathbf{x}^*) \mathbf{x} > 0, \mathbf{y}^\top H_f(\mathbf{x}^*) \mathbf{y} < 0$ ou, de façon équivalente, $\exists i, j = 1, \dots, n : \lambda_i(H_f(\mathbf{x}^*)) > 0$ et $\lambda_j(H_f(\mathbf{x}^*)) < 0$) alors \mathbf{x}^* est un point selle. Lorsque $n = 2$, une condition suffisante pour que $H_f(\mathbf{x}^*)$ soit indéfinie est que $\det(H_f(\mathbf{x}^*)) < 0$.
- Si $H_f(\mathbf{x}^*)$ est *seulement* semi-définie positive (ou négative), on ne peut pas conclure sur la nature du point stationnaire à partir des résultats précédents.

Exemple 7.10. Considérons les trois fonctions $f_1, f_2, f_3 : \mathbb{R}^2 \rightarrow \mathbb{R}$

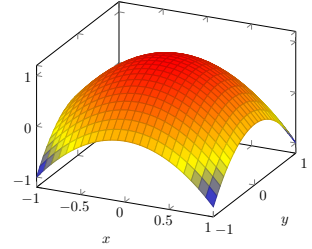
$$f_1(x, y) = x^2 + y^2, \quad f_2(x, y) = 1 - x^2 - y^2, \quad f_3(x, y) = x^2 - y^2.$$

Pour toutes les trois fonctions, le seul point stationnaire est $\mathbf{x}^* = (0, 0)$. Étudions sa nature dans les trois cas.

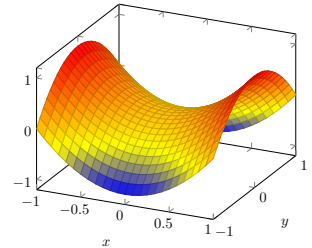
Fonction f_1 : $H_{f_1}(0, 0) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ définie positive



Fonction f_2 : $H_{f_2}(0, 0) = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$ définie négative



Fonction f_3 : $H_{f_3}(0, 0) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$ indéfinie



Exercice 7.11. Trouver les extrema locaux de la fonction $f(x, y) = x^3 + y^3 - 3x - 12y + 1$ et les caractériser. Est-ce que cette fonction admet un maximum / minimum global ?

Exercice 7.12. Montrer que la fonction $f(x, y) = 3x^2 + 3y^2 - 2xy - 8(x + y - 1)$ a un seul point de minimum local qui est aussi un minimum global.

Si $H_f(\mathbf{x}^*)$ est seulement semi-définie positive (ou négative), on ne peut pas conclure que \mathbf{x}^* est un point de minimum local (ou maximum local). De plus, même si $g_{\mathbf{v}}(t) = f(\mathbf{x}^* + t\mathbf{v})$ a un minimum local en $t = 0$ pour tout $\mathbf{v} \in \mathbb{R}^n$, on ne peut pas conclure que f admet un minimum local en \mathbf{x}^* comme l'exemple suivant le montre.

Exemple 7.13. *Considérons la fonction $f(x, y) = y^2 - 3x^2y + 2x^4$. On a*

$$\nabla f(x, y) = \begin{bmatrix} -6xy + 8x^3 \\ 2y - 3x^2 \end{bmatrix}, \quad H_f(x, y) = \begin{bmatrix} -6y + 24x^2 & -6x \\ -6x & 2 \end{bmatrix}.$$

Donc

$$\nabla f(x, y) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \begin{cases} -6xy + 8x^3 = 0 \\ y = \frac{3}{2}x^2 \end{cases} \iff (x, y) = (0, 0)$$

et $P = (0, 0)$ est le seul point stationnaire. De plus $H_f(0, 0) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$ est semi-définie positive. Pour toute droite $\mathbf{x} = (x, y) = t\mathbf{v}$, $t \in \mathbb{R}$, $\mathbf{v} \neq (1, 0)$, on a

$$\mathbf{v}^\top H_f(0, 0) \mathbf{v} = (v_1, v_2) H_f(0, 0) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 2v_2^2 > 0,$$

donc la fonction $g_{\mathbf{v}}(t) = f(t\mathbf{v})$ a un minimum local strict en $t = 0$. De plus pour $\mathbf{v} = (1, 0)$, $g_{\mathbf{v}}(t) = f(t(1, 0)) = 2t^4$ a un minimum local strict en $t = 0$, donc $\forall \mathbf{v} \in \mathbb{R}^2$, $g_{\mathbf{v}}(t) = f(t\mathbf{v})$ a un minimum local strict $t = 0$. Toutefois, si on prend $x = t$ et $y = \frac{3}{2}t^2$ on a

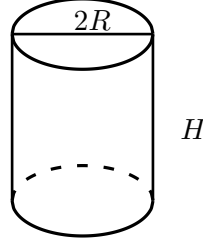
$$f(t, \frac{3}{2}t^2) = \frac{9}{4}t^4 - \frac{9}{2}t^4 + 2t^4 = -\frac{1}{4}t^4$$

qui a un maximum local strict en $t = 0$! Donc $(0, 0)$ n'est pas un point d'extremum local de f .

7.2 Extrema liés

Il est souvent le cas dans les applications, qu'on cherche à trouver le minimum (ou maximum) d'une fonction $\min_{\mathbf{z} \in E} f(\mathbf{z})$ mais les variables $\mathbf{z} = (z_1, \dots, z_n)$ ne peuvent pas être choisies arbitrairement et sont liées par des contraintes.

Exemple 7.14. *Considérons une canette de forme cylindrique. On veut trouver la forme optimale qui minimise la surface (qui requiert le minimum de matériel) tout en gardant un volume constant. Soit $\Sigma_{R,H} = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 \leq R^2, 0 \leq z \leq H\}$ un cylindre plein de hauteur $H > 0$ et de rayon $R > 0$.*



Ce cylindre a un volume $V(R, H) = \pi R^2 H$ et une surface $S(R, H) = 2\pi R^2 + 2\pi RH$. Le problème d'optimisation prend la forme suivante : pour $\bar{V} > 0$ donné, discuter

$$\min_{R, H > 0} S(R, H) \quad \text{sous la contrainte } V(R, H) = \bar{V}.$$

Ceci est un problème de minimisation sous contrainte.

Soit $E \subset \mathbb{R}^n$ ouvert non vide et $f, g : E \rightarrow \mathbb{R}$ de classe C^1 . On pose le problème de minimisation (maximisation) sous contrainte suivant

$$\min_{\mathbf{z} \in E} f(\mathbf{z}) \quad \text{sous la contrainte } g(\mathbf{z}) = 0.$$

Si on dénote par $\Sigma_g = \{\mathbf{z} \in E : g(\mathbf{z}) = 0\}$ l'ensemble des points qui satisfont la contrainte (appelé aussi *ensemble faisable* ou *admissible*), le problème de minimisation sous contrainte peut s'écrire de façon équivalente comme

$$\min_{\mathbf{z} \in \Sigma_g} f(\mathbf{z}).$$

Déterminer si le minimum est fini et effectivement atteint fait partie de la discussion.

Définition 7.15. On dit que $\mathbf{z}^* \in \Sigma_g$ est un point de minimum local de f sur Σ_g si

$$\exists \delta > 0 \quad \forall \mathbf{z} \in B(\mathbf{z}^*, \delta) \cap \Sigma_g \quad f(\mathbf{z}^*) \leq f(\mathbf{z})$$

Le minimum est strict si l'inégalité est stricte, dans le même sens qu'expliqué plus haut.

De la même façon, on dit que $\mathbf{z}^* \in \Sigma_g$ est un point de maximum local de f sur Σ_g si

$$\exists \delta > 0 \quad \forall \mathbf{z} \in B(\mathbf{z}^*, \delta) \cap \Sigma_g \quad f(\mathbf{z}^*) \geq f(\mathbf{z}).$$

Le maximum est strict si l'inégalité est stricte. On utilise aussi la terminologie de minimum/maximum (strict ou non) lié.

On se pose la question de caractériser les points d'extremum (minimum/maximum) liés. Voyons quelques exemples :

Exemple 7.16. On considère le problème de minimisation suivante :

$$\min_{(x, y) \in \mathbb{R}^2} x^2 + y^2 \quad \text{sous la contrainte } x + y - 1 = 0.$$

La fonction $f(x, y) = x^2 + y^2$ à minimiser est convexe et a un minimum global en $(0, 0)$. Toutefois, ce point ne satisfait pas la contrainte. Le minimum lié est caractérisé par le fait que ∇f est perpendiculaire à la contrainte au point de minimum lié. Voir la figure 7.2 (gauche) pour une interprétation graphique.

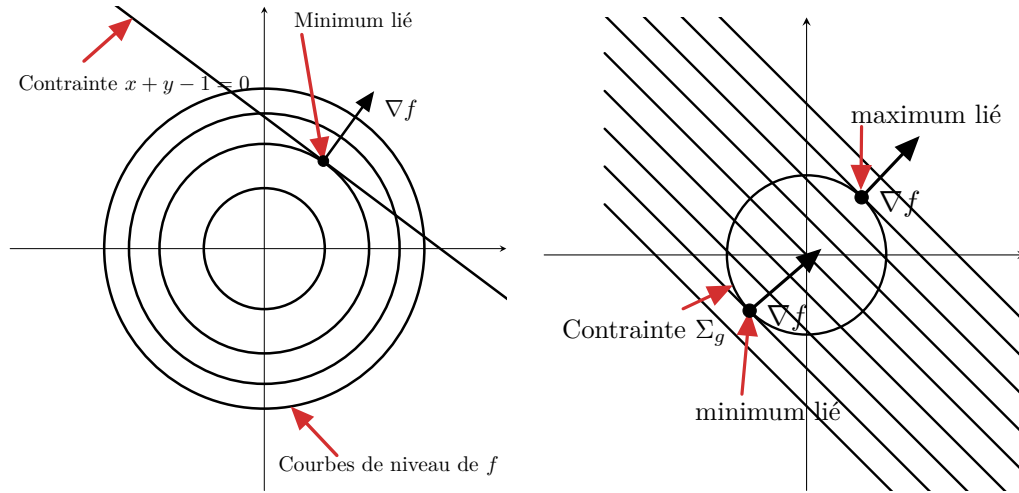


FIGURE 7.2 – Gauche : Problème de minimisation, Exemple 7.16. Droite : Problème de minimisation, Exemple 7.17.

Exemple 7.17. On considère le problème de minimisation suivante :

$$\min_{(x,y) \in \mathbb{R}^2} x + y \quad \text{sous la contrainte } x^2 + y^2 - 1 = 0.$$

La fonction $f(x, y) = x + y$ n'a pas de minimum ou maximum sur \mathbb{R}^2 . Toutefois, l'ensemble $\Sigma_g = \{(x, y) : x^2 + y^2 - 1 = 0\}$ est compact, donc f étant continue, elle atteint son maximum et minimum sur Σ_g . On voit encore qu'aussi bien au point de minimum lié qu'au point de maximum lié, le vecteur ∇f est perpendiculaire à la courbe Σ_g (c'est-à-dire, orthogonal au plan tangent).

Soit $g(x, y) = x^2 + y^2 - 1$ et x_M le point de maximum lié. Puisque $\nabla g = (2y, 2x)^\top \neq \mathbf{0}$, $\forall (x, y) \in \Sigma_g$, l'équation $g = 0$ définit implicitement une fonction $y = \phi(x)$ ou $x = \phi(y)$, et, autour de chaque point $\mathbf{z} \in \Sigma_g$, l'ensemble Σ_g peut être représenté par le graphe d'une fonction. De plus, le vecteur $\nabla g(\mathbf{z})$ est normal au plan tangent $\Pi_{\mathbf{z}}(\Sigma)$ à Σ_g en \mathbf{z} .

Puisqu'au point du maximum lié on a que $\nabla f(\mathbf{x}_M)$ est aussi un vecteur normal au plan tangent $\Pi_{\mathbf{x}_M}(\Sigma)$, il s'ensuit que $\nabla f(\mathbf{x}_M) \parallel \nabla g(\mathbf{x}_M)$, c'est-à-dire $\exists \lambda \in \mathbb{R} : \nabla f(\mathbf{x}_M) = \lambda \nabla g(\mathbf{x}_M)$. Ceci est en fait une condition nécessaire pour avoir un extremum lié comme le théorème suivant le montre.

Théorème 7.18 (Condition nécessaire d'optimalité). Soit $n \geq 2$, $E \subset \mathbb{R}^n$ ouvert non vide, $f, g : E \rightarrow \mathbb{R}$ deux fonctions de classe $C^1(E)$ et $\mathbf{z}^* \in \Sigma_g = \{\mathbf{z} \in E : g(\mathbf{z}) = 0\}$ un point d'extremum local de f sur Σ_g . Alors, si $\nabla g(\mathbf{z}^*) \neq \mathbf{0}$, il existe $\lambda^* \in \mathbb{R}$ tel que $\nabla f(\mathbf{z}^*) = \lambda^* \nabla g(\mathbf{z}^*)$.

Démonstration. Puisque $\nabla g(\mathbf{z}^*) \neq \mathbf{0}$, il y a au moins une composante non nulle, soit $\frac{\partial g}{\partial z_j}(\mathbf{z}^*) \neq 0$. Pour se fixer les idées, supposons que $j = n$, le cas général se traitant de la même manière en permutant le rôle des coordonnées. Notons $y = z_n$ et $\mathbf{x} = (z_1, \dots, z_{n-1}) \in \mathbb{R}^{n-1}$, de sorte que tout point $\mathbf{z} \in E$ puisse s'écrire comme $\mathbf{z} = (\mathbf{x}, y)$; en particulier

$\mathbf{z}^* = (\mathbf{x}^*, y^*)$. Par le théorème des fonctions implicites, il existe $\delta > 0$ et une unique fonction $\phi : B(\mathbf{x}^*, \delta) \subset \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ telle que $\phi(\mathbf{x}^*) = y^*$, $(\mathbf{x}, \phi(\mathbf{x})) \in E$ et $g(\mathbf{x}, \phi(\mathbf{x})) = 0$ sur $B(\mathbf{x}^*, \delta)$, et le graphe de ϕ coïncide avec Σ_g dans un voisinage de \mathbf{z}^* .

Alors la fonction $\tilde{f}(\mathbf{x}) = f(\mathbf{x}, \phi(\mathbf{x}))$ admet un extremum local (libre) en \mathbf{x}^* et $\nabla \tilde{f}(\mathbf{x}^*) = \mathbf{0}$. Donc

$$0 = \frac{\partial \tilde{f}}{\partial x_i}(\mathbf{x}^*) = \frac{\partial f}{\partial x_i}(\mathbf{x}^*, \phi(\mathbf{x}^*)) + \frac{\partial f}{\partial y}(\mathbf{x}^*, \phi(\mathbf{x}^*)) \frac{\partial \phi}{\partial x_i}(\mathbf{x}^*), \quad \forall i = 1, \dots, n-1.$$

D'autre part,

$$\frac{\partial \phi}{\partial x_i}(\mathbf{x}^*) = -\frac{\frac{\partial g}{\partial x_i}(\mathbf{z}^*)}{\frac{\partial g}{\partial y}(\mathbf{z}^*)}.$$

Par conséquent, on a

$$\frac{\partial f}{\partial x_i}(\mathbf{z}^*) - \frac{\partial f}{\partial y}(\mathbf{z}^*) \frac{\frac{\partial g}{\partial x_i}(\mathbf{z}^*)}{\frac{\partial g}{\partial y}(\mathbf{z}^*)} = 0, \quad \forall i = 1, \dots, n-1.$$

Si on pose $\lambda^* = \frac{\frac{\partial f}{\partial y}(\mathbf{z}^*)}{\frac{\partial g}{\partial y}(\mathbf{z}^*)}$, alors

$$\frac{\partial f}{\partial x_i}(\mathbf{z}^*) = \lambda^* \frac{\partial g}{\partial x_i}(\mathbf{z}^*), \quad \forall i = 1, \dots, n-1$$

et, par définition, $\frac{\partial f}{\partial y}(\mathbf{z}^*) = \lambda^* \frac{\partial g}{\partial y}(\mathbf{z}^*)$. On a donc bien

$$\nabla f(\mathbf{z}^*) = \lambda^* \nabla g(\mathbf{z}^*).$$

□

D'après le théorème, une condition nécessaire pour que $\mathbf{z}^* \in \Sigma_g$, avec $\nabla g(\mathbf{z}^*) \neq \mathbf{0}$, soit un point d'extremum lié de f sur Σ_g est que $(\mathbf{z}^*, \lambda^*)$ soit solution du système

$$\begin{cases} \nabla f(\mathbf{z}) = \lambda \nabla g(\mathbf{z}) \\ g(\mathbf{z}) = 0 \end{cases} \quad (7.1)$$

de $n+1$ équations à $n+1$ inconnues $(\mathbf{z}, \lambda) = (z_1, \dots, z_n, \lambda)$. On remarque, en particulier, qu'un point \mathbf{z}^* d'extremum lié de f sur Σ_g n'est généralement pas un point stationnaire de f car $\nabla f(\mathbf{z}^*) = \lambda^* \nabla g(\mathbf{z}^*) \neq \mathbf{0}$ si $\lambda^* \neq 0$.

Exemple 7.19. On cherche les extrema de $f(x, y) = x+y$, liés par la contrainte $g(x, y) = 0$, où $g(x, y) = x^2 + y^2 - 1$. On résout d'abord le système de 3 équations à trois inconnues (x, y, λ) :

$$\begin{cases} \nabla f(x, y) = \lambda \nabla g(x, y) \\ g(x, y) = 0 \end{cases} \implies \begin{cases} \frac{\partial f}{\partial x}(x, y) = \lambda \frac{\partial g}{\partial x}(x, y) \\ \frac{\partial f}{\partial y}(x, y) = \lambda \frac{\partial g}{\partial y}(x, y) \\ g(x, y) = 0 \end{cases} \implies \begin{cases} 1 = 2\lambda x \\ 1 = 2\lambda y \\ x^2 + y^2 = 1 \end{cases}$$

On remarque que $\lambda = 0$ n'est pas une solution. Alors, les premières deux équations donnent $x = y = \frac{1}{2\lambda}$ et, de la troisième équation, on obtient $\lambda^2 = \frac{1}{2}$ qui admet les deux solutions $\lambda^* = \pm \frac{1}{\sqrt{2}}$. On a alors deux points candidats :

$$\mathbf{P}_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \quad \mathbf{P}_2 = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right).$$

Puisque l'ensemble $\Sigma_g = \{(x, y) : g(x, y) = 0\}$ est compact et f est continue, alors f admet un maximum et un minimum sur Σ_g et nécessairement \mathbf{P}_1 ou \mathbf{P}_2 doit être un point de maximum lié et l'autre un point de minimum lié. Par évaluation directe de f en \mathbf{P}_1 et \mathbf{P}_2 on conclut :

$$\begin{aligned} \mathbf{P}_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \text{ est un point de maximum lié, } & \max_{(x,y) \in \Sigma_g} f(x, y) = f(\mathbf{P}_1) = \frac{2}{\sqrt{2}}, \\ \mathbf{P}_2 = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) \text{ est un point de minimum lié, } & \min_{(x,y) \in \Sigma_g} f(x, y) = f(\mathbf{P}_2) = -\frac{2}{\sqrt{2}}. \end{aligned}$$

7.3 Méthode des multiplicateurs de Lagrange

Le système (7.1) donne des conditions nécessaires d'optimalité. Il y a une autre façon d'obtenir ce système qui utilise la fonction de Lagrange (ou lagrangienne)

$$\mathcal{L} : E \times \mathbb{R} \rightarrow \mathbb{R}, \quad \mathcal{L}(\mathbf{z}, \lambda) = f(\mathbf{z}) - \lambda g(\mathbf{z}).$$

La variable $\lambda \in \mathbb{R}$ est appelée dans ce cas le *multiplicateur de Lagrange*.

Si $\mathbf{z}^* \in \Sigma_g = \{\mathbf{z} \in E : g(\mathbf{z}) = 0\}$ est un point d'extremum lié de f sur Σ_g avec $\nabla g(\mathbf{z}^*) \neq \mathbf{0}$, alors d'après le théorème 7.18, il existe $\lambda^* \in \mathbb{R}$ tel que $(\mathbf{z}^*, \lambda^*)$ est solution de (7.1). On vérifie facilement que ceci est équivalent à dire que $(\mathbf{z}^*, \lambda^*)$ est un point stationnaire de la fonction de Lagrange. En effet

$$\nabla \mathcal{L}(\mathbf{z}^*, \lambda^*) = \mathbf{0} \quad \Longleftrightarrow \quad \begin{cases} \nabla_{\mathbf{z}} \mathcal{L}(\mathbf{z}^*, \lambda^*) = \mathbf{0} \\ \frac{\partial}{\partial \lambda} \mathcal{L}(\mathbf{z}^*, \lambda^*) = 0 \end{cases} \quad \Longleftrightarrow \quad \begin{cases} \nabla f(\mathbf{z}^*) = \lambda^* \nabla g(\mathbf{z}^*) \\ g(\mathbf{z}^*) = 0. \end{cases}$$

Donc pour trouver les extrema liés de f sur Σ_g , il faut d'abord chercher les points stationnaires de \mathcal{L} .

7.4 Extrema sous contraintes multiples

Dans les sections précédentes on a considéré le cas où les variables $\mathbf{z} = (z_1, \dots, z_n)$ sont liées par une contrainte $g(\mathbf{z}) = 0$. Les arguments présentés se généralisent au cas où les variables $\mathbf{z} = (z_1, \dots, z_n)$ sont liées par plusieurs contraintes.

Soit $E \subset \mathbb{R}^n$ ouvert non vide et $f, g_1, g_2, \dots, g_m : E \rightarrow \mathbb{R}$ de classe $C^1(E)$, avec $m < n$. On pose le problème de minimisation sous contraintes multiples suivant

$$\min_{\mathbf{z} \in E} f(\mathbf{z}) \quad \text{sous les contraintes} \quad g_i(\mathbf{z}) = 0, \quad i = 1, \dots, m. \quad (7.2)$$

Soit $\mathbf{g} = (g_1, \dots, g_m) : E \rightarrow \mathbb{R}^m$ et $\Sigma_{\mathbf{g}} = \{\mathbf{z} \in E : \mathbf{g}(\mathbf{z}) = \mathbf{0}\} = \{\mathbf{z} \in E : g_i(\mathbf{z}) = 0, \forall i = 1, \dots, m\}$ l'ensemble faisable. Alors le problème (7.2) est équivalent à

$$\min_{\mathbf{z} \in \Sigma_{\mathbf{g}}} f(\mathbf{z}).$$

Le théorème 7.18 sur les conditions nécessaires d'optimalité se généralise de la façon suivante.

Théorème 7.20 (condition nécessaire d'optimalité – contraintes multiples). *Soit $E \subset \mathbb{R}^n$ ouvert non vide, $f, g_1, \dots, g_m \in C^1(E)$ et $\mathbf{z}^* \in \Sigma_{\mathbf{g}} = \{\mathbf{z} \in E : g_i(\mathbf{z}) = 0, i = 1, \dots, m\}$ un point d'extremum local lié de f sur $\Sigma_{\mathbf{g}}$ (avec $m < n$). Si $\text{Rang}(D\mathbf{g}(\mathbf{z}^*)) = m$, c'est-à-dire si les vecteurs $\{\nabla g_1(\mathbf{z}^*), \dots, \nabla g_m(\mathbf{z}^*)\}$ sont linéairement indépendants, alors il existe $\boldsymbol{\lambda}^* = (\lambda_1^*, \dots, \lambda_m^*) \in \mathbb{R}^m$ tel que*

$$\nabla f(\mathbf{z}^*) = \sum_{i=1}^m \lambda_i^* \nabla g_i(\mathbf{z}^*)$$

ou, de façon équivalente, $(\mathbf{z}^*, \boldsymbol{\lambda}^*) \in E \times \mathbb{R}^m$ est un point stationnaire de la fonction de Lagrange $\mathcal{L} : E \times \mathbb{R}^m \rightarrow \mathbb{R}$, $\mathcal{L}(\mathbf{z}, \boldsymbol{\lambda}) = f(\mathbf{z}) - \boldsymbol{\lambda} \cdot \mathbf{g}(\mathbf{z}) = f(\mathbf{z}) - \sum_{i=1}^m \lambda_i g_i(\mathbf{z})$ i.e.

$$\nabla_{(\mathbf{z}, \boldsymbol{\lambda})} \mathcal{L}(\mathbf{z}^*, \boldsymbol{\lambda}^*) = \mathbf{0}.$$

Démonstration. Puisque $\text{Rang}(D\mathbf{g}(\mathbf{z}^*)) = m$, il existe m colonnes linéairement indépendantes de la matrice Jacobienne $D\mathbf{g}(\mathbf{z}^*)$. Soient (i_1, \dots, i_m) ces colonnes. Pour se fixer les idées, supposons que ce soit les m dernières : $(i_1, \dots, i_m) = (n-m+1, \dots, n)$, le cas général se traitant de la même manière en permutant le rôle des n coordonnées. Notons $\mathbf{y} = (z_{n-m+1}, \dots, z_n)$ et $\mathbf{x} = (z_1, \dots, z_{n-m})$ les variables restantes, de sorte que $\mathbf{z} = (\mathbf{x}, \mathbf{y})$ et, en particulier, $\mathbf{z}^* = (\mathbf{x}^*, \mathbf{y}^*)$. De plus, nous décomposons la matrice jacobienne en $D\mathbf{g}(\mathbf{z}^*) = [D_{\mathbf{x}}\mathbf{g}(\mathbf{z}^*) | D_{\mathbf{y}}\mathbf{g}(\mathbf{z}^*)]$ où

$$D_{\mathbf{y}}\mathbf{g}(\mathbf{z}^*) = \begin{pmatrix} \frac{\partial g_1}{\partial z_{n-m+1}}(\mathbf{z}^*) & \dots & \frac{\partial g_1}{\partial z_n}(\mathbf{z}^*) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial z_{n-m+1}}(\mathbf{z}^*) & \dots & \frac{\partial g_m}{\partial z_n}(\mathbf{z}^*) \end{pmatrix} \in \mathbb{R}^{m \times m}$$

est inversible. Puisque $\det(D_{\mathbf{y}}\mathbf{g}(\mathbf{z}^*)) \neq 0$, on peut appliquer le théorème des fonctions implicites. Donc il existe un ouvert $U = B(\mathbf{x}^*, \delta) \subset \mathbb{R}^{n-m}$, un ouvert $V \subset E \subset \mathbb{R}^n$ contenant \mathbf{z}^* et une fonction $\phi : U \rightarrow \mathbb{R}^m$ de classe $C^1(U)$ tels que

- $\mathbf{y}^* = \phi(\mathbf{x}^*)$ et, pour tout $\mathbf{x} \in U$, $(\mathbf{x}, \phi(\mathbf{x})) \in V$ et $\mathbf{g}(\mathbf{x}, \phi(\mathbf{x})) = \mathbf{0}$;
- $\Sigma_{\mathbf{g}} \cap V = \mathcal{G}(\phi)$
- $D\phi(\mathbf{x}^*) = -(D_{\mathbf{y}}\mathbf{g}(\mathbf{z}^*))^{-1} D_{\mathbf{x}}\mathbf{g}(\mathbf{z}^*)$.

Introduisons $\tilde{f}(\mathbf{x}) = f(\mathbf{x}, \phi(\mathbf{x}))$, $\mathbf{x} \in U$. Alors \mathbf{x}^* est un extremum local de \tilde{f} sur U et $D\tilde{f}(\mathbf{x}^*) = \mathbf{0}$. On a

$$\begin{aligned} \mathbf{0} &= D\tilde{f}(\mathbf{x}^*) = D_{\mathbf{x}}f(\mathbf{z}^*) + D_{\mathbf{y}}f(\mathbf{z}^*) \cdot D\phi(\mathbf{x}^*) \\ &= D_{\mathbf{x}}f(\mathbf{z}^*) - D_{\mathbf{y}}f(\mathbf{z}^*) D_{\mathbf{y}}\mathbf{g}(\mathbf{z}^*)^{-1} D_{\mathbf{x}}\mathbf{g}(\mathbf{z}^*). \end{aligned}$$

Si on pose $\lambda^* = \underbrace{D_{\mathbf{y}}f(\mathbf{z}^*)}_{\in \mathbb{R}^{1 \times m}} \underbrace{D_{\mathbf{y}}\mathbf{g}(\mathbf{z}^*)^{-1}}_{\in \mathbb{R}^{m \times m}} \in \mathbb{R}^{1 \times m}$ (vecteur ligne), on a pour $i = 1, \dots, n - m$

$$0 = \frac{\partial f}{\partial x_i}(\mathbf{z}^*) - \lambda^* D_{x_i} \mathbf{g}(\mathbf{z}^*) = \frac{\partial f}{\partial x_i}(\mathbf{z}^*) - \sum_{j=1}^m \lambda_j^* \frac{\partial g_j}{\partial x_i}(\mathbf{z}^*)$$

et, par définition, $D_{\mathbf{y}}f(\mathbf{z}^*) = \lambda^* D_{\mathbf{y}}\mathbf{g}(\mathbf{z}^*)$ qui implique

$$\frac{\partial f}{\partial y_i}(\mathbf{z}^*) = \sum_{j=1}^m \lambda_j^* \frac{\partial g_j}{\partial y_i}(\mathbf{z}^*) \quad i = 1, \dots, m.$$

Donc finalement

$$\frac{\partial f}{\partial z_i}(\mathbf{z}^*) = \sum_{j=1}^m \lambda_j^* \frac{\partial g_j}{\partial z_i}(\mathbf{z}^*), \quad i = 1, \dots, n$$

que l'on peut écrire comme

$$\nabla f(\mathbf{z}^*) = \sum_{j=1}^m \lambda_j^* \nabla g_j(\mathbf{z}^*)$$

ou encore

$$\nabla_{(\mathbf{z}, \lambda)} \mathcal{L}(\mathbf{z}^*, \lambda^*) = \mathbf{0}.$$

□

Exercice 7.21. Chercher les extrema liés de $f(x, y, z) = x + y + z$ sous les contraintes $g_1(x, y, z) = x^2 + y^2 - 2 = 0$, $g_2(x, y, z) = x + z - 1 = 0$.

7.5 Conditions suffisantes

On mentionne ici sans démonstration (laissée comme exercice) des conditions suffisantes pour avoir un extremum local lié.

Théorème 7.22. Soit $E \subset \mathbb{R}^n$ ouvert non vide, $f, g_1, \dots, g_m \in C^2(E)$ et $(\mathbf{z}^*, \lambda^*) \in E \times \mathbb{R}^m$ un point stationnaire de la fonction de Lagrange $\mathcal{L}(\mathbf{z}, \lambda) = f(\mathbf{z}) - \sum_{i=1}^m \lambda_i g_i(\mathbf{z})$ (c'est-à-dire $\nabla \mathcal{L}(\mathbf{z}^*, \lambda^*) = \mathbf{0}$). On note $\mathbf{g} = (g_1, \dots, g_m) : E \rightarrow \mathbb{R}^m$ et $\Sigma_{\mathbf{g}} = \{\mathbf{z} \in E : \mathbf{g}(\mathbf{z}) = \mathbf{0}\}$ l'ensemble faisable. Supposons encore $\text{Rang}(D\mathbf{g}(\mathbf{z}^*)) = m$ et considérons l'espace vectoriel tangent à $\Sigma_{\mathbf{g}}$ en \mathbf{z}^* :

$$T_{\mathbf{z}^*}(\Sigma_{\mathbf{g}}) = \{\mathbf{w} \in \mathbb{R}^n : D\mathbf{g}(\mathbf{z}^*) \cdot \mathbf{w} = \mathbf{0}\}.$$

Si

$$\forall \mathbf{w} \in T_{\mathbf{z}^*}(\Sigma_{\mathbf{g}}) \setminus \{\mathbf{0}\} \quad \mathbf{w}^\top \left(H_f(\mathbf{z}^*) - \sum_{i=1}^m \lambda_i^* H_{g_i}(\mathbf{z}^*) \right) \mathbf{w} > 0,$$

alors \mathbf{z}^* est un point de minimum local de f sur $\Sigma_{\mathbf{g}}$.

La condition $D\mathbf{g}(\mathbf{z}^*) \cdot \mathbf{w} = \mathbf{0}$ dans le théorème précédent implique $(\nabla g_i(\mathbf{z}^*))^\top \cdot \mathbf{w} = 0$, $\forall i = 1, \dots, m$. Donc les directions \mathbf{w} sont des directions orthogonales à tous les vecteurs $\nabla g_i(\mathbf{z}^*)$ et donc *tangentes* à la contrainte $\mathbf{g}(\mathbf{x}) = \mathbf{0}$.

Si on note $H_{\mathcal{L}}^{(\mathbf{z})}(\mathbf{z}, \boldsymbol{\lambda})$ la matrice Hessienne de \mathcal{L} calculée uniquement par rapport aux variables \mathbf{z} :

$$\left(H_{\mathcal{L}}^{(\mathbf{z})}(\mathbf{z}, \boldsymbol{\lambda})\right)_{ij} = \frac{\partial^2 \mathcal{L}}{\partial z_i \partial z_j}(\mathbf{z}, \boldsymbol{\lambda}) = \frac{\partial^2 f}{\partial z_i \partial z_j}(\mathbf{x}) - \sum_{\ell=1}^m \lambda_\ell \frac{\partial^2 g_\ell}{\partial z_i \partial z_j}(\mathbf{x}),$$

on voit que $H_{\mathcal{L}}^{(\mathbf{z})}(\mathbf{z}, \boldsymbol{\lambda}) = H_f(\mathbf{z}) - \sum_{\ell=1}^m \lambda_\ell H_{g_\ell}(\mathbf{z})$. Le théorème précédent nous dit ainsi que, si

$$\forall \mathbf{w} \in T_{\mathbf{z}^*}(\Sigma_{\mathbf{g}}) \setminus \{\mathbf{0}\} \quad \mathbf{w}^\top H_{\mathcal{L}}^{(\mathbf{z})}(\mathbf{z}^*, \boldsymbol{\lambda}^*) \mathbf{w} > 0,$$

alors \mathbf{z}^* est un point de minimum local lié. Il suffit de vérifier la positivité de la forme quadratique $\mathbf{w}^\top H_{\mathcal{L}}^{(\mathbf{z})}(\mathbf{z}^*, \boldsymbol{\lambda}^*) \mathbf{w}$ (et non pas de $\mathbf{w}^\top H_f(\mathbf{z}^*) \mathbf{w}$!) uniquement pour les directions tangentes à la contrainte $\mathbf{g}(\mathbf{z}) = \mathbf{0}$.

Chapitre 8

Intégrale multiple au sens de Riemann

Soit $E \subset \mathbb{R}^n$ borné non vide et $f : E \rightarrow \mathbb{R}$ bornée. On veut définir l'intégrale de f sur E , noté

$$\int_E f(\mathbf{x}) d\mathbf{x}.$$

On parle d'*intégrale double* lorsque $E \subset \mathbb{R}^2$, d'*intégrale triple* lorsque $E \subset \mathbb{R}^3$ et, plus généralement, d'*intégrale multiple* pour $n > 1$. On commence notre étude par le cas où le domaine E est un *pavé de \mathbb{R}^n* , i.e. un rectangle fermé en dimension $n = 2$ et un hyper-rectangle fermé en dimension $n > 2$.

8.1 Pavés de \mathbb{R}^n

On commence par donner la définition de *pavé* ainsi que de *partition* d'un pavé et *raffinement* d'une partition.

Définition 8.1 (Pavé). *On appelle pavé tout ensemble $R \subset \mathbb{R}^n$ de la forme $R = [a_1, b_1] \times \dots \times [a_n, b_n]$ où $a_j \leq b_j$, $j = 1, \dots, n$ sont des nombres réels. Le volume de R , noté $\text{Vol}(R)$, est défini comme*

$$\text{Vol}(R) = \prod_{j=1}^n (b_j - a_j).$$

On dit que R est un pavé dégénéré s'il existe un ou plusieurs $k = 1, \dots, n$ tels que $a_k = b_k$. Dans ce cas on a $\text{Vol}(R) = 0$.

Définition 8.2 (Partition). *On appelle partition d'un pavé $R \subset \mathbb{R}^n$, une collection finie \mathcal{P} de pavés tels que $\bigcup_{Q \in \mathcal{P}} Q = R$ et, pour tout $Q, Q' \in \mathcal{P}$, $Q \neq Q'$, on a $\overset{\circ}{Q} \cap \overset{\circ}{Q'} = \emptyset$.*

La figure 8.1(gauche) montre un exemple de partition d'un pavé.

Définition 8.3. *Une partition \mathcal{P} d'un pavé $R = [a_1, b_1] \times \dots \times [a_n, b_n]$ sera dite tensorielle s'il existe, pour tout $j \in \{1, \dots, n\}$,*

$$a_j = t_j^0 \leq \dots \leq t_j^{N_j} = b_j \quad \text{dans } \mathbb{R}, \quad N_j \in \mathbb{N}^*,$$

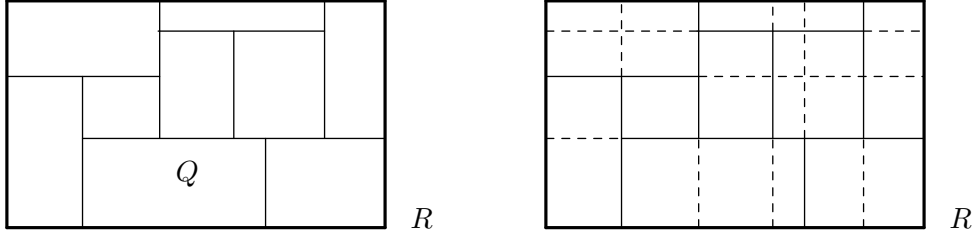


FIGURE 8.1 – Exemple d’une partition d’un pavé de \mathbb{R}^2 (gauche) et d’un possible raffinement de cette partition (droite).

tels que

$$\mathcal{P} = \left\{ [t_1^{\alpha_1}, t_1^{1+\alpha_1}] \times \dots \times [t_n^{\alpha_n}, t_n^{1+\alpha_n}] : 0 \leq \alpha_1 \leq N_1 - 1, \dots, 0 \leq \alpha_n \leq N_n - 1 \right\}.$$

On notera alors $\mathcal{P} = (t_1^0, \dots, t_1^{N_1}) \otimes \dots \otimes (t_n^0, \dots, t_n^{N_n})$. Le contexte devrait permettre de ne pas confondre des indices supérieurs comme ici avec des puissances.

Définition 8.4 (Raffinement d’une partition). Soit $\mathcal{P}, \mathcal{P}'$ deux partitions d’un pavé $R \subset \mathbb{R}^n$. On dit que \mathcal{P}' est un raffinement de \mathcal{P} si, pour tout $Q \in \mathcal{P}$, la collection $\mathcal{P}'_Q = \{Q' \in \mathcal{P}' : Q' \subset Q\}$ est une partition du pavé Q .

La figure 8.1(droite) montre un exemple d’un raffinement d’une partition. À partir de la définition 8.4 on peut montrer les propriétés suivantes des pavés d’un raffinement \mathcal{P}' d’une partition \mathcal{P} .

- Si $Q' \subset Q_1$ et $Q' \subset Q_2$ avec $Q' \in \mathcal{P}'$ et $Q_1 \neq Q_2$ dans \mathcal{P} , alors $\text{Vol}(Q') = 0$.
- Si $Q' \in \mathcal{P}'$ n’est inclus dans aucun $Q \in \mathcal{P}$, alors $\text{Vol}(Q') = 0$.
- \mathcal{P}' est un raffinement de \mathcal{P} si et seulement si pour tout $Q' \in \mathcal{P}'$ non dégénéré il existe $Q \in \mathcal{P}$ tel que $Q' \subset Q$.
- Si le raffinement \mathcal{P}' de \mathcal{P} est une partition tensorielle, alors, pour tout $Q \in \mathcal{P}$, $\mathcal{P}'_Q = \{Q' \in \mathcal{P}' : Q' \subset Q\}$ est une partition tensorielle de Q .

On a aussi les résultats suivants.

Lemme 8.5. Soit $\mathcal{P}, \mathcal{P}'$ deux partitions d’un pavé $R \subset \mathbb{R}^n$. Alors il existe une partition tensorielle \mathcal{P}'' qui est un raffinement à la fois de \mathcal{P} et \mathcal{P}' . En particulier, toute partition \mathcal{P} admet un raffinement tensoriel.

Démonstration. Notons $R = [a_1, b_1] \times \dots \times [a_n, b_n]$, et soit K et K' les nombres de pavés dans \mathcal{P} et \mathcal{P}' . Notons aussi

$$\mathcal{P} \cup \mathcal{P}' = \left\{ [a_1^i, b_1^i] \times \dots \times [a_n^i, b_n^i] : 1 \leq i \leq K + K' \right\}$$

Pour chaque $j \in \{1, \dots, n\}$, ordonnons $a_j^1, b_j^1, a_j^2, b_j^2, \dots, a_j^{K+K'}, b_j^{K+K'}$:

$$\{a_j^1, b_j^1, a_j^2, b_j^2, \dots, a_j^{K+K'}, b_j^{K+K'}\} = \{t_j^0, t_j^1, \dots, t_j^{2(K+K')-1}\} \quad \text{avec} \quad t_j^0 \leq t_j^1 \leq \dots \leq t_j^{2(K+K')-1}.$$

Alors

$$\mathcal{P}'' = (t_1^0, \dots, t_1^{2(K+K')-1}) \otimes \dots \otimes (t_n^0, \dots, t_n^{2(K+K')-1})$$

est un raffinement commun. □

Lemme 8.6. *Soit \mathcal{P} une partition d'un pavé $R \subset \mathbb{R}^n$. Alors*

$$\text{Vol}(R) = \sum_{Q \in \mathcal{P}} \text{Vol}(Q)$$

Démonstration. Supposons d'abord que \mathcal{P} est tensorielle :

$$\mathcal{P} = (t_1^0, \dots, t_1^{N_1}) \otimes \dots \otimes (t_n^0, \dots, t_n^{N_n}).$$

Alors

$$\begin{aligned} \sum_{Q \in \mathcal{P}} \text{Vol}(Q) &= \sum_{\alpha_1=0}^{N_1-1} \dots \sum_{\alpha_n=0}^{N_n-1} (t_1^{1+\alpha_1} - t_1^{\alpha_1}) \dots (t_n^{1+\alpha_n} - t_n^{\alpha_n}) \\ &= \sum_{\alpha_1=0}^{N_1-1} (t_1^{1+\alpha_1} - t_1^{\alpha_1}) \dots \sum_{\alpha_n=0}^{N_n-1} (t_n^{1+\alpha_n} - t_n^{\alpha_n}) = (t_1^{N_1} - t_1^0) \dots (t_n^{N_n} - t_n^0) = \text{Vol}(R). \end{aligned}$$

Passons au cas général et appliquons le lemme 8.5 à $\mathcal{P}' = \mathcal{P}$, ce qui donne un raffinement \mathcal{P}'' de \mathcal{P} qui est une partition tensorielle. On a donc

$$\begin{aligned} \text{Vol}(R) &= \sum_{Q'' \in \mathcal{P}''} \text{Vol}(Q'') \quad (\text{cf début de la preuve}) \\ &= \sum_{Q \in \mathcal{P}} \left(\sum_{Q'' \in \mathcal{P}'', Q'' \subset Q} \text{Vol}(Q'') \right) \quad (\text{car } \mathcal{P}'' \text{ est un raffinement de } \mathcal{P}) \\ &= \sum_{Q \in \mathcal{P}} \text{Vol}(Q). \end{aligned}$$

Pour la dernière égalité on a utilisé le fait que $\mathcal{P}_Q'' = \{Q'' \in \mathcal{P}'' : Q'' \subset Q\}$ est une partition tensorielle de Q ; voir une remarque précédente. □

8.2 Fonctions intégrables au sens de Riemann sur un pavé

Soit $R \subset \mathbb{R}^n$ un pavé et $f : R \rightarrow \mathbb{R}$ une fonction *bornée* (pas nécessairement continue).

Définition 8.7 (Sommes de Darboux). *Soit \mathcal{P} une partition de R . On définit*

$$\begin{aligned} \text{somme inférieure :} \quad \underline{S}(f, \mathcal{P}) &= \sum_{Q \in \mathcal{P}} \left(\inf_{\mathbf{x} \in Q} f(\mathbf{x}) \right) \text{Vol}(Q), \\ \text{somme supérieure :} \quad \overline{S}(f, \mathcal{P}) &= \sum_{Q \in \mathcal{P}} \left(\sup_{\mathbf{x} \in Q} f(\mathbf{x}) \right) \text{Vol}(Q). \end{aligned}$$

On remarque que, puisque f est bornée et \mathcal{P} est une partition finie, les quantités $\underline{S}(f, \mathcal{P})$ et $\overline{S}(f, \mathcal{P})$ sont bien définies. Le lemme suivant montre aussi qu'une somme inférieure est toujours plus petite qu'une somme supérieure.

Lemme 8.8. *Soit \mathcal{P} une partition de R , et \mathcal{P}'' un raffinement de \mathcal{P} . Alors,*

$$\underline{S}(f, \mathcal{P}) \leq \underline{S}(f, \mathcal{P}'') \leq \overline{S}(f, \mathcal{P}'') \leq \overline{S}(f, \mathcal{P}).$$

De plus, pour toute partition $\mathcal{P}, \mathcal{P}'$ de R on a $\underline{S}(f, \mathcal{P}) \leq \overline{S}(f, \mathcal{P}')$.

Démonstration. Puisque \mathcal{P}'' est un raffinement de \mathcal{P} , on a

$$\forall Q \in \mathcal{P} \quad Q = \bigcup_{\substack{Q'' \in \mathcal{P}'' \\ Q'' \subset Q}} Q'' \quad \text{et} \quad \text{Vol}(Q) = \sum_{\substack{Q'' \in \mathcal{P}'' \\ Q'' \subset Q}} \text{Vol}(Q'').$$

Donc

$$\begin{aligned} \underline{S}(f, \mathcal{P}) &= \sum_{Q \in \mathcal{P}} \left(\inf_{\mathbf{x} \in Q} f(\mathbf{x}) \right) \text{Vol}(Q) = \sum_{Q \in \mathcal{P}} \left(\inf_{\mathbf{x} \in Q} f(\mathbf{x}) \right) \sum_{\substack{Q'' \in \mathcal{P}'' \\ Q'' \subset Q}} \text{Vol}(Q'') \\ &\leq \sum_{Q \in \mathcal{P}} \sum_{\substack{Q'' \in \mathcal{P}'' \\ Q'' \subset Q}} \left(\inf_{\mathbf{x} \in Q''} f(\mathbf{x}) \right) \text{Vol}(Q'') = \underline{S}(f, \mathcal{P}''). \end{aligned}$$

De la même façon, on prouve $\overline{S}(f, \mathcal{P}'') \leq \overline{S}(f, \mathcal{P})$. Enfin, l'inégalité $\underline{S}(f, \mathcal{P}'') \leq \overline{S}(f, \mathcal{P}'')$ est immédiate.

Puisque toute paire de partitions $\mathcal{P}, \mathcal{P}'$ de R admet un raffinement commun, le résultat précédent implique immédiatement $\underline{S}(f, \mathcal{P}) \leq \underline{S}(f, \mathcal{P}')$. \square

On est maintenant en mesure de donner la définition suivante de fonction intégrable au sens de Riemann.

Définition 8.9 (Fonction intégrable au sens de Riemann). *Soit $R \subset \mathbb{R}^n$ un pavé et $f : R \rightarrow \mathbb{R}$ une fonction bornée. On appelle*

$$\begin{aligned} \text{intégrale de Riemann supérieure} \quad & \overline{\int}_R f(\mathbf{x}) d\mathbf{x} := \inf \{ \overline{S}(f, \mathcal{P}), \mathcal{P} \text{ partition de } R \}, \\ \text{intégrale de Riemann inférieure} \quad & \underline{\int}_R f(\mathbf{x}) d\mathbf{x} := \sup \{ \underline{S}(f, \mathcal{P}), \mathcal{P} \text{ partition de } R \}. \end{aligned}$$

On dit que f est intégrable au sens de Riemann (ou “Riemann-intégrable” ou simplement “intégrable”) si

$$\underline{\int}_R f(\mathbf{x}) d\mathbf{x} = \overline{\int}_R f(\mathbf{x}) d\mathbf{x}.$$

Dans ce cas on note

$$\int_R f(\mathbf{x}) d\mathbf{x} := \underline{\int}_R f(\mathbf{x}) d\mathbf{x} = \overline{\int}_R f(\mathbf{x}) d\mathbf{x}.$$

On note $\mathcal{R}(R)$ l'ensemble des fonctions $f : R \rightarrow \mathbb{R}$ bornées et intégrables au sens de Riemann.

Puisque dans la définition ci-dessus on a supposé f bornée, pour toute partition \mathcal{P} de R donnée, on a

$$-\infty < \underline{S}(f, \mathcal{P}) \leq \int_{\underline{R}} f(\mathbf{x}) d\mathbf{x} \leq \int_R f(\mathbf{x}) d\mathbf{x} \leq \overline{S}(f, \mathcal{P}) < +\infty.$$

On voit donc que $\int_{\underline{R}} f(\mathbf{x}) d\mathbf{x}$ et $\int_R f(\mathbf{x}) d\mathbf{x}$ existent et sont finies. Une autre caractérisation des fonctions intégrables au sens de Riemann est la suivante.

Lemme 8.10. *Soit $R \subset \mathbb{R}^n$ un pavé et $f : R \rightarrow \mathbb{R}$ bornée. Alors f est intégrable au sens de Riemann si et seulement si pour tout $\epsilon > 0$ il existe une partition \mathcal{P}_ϵ de R telle que*

$$\overline{S}(f, \mathcal{P}_\epsilon) - \underline{S}(f, \mathcal{P}_\epsilon) < \epsilon.$$

Démonstration. « \Leftarrow » : Supposons que, pour tout $\epsilon > 0$, il existe une partition \mathcal{P}_ϵ de R telle que $\overline{S}(f, \mathcal{P}_\epsilon) - \underline{S}(f, \mathcal{P}_\epsilon) < \epsilon$. Alors

$$\int_R f(\mathbf{x}) d\mathbf{x} \leq \overline{S}(f, \mathcal{P}_\epsilon), \quad \int_{\underline{R}} f(\mathbf{x}) d\mathbf{x} \geq \underline{S}(f, \mathcal{P}_\epsilon)$$

et

$$\int_R f(\mathbf{x}) d\mathbf{x} - \int_{\underline{R}} f(\mathbf{x}) d\mathbf{x} \leq \overline{S}(f, \mathcal{P}_\epsilon) - \underline{S}(f, \mathcal{P}_\epsilon) < \epsilon.$$

Comme $\epsilon > 0$ est arbitraire, ceci implique $\int_R f(\mathbf{x}) d\mathbf{x} = \int_{\underline{R}} f(\mathbf{x}) d\mathbf{x}$ et par conséquent, f est intégrable au sens de Riemann.

« \Rightarrow » : Supposons f intégrable au sens de Riemann. Par définition de sup et inf on a que $\forall \epsilon > 0$

$$\begin{aligned} \exists \mathcal{P}_\epsilon \text{ partition de } R : \quad \underline{S}(f, \mathcal{P}_\epsilon) &> \int_{\underline{R}} f(\mathbf{x}) d\mathbf{x} - \frac{\epsilon}{2} = \int_R f(\mathbf{x}) d\mathbf{x} - \frac{\epsilon}{2}, \\ \exists \mathcal{P}'_\epsilon \text{ partition de } R : \quad \overline{S}(f, \mathcal{P}'_\epsilon) &< \int_R f(\mathbf{x}) d\mathbf{x} + \frac{\epsilon}{2} = \int_{\underline{R}} f(\mathbf{x}) d\mathbf{x} + \frac{\epsilon}{2}. \end{aligned}$$

Soit \mathcal{P}''_ϵ un raffinement commun de \mathcal{P}_ϵ et \mathcal{P}'_ϵ . Alors

$$\begin{aligned} \underline{S}(f, \mathcal{P}''_\epsilon) &\geq \underline{S}(f, \mathcal{P}_\epsilon) > \int_{\underline{R}} f(\mathbf{x}) d\mathbf{x} - \frac{\epsilon}{2}, \\ \overline{S}(f, \mathcal{P}''_\epsilon) &\leq \overline{S}(f, \mathcal{P}'_\epsilon) < \int_{\underline{R}} f(\mathbf{x}) d\mathbf{x} + \frac{\epsilon}{2}, \end{aligned}$$

ce qui implique

$$\overline{S}(f, \mathcal{P}''_\epsilon) - \underline{S}(f, \mathcal{P}''_\epsilon) < \epsilon.$$

□

À partir de cette caractérisation, on montre facilement le résultat suivant :

Théorème 8.11. *Soit $R \subset \mathbb{R}^n$ un pavé et $f : R \rightarrow \mathbb{R}$ une fonction continue. Alors f est intégrable au sens de Riemann.*

Démonstration. R est compact et f est continue sur R , donc uniformément continue. Ceci implique que

$$\forall \epsilon > 0 \exists \delta_\epsilon > 0 \forall \mathbf{x}, \mathbf{y} \in R \quad (\|\mathbf{x} - \mathbf{y}\| < \delta_\epsilon \Rightarrow |f(\mathbf{x}) - f(\mathbf{y})| < \frac{\epsilon}{1 + \text{Vol}(R)}),$$

où on utilise dans cette démonstration la norme euclidienne de \mathbb{R}^n . Soit \mathcal{P}_ϵ une partition de R telle que

$$\forall Q \in \mathcal{P}_\epsilon \quad \text{diam}(Q) = \max_{\mathbf{x}, \mathbf{y} \in Q} \|\mathbf{x} - \mathbf{y}\| < \delta_\epsilon.$$

Une telle partition existe toujours. Il suffit de prendre la partition tensorielle de $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$ suivante : $\mathcal{P}_\epsilon = (t_1^0, \dots, t_1^N) \otimes \cdots \otimes (t_n^0, \dots, t_n^N)$ avec

$$t_j^i = a_j + (b_j - a_j) \frac{i}{N_\epsilon}, \quad 0 \leq i \leq N_\epsilon, \quad 1 \leq j \leq n,$$

et $N \in \mathbb{N}^*$ choisi tel que $N_\epsilon > \sqrt{\sum_{i=1}^n (b_i - a_i)^2} / \delta_\epsilon$ de telle sorte que pour tout $Q \in \mathcal{P}_\epsilon$ et $\mathbf{x}, \mathbf{y} \in Q$ on a $\|\mathbf{x} - \mathbf{y}\| < \delta_\epsilon$. On a alors

$$\begin{aligned} \overline{S}(f, \mathcal{P}_\epsilon) - \underline{S}(f, \mathcal{P}_\epsilon) &= \sum_{Q \in \mathcal{P}_\epsilon} \left(\max_{\mathbf{x} \in Q} f(\mathbf{x}) - \min_{\mathbf{x} \in Q} f(\mathbf{x}) \right) \text{Vol}(Q) \\ &\leq \frac{\epsilon}{1 + \text{Vol}(R)} \sum_{Q \in \mathcal{P}_\epsilon} \text{Vol}(Q) \\ &< \epsilon, \end{aligned}$$

ce qui montre que f est intégrable au sens de Riemann. \square

Grâce à ce résultat, on a immédiatement que $C^0(R) \subset \mathcal{R}(R)$. D'autre part, si R est dégénéré et $f : R \rightarrow \mathbb{R}$ est bornée, alors f est toujours intégrable au sens de Riemann et d'intégrale nulle.

Exercice 8.12. Dire si les fonctions suivantes définies sur $[0, 1]^2$ sont intégrables au sens de Riemann ou non :

$$\begin{aligned} f_1(x, y) &= \begin{cases} 1, & (x, y) \in [0, 1]^2 \cap \mathbb{Q}^2 \\ 0 & \text{autrement} \end{cases} & f_2(x, y) &= \begin{cases} 1, & (x, y) \in [0, 1]^2, x \leq \frac{1}{2} \\ 0 & \text{autrement} \end{cases} \\ f_3(x, y) &= \begin{cases} \frac{x}{\sqrt{x^2 + y^2}}, & (x, y) \in [0, 1]^2 \setminus \{(0, 0)\} \\ 0 & (x, y) = (0, 0) \end{cases} \end{aligned}$$

On conclut cette section par un résultat qui sera utile par la suite.

Lemme 8.13. Soit deux pavés R et \hat{R} dans \mathbb{R}^n tels que $R \subset \text{int}(\hat{R})$, soit $f : R \rightarrow \mathbb{R}$ une fonction bornée et notons $\hat{f} : \hat{R} \rightarrow \mathbb{R}$ le prolongement par zéro de f sur \hat{R} :

$$\hat{f}(\mathbf{x}) = f(\mathbf{x}) \text{ pour tout } \mathbf{x} \in R \text{ et } \hat{f}(\mathbf{x}) = 0 \text{ pour tout } \mathbf{x} \in \hat{R} \setminus R.$$

Alors f est Riemann-intégrable ssi \hat{f} l'est, auquel cas $\int_R f(\mathbf{x}) d\mathbf{x} = \int_{\hat{R}} \hat{f}(\mathbf{x}) d\mathbf{x}$.

Démonstration. Supposons d'abord f Riemann-intégrable et soit $\epsilon > 0$. Choisissons une partition \mathcal{P}_ϵ de R telle que

$$\overline{S}(f, \mathcal{P}_\epsilon) - \underline{S}(f, \mathcal{P}_\epsilon) < \epsilon.$$

Soit encore un pavé R_ϵ tel que

$$R \subset \text{int}(R_\epsilon) \subset R_\epsilon \subset \text{int}(\widehat{R})$$

et

$$\text{Vol}(R_\epsilon) - \text{Vol}(R) < \frac{\epsilon}{2M} \quad \text{avec} \quad M > \sup_{\mathbf{x} \in R} |f(\mathbf{x})|.$$

Soit la partition tensorielle $\widehat{\mathcal{P}}_\epsilon$ de \widehat{R} définie à partir de toutes les composantes (éventuellement répétées) de tous les sommets des pavés dans \mathcal{P}_ϵ et des sommets de R_ϵ et \widehat{R} . Alors aucun $Q \in \widehat{\mathcal{P}}_\epsilon$ ne contient à la fois un point de $\overset{\circ}{R}$ et un point de $\widehat{R} \setminus R$. De même aucun $Q \in \widehat{\mathcal{P}}_\epsilon$ ne contient à la fois un point de $\overset{\circ}{R}_\epsilon$ et un point de $\widehat{R} \setminus R_\epsilon$. De plus $\{Q \in \widehat{\mathcal{P}}_\epsilon : Q \subset R\}$ est une partition de R qui raffine \mathcal{P}_ϵ . On a donc, où $Q \in \widehat{\mathcal{P}}_\epsilon$,

$$\begin{aligned} \overline{S}(\widehat{f}, \widehat{\mathcal{P}}_\epsilon) &= \left(\sum_{Q \subset R} + \sum_{Q \subset R_\epsilon \setminus \overset{\circ}{R}} + \sum_{Q \subset \widehat{R} \setminus \overset{\circ}{R}_\epsilon} \right) \sup_{\mathbf{x} \in Q} \widehat{f}(\mathbf{x}) \text{Vol}(Q) \\ &\leq \overline{S}(f, \mathcal{P}_\epsilon) + \sum_{Q \subset R_\epsilon \setminus \overset{\circ}{R}} M \text{Vol}(Q) + \sum_{Q \subset \widehat{R} \setminus \overset{\circ}{R}_\epsilon} 0 \leq \overline{S}(f, \mathcal{P}_\epsilon) + \frac{M\epsilon}{2M} = \overline{S}(f, \mathcal{P}_\epsilon) + \frac{\epsilon}{2}. \end{aligned}$$

De même

$$\underline{S}(\widehat{f}, \widehat{\mathcal{P}}_\epsilon) \geq \underline{S}(f, \mathcal{P}_\epsilon) - \frac{\epsilon}{2}$$

et donc

$$\overline{S}(\widehat{f}, \widehat{\mathcal{P}}_\epsilon) - \underline{S}(\widehat{f}, \widehat{\mathcal{P}}_\epsilon) \leq \overline{S}(f, \mathcal{P}_\epsilon) - \underline{S}(f, \mathcal{P}_\epsilon) + \epsilon < 2\epsilon.$$

Comme $\epsilon > 0$ est arbitraire, \widehat{f} est Riemann-intégrable. De plus, en considérant de nouveau $\epsilon > 0$, ce qui précède donne

$$\underline{S}(f, \mathcal{P}_\epsilon) - \frac{\epsilon}{2} \leq \int_{\widehat{R}} \widehat{f}(\mathbf{x}) d\mathbf{x} \leq \overline{S}(f, \mathcal{P}_\epsilon) + \frac{\epsilon}{2}$$

et, par ailleurs,

$$\underline{S}(f, \mathcal{P}_\epsilon) - \frac{\epsilon}{2} \leq \underline{S}(f, \mathcal{P}_\epsilon) \leq \int_R f(\mathbf{x}) d\mathbf{x} \leq \overline{S}(f, \mathcal{P}_\epsilon) \leq \overline{S}(f, \mathcal{P}_\epsilon) + \frac{\epsilon}{2}.$$

D'où

$$\left| \int_{\widehat{R}} \widehat{f}(\mathbf{x}) d\mathbf{x} - \int_R f(\mathbf{x}) d\mathbf{x} \right| < 2\epsilon$$

et, puisque $\epsilon > 0$ est arbitraire, les deux intégrales sont égales.

Réciproquement, supposons \widehat{f} Riemann-intégrable et soit $\epsilon > 0$. Choisissons une partition $\widehat{\mathcal{P}}_\epsilon$ de \widehat{R} telle que

$$\overline{S}(\widehat{f}, \widehat{\mathcal{P}}_\epsilon) - \underline{S}(\widehat{f}, \widehat{\mathcal{P}}_\epsilon) < \epsilon.$$

Soit encore un pavé R_ϵ tel que

$$R \subset \text{int}(R_\epsilon) \subset R_\epsilon \subset \text{int}(\hat{R})$$

et

$$\text{Vol}(R_\epsilon) - \text{Vol}(R) < \frac{\epsilon}{2M} \quad \text{avec} \quad M > \sup_{\mathbf{x} \in R} |f(\mathbf{x})|.$$

Soit la partition tensorielle $\hat{\mathcal{P}}'_\epsilon$ de \hat{R} définie à partir de toutes les composantes (éventuellement répétées) de tous les sommets des pavés dans $\hat{\mathcal{P}}_\epsilon$ et des sommets de R et R_ϵ . Alors $\mathcal{P}_\epsilon = \{Q \in \hat{\mathcal{P}}'_\epsilon : Q \subset R\}$ est une partition de R . On a, où $Q \in \hat{\mathcal{P}}'_\epsilon$,

$$\begin{aligned} \bar{S}(\hat{f}, \hat{\mathcal{P}}'_\epsilon) &= \left(\sum_{Q \in \mathcal{P}_\epsilon} + \sum_{Q \subset R_\epsilon \setminus \hat{R}} + \sum_{Q \subset \hat{R} \setminus \hat{R}_\epsilon} \right) \sup_{\mathbf{x} \in Q} \hat{f}(\mathbf{x}) \text{Vol}(Q) \\ &\geq \bar{S}(f, \mathcal{P}_\epsilon) - \frac{M\epsilon}{2M} + 0 = \bar{S}(f, \mathcal{P}_\epsilon) - \frac{\epsilon}{2}. \end{aligned}$$

De même

$$\underline{S}(\hat{f}, \hat{\mathcal{P}}'_\epsilon) \leq \underline{S}(f, \mathcal{P}_\epsilon) + \frac{\epsilon}{2}$$

et donc

$$\bar{S}(f, \mathcal{P}_\epsilon) - \underline{S}(f, \mathcal{P}_\epsilon) \leq \bar{S}(\hat{f}, \hat{\mathcal{P}}'_\epsilon) - \underline{S}(\hat{f}, \hat{\mathcal{P}}'_\epsilon) + \epsilon \leq \bar{S}(\hat{f}, \hat{\mathcal{P}}_\epsilon) - \underline{S}(\hat{f}, \hat{\mathcal{P}}_\epsilon) + \epsilon < 2\epsilon.$$

Comme $\epsilon > 0$ est arbitraire, f est Riemann-intégrable et on conclut que $\int_R f(\mathbf{x}) d\mathbf{x} = \int_{\hat{R}} \hat{f}(\mathbf{x}) d\mathbf{x}$ par le même argument qu'auparavant. \square

8.3 Intégrales itérées sur un pavé et théorème de Fubini

On montre ici une procédure simple pour calculer l'intégrale d'une fonction sur un pavé. Elle se base sur le résultat suivant qui prend le nom de formule des intégrales itérées.

Théorème 8.14 (de Fubini). *Soit $R \subset \mathbb{R}^{n+m}$ un pavé de la forme $R = R^{(1)} \times R^{(2)}$, $R^{(1)} \subset \mathbb{R}^n$, $R^{(2)} \subset \mathbb{R}^m$ et $f : R \rightarrow \mathbb{R}$ une fonction bornée et intégrable au sens de Riemann, $f = f(\mathbf{x}, \mathbf{y})$, $\mathbf{x} \in R^{(1)}$, $\mathbf{y} \in R^{(2)}$. Si $\forall \mathbf{y} \in R^{(2)}$ la fonction $f(\cdot, \mathbf{y}) : R^{(1)} \rightarrow \mathbb{R}$ est intégrable au sens de Riemann, alors la fonction $\mathbf{y} \mapsto G(\mathbf{y}) = \int_{R^{(1)}} f(\mathbf{x}, \mathbf{y}) d\mathbf{x}$, $\mathbf{y} \in R^{(2)}$, est aussi intégrable au sens de Riemann et*

$$\int_R f(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y} = \int_{R^{(2)}} G(\mathbf{y}) d\mathbf{y} = \int_{R^{(2)}} \left(\int_{R^{(1)}} f(\mathbf{x}, \mathbf{y}) d\mathbf{x} \right) d\mathbf{y},$$

où $\int_R f(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y}$ est une notation pour l'intégrale de f sur R .

Réciproquement, si $\forall \mathbf{x} \in R^{(1)}$ la fonction $f(\mathbf{x}, \cdot) : R^{(2)} \rightarrow \mathbb{R}$ est intégrable au sens de Riemann, alors la fonction $\mathbf{x} \mapsto F(\mathbf{x}) = \int_{R^{(2)}} f(\mathbf{x}, \mathbf{y}) d\mathbf{y}$, $\mathbf{x} \in R^{(1)}$, est aussi intégrable au sens de Riemann et

$$\int_R f(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y} = \int_{R^{(1)}} F(\mathbf{x}) d\mathbf{x} = \int_{R^{(1)}} \left(\int_{R^{(2)}} f(\mathbf{x}, \mathbf{y}) d\mathbf{y} \right) d\mathbf{x},$$

Démonstration. Puisque f est intégrable, $\forall \epsilon > 0$ il existe une partition \mathcal{P}_ϵ de R telle que $\bar{S}(f, \mathcal{P}_\epsilon) - \underline{S}(f, \mathcal{P}_\epsilon) < \epsilon$. Cette partition admet toujours un raffinement \mathcal{P}'_ϵ qui est une partition tensorielle :

$$\mathcal{P}'_\epsilon = (t_1^0, \dots, t_1^{N_1}) \otimes \dots \otimes (t_{n+m}^0, \dots, t_{n+m}^{N_{n+m}}).$$

Soit

$$\mathcal{P}_\epsilon^{(1)} = (t_1^0, \dots, t_1^{N_1}) \otimes \dots \otimes (t_n^0, \dots, t_n^{N_n}) \quad \text{et} \quad \mathcal{P}_\epsilon^{(2)} = (t_{n+1}^0, \dots, t_{n+1}^{N_{n+1}}) \otimes \dots \otimes (t_{n+m}^0, \dots, t_{n+m}^{N_{n+m}}),$$

des partitions de respectivement $R^{(1)}$ et $R^{(2)}$. Alors, pour tout $\mathbf{y} \in R^{(2)}$,

$$\underline{S}(f(\cdot, \mathbf{y}), \mathcal{P}_\epsilon^{(1)}) \leq \int_{R^{(1)}} f(\mathbf{x}, \mathbf{y}) d\mathbf{x} = G(\mathbf{y}) \leq \bar{S}(f(\cdot, \mathbf{y}), \mathcal{P}_\epsilon^{(1)}).$$

Donc

$$\begin{aligned} \underline{S}(G, \mathcal{P}_\epsilon^{(2)}) &= \sum_{Q^{(2)} \in \mathcal{P}_\epsilon^{(2)}} \left(\inf_{\mathbf{y} \in Q^{(2)}} G(\mathbf{y}) \right) \text{Vol}(Q^{(2)}) \\ &\geq \sum_{Q^{(2)} \in \mathcal{P}_\epsilon^{(2)}} \left(\inf_{\mathbf{y} \in Q^{(2)}} \underline{S}(f(\cdot, \mathbf{y}), \mathcal{P}_\epsilon^{(1)}) \right) \text{Vol}(Q^{(2)}) \\ &= \sum_{Q^{(2)} \in \mathcal{P}_\epsilon^{(2)}} \inf_{\mathbf{y} \in Q^{(2)}} \left(\sum_{Q^{(1)} \in \mathcal{P}_\epsilon^{(1)}} \inf_{\mathbf{x} \in Q^{(1)}} f(\mathbf{x}, \mathbf{y}) \text{Vol}(Q^{(1)}) \right) \text{Vol}(Q^{(2)}) \\ &\geq \sum_{Q^{(2)} \in \mathcal{P}_\epsilon^{(2)}} \sum_{Q^{(1)} \in \mathcal{P}_\epsilon^{(1)}} \left(\inf_{(\mathbf{x}, \mathbf{y}) \in Q^{(1)} \times Q^{(2)}} f(\mathbf{x}, \mathbf{y}) \right) \text{Vol}(Q^{(1)}) \text{Vol}(Q^{(2)}) \\ &\geq \underline{S}(f, \mathcal{P}_\epsilon). \end{aligned}$$

De façon similaire, on montre que $\bar{S}(G, \mathcal{P}_\epsilon^{(2)}) \leq \bar{S}(f, \mathcal{P}_\epsilon)$ et donc

$$\bar{S}(G, \mathcal{P}_\epsilon^{(2)}) - \underline{S}(G, \mathcal{P}_\epsilon^{(2)}) \leq \bar{S}(f, \mathcal{P}_\epsilon) - \underline{S}(f, \mathcal{P}_\epsilon) < \epsilon,$$

ce qui implique que G est intégrable au sens de Riemann sur $R^{(2)}$. De plus,

$$\underline{S}(f, \mathcal{P}_\epsilon) \leq \underline{S}(G, \mathcal{P}_\epsilon^{(2)}) \leq \int_{R^{(2)}} G(\mathbf{y}) d\mathbf{y} \leq \bar{S}(G, \mathcal{P}_\epsilon^{(2)}) \leq \bar{S}(f, \mathcal{P}_\epsilon)$$

et

$$\underline{S}(f, \mathcal{P}_\epsilon) \leq \int_R f(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y} \leq \bar{S}(f, \mathcal{P}_\epsilon)$$

avec $\bar{S}(f, \mathcal{P}_\epsilon) - \underline{S}(f, \mathcal{P}_\epsilon) < \epsilon$; par l'arbitrarité de ϵ , on conclut

$$\int_{R^{(2)}} G(\mathbf{y}) d\mathbf{y} = \int_R f(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y}.$$

□

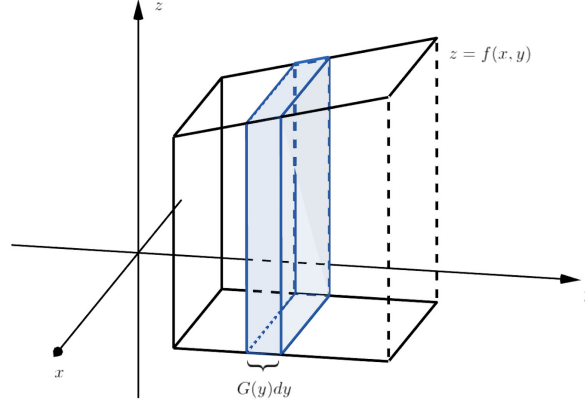


FIGURE 8.2 – Interprétation géométrique de la formule des intégrales itérées

Corollaire 8.15. Soit $R = R^{(1)} \times R^{(2)} \subset \mathbb{R}^{n+m}$ et $f \in C^0(R)$. Alors

- $\forall \mathbf{y} \in R^{(2)}, G(\mathbf{y}) = \int_{R^{(1)}} f(\mathbf{x}, \mathbf{y}) d\mathbf{x}$ existe,
- $\forall \mathbf{x} \in R^{(1)}, F(\mathbf{x}) = \int_{R^{(2)}} f(\mathbf{x}, \mathbf{y}) d\mathbf{y}$ existe,
- $\int_{R^{(2)}} G(\mathbf{y}) d\mathbf{y} = \int_{R^{(1)}} F(\mathbf{x}) d\mathbf{x} = \int_R f(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y}$.

Détaillons ce résultat en dimension $n = 2$: $R = [a, b] \times [c, d]$, $f : R \rightarrow \mathbb{R}$ continue. Alors

$$\int_R f(x, y) dx dy = \int_c^d \left(\int_a^b f(x, y) dx \right) dy = \int_a^b \left(\int_c^d f(x, y) dy \right) dx$$

Cette formule prend le nom de formule des *intégrales itérées*. La Figure 8.2 donne une interprétation géométrique.

Exemple 8.16. Soit $R = [1, 2] \times [0, 1]$, et $f(x, y) = \frac{1}{x^3} e^{y/x} : R \rightarrow \mathbb{R}$. Clairement, f est continue sur R (fermé). On peut donc calculer l'intégrale double de f par la formule des intégrales itérées :

$$\int_R f(x, y) dx dy = \underbrace{\int_1^2 \left(\int_0^1 f(x, y) dy \right) dx}_{(A)} = \underbrace{\int_0^1 \left(\int_1^2 f(x, y) dx \right) dy}_{(B)}.$$

Utilisons (A) :

$$\begin{aligned} \int_1^2 \left(\int_0^1 \frac{1}{x^3} e^{y/x} dy \right) dx &= \int_1^2 \frac{1}{x^3} x e^{y/x} \Big|_{y=0}^{y=1} dx \\ &= \int_1^2 \frac{1}{x^2} (e^{1/x} - 1) dx = \int_1^2 \frac{1}{x^2} e^{1/x} dx - \int_1^2 \frac{1}{x^2} dx \\ &= - \int_1^{1/2} e^t dt + \frac{1}{x} \Big|_1^2 = e - \sqrt{e} - \frac{1}{2}. \end{aligned}$$

On a donc $\int_R f(x, y) dx dy = e - \sqrt{e} - \frac{1}{2}$.

8.4 Intégrale de Riemann sur un ensemble quelconque

On souhaite maintenant, généraliser la définition d'intégrale au sens de Riemann d'une fonction bornée $f : E \rightarrow \mathbb{R}$ sur un sous-ensemble $E \subset \mathbb{R}^n$ quelconque borné. L'idée est simple : puisque E est borné, il existe un pavé $R \supset E$ qui le contient. On prolonge alors par zéro la fonction f et on l'intègre sur le pavé R .

Définition 8.17 (Intégrale de Riemann sur un ensemble quelconque). *Soit $E \subset \mathbb{R}^n$ borné, $f : E \rightarrow \mathbb{R}$ bornée et $R \subset \mathbb{R}^n$ un pavé contenant E . On dit que f est intégrable au sens de Riemann si la fonction $\tilde{f} : R \rightarrow \mathbb{R}$ définie par*

$$\tilde{f}(\mathbf{x}) = f(\mathbf{x}), \quad \text{si } \mathbf{x} \in E, \quad \tilde{f}(\mathbf{x}) = 0, \quad \text{si } \mathbf{x} \in R \setminus E$$

est intégrable au sens de Riemann. Dans ce cas, on définit l'intégrale de f sur E par

$$\int_E f(\mathbf{x}) d\mathbf{x} = \int_R \tilde{f}(\mathbf{x}) d\mathbf{x}. \quad (8.1)$$

On note $\mathcal{R}(E)$ l'ensemble de fonctions $f : E \rightarrow \mathbb{R}$ intégrables au sens de Riemann sur E .

Il est parfois commode de considérer $E = \emptyset$. Dans ce cas \tilde{f} est nulle sur R et donc Riemann-intégrable avec $\int_R \tilde{f}(\mathbf{x}) d\mathbf{x} = 0$. On conviendra donc que $\int_E f(\mathbf{x}) d\mathbf{x} = 0$ dès que $E = \emptyset$.

Il est important de vérifier que la définition de l'intégrale en (8.1) ne dépende pas du choix de R .

Lemme 8.18. *Soit $E \subset \mathbb{R}^n$ borné, $f : E \rightarrow \mathbb{R}$ bornée et $\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R}$ le prolongement de f par zéro au dehors de E . S'il existe un pavé $R \supset E$ tel que la fonction $\tilde{f} : R \rightarrow \mathbb{R}$ est intégrable, alors pour tout autre pavé $R' \supset E$, la fonction $\tilde{f} : R' \rightarrow \mathbb{R}$ est intégrable et on a $\int_R \tilde{f}(\mathbf{x}) d\mathbf{x} = \int_{R'} \tilde{f}(\mathbf{x}) d\mathbf{x}$.*

Démonstration. Soit $R \supset E$ un pavé tel que la fonction $\tilde{f} : R \rightarrow \mathbb{R}$ est intégrable au sens de Riemann, soit $R' \supset E$ un autre pavé et soit \hat{R} un troisième pavé dont l'intérieur contient aussi bien R que R' , i.e. $\hat{R} \supset R \cup R'$.

Par le lemme 8.13 appliqué à $\tilde{f} : R \rightarrow \mathbb{R}$ et $\tilde{f} : \hat{R} \rightarrow \mathbb{R}$, on sait que $\tilde{f} : \hat{R} \rightarrow \mathbb{R}$ est Riemann-intégrable et $\int_{\hat{R}} \tilde{f} d\mathbf{x} = \int_R \tilde{f} d\mathbf{x}$. Par le même lemme appliqué à $\tilde{f} : R' \rightarrow \mathbb{R}$ et $\tilde{f} : \hat{R} \rightarrow \mathbb{R}$, on sait ensuite que $\tilde{f} : R' \rightarrow \mathbb{R}$ est Riemann-intégrable et $\int_{R'} \tilde{f} d\mathbf{x} = \int_{\hat{R}} \tilde{f} d\mathbf{x} = \int_R \tilde{f} d\mathbf{x}$. \square

L'intégrale multiple au sens de Riemann a les propriétés suivantes, analogues à celles de l'intégrale en dimension $n = 1$.

Théorème 8.19 (Propriétés de l'intégrale de Riemann).

- (i) $\mathcal{R}(E)$ est un espace vectoriel et l'intégrale de Riemann est linéaire, c'est-à-dire, si $f, g \in \mathcal{R}(E)$ et $\alpha, \beta \in \mathbb{R}$, alors $\alpha f + \beta g \in \mathcal{R}(E)$ et

$$\int_E (\alpha f + \beta g) = \alpha \int_E f + \beta \int_E g.$$

(ii) Si $f, g \in \mathcal{R}(E)$ et $f(\mathbf{x}) \leq g(\mathbf{x}) \forall \mathbf{x} \in E$ alors

$$\int_E f(\mathbf{x}) d\mathbf{x} \leq \int_E g(\mathbf{x}) d\mathbf{x}.$$

(iii) Si $f \in \mathcal{R}(E)$, alors $|f|, f_+, f_- \in \mathcal{R}(E)$, où $f_+ = \max\{f, 0\}$ et $f_- = \max\{-f, 0\}$; de plus

$$\left| \int_E f(\mathbf{x}) d\mathbf{x} \right| \leq \int_E |f(\mathbf{x})| d\mathbf{x}.$$

(iv) Si $f, g \in \mathcal{R}(E)$ alors $fg \in \mathcal{R}(E)$ (i.e. $\mathcal{R}(E)$ est une algèbre).

Démonstration. Soit $R \subset \mathbb{R}^n$ un pavé contenant E et $\tilde{f}, \tilde{g} : R \rightarrow \mathbb{R}$ les prolongements de f et g par zéro en dehors de E .

La démonstration des points (i) et (ii) est laissée comme exercice aux étudiants.

(iii) Pour $\epsilon > 0$, soit \mathcal{P}_ϵ une partition de R telle que $\overline{S}(\tilde{f}, \mathcal{P}_\epsilon) - \underline{S}(\tilde{f}, \mathcal{P}_\epsilon) < \epsilon$. Alors, $\forall Q \in \mathcal{P}_\epsilon$, on a

$$\sup_Q \tilde{f}_+ - \inf_Q \tilde{f}_+ \leq \sup_Q \tilde{f} - \inf_Q \tilde{f}.$$

En effet :

- si $\sup_Q \tilde{f} \geq \inf_Q \tilde{f} \geq 0$, alors $\tilde{f}_+(\mathbf{x}) = \tilde{f}(\mathbf{x}) \quad \forall \mathbf{x} \in Q$ et on a égalité;
- si $\inf_Q \tilde{f} \leq \sup_Q \tilde{f} \leq 0$, alors $\sup_Q \tilde{f}_+ - \inf_Q \tilde{f}_+ = 0 \leq \sup_Q \tilde{f} - \inf_Q \tilde{f}$;
- si $\inf_Q \tilde{f} \leq 0 \leq \sup_Q \tilde{f}$, alors, pour tout $\mathbf{x} \in Q$, $\tilde{f}_+(\mathbf{x}) = \tilde{f}(\mathbf{x}) + \tilde{f}_-(\mathbf{x}) \leq \tilde{f}(\mathbf{x}) + \sup_Q \tilde{f}_- = \tilde{f}(\mathbf{x}) - \inf_Q \tilde{f}$ et

$$\sup_Q \tilde{f}_+ - \inf_Q \tilde{f}_+ = \sup_Q \tilde{f}_+ \leq \sup_Q \tilde{f} - \inf_Q \tilde{f}.$$

Il s'ensuit que $\overline{S}(\tilde{f}_+, \mathcal{P}_\epsilon) - \underline{S}(\tilde{f}_+, \mathcal{P}_\epsilon) \leq \overline{S}(\tilde{f}, \mathcal{P}_\epsilon) - \underline{S}(\tilde{f}, \mathcal{P}_\epsilon) < \epsilon$ et $\tilde{f}_+ \in \mathcal{R}(R)$, donc $f_+ \in \mathcal{R}(E)$. Comme $f_- = (-f)_+$ on a aussi $f_- \in \mathcal{R}(E)$ et $|f| = f_+ + f_- \in \mathcal{R}(E)$. Enfin,

$$\begin{aligned} \text{puisque } f \leq |f|, \quad \text{on a } \int_E f(\mathbf{x}) d\mathbf{x} &\leq \int_E |f(\mathbf{x})| d\mathbf{x}, \\ \text{puisque } -f \leq |f|, \quad \text{on a } -\int_E f(\mathbf{x}) d\mathbf{x} &\leq \int_E |f(\mathbf{x})| d\mathbf{x}, \end{aligned}$$

ce qui implique que $|\int_E f(\mathbf{x}) d\mathbf{x}| \leq \int_E |f(\mathbf{x})| d\mathbf{x}$.

(iv) Supposons d'abord $f, g \geq 0$ et $f(\mathbf{x}) \leq M, g(\mathbf{x}) \leq M \quad \forall \mathbf{x} \in E$ avec $M \in]0, +\infty[$ (M existe car f et g sont bornées). Pour tout pavé $Q \subset R$ on a

$$\begin{aligned} \sup_Q \tilde{f} \cdot \tilde{g} - \inf_Q \tilde{f} \cdot \tilde{g} &\leq \sup_Q \tilde{f} \cdot \sup_Q \tilde{g} - \inf_Q \tilde{f} \cdot \inf_Q \tilde{g} \\ &\leq (\sup_Q \tilde{f} - \inf_Q \tilde{f}) \sup_Q \tilde{g} + \inf_Q \tilde{f} (\sup_Q \tilde{g} - \inf_Q \tilde{g}) \\ &\leq M(\sup_Q \tilde{f} - \inf_Q \tilde{f}) + M(\sup_Q \tilde{g} - \inf_Q \tilde{g}). \end{aligned}$$

Soit maintenant $\epsilon > 0$ et \mathcal{P}_ϵ une partition de R telle que

$$\overline{S}(\tilde{f}, \mathcal{P}_\epsilon) - \underline{S}(\tilde{f}, \mathcal{P}_\epsilon) < \frac{\epsilon}{2M}, \quad \overline{S}(\tilde{g}, \mathcal{P}_\epsilon) - \underline{S}(\tilde{g}, \mathcal{P}_\epsilon) < \frac{\epsilon}{2M}.$$

Alors

$$\overline{S}(\tilde{f}\tilde{g}, \mathcal{P}_\epsilon) - \underline{S}(\tilde{f}\tilde{g}, \mathcal{P}_\epsilon) \leq M(\overline{S}(\tilde{f}, \mathcal{P}_\epsilon) - \underline{S}(\tilde{f}, \mathcal{P}_\epsilon)) + M(\overline{S}(\tilde{g}, \mathcal{P}_\epsilon) - \underline{S}(\tilde{g}, \mathcal{P}_\epsilon)) < \epsilon$$

et donc $\tilde{f}\tilde{g} \in \mathcal{R}(R)$ ce qui implique que $fg \in \mathcal{R}(E)$. Si f n'est pas positive ou g n'est pas positive (ou ni l'une ni l'autre), on a $fg = (f_+ - f_-)(g_+ - g_-) = f_+g_+ - f_+g_- - f_-g_+ + f_-g_-$ et chaque terme $f_\pm, g_\pm \in \mathcal{R}(E)$, donc $fg \in \mathcal{R}(E)$. \square

8.5 Ensembles mesurables au sens de Jordan

Il est clair que si E est un sous-ensemble quelconque borné de \mathbb{R}^n et $f : E \rightarrow \mathbb{R}$ une fonction quelconque bornée, le fait que f soit intégrable sur E au sens de Riemann dépendra aussi bien des propriétés de f que des propriétés de l'ensemble E . En particulier, on peut se poser la question si une fonction $f : E \rightarrow \mathbb{R}$ continue est intégrable. La réponse à cette question n'est pas toujours affirmative et dépend des propriétés de E , notamment qu'il soit *mesurable au sens de Jordan*, notion qu'on va introduire dans cette section.

On introduit la fonction caractéristique de l'ensemble E , notée $\mathbf{1}_E : \mathbb{R}^n \rightarrow \mathbb{R}$ et définie par

$$\mathbf{1}_E(\mathbf{x}) = 1 \text{ si } \mathbf{x} \in E, \quad \mathbf{1}_E(\mathbf{x}) = 0 \text{ si } \mathbf{x} \notin E.$$

On peut se poser la question de savoir pour quels sous-ensembles $E \subset \mathbb{R}^n$ bornés la fonction $\mathbf{1}_E$ est intégrable au sens de Riemann.

Définition 8.20 (Ensemble mesurable au sens de Jordan). *Soit $E \subset \mathbb{R}^n$ borné. On dit que E est mesurable au sens de Jordan (ou “Jordan-mesurable” ou simplement “mesurable”) si $\mathbf{1}_E \in \mathcal{R}(E)$. Dans ce cas on pose*

$$\text{Vol}(E) = \int_E \mathbf{1}_E(\mathbf{x}) d\mathbf{x}.$$

Un ensemble borné $E \subset \mathbb{R}^n$ est dit négligeable s'il est mesurable au sens de Jordan et $\text{Vol}(E) = 0$.

Une caractérisation équivalente d'ensemble Jordan-mesurable est la suivante.

Lemme 8.21. *Soit $E \subset \mathbb{R}^n$ borné et $R \subset \mathbb{R}^n$ un pavé contenant E . Alors E est Jordan-mesurable si et seulement, pour tout $\epsilon > 0$, il existe une partition \mathcal{P}_ϵ de R telle que*

$$\sum_{\substack{Q \in \mathcal{P}_\epsilon \\ Q \cap E \neq \emptyset, \quad Q \cap (\mathbb{R}^n \setminus E) \neq \emptyset}} \text{Vol}(Q) < \epsilon.$$

Démonstration. Par le Lemme 8.10, E est Jordan-mesurable si, et seulement si, pour tout $\epsilon > 0$, il existe une partition \mathcal{P}_ϵ de R telle que $\overline{S}(\mathbf{1}_E, \mathcal{P}_\epsilon) - \underline{S}(\mathbf{1}_E, \mathcal{P}_\epsilon) < \epsilon$. Or

$$\overline{S}(\mathbf{1}_E, \mathcal{P}_\epsilon) = \sum_{\substack{Q \in \mathcal{P}_\epsilon \\ Q \cap E \neq \emptyset}} \text{Vol}(Q)$$

et

$$\underline{S}(\mathbb{1}_E, \mathcal{P}_\epsilon) = \sum_{\substack{Q \in \mathcal{P}_\epsilon \\ Q \cap (\mathbb{R}^n \setminus E) = \emptyset}} \text{Vol}(Q).$$

□

De façon similaire, on peut donner une caractérisation équivalente d'ensemble négligeable.

Lemme 8.22. *Un ensemble borné $E \subset \mathbb{R}^n$ est négligeable si et seulement si $\forall \epsilon > 0$ il existe $K \in \mathbb{N}^*$ et une collection finie de pavés R_1, \dots, R_K tels que $E \subset \bigcup_{i=1}^K R_i$ et $\sum_{i=1}^K \text{Vol}(R_i) < \epsilon$.*

Démonstration. Choisissons un pavé $R \subset \mathbb{R}^n$ tel que $E \subset \overset{\circ}{R}$.

Supposons d'abord que $\mathbb{1}_E$ est Riemann-intégrable sur R et $\int_R \mathbb{1}_E(\mathbf{x}) d\mathbf{x} = 0$. Rappelons qu'alors $\int_R \mathbb{1}_E(\mathbf{x}) d\mathbf{x}$ est l'infimum des sommes de Darboux supérieures sur toutes les partitions de R . Soit $\epsilon > 0$. Il existe donc une partition $\mathcal{P}_\epsilon = \{Q_i \subset R : 1 \leq i \leq L\}$ de R telle que $\overline{S}(\mathbb{1}_E, \mathcal{P}_\epsilon) < \epsilon$. Or

$$\overline{S}(\mathbb{1}_E, \mathcal{P}_\epsilon) = \sum_{i \in \{1, \dots, L\}, Q_i \cap E \neq \emptyset} \text{Vol}(Q_i) < \epsilon$$

et $E \subset \bigcup_{i \in \{1, \dots, L\}, Q_i \cap E \neq \emptyset} Q_i$. Ceci termine la première partie de la preuve.

Soit $\epsilon > 0$. Soit aussi des pavés R_1, \dots, R_K tels que $E \subset \bigcup_{i=1}^K R_i$ et $\sum_{i=1}^K \text{Vol}(R_i) < \epsilon$. Sans perte de généralité, on peut supposer que $R_i \subset R$ pour tout $1 \leq i \leq K$ et que $\forall \mathbf{x} \in E \exists i \in \{1, \dots, K\} \mathbf{x} \in \overset{\circ}{R}_i$.

Soit la partition tensorielle \mathcal{P}_ϵ de R définie à partir de toutes les composantes de tous les sommets de R_1, \dots, R_K, R . Alors

$$\forall i \in \{1, \dots, K\} \quad R_i = \bigcup \{Q \in \mathcal{P}_\epsilon : Q \subset R_i\},$$

$\underline{S}(\mathbb{1}_E, \mathcal{P}_\epsilon) \geq 0$ et

$$\overline{S}(\mathbb{1}_E, \mathcal{P}_\epsilon) = \sum_{Q \in \mathcal{P}_\epsilon, Q \cap E \neq \emptyset} \text{Vol}(Q) \leq \sum_{i=1}^K \sum_{Q \in \mathcal{P}_\epsilon, Q \subset R_i} \text{Vol}(Q) = \sum_{i=1}^K \text{Vol}(R_i) < \epsilon,$$

où on a utilisé que

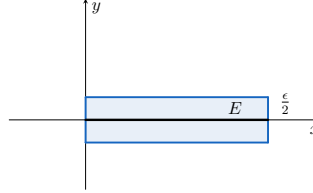
$$\forall Q \in \mathcal{P}_\epsilon \quad (Q \cap E \neq \emptyset \Rightarrow \exists i \in \{1, \dots, K\} Q \subset R_i).$$

Comme $\epsilon > 0$ est arbitraire, $\mathbb{1}_E$ est intégrable sur R et $\int_R \mathbb{1}_E(\mathbf{x}) d\mathbf{x} = 0$. □

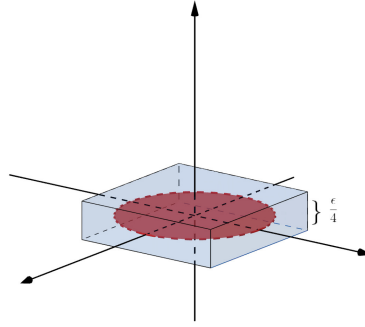
Remarque 8.23. *Il résulte de ce lemme que toute union finie d'ensembles négligeables est un ensemble négligeable, et que tout sous-ensemble d'un ensemble négligeable est aussi négligeable.*

Exemple 8.24. *Considérons les ensembles suivants :*

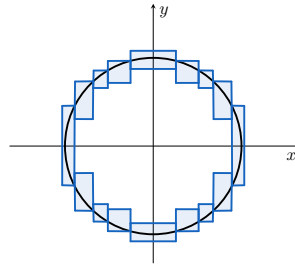
1. $E = \{(x, y) \in \mathbb{R}^2 : x \in [0, 1], y = 0\}$. Alors E est un pavé de volume nul et donc E est négligeable. Dans le critère du lemme, on peut choisir $K = 1$ et $R_1 = E$. Si on souhaite, pour une raison ou une autre, une collection finie de pavés non dégénérés, on s'en sort dans cet exemple en choisissant un seul pavé $R = [0, 1] \times [-\frac{\epsilon}{4}, \frac{\epsilon}{4}]$ ($K = 1$), et on a bien $E \subset R$ avec $\text{Vol}(R) < \epsilon$.



2. $E = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 \leq 1, z = 0\}$. Soit $R = [-1, 1] \times [-1, 1] \times [-\frac{\epsilon}{8}, \frac{\epsilon}{8}]$, alors $E \subset R$, $\text{Vol}(R) = \epsilon$ donc E est négligeable.



3. $E = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ est négligeable.



L'important résultat suivant caractérise les ensembles mesurables au sens de Jordan.

Théorème 8.25. *Un ensemble borné $E \subset \mathbb{R}^n$ est mesurable au sens de Jordan si et seulement si $\partial E = \bar{E} \setminus \overset{\circ}{E}$ est négligeable.*

Démonstration. Soit $R \subset \mathbb{R}^n$ un pavé tel que $E \subset R$.

« \Rightarrow » (E mesurable $\Rightarrow \partial E$ négligeable) Soit $\epsilon > 0$. Par le Lemme 8.21, il existe une partition \mathcal{P}_ϵ de R telle que $\sum_{Q \in \mathcal{P}_\epsilon} \text{Vol}(Q) < \epsilon$, où

$$\mathcal{E} = \{Q \in \mathcal{P}_\epsilon : Q \cap E \neq \emptyset, Q \cap E^c \neq \emptyset\}.$$

Pour tout $\mathbf{x} \in \partial E$, il existe au moins un pavé $Q \in \mathcal{P}_\epsilon$ qui le contient. Alors soit $\mathbf{x} \in \partial Q$, soit $\mathbf{x} \in \overset{\circ}{Q}$, mais dans ce dernier cas on a nécessairement $Q \in \mathcal{E}$. On a donc

$$\partial E \subset \left(\bigcup_{Q \in \mathcal{E}} Q \right) \cup \left(\bigcup_{Q \in \mathcal{P}_\epsilon} \partial Q \right).$$

Le bord de tout pavé étant une union finie de pavés de volume nul, ∂E est inclus dans l'union d'une famille finie de pavés (éventuellement dégénérés) dont la somme des volumes est $< \epsilon$. Comme $\epsilon > 0$ est arbitraire, ∂E est négligeable.

« \Leftarrow » (∂E négligeable $\Rightarrow E$ mesurable) Puisque $\mathbb{1}_{\partial E} \in \mathcal{R}(R)$ et $\int_R \mathbb{1}_{\partial E}(\mathbf{x}) d\mathbf{x} = 0$, $\forall \epsilon > 0$ il existe une partition \mathcal{P}_ϵ de R telle que $\overline{S}(\mathbb{1}_{\partial E}, \mathcal{P}_\epsilon) < \epsilon$. Mais pour $Q \in \mathcal{P}_\epsilon$, $\sup_Q \mathbb{1}_{\partial E} = 1$ si et seulement si $Q \cap \partial E \neq \emptyset$. D'où

$$\sum_{Q \in \mathcal{P}_\epsilon, Q \cap \partial E \neq \emptyset} \text{Vol}(Q) < \epsilon.$$

Observons que si un certain $Q \in \mathcal{P}_\epsilon$ rencontre à la fois E et E^c , alors $Q \cap \partial E \neq \emptyset$. Pour le voir, considérons $\mathbf{x} \in Q \cap E^c$ et $\mathbf{y} \in Q \cap E$. L'ensemble $\{t \in [0, 1] : (1-t)\mathbf{x} + t\mathbf{y} \in E\}$ n'est pas vide (il contient $t = 1$) ; soit $t_0 \in [0, 1]$ son infimum. Alors $t_0\mathbf{x} + (1-t_0)\mathbf{y} \in Q \cap \partial E$ et donc $Q \cap \partial E \neq \emptyset$.

On obtient ainsi

$$\overline{S}(\mathbb{1}_E, \mathcal{P}_\epsilon) - \underline{S}(\mathbb{1}_E, \mathcal{P}_\epsilon) = \sum_{\substack{Q \in \mathcal{P}_\epsilon \\ Q \cap E \neq \emptyset, Q \cap E^c \neq \emptyset}} \text{Vol}(Q) \leq \sum_{\substack{Q \in \mathcal{P}_\epsilon \\ Q \cap \partial E \neq \emptyset}} \text{Vol}(Q) < \epsilon.$$

Comme $\epsilon > 0$ est arbitraire, ceci prouve que $\mathbb{1}_E$ est Riemann-intégrable, et donc E est Jordan-mesurable. \square

Corollaire 8.26. Soient $E, F \subset \mathbb{R}^n$ des ensembles bornés et mesurables au sens de Jordan. Alors

$$E \cap F, E \cup F, E \setminus F, \overset{\circ}{E}, \overline{E}$$

sont mesurables.

Démonstration. Soit $R \subset \mathbb{R}^n$ un pavé contenant $E \cup F$. Par hypothèse $\mathbb{1}_E, \mathbb{1}_F \in \mathcal{R}(R)$. On a

- $\mathbb{1}_{E \cap F} = \mathbb{1}_E \cdot \mathbb{1}_F \in \mathcal{R}(R)$;
- $\mathbb{1}_{E \cup F} = \mathbb{1}_E + \mathbb{1}_F - \mathbb{1}_{E \cap F} \in \mathcal{R}(R)$;
- $\mathbb{1}_{E \setminus F} = \mathbb{1}_E - \mathbb{1}_{E \cap F} \in \mathcal{R}(R)$;
- $\mathbb{1}_{\overset{\circ}{E}} = \mathbb{1}_E - \mathbb{1}_{E \cap \partial E} \in \mathcal{R}(R)$;
- $\mathbb{1}_{\overline{E}} = \mathbb{1}_{E \cup \partial E} \in \mathcal{R}(R)$.

\square

Définition 8.27. On note $\mathcal{J}(\mathbb{R}^n)$ la collection des sous-ensembles de \mathbb{R}^n compacts et mesurables au sens de Jordan. Plus généralement, pour $E \subset \mathbb{R}^n$, on note $\mathcal{J}(E)$ la collection des sous-ensembles de E qui sont compacts et mesurables au sens de Jordan.

8.6 Caractérisation des fonctions intégrables

On souhaite maintenant donner une caractérisation des fonctions intégrables au sens de Riemann sur un domaine quelconque. On commence par l'important résultat suivant :

Théorème 8.28. *Soit un pavé $R \subset \mathbb{R}^n$ et une fonction bornée $f : R \rightarrow \mathbb{R}$. Si l'ensemble des points de discontinuité de f est négligeable, alors f est Riemann-intégrable.*

Démonstration. Soit $M = \sup_{\mathbf{x} \in R} |f(\mathbf{x})|$, notons par N l'ensemble des points de discontinuité de f et fixons $\epsilon > 0$. Comme N est négligeable, il existe une partition \mathcal{P}_ϵ de R telle que

$$\overline{S}(\mathbf{1}_N, \mathcal{P}_\epsilon) = \sum_{Q \in \mathcal{P}_\epsilon, Q \cap N \neq \emptyset} \text{Vol}(Q) < \frac{\epsilon}{1 + 4M}.$$

Soit l'ensemble compact

$$K = \cup \{Q \in \mathcal{P}_\epsilon : Q \cap N = \emptyset\}$$

(union finie de fermés bornés). Comme $f : K \rightarrow \mathbb{R}$ est continue sur le compact K , elle est uniformément continue sur K . Il existe donc $\delta_\epsilon > 0$ tel que

$$\forall \mathbf{x}, \mathbf{y} \in K \quad (\|\mathbf{x} - \mathbf{y}\| < \delta_\epsilon \Rightarrow |f(\mathbf{x}) - f(\mathbf{y})| < \frac{\epsilon}{1 + 2 \text{Vol}(R)})$$

(si $K = \emptyset$, on peut choisir $\delta_\epsilon > 0$ librement). Quitte à prendre un raffinement de \mathcal{P}_ϵ , nous pouvons supposer que

$$\forall Q \in \mathcal{P}_\epsilon \quad (Q \subset K \Rightarrow \text{diam}(Q) < \delta_\epsilon).$$

On obtient alors

$$\begin{aligned} \overline{S}(f, \mathcal{P}_\epsilon) - \underline{S}(f, \mathcal{P}_\epsilon) &= \sum_{Q \in \mathcal{P}_\epsilon} \left(\sup_Q f - \inf_Q f \right) \text{Vol}(Q) \\ &= \left(\sum_{Q \in \mathcal{P}_\epsilon, Q \subset K} + \sum_{Q \in \mathcal{P}_\epsilon, Q \cap N \neq \emptyset} \right) \left(\sup_Q f - \inf_Q f \right) \text{Vol}(Q) \\ &\leq \frac{\epsilon}{1 + 2 \text{Vol}(R)} \text{Vol}(R) + 2M \frac{\epsilon}{1 + 4M} < \epsilon. \end{aligned}$$

Comme $\epsilon > 0$ est arbitraire, ceci prouve que f est Riemann-intégrable. \square

Continuons avec un autre résultat important.

Théorème 8.29. *Soit $E \subset \mathbb{R}^n$ borné et mesurable. Si $f : E \rightarrow \mathbb{R}$ est bornée sur E et continue sur \mathring{E} , alors f est intégrable sur E au sens de Riemann.*

Démonstration. Soit R un pavé contenant E et \tilde{f} le prolongement de f par zéro en dehors de E . Comme E est Jordan-mesurable, ∂E est négligeable. De plus l'ensemble \tilde{N} des points de discontinuité de \tilde{f} est inclus dans ∂E , et donc \tilde{N} est négligeable. Par le théorème précédent, $\tilde{f} \in \mathcal{R}(R)$ et donc f est intégrable sur E au sens de Riemann. \square

Comme cas particulier, on obtient le théorème suivant :

Théorème 8.30. *Soit $E \in \mathcal{J}(\mathbb{R}^n)$ (compact et mesurable au sens de Jordan) et $f : E \rightarrow \mathbb{R}$ continue. Alors f est intégrable sur E au sens de Riemann.*

Ce dernier théorème implique, en particulier, que $C^0(E) \subset \mathcal{R}(E)$ si $E \in \mathcal{J}(\mathbb{R}^n)$.

On conclut cette section en présentant d'autres propriétés de l'intégrale de Riemann, au delà de celles déjà énoncées dans le théorème 8.19.

Théorème 8.31 (Propriétés de l'intégrale de Riemann – suite du théorème 8.19).

(v) *Soit $E \subset \mathbb{R}^n$ borné et mesurable, et $f : E \rightarrow \mathbb{R}$ bornée. Alors*

$$\inf_{\mathbf{x} \in E} f(\mathbf{x}) \text{Vol}(E) \leq \int_{-R} \tilde{f}(\mathbf{x}) d\mathbf{x} \leq \int_R \tilde{f}(\mathbf{x}) d\mathbf{x} \leq \sup_{\mathbf{x} \in E} f(\mathbf{x}) \text{Vol}(E)$$

pour tout pavé R contenant E et le prolongement $\tilde{f} : R \rightarrow \mathbb{R}$ de f par la valeur 0. En particulier, si $f \in \mathcal{R}(E)$,

$$\inf_{\mathbf{x} \in E} f(\mathbf{x}) \text{Vol}(E) \leq \int_R \tilde{f}(\mathbf{x}) d\mathbf{x} = \int_E f(\mathbf{x}) d\mathbf{x} \leq \sup_{\mathbf{x} \in E} f(\mathbf{x}) \text{Vol}(E).$$

(vi) *Soit $E \in \mathcal{J}(\mathbb{R}^n)$ connexe par arcs et $f \in C^0(E)$. Si $\text{Vol}(E) \neq 0$, alors*

$$\exists \mathbf{x}_0 \in E \quad \frac{1}{\text{Vol}(E)} \int_E f(\mathbf{x}) d\mathbf{x} = f(\mathbf{x}_0).$$

Cette dernière égalité s'appelle le théorème de la moyenne.

Démonstration. (v) Soit $M = 1 + \sup_{\mathbf{x} \in E} |f(\mathbf{x})|$ et fixons $\epsilon > 0$. Il existe une partition \mathcal{P}_ϵ de R telle que

$$\overline{S}(\mathbf{1}_E, \mathcal{P}_\epsilon) - \underline{S}(\mathbf{1}_E, \mathcal{P}_\epsilon) < \frac{\epsilon}{M}$$

et donc

$$\sum_{\substack{Q \in \mathcal{P}_\epsilon \\ Q \cap E \neq \emptyset, Q \cap (\mathbb{R}^n \setminus E) \neq \emptyset}} \text{Vol}(Q) < \frac{\epsilon}{M}.$$

On a donc, où $Q \in \mathcal{P}_\epsilon$,

$$\begin{aligned} \overline{S}(\tilde{f}, \mathcal{P}_\epsilon) &= \left(\sum_{Q \subset E} + \sum_{Q \cap E \neq \emptyset, Q \cap E^c \neq \emptyset} + \sum_{Q \subset E^c} \right) \sup_{\mathbf{x} \in Q} \tilde{f}(\mathbf{x}) \text{Vol}(Q) \\ &\leq \sum_{Q \subset E} \sup_{\mathbf{x} \in E} f(\mathbf{x}) \text{Vol}(Q) + \sum_{Q \cap E \neq \emptyset, Q \cap E^c \neq \emptyset} M \text{Vol}(Q) + 0 < \sup_{\mathbf{x} \in E} f(\mathbf{x}) \underline{S}(\mathbf{1}_E, \mathcal{P}_\epsilon) + M \frac{\epsilon}{M} + 0 \\ &\leq \sup_{\mathbf{x} \in E} f(\mathbf{x}) \text{Vol}(E) + M |\underline{S}(\mathbf{1}_E, \mathcal{P}_\epsilon) - \text{Vol}(E)| + \epsilon < \sup_{\mathbf{x} \in E} f(\mathbf{x}) \text{Vol}(E) + 2\epsilon \end{aligned}$$

car

$$\underline{S}(\mathbf{1}_E, \mathcal{P}_\epsilon) \leq \text{Vol}(E) \leq \overline{S}(\mathbf{1}_E, \mathcal{P}_\epsilon) \quad \text{avec} \quad \overline{S}(\mathbf{1}_E, \mathcal{P}_\epsilon) - \underline{S}(\mathbf{1}_E, \mathcal{P}_\epsilon) < \frac{\epsilon}{M}.$$

De même,

$$\underline{S}(\tilde{f}, \mathcal{P}_\epsilon) > \inf_{\mathbf{x} \in E} f(\mathbf{x}) \text{Vol}(E) - 2\epsilon.$$

D'où

$$\overline{\int_R \tilde{f}(\mathbf{x}) d\mathbf{x}} < \sup_{\mathbf{x} \in E} f(\mathbf{x}) \text{Vol}(E) + 2\epsilon$$

et

$$\underline{\int_R \tilde{f}(\mathbf{x}) d\mathbf{x}} > \inf_{\mathbf{x} \in E} f(\mathbf{x}) \text{Vol}(E) - 2\epsilon.$$

Comme $\epsilon > 0$ est arbitraire, on obtient la conclusion voulue.

(vi) Par (v) et le fait que f est continue sur un compact non vide, on a

$$\min_E f \leq \frac{1}{\text{Vol}(E)} \int_E f(\mathbf{x}) d\mathbf{x} \leq \max_E f.$$

Or, puisque E est compact et connexe par arcs,

$$\text{Im}(f) = \left[\min_{\mathbf{x} \in E} f(\mathbf{x}), \max_{\mathbf{x} \in E} f(\mathbf{x}) \right].$$

Donc il existe $\mathbf{x}_0 \in E$ tel que $f(\mathbf{x}_0) = \frac{1}{\text{Vol}(E)} \int_E f(\mathbf{x}) d\mathbf{x}$. □

Voici un corollaire très utile :

Corollaire 8.32. *Soit $E \subset \mathbb{R}^n$ borné et négligeable, et $f : E \rightarrow \mathbb{R}$ bornée. Alors $f \in \mathcal{R}(E)$ et $\int_E f(\mathbf{x}) d\mathbf{x} = 0$.*

Démonstration. En effet

$$0 = \inf_{\mathbf{x} \in E} f(\mathbf{x}) \text{Vol}(E) \leq \underline{\int_R \tilde{f}(\mathbf{x}) d\mathbf{x}} \leq \overline{\int_R \tilde{f}(\mathbf{x}) d\mathbf{x}} \leq \sup_{\mathbf{x} \in E} f(\mathbf{x}) \text{Vol}(E) = 0$$

pour tout pavé R contenant E et le prolongement $\tilde{f} : R \rightarrow \mathbb{R}$ de f par la valeur 0. □

Le théorème suivant montre l'additivité de l'intégrale par rapport au domaine d'intégration.

Théorème 8.33. *Soient $E_1, E_2 \subset \mathbb{R}^n$ bornés, tels que $E_1 \cap E_2$ est négligeable et notons $E = E_1 \cup E_2$. Si $f : E \rightarrow \mathbb{R}$ bornée est telle que ses restrictions $f|_{E_i} \in \mathcal{R}(E_i)$, $i = 1, 2$, alors $f \in \mathcal{R}(E)$ et*

$$\int_E f(\mathbf{x}) d\mathbf{x} = \int_{E_1} f(\mathbf{x}) d\mathbf{x} + \int_{E_2} f(\mathbf{x}) d\mathbf{x}. \quad (8.2)$$

Inversement, si E, E_1, E_2 sont mesurables (avec $E_1 \cap E_2$ négligeable) et $f \in \mathcal{R}(E)$, alors $f|_{E_i} \in \mathcal{R}(E_i)$, $i = 1, 2$ et on a encore (8.2).

Démonstration. Notons par $\mathbb{1}_{E_i} : E \rightarrow \mathbb{R}$ la fonction caractéristique de E_i pour $i \in \{1, 2\}$, et par $\mathbb{1}_{E_1 \cap E_2} : E \rightarrow \mathbb{R}$ la fonction caractéristique de $E_1 \cap E_2$. Supposons $f|_{E_i} \in \mathcal{R}(E_i)$, $i = 1, 2$. Alors les fonctions $f \cdot \mathbb{1}_{E_i}$ sont intégrables sur E et $\int_E f(\mathbf{x}) \cdot \mathbb{1}_{E_i}(\mathbf{x}) d\mathbf{x} = \int_{E_i} f(\mathbf{x}) d\mathbf{x}$. De plus $f|_{E_1 \cap E_2}$ est intégrable sur $E_1 \cap E_2$ et $\int_{E_1 \cap E_2} f|_{E_1 \cap E_2}(\mathbf{x}) d\mathbf{x} = 0$, car $E_1 \cap E_2$ est négligeable. D'où la fonction $f \cdot \mathbb{1}_{E_1 \cap E_2}$ est intégrable sur E et $\int_E f(\mathbf{x}) \cdot \mathbb{1}_{E_1 \cap E_2}(\mathbf{x}) d\mathbf{x} = 0$. Puisque $f = f \cdot \mathbb{1}_{E_1} + f \cdot \mathbb{1}_{E_2} - f \cdot \mathbb{1}_{E_1 \cap E_2}$ on conclut que $f \in \mathcal{R}(E)$ et vérifie (8.2).

Soit maintenant $f \in \mathcal{R}(E)$. Puisque les fonctions $\mathbb{1}_{E_i}$, $i = 1, 2$, sont aussi intégrables sur E (car E_1, E_2 sont mesurables), on a $f \cdot \mathbb{1}_{E_i} \in \mathcal{R}(E)$ (produit de fonctions intégrables), ce qui équivaut à $f|_{E_i} \in \mathcal{R}(E_i)$ et on a encore (8.2). \square

On remarque que dans la deuxième partie du théorème précédent, sans hypothèse de mesurabilité de E , E_1 et E_2 , le résultat n'est pas forcément vrai. Il suffit de prendre la fonction constante 1 sur $E = [0, 1]$ et prendre $E_1 = \mathbb{Q} \cap [0, 1]$ et $E_2 = [0, 1] \setminus \mathbb{Q}$ pour s'en convaincre.

8.7 Généralisation de la formule des intégrales itérées

On a déjà rencontré la formule des intégrales itérées pour le calcul de l'intégrale d'une fonction sur un pavé. On va maintenant la généraliser à des domaines de formes plus complexes.

Définition 8.34. On dit que $E \subset \mathbb{R}^{n+1}$ est un domaine simple par rapport à la variable y s'il existe un ensemble $K \in \mathcal{J}(\mathbb{R}^n)$ (compact et mesurable) et deux fonctions $g, h : K \rightarrow \mathbb{R}$ continues avec $g(\mathbf{x}) \leq h(\mathbf{x})$, $\forall \mathbf{x} \in K$, telles que E ait la forme

$$E = \{(\mathbf{x}, y) \in \mathbb{R}^{n+1}, \mathbf{x} \in K, g(\mathbf{x}) \leq y \leq h(\mathbf{x})\}.$$

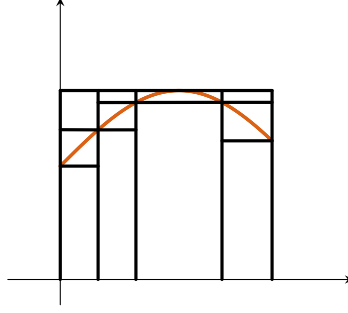
Théorème 8.35. Un domaine simple $E = \{(\mathbf{x}, y) \in \mathbb{R}^{n+1}, \mathbf{x} \in K, g(\mathbf{x}) \leq y \leq h(\mathbf{x})\}$ avec $K \in \mathcal{J}(\mathbb{R}^n)$, et $g, h : K \rightarrow \mathbb{R}$ continues telles que $g(\mathbf{x}) \leq h(\mathbf{x})$, $\forall \mathbf{x} \in K$, est mesurable au sens de Jordan et $\text{Vol}(E) = \int_K (h(\mathbf{x}) - g(\mathbf{x})) d\mathbf{x}$.

Remarque. Pour $j \in \{1, \dots, n\}$ fixé, il y a une version où les rôles des variables x_j et y sont échangés. Ainsi, par exemple, g et h délimitent x_j en fonction de $(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n, y)$.

Démonstration. On considère d'abord un domaine de la forme $E = \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} : \mathbf{x} \in K, 0 \leq y \leq h(\mathbf{x})\}$, avec $h(\mathbf{x}) \geq 0$, $\forall \mathbf{x} \in K$. Soit R un pavé non dégénéré tel que $K \subset R$, et $\tilde{R} = R \times [0, M]$ un pavé qui contient E , où $M = \max_{\mathbf{x} \in K} h(\mathbf{x}) + 1$.

Puisque h est continue et $K \in \mathcal{J}(\mathbb{R}^n)$, h est intégrable sur K au sens de Riemann. Soit $\epsilon \in]0, \text{Vol}(R)[$. Il existe une partition $\mathcal{P}_\epsilon = \{R_i\}_{i=1}^L$ de R telle que $\bar{S}(\tilde{h}, \mathcal{P}_\epsilon) - \underline{S}(\tilde{h}, \mathcal{P}_\epsilon) < \epsilon$, où on a noté \tilde{h} le prolongement par zéro de h en-dehors de K . Notons $m_i = \inf_{\mathbf{x} \in R_i} \tilde{h}(\mathbf{x})$ et $M_i = \sup_{\mathbf{x} \in R_i} \tilde{h}(\mathbf{x})$, $\forall i = 1, \dots, L$, et considérons la partition suivante de \tilde{R} (voir Figure ci-dessous)

$$\tilde{\mathcal{P}} = \bigcup_{i=1}^L \left\{ R_i \times [0, m_i], R_i \times \left[m_i, M_i + \frac{\epsilon}{\text{Vol}(R)} \right], \left[M_i + \frac{\epsilon}{\text{Vol}(R)}, M \right] \right\}.$$



On a

$$\begin{aligned}\underline{S}(\mathbb{1}_E, \tilde{\mathcal{P}}) &= \sum_{i: R_i \subset K} \text{Vol}(R_i \times [0, m_i]) = \sum_{i: R_i \subset K} m_i \text{Vol}(R_i) = \underline{S}(\tilde{h}, \mathcal{P}_\epsilon), \\ \overline{S}(\mathbb{1}_E, \tilde{\mathcal{P}}) &= \sum_{i: R_i \cap K \neq \emptyset} \text{Vol}(R_i \times [0, m_i]) + \sum_{i: R_i \cap K \neq \emptyset} \text{Vol}\left(R_i \times \left[m_i, M_i + \frac{\epsilon}{\text{Vol}(R)}\right]\right) \\ &= \sum_{i: R_i \cap K \neq \emptyset} \text{Vol}(R_i) \left(M_i + \frac{\epsilon}{\text{Vol}(R)}\right) \leq \sum_{i: R_i \cap K \neq \emptyset} \text{Vol}(R_i) M_i + \epsilon = \overline{S}(\tilde{h}, \mathcal{P}_\epsilon) + \epsilon,\end{aligned}$$

ce qui implique que $\overline{S}(\mathbb{1}_E, \tilde{\mathcal{P}}) - \underline{S}(\mathbb{1}_E, \tilde{\mathcal{P}}) < 2\epsilon$ et donc, puisque $\epsilon \in]0, \text{Vol}(R)[$ est arbitraire, $\mathbb{1}_E \in \mathcal{R}(\tilde{R})$. De plus, puisque $\epsilon > 0$ peut être pris aussi petit qu'on veut et $\underline{S}(\tilde{h}, \mathcal{P}_\epsilon) = \underline{S}(\mathbb{1}_E, \tilde{\mathcal{P}}) \leq \overline{S}(\mathbb{1}_E, \tilde{\mathcal{P}}) \leq \overline{S}(\tilde{h}, \mathcal{P}_\epsilon) + \epsilon$, on conclut que

$$\text{Vol}(E) = \int_{\tilde{R}} \mathbb{1}_E(\mathbf{x}, y) d\mathbf{x} dy = \int_R \tilde{h}(\mathbf{x}) d\mathbf{x} = \int_K h(\mathbf{x}) d\mathbf{x}.$$

Par les mêmes arguments, on vérifie facilement qu'un domaine de la forme $E = \{(\mathbf{x}, y) : \mathbf{x} \in K, c \leq y \leq h(\mathbf{x})\}$, avec $c \leq \min_{\mathbf{x} \in K} h(\mathbf{x})$, est aussi mesurable.

Considérons maintenant le cas générale : $E = \{(\mathbf{x}, y) : \mathbf{x} \in K, g(\mathbf{x}) \leq y \leq h(\mathbf{x})\}$. Soit $c = \min_{\mathbf{x} \in K} g(\mathbf{x})$ et

$$\begin{aligned}E_h &= \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} : \mathbf{x} \in K, c \leq y \leq h(\mathbf{x})\}, \\ E_g &= \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} : \mathbf{x} \in K, c \leq y \leq g(\mathbf{x})\}.\end{aligned}$$

E_h et E_g sont mesurables, par les arguments ci-dessus, et

$$\text{Vol}(E_h) = \int_K (h(\mathbf{x}) - c) d\mathbf{x} \quad \text{et} \quad \text{Vol}(E_g) = \int_K (g(\mathbf{x}) - c) d\mathbf{x}.$$

D'autre part $E_h = E_g \cup E$ et $E_g \cap E = \mathcal{G}(g)$ (le graphe de g). Or $\mathcal{G}(g) \subset \partial E_g$ et ∂E_g est négligeable car E_g est Jordan-mesurable. Donc $\mathcal{G}(g)$ est aussi négligeable. Ainsi $E = (E_h \setminus E_g) \cup \mathcal{G}(g)$ est Jordan-mesurable, $\text{Vol}(E_h) = \text{Vol}(E_g) + \text{Vol}(E)$ et

$$\text{Vol}(E) = \text{Vol}(E_h) - \text{Vol}(E_g) = \int_K (h(\mathbf{x}) - g(\mathbf{x})) d\mathbf{x}.$$

□

Exemple 8.36. Considérons le triangle $T = \{(x, y) \in \mathbb{R}^2, 0 \leq x \leq 1, 0 \leq y \leq 1 - x\}$. Si on note $K = [0, 1]$ et $h(x) = 1 - x$, on a donc $T = \{(x, y) \in \mathbb{R}^2, x \in K, 0 \leq y \leq h(x)\}$ et l'aire de T peut se calculer par la formule

$$\text{Vol}(T) = \int_K h(x) dx = \int_0^1 (1 - x) dx = \frac{1}{2}.$$

Exemple 8.37. Considérons la boule unitaire

$$E = \{(x, y) : x^2 + y^2 \leq 1\} = \{(x, y) : -1 \leq x \leq 1, -\sqrt{1 - x^2} \leq y \leq \sqrt{1 - x^2}\}.$$

Alors, l'aire de la boule peut être calculée par la formule

$$\text{Vol}(E) = \int_{-1}^1 2\sqrt{1 - x^2} dx = \int_{-\pi/2}^{\pi/2} 2 \cos^2 \theta d\theta = \int_{-\pi/2}^{\pi/2} (\cos 2\theta + 1) d\theta = \pi.$$

Théorème 8.38 (Intégrales itérées). Soit $E = \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} : \mathbf{x} \in K, g(\mathbf{x}) \leq y \leq h(\mathbf{x})\}$ un domaine simple, avec $K \in \mathcal{J}(\mathbb{R}^n)$, $g, h : K \rightarrow \mathbb{R}$ continues avec $g \leq h$ sur K , et $f : E \rightarrow \mathbb{R}$ une fonction continue. Alors f est intégrable sur E et

$$\int_E f(\mathbf{x}, y) d\mathbf{x} dy = \int_K \left(\int_{g(\mathbf{x})}^{h(\mathbf{x})} f(\mathbf{x}, y) dy \right) d\mathbf{x}.$$

Remarque. Pour $j \in \{1, \dots, n\}$ fixé, il y a une version où les rôles des variables x_j et y sont échangés.

Démonstration. Soient R un pavé contenant K , $m \leq \min_K g$, $M \geq \max_K h$ et $\tilde{R} = R \times [m, M] \supset E$. Par le théorème 8.35, E est mesurable et par le théorème 8.30 $f \in \mathcal{R}(E)$, c'est-à-dire \tilde{f} est intégrable sur \tilde{R} , où $\tilde{f} : \tilde{R} \rightarrow \mathbb{R}$ est le prolongement de f par la valeur 0. De plus, $\forall \mathbf{x} \in R$, la fonction (de y) $\tilde{f}(\mathbf{x}, \cdot)$ est intégrable sur $[m, M]$, car elle est continue par morceaux. Donc $\tilde{F}(\mathbf{x}) = \int_m^M \tilde{f}(\mathbf{x}, y) dy$ existe pour tout $\mathbf{x} \in R$ et, par le théorème de Fubini 8.14, $\tilde{F} \in \mathcal{R}(R)$ et $\int_{\tilde{R}} \tilde{f}(\mathbf{x}, y) d\mathbf{x} dy = \int_R \tilde{F}(\mathbf{x}) d\mathbf{x}$. Comme $\tilde{F} : R \rightarrow \mathbb{R}$ est le prolongement par la valeur 0 de la fonction $K \ni \mathbf{x} \mapsto F(\mathbf{x}) = \int_{g(\mathbf{x})}^{h(\mathbf{x})} f(\mathbf{x}, y) dy$, cette dernière fonction est intégrable sur K et

$$\begin{aligned} \int_E f(\mathbf{x}, y) d\mathbf{x} dy &= \int_{\tilde{R}} \tilde{f}(\mathbf{x}, y) d\mathbf{x} dy \\ &= \int_R \tilde{F}(\mathbf{x}) d\mathbf{x} = \int_K F(\mathbf{x}) d\mathbf{x} = \int_K \left(\int_{g(\mathbf{x})}^{h(\mathbf{x})} f(\mathbf{x}, y) dy \right) d\mathbf{x}. \end{aligned}$$

□

Exercice 8.39. Calculer $\int_T x dx dy dz$ où $T = \{(x, y, z) : 0 \leq x \leq 1, 0 \leq y \leq 1 - x, 0 \leq z \leq 1 - x - y\}$ est un simplexe de \mathbb{R}^3 .

8.8 Changement de variables dans les intégrales de Riemann

Soit $F \subset \mathbb{R}^n$ borné et mesurable, et $f : F \rightarrow \mathbb{R}$ une fonction bornée et intégrable sur F . Notons $I = \int_F f(\mathbf{x})d\mathbf{x}$. On introduit maintenant un changement de variables

$$\mathbf{x} = \boldsymbol{\psi}(\mathbf{u}), \quad \boldsymbol{\psi} : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^n,$$

tel que $F = \boldsymbol{\psi}(E)$. On se pose les questions suivantes : i) l'ensemble E est-il mesurable ? ii) La fonction $\tilde{f}(\mathbf{u}) = f(\boldsymbol{\psi}(\mathbf{u})) : E \rightarrow \mathbb{R}$ est-elle intégrable sur E ? iii) Comment l'intégrale I se transforme suite au changement de variables ?

Rappelons d'abord le cas $n = 1$. Soit $F = [\alpha, \beta] \subset \mathbb{R}$ un intervalle, $f : F \rightarrow \mathbb{R}$ une fonction continue et $\psi : E = [a, b] \rightarrow \mathbb{R}$ une fonction de classe C^1 (jusqu'au bord) telle que $\psi(a) = \alpha$, $\psi(b) = \beta$ et $\psi(E) = F$ (avec $a < b$ et $\alpha < \beta$). On ne suppose pas que ψ soit injective. Alors la fonction $u \mapsto f(\psi(u))\psi'(u)$, $u \in E$, est aussi continue et

$$\int_{\alpha}^{\beta} f(x)dx = \int_a^b f(\psi(u))\psi'(u)du.$$

Ce résultat se montre à partir du théorème fondamental du calcul intégral. Soit $G(t) = \int_{\alpha}^{\psi(t)} f(x)dx$, alors $G'(t) = f(\psi(t))\psi'(t)$ et

$$\int_{\alpha=\psi(a)}^{\beta=\psi(b)} f(x)dx = G(b) - G(a) = \int_a^b f(\psi(u))\psi'(u)du.$$

On remarque, en particulier, que pour obtenir ce résultat, on n'a pas demandé que la fonction ψ soit une bijection entre E et F . En effet, le résultat reste vrai même si la fonction ψ n'est pas une bijection.

En dimension $n > 1$, on n'a plus un théorème fondamental du calcul intégral, donc on cherche un résultat qui soit valable sous des conditions un peu plus fortes et qui soit généralisable au cas $n > 1$. Pour $n = 1$, on va demander que la transformation $\psi : E \rightarrow F$ soit un *difféomorphisme* (donc une bijection de classe C^1 avec application inverse de classe C^1), où $E = [a, b]$ avec $a < b$ dans \mathbb{R} . Alors, nécessairement, $\psi' \neq 0$ sur E , $F = [\psi(a), \psi(b)]$ si ψ est strictement croissante, $F = [\psi(b), \psi(a)]$ si ψ est strictement décroissante, et on a :

1. Si $\psi'(u) > 0$, $\forall u \in [a, b]$, alors ψ est strictement croissante et $\psi(a) < \psi(u) < \psi(b)$, $\forall u \in]a, b[$. Donc

$$\int_F f(x)dx = \int_{\psi(a)}^{\psi(b)} f(x)dx = \int_a^b f(\psi(u))\psi'(u)du = \int_a^b f(\psi(u))|\psi'(u)|du.$$

2. Si $\psi'(u) < 0$, $\forall u \in [a, b]$, alors ψ est strictement décroissante et $\psi(b) < \psi(u) < \psi(a)$, $\forall u \in]a, b[$. Donc

$$\begin{aligned} \int_F f(x)dx &= \int_{\psi(b)}^{\psi(a)} f(x)dx = - \int_{\psi(a)}^{\psi(b)} f(x)dx \\ &= - \int_a^b f(\psi(u))\psi'(u)du = \int_a^b f(\psi(u))|\psi'(u)|du. \end{aligned}$$

Donc dans les deux cas on a

$$\int_F f(x)dx = \int_E f(\psi(u))|\psi'(u)|du.$$

On note que $|\psi'(u)| = \left| \frac{dx}{du} \right|$ représente le changement de l'élément infinitésimal de longueur. C'est bien cette formule qui se généralise en dimension $n > 1$. On a l'important résultat suivant qu'on ne va pas démontrer dans ce cours.

Théorème 8.40 (Changement des variables d'intégration). *Soient $U, V \subset \mathbb{R}^n$ ouverts tels que, pour tout $r > 0$, $U \cap B(\mathbf{0}, r)$ et $V \cap B(\mathbf{0}, r)$ sont Jordan-mesurables. Soit $\psi : U \rightarrow V$ un difféomorphisme de classe C^1 tel que toutes les composantes de ψ et de $D\psi \in \mathbb{R}^{n \times n}$ sont bornés sur tout sous-ensemble borné de U . Soit encore un borné non-vide $E \subset U$ et $F = \psi(E) \subset V$, qui est aussi borné non-vide. Alors*

1. *E est Jordan-mesurable si et seulement si F l'est ;*
2. *si E est Jordan-mesurable et $f : F = \psi(E) \rightarrow \mathbb{R}$ est continue et bornée, alors $f \in \mathcal{R}(F)$ et*

$$\int_F f(\mathbf{x})d\mathbf{x} = \int_E f(\psi(\mathbf{u})) |\det D\psi(\mathbf{u})| d\mathbf{u}, \quad (8.3)$$

où apparaît la valeur absolue d'un déterminant dans le membre de droite.

Notons que $\tilde{f} = f \circ \psi : E \rightarrow \mathbb{R}$ est continue (composition de fonctions continues) et bornée, donc $\tilde{f} \in \mathcal{R}(E)$ si E est mesurable. De plus on a fait l'hypothèse que toutes les composantes de $D\psi$ sont bornées sur E . Puisque elles sont aussi continues (car $\psi \in C^1(U)$ et $E \subset U$), donc intégrables sur E si celui ci est mesurable, on a que $J_\psi = \det D\psi \in \mathcal{R}(E)$ (produit de fonctions intégrables) et l'intégrale à droite de (8.3) existe.

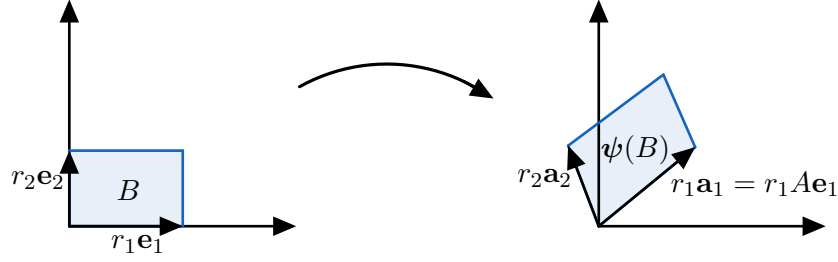
Le terme $J_\psi(u) = \det D\psi(\mathbf{u})$ dans (8.3), souvent appelé le *jacobien* de ψ , représente le *changement infinitésimal de volume* par le difféomorphisme ψ . En effet, soit $\mathbf{u}_0 \in U$ et $r > 0$ suffisamment petit tel que le pavé $\bar{B}_r = \{\mathbf{u} \in \mathbb{R}^n : \|\mathbf{u} - \mathbf{u}_0\|_\infty \leq r\}$ soit contenu dans U . Grâce au théorème 8.40 on a $\psi(B_r)$ est mesurable puisque B_r l'est et

$$\min_{\mathbf{u} \in \bar{B}_r} |J_\psi(\mathbf{u})| \text{Vol}(B_r) \leq \text{Vol}(\psi(B_r)) = \int_{B_r} |J_\psi(\mathbf{u})| d\mathbf{u} \leq \max_{\mathbf{u} \in \bar{B}_r} |J_\psi(\mathbf{u})| \text{Vol}(B_r).$$

En prenant la limite pour $r \rightarrow 0$ et grâce à la continuité de $|J_\psi|$ on voit que

$$\lim_{r \rightarrow 0} \frac{\text{Vol}(\psi(B_r))}{\text{Vol}(B_r)} = |J_\psi(\mathbf{u}_0)|.$$

Que le changement infinitésimal de volume en \mathbf{u}_0 soit donné par $|\det D\psi(\mathbf{u}_0)|$ ne devrait pas surprendre. Considérons par exemple une application affine en dimension $n = 2$: $\mathbf{x} = \psi(\mathbf{u}) = A\mathbf{u} + \mathbf{b}$, avec $A \in \mathbb{R}^{2 \times 2}$, $\mathbf{b} \in \mathbb{R}^2$ et notons $(\mathbf{a}_1, \mathbf{a}_2)$ les deux colonnes de A . Alors, un rectangle B_{r_1, r_2} de sommets $\{\mathbf{u}_0, \mathbf{u}_0 + r_1 \mathbf{e}_1, \mathbf{u}_0 + r_2 \mathbf{e}_2, \mathbf{u}_0 + r_1 \mathbf{e}_1 + r_2 \mathbf{e}_2\}$ est transformé en un parallélogramme $\psi(B_{r_1, r_2})$ de sommets $\{\mathbf{x}_0, \mathbf{x}_0 + r_1 \mathbf{a}_1, \mathbf{x}_0 + r_2 \mathbf{a}_2, \mathbf{x}_0 + r_1 \mathbf{a}_1 + r_2 \mathbf{a}_2\}$ (voir figure ci-dessous),



où $\mathbf{x}_0 = \psi(\mathbf{u}_0) = A\mathbf{u}_0 + \mathbf{b}$ et

$$\text{Vol}(\psi(B_{r_1, r_2})) = |r_1 \mathbf{a}_1 \times r_2 \mathbf{a}_2| = r_1 r_2 |\det A| = |\det A| \text{Vol}(B_{r_1, r_2}) = |\det D\psi(\mathbf{u}_0)| \text{Vol}(B_{r_1, r_2}).$$

On peut montrer que le même résultat est vrai en dimension n quelconque pour une application affine.

Considérons maintenant une transformation non linéaire $\mathbf{x} = \psi(\mathbf{u})$. Autour d'un point $\mathbf{x}_0 = \psi(\mathbf{u}_0)$ on peut écrire un développement limité à l'ordre 1

$$\begin{aligned} \mathbf{x} &= \psi(\mathbf{u}_0) + D\psi(\mathbf{u}_0) \cdot (\mathbf{u} - \mathbf{u}_0) + \mathbf{R}(\mathbf{u}) \\ &= D\psi(\mathbf{u}_0)\mathbf{u} + (\mathbf{x}_0 - D\psi(\mathbf{u}_0)\mathbf{u}_0) + \mathbf{R}(\mathbf{u}). \end{aligned}$$

Négligeant le reste $\mathbf{R}(\mathbf{u})$ on s'attend donc à ce que le facteur du changement infinitésimal de volume soit encore donné par $|\det D\psi(\mathbf{u}_0)|$.

8.9 Quelques changements de variables usuels

8.9.1 Changement en coordonnées polaires

Considérons le changement de variables

$$\begin{pmatrix} x \\ y \end{pmatrix} = \psi(\rho, \theta) = \begin{pmatrix} \rho \cos \theta \\ \rho \sin \theta \end{pmatrix},$$

qui est un difféomorphisme entre les ouverts $U =]0, +\infty[\times]-\pi, \pi[$ et $V = \mathbb{R}^2 \setminus \{(x, y) : y = 0, x \leq 0\}$. En particulier,

$$D\psi(\rho, \theta) = \begin{pmatrix} \cos \theta & -\rho \sin \theta \\ \sin \theta & \rho \cos \theta \end{pmatrix}, \quad J_\psi(\rho, \theta) = \det D\psi(\rho, \theta) = \rho > 0 \text{ sur } U.$$

Soit maintenant $F \subset V$ borné et mesurable, et $f : F \rightarrow \mathbb{R}$ continue et bornée, donc intégrable. Notons $\tilde{f}(\rho, \theta) = f \circ \psi(\rho, \theta) = f(\rho \cos \theta, \rho \sin \theta)$. Puisque ψ^{-1} est bornée sur tout ensemble borné, on a que $E = \psi^{-1}(F)$ est borné, ψ et $D\psi$ sont bornées sur E et, grâce au théorème 8.40, E est mesurable et

$$\int_F f(x, y) dx dy = \int_E \tilde{f}(\rho, \theta) \rho d\rho d\theta.$$

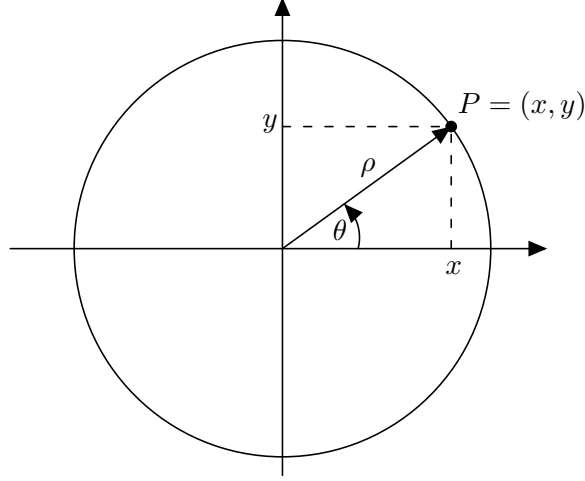


FIGURE 8.3 – Changement en coordonnées polaires

Exemple 8.41. On souhaite calculer $\int_F f(x, y) dx dy$ où $F = \{(x, y) : 1 \leq x^2 + y^2 \leq 4, x \geq 0, y \geq 0\} \subset V$ et $f(x, y) = \frac{1}{1+x^2+y^2}$, en utilisant le changement en coordonnées polaires.

On a $E = \psi^{-1}(F) = \{(\rho, \theta) : 1 \leq \rho \leq 2, 0 \leq \theta \leq \frac{\pi}{2}\} \subset U$. De plus, toutes les hypothèses du théorème 8.40 sont vérifiées et on peut calculer l'intégrale par la formule (8.3)

$$\begin{aligned} \int_F f(x, y) dx dy &= \int_E \tilde{f}(\rho, \theta) \rho d\rho d\theta = \int_1^2 \left(\int_0^{\pi/2} \frac{1}{1+\rho^2} d\theta \right) \rho d\rho \\ &= \int_1^2 \frac{\pi}{2} \frac{\rho}{1+\rho^2} d\rho = \frac{\pi}{4} \log(1+\rho^2) \Big|_1^2 = \frac{\pi}{4} \log \frac{5}{2}. \end{aligned}$$

Exemple 8.42. On souhaite calculer $\int_F f(x, y) dx dy$ où $F = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ et $f(x, y) = \frac{1}{1+x^2+y^2}$. L'intégrale existe car F est compact et mesurable, et f est continue.

Considérons de nouveau le changement en coordonnées polaires. Dans ce cas, $F \not\subset V$ ce qui pose un problème. Toutefois, l'ensemble $G = \{(x, y) : y = 0, -1 \leq x \leq 0\}$ est négligeable, donc $\tilde{F} = F \setminus G \subset V$ est mesurable et

$$\int_F f(x, y) dx dy = \int_{\tilde{F}} f(x, y) dx dy.$$

On peut alors appliquer la formule (8.3) à cette dernière intégrale. On a $\tilde{E} = \psi^{-1}(\tilde{F}) = \{(\rho, \theta) : 0 < \rho \leq 1, -\pi < \theta < \pi\} \subset U$ et

$$\begin{aligned} \int_F f(x, y) dx dy &= \int_{\tilde{F}} f(x, y) dx dy = \int_{\tilde{E}} \tilde{f}(\rho, \theta) \rho d\rho d\theta \\ &= \int_0^1 \left(\int_{-\pi}^{\pi} \frac{1}{1+\rho^2} d\theta \right) \rho d\rho = \pi \log 2. \end{aligned}$$

Observez que $([0, 1] \times [-\pi, \pi]) \setminus \tilde{E}$ est négligeable et donc les intégrales sur $[0, 1] \times [-\pi, \pi]$ et sur \tilde{E} sont égales.

8.9.2 Changement en coordonnées cylindriques

Considérons le changement de variables

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \psi(\rho, \theta, \zeta) = \begin{pmatrix} \rho \cos \theta \\ \rho \sin \theta \\ \zeta \end{pmatrix},$$

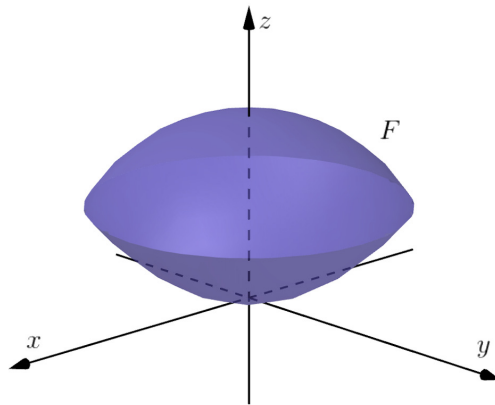
qui est un difféomorphisme entre les ouverts $U =]0, +\infty[\times]-\pi, \pi[\times \mathbb{R}$ et $V = \mathbb{R}^3 \setminus \{(x, y, z) : x \leq 0, y = 0, z \in \mathbb{R}\}$ avec

$$\det D\psi(\rho, \theta, \zeta) = \det \begin{pmatrix} \cos \theta & -\rho \sin \theta & 0 \\ \sin \theta & \rho \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \rho > 0 \text{ sur } U.$$

Comme pour le changement en coordonnées polaires, si $F \subset V$ est borné et mesurable, alors $E = \psi^{-1}(F)$ est aussi borné et $\psi, D\psi$ sont bornés sur E . Donc, si $f : F \rightarrow \mathbb{R}$ est continue et bornée et on note $\tilde{f} = f \circ \psi$, on peut appliquer le théorème 8.40 : E est mesurable et

$$\int_F f(x, y, z) dx dy dz = \int_E \tilde{f}(\rho, \theta, \zeta) \rho d\rho d\theta d\zeta.$$

Exercice 8.43. Calculer $\text{Vol}(F)$ où $F = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 4, z \geq \frac{1}{3}(x^2 + y^2)\}$ (voir figure ci-dessous).



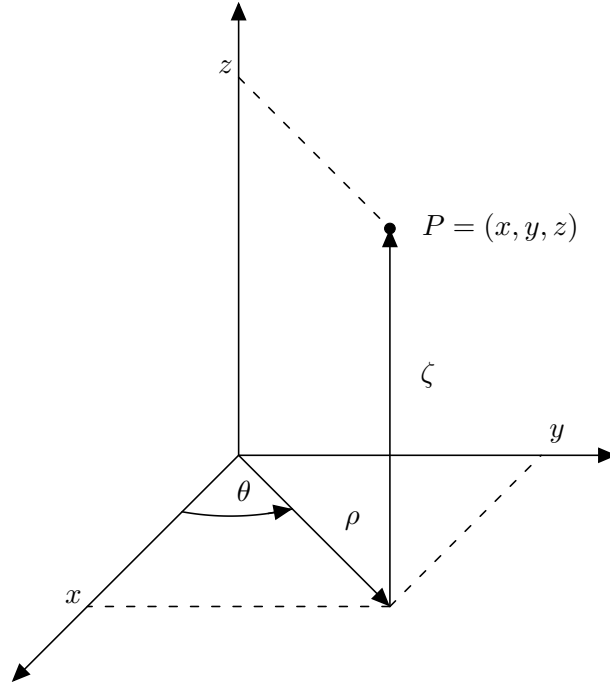


FIGURE 8.4 – Changement en coordonnées cylindriques

8.9.3 Changement en coordonnées sphériques

Considérons le changement de variables

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \psi(\rho, \theta, \varphi) = \begin{pmatrix} \rho \cos \theta \sin \varphi \\ \rho \sin \theta \sin \varphi \\ \rho \cos \varphi \end{pmatrix},$$

qui est un difféomorphisme entre les ouverts $U =]0, +\infty[\times]-\pi, \pi[\times]0, \pi[$ et $V = \mathbb{R}^3 \setminus \{(x, y, z) : x \leq 0, y = 0, z \in \mathbb{R}\}$, avec

$$\det D\psi(\rho, \theta, \varphi) = \det \begin{pmatrix} \cos \theta \sin \varphi & -\rho \sin \theta \sin \varphi & \rho \cos \theta \cos \varphi \\ \sin \theta \sin \varphi & \rho \cos \theta \sin \varphi & \rho \sin \theta \cos \varphi \\ \cos \varphi & 0 & -\rho \sin \varphi \end{pmatrix} = -\rho^2 \sin \varphi < 0 \text{ sur } U.$$

Comme pour le changement en coordonnées cylindriques ou polaires, si $F \subset V$ est borné et mesurable, et $f : F \rightarrow \mathbb{R}$ est continue et bornée, notant $\tilde{f} = f \circ \psi$ et $E = \psi^{-1}(F)$, on peut appliquer le théorème 8.40 : E est bornée et mesurable, et

$$\int_F f(x, y, z) dx dy dz = \int_E \tilde{f}(\rho, \theta, \varphi) \rho^2 \sin \phi d\rho d\theta d\varphi.$$

Exercice 8.44. Calculer le volume de la sphère de rayon $R > 0$: $F = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq R^2\}$.

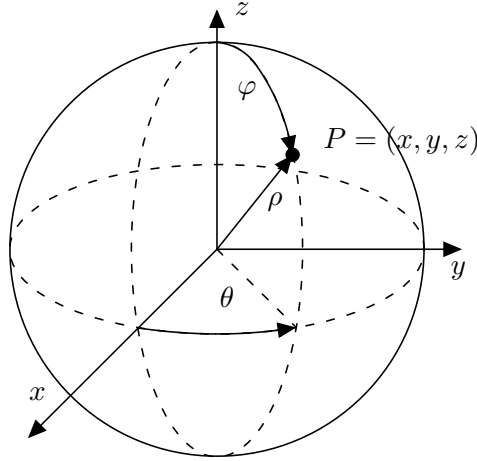


FIGURE 8.5 – Changement en coordonnées sphériques

8.10 Intégrale de Riemann généralisée

Jusqu'à présent on a défini l'intégrale d'une fonction $f : E \rightarrow \mathbb{R}$ bornée sur un ensemble borné $E \subset \mathbb{R}^n$ par un prolongement $\tilde{f} : R \rightarrow \mathbb{R}$ de f par zéro sur un pavé $R \supset E$:

$$\int_E f(\mathbf{x}) d\mathbf{x} = \int_R \tilde{f}(\mathbf{x}) d\mathbf{x}.$$

On souhaite ici généraliser la définition de l'intégrale de Riemann aux cas où f ou E ne sont pas bornés. Rappelons que, pour $E \subset \mathbb{R}^n$, on note $\mathcal{J}(E)$ la collection des sous-ensembles de E compacts et mesurables au sens de Jordan, et commençons par la définition suivante :

Définition 8.45 (Fonction absolument intégrable). *Soit $E \subset \mathbb{R}^n$ ouvert non-vide et $f : E \rightarrow \mathbb{R}$ (les deux pas forcément bornés). Soit $\{K_j, j \in \mathbb{N}\}$ une suite de sous-ensembles non-vides tels que :*

- $K_j \in \mathcal{J}(E), \forall j \in \mathbb{N}$;
- $K_j \subset \overset{\circ}{K}_{j+1}$;
- $\bigcup_{j \in \mathbb{N}} K_j = E$.

Soit f bornée et intégrable au sens de Riemann sur chaque K_j ($\int_{K_j} f(\mathbf{x}) d\mathbf{x}$ existe ainsi que $\int_{K_j} |f(\mathbf{x})| d\mathbf{x}$). On dit que f est absolument intégrable sur E si $\lim_{j \rightarrow \infty} \int_{K_j} |f(\mathbf{x})| d\mathbf{x}$ existe et est finie. Dans ce cas on pose

$$\int_E f(\mathbf{x}) d\mathbf{x} = \lim_{j \rightarrow \infty} \int_{K_j} f(\mathbf{x}) d\mathbf{x}.$$

On remarque que pour tout ensemble $E \subset \mathbb{R}^n$ on peut toujours trouver une suite de sous-ensembles $\{K_j\}_{j \in \mathbb{N}}$ qui satisfait les propriétés de la définition ci-dessus. Il suffit en effet de considérer, par exemple, la suite $\{K_j, j \geq j_0\}$ d'ensembles compacts

$$K_j = \{\mathbf{x} \in E : \|\mathbf{x}\| \leq j, \|\mathbf{x} - \mathbf{y}\| \geq \frac{1}{j}, \forall \mathbf{y} \notin E\},$$

où $j_0 \geq 1$ est tel que $K_{j_0} \neq \emptyset$.

On remarque aussi que si $\lim_{j \rightarrow \infty} \int_{K_j} |f(\mathbf{x})| d\mathbf{x}$ existe (sous-entendu : et est finie), alors $\lim_{j \rightarrow \infty} \int_{K_j} f(\mathbf{x}) d\mathbf{x}$ existe aussi. En effet, les suites $\int_{K_j} f_+(\mathbf{x}) d\mathbf{x}$ et $\int_{K_j} f_-(\mathbf{x}) d\mathbf{x}$ sont croissantes et bornées par $\lim_{j \rightarrow \infty} \int_{K_j} |f(\mathbf{x})| d\mathbf{x}$. Donc $\lim_{j \rightarrow \infty} \int_{K_j} f_+(\mathbf{x}) d\mathbf{x}$ et $\lim_{j \rightarrow \infty} \int_{K_j} f_-(\mathbf{x}) d\mathbf{x}$ existent et aussi $\lim_{j \rightarrow \infty} \int_{K_j} f(\mathbf{x}) d\mathbf{x} = \lim_{j \rightarrow \infty} \int_{K_j} f_+(\mathbf{x}) d\mathbf{x} - \lim_{j \rightarrow \infty} \int_{K_j} f_-(\mathbf{x}) d\mathbf{x}$.

Le théorème suivant assure que la valeur de l'intégrale $\int_E f(\mathbf{x}) d\mathbf{x}$ dans la définition précédente ne dépend pas du choix de la suite $\{K_j\}_{j \in \mathbb{N}}$.

Théorème 8.46. *Si $f : E \rightarrow \mathbb{R}$ est absolument intégrable selon la définition (8.45), alors pour toute autre suite $\{K'_j\}_{j \in \mathbb{N}}$ telle que $K'_j \in \mathcal{J}(E)$, $K'_j \subset \mathring{K}'_{j+1}$, $\forall j \in \mathbb{N}$, $\bigcup_{j \in \mathbb{N}} K'_j = E$, on a que f est bornée et intégrable au sens de Riemann sur chaque K'_j et*

$$\int_E f(\mathbf{x}) d\mathbf{x} = \lim_{j \rightarrow \infty} \int_{K'_j} f(\mathbf{x}) d\mathbf{x}.$$

Démonstration. Fixons $j \in \mathbb{N}$ et $K'_j \in \mathcal{J}(E)$. Clairement $K'_j \subset E = \bigcup_{m \in \mathbb{N}} K_m \subset \bigcup_{m \in \mathbb{N}} \mathring{K}_{m+1}$ donc $\{\mathring{K}_m\}_{m \in \mathbb{N}}$ est un recouvrement ouvert de K'_j . Puisque K'_j est compact, on peut extraire un recouvrement fini $K'_j \subset \mathring{K}_{i_1} \cup \dots \cup \mathring{K}_{i_\ell}$, et puisque les \mathring{K}_m sont emboîtés, $K'_j \subset \mathring{K}_{N_j}$ avec $N_j = \max\{i_1, \dots, i_\ell\}$. Puisque f est intégrable sur K_{N_j} et K'_j est mesurable, on a que f est intégrable sur K'_j , $\forall j \in \mathbb{N}$. De plus,

$$\begin{aligned} \int_{K'_j} f_+(\mathbf{x}) d\mathbf{x} &\leq \int_{K_{N_j}} f_+(\mathbf{x}) d\mathbf{x} \leq \lim_{m \rightarrow \infty} \int_{K_m} f_+(\mathbf{x}) d\mathbf{x} < +\infty, \\ \int_{K'_j} f_-(\mathbf{x}) d\mathbf{x} &\leq \int_{K_{N_j}} f_-(\mathbf{x}) d\mathbf{x} \leq \lim_{m \rightarrow \infty} \int_{K_m} f_-(\mathbf{x}) d\mathbf{x} < +\infty. \end{aligned}$$

Donc $\lim_{j \rightarrow \infty} \int_{K'_j} f_\pm(\mathbf{x}) d\mathbf{x}$ existent et

$$\lim_{j \rightarrow \infty} \int_{K'_j} f_+(\mathbf{x}) d\mathbf{x} \leq \lim_{m \rightarrow \infty} \int_{K_m} f_+(\mathbf{x}) d\mathbf{x}, \quad \lim_{j \rightarrow \infty} \int_{K'_j} f_-(\mathbf{x}) d\mathbf{x} \leq \lim_{m \rightarrow \infty} \int_{K_m} f_-(\mathbf{x}) d\mathbf{x}.$$

De la même façon on peut montrer que $\forall j \in \mathbb{N}$, $\exists N'_j : K_j \subset \mathring{K}'_{N'_j}$ et

$$\begin{aligned} \lim_{j \rightarrow \infty} \int_{K_j} f_+(\mathbf{x}) d\mathbf{x} &\leq \lim_{j \rightarrow \infty} \int_{K'_j} f_+(\mathbf{x}) d\mathbf{x}, \\ \lim_{j \rightarrow \infty} \int_{K_j} f_-(\mathbf{x}) d\mathbf{x} &\leq \lim_{j \rightarrow \infty} \int_{K'_j} f_-(\mathbf{x}) d\mathbf{x}, \end{aligned}$$

ce qui implique

$$\lim_{j \rightarrow \infty} \int_{K'_j} f_+(\mathbf{x}) d\mathbf{x} = \lim_{j \rightarrow \infty} \int_{K_j} f_+(\mathbf{x}) d\mathbf{x}, \quad \lim_{j \rightarrow \infty} \int_{K'_j} f_-(\mathbf{x}) d\mathbf{x} = \lim_{j \rightarrow \infty} \int_{K_j} f_-(\mathbf{x}) d\mathbf{x}$$

et

$$\begin{aligned} \lim_{j \rightarrow \infty} \int_{K'_j} f(\mathbf{x}) d\mathbf{x} &= \lim_{j \rightarrow \infty} \int_{K'_j} f_+(\mathbf{x}) d\mathbf{x} - \lim_{j \rightarrow \infty} \int_{K'_j} f_-(\mathbf{x}) d\mathbf{x} \\ &= \lim_{j \rightarrow \infty} \int_{K_j} f_+(\mathbf{x}) d\mathbf{x} - \lim_{j \rightarrow \infty} \int_{K_j} f_-(\mathbf{x}) d\mathbf{x} = \int_E f(\mathbf{x}) d\mathbf{x}. \end{aligned}$$

□

Exemple 8.47. On veut vérifier si on peut définir $\int_E f(x, y) dx dy$ où $E = B((0, 0), 1) \setminus \{(0, 0)\}$ et

$$f(x, y) = \frac{1}{\sqrt{x^2 + y^2}} : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}.$$

Considérons la suite de sous-ensembles $K_j = \{(x, y) \in \mathbb{R}^2 : \frac{1}{j+2} \leq x^2 + y^2 \leq 1 - \frac{1}{j+2}\}$, $j \in \mathbb{N}$, qui satisfait les bonnes propriétés : $K_j \in \mathcal{J}(E)$, $K_j \subset K_{j+1}$, $\bigcup_{j \in \mathbb{N}} K_j = E$. On a

$$\begin{aligned} \int_{K_j} |f(x, y)| dx dy &= \int_{K_j} f(x, y) dx dy = \int_{\frac{1}{\sqrt{j+2}}}^{\sqrt{1-\frac{1}{j+2}}} \left(\int_{-\pi}^{\pi} \frac{1}{\rho} \cdot \rho d\theta \right) d\rho \\ &= 2\pi \left(\sqrt{1 - \frac{1}{j+2}} - \frac{1}{\sqrt{j+2}} \right) \leq 2\pi, \quad \forall j \geq 0. \end{aligned}$$

Donc f est absolument intégrable sur E et

$$\int_E f(x, y) dx dy = \lim_{j \rightarrow \infty} \int_{K_j} f(x, y) dx dy = 2\pi.$$

L'exemple précédent montre que la fonction $\frac{1}{r}$, avec $r = \|\mathbf{x}\|$, est (absolument) intégrable pour $\mathbf{x} \in \mathbb{R}^n$, $n = 2$, dans un voisinage de zéro. Par un calcul similaire on trouve qu'elle est absolument intégrable pour tout $n \geq 2$. Toutefois, elle n'est pas intégrable pour $n = 1$. Plus généralement, on a que la fonction $f(\mathbf{x}) = \|\mathbf{x}\|^\alpha$ est absolument intégrable sur $E = B(\mathbf{0}, 1) \setminus \{\mathbf{0}\}$ pour tout $\alpha > -n$.

Exemple 8.48. On veut vérifier si on peut définir $\int_E f(x, y) dx dy$ où $E = B((0, 0), 1) \setminus \{(0, 0)\}$ et

$$f(x, y) = \frac{x}{(x^2 + y^2)^2} : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}.$$

De nouveau, on prend $K_j = \{(x, y) : \frac{1}{j} \leq x^2 + y^2 \leq 1 - \frac{1}{j}\}$, $j \geq 2$. On a

$$\begin{aligned} \int_{K_j} |f(x, y)| dx dy &= \int_{\frac{1}{\sqrt{j}}}^{\sqrt{1-\frac{1}{j}}} \int_{-\pi}^{\pi} \frac{\rho |\cos \theta|}{\rho^4} \rho d\rho d\theta \\ &= 4 \left(-\frac{1}{\rho} \right) \Big|_{\frac{1}{\sqrt{j}}}^{\sqrt{1-\frac{1}{j}}} = 4\sqrt{j} - \frac{4}{\sqrt{1-\frac{1}{j}}} \xrightarrow{j \rightarrow \infty} +\infty. \end{aligned}$$

Donc la fonction f n'est pas absolument intégrable.

Attention : Si on n'avait pas mis la valeur absolue, alors $\int_{K_j} f(x, y) dx dy = 0$, $\forall j \geq 2$ et donc $\lim_{j \rightarrow \infty} \int_{K_j} f(x, y) dx dy = 0$. Le problème est que cette limite dépend du choix de la suite $\{K_j\}$ car la fonction f n'est pas absolument intégrable !

Chapitre 9

Équations différentielles ordinaires

Soient I un intervalle ouvert, E un ouvert inclus dans \mathbb{R} et $f : I \times E \rightarrow \mathbb{R}$ une fonction continue. Une fonction $u : I \rightarrow E$ différentiable, qui satisfait l'équation

$$u'(t) = f(t, u(t)), \quad \forall t \in I, \quad (9.1)$$

est appelée une *intégrale* ou une *solution globale* de l'équation (9.1) qui, à son tour, est appelée *Équation Différentielle Ordinaire (EDO) scalaire du premier ordre* car elle introduit une relation entre la valeur de la fonction (scalaire) inconnue $u(t)$ et la valeur de sa dérivée première en tout $t \in I$.

L'adjectif « ordinaire » fait référence au fait que l'inconnue est une fonction d'une seule variable t , donc sa dérivée est une dérivée ordinaire. Ceci est pour distinguer du cas où l'inconnue est une fonction de plusieurs variables (t, x_1, \dots, x_m) et l'équation fait intervenir les dérivées partielles de u . Dans ce dernier cas, on parle d'une *Équation aux Dérivées Partielles* (EDP), qu'on ne va pas traiter dans ce cours.

Exemple 9.1. Soit $I, E = \mathbb{R}$ et $f(t, x) = t^2 + x$, ce qui va définir l'équation différentielle ordinaire du premier ordre

$$u'(t) = f(t, u(t)) = t^2 + u(t), \quad \forall t \in \mathbb{R}.$$

On vérifie facilement que $u(t) = Ce^t - (t^2 + 2t + 2)$, avec $C \in \mathbb{R}$ arbitraire, est une intégrale de l'équation.

L'exemple précédent montre que la solution d'une EDO scalaire du premier ordre dépend généralement d'un paramètre arbitraire, mais les exceptions sont courantes. On a donc ici une famille infinie de solutions $u(\cdot, C) : I \rightarrow \mathbb{R}$ paramétrées par un paramètre réel.

Définition 9.2. On appelle intégrale générale de l'équation différentielle (9.1) l'ensemble de toutes les solutions globales de (9.1).

La notion d'équation différentielle se généralise facilement à une dimension n quelconque. Dans ce cas, on parle d'une EDO vectorielle du premier ordre ou bien d'un *système d'EDO du premier ordre*. Soit $I \subset \mathbb{R}$ un intervalle ouvert, $E \subset \mathbb{R}^n$ un ouvert et $\mathbf{f} : I \times E \rightarrow \mathbb{R}^n$

une fonction à valeurs dans \mathbb{R}^n continue. On considère l'EDO vectorielle

$$\mathbf{u}'(t) = \mathbf{f}(t, \mathbf{u}(t)) \iff \begin{cases} u_1'(t) = f_1(t, u_1(t), \dots, u_n(t)) \\ \vdots \\ u_n'(t) = f_n(t, u_1(t), \dots, u_n(t)) \end{cases} \quad \forall t \in I, \quad (9.2)$$

dont la solution sera une fonction vectorielle $\mathbf{u} : I \rightarrow E \subset \mathbb{R}^n$ différentiable, $t \mapsto \mathbf{u}(t) = (u_1(t), \dots, u_n(t))$. On verra par la suite que, généralement, l'intégrale générale de (9.2) est une famille de fonctions $\mathbf{u}(\cdot, C_1, \dots, C_n) : I \rightarrow E$ paramétrée par n paramètres réels (mais de nouveau les exceptions sont courantes).

On peut aussi introduire des équations différentielles d'ordres supérieurs à un. Par exemple, soit $I \subset \mathbb{R}$ un intervalle ouvert, $E \subset \mathbb{R}^n$ un ouvert et $f : I \times E \rightarrow \mathbb{R}$ une fonction scalaire continue. Considérons l'équation différentielle d'ordre n

$$u^{(n)}(t) = f(t, u(t), u'(t), \dots, u^{(n-1)}(t)), \quad \forall t \in I, \quad (9.3)$$

dont la solution est une fonction $u : I \rightarrow \mathbb{R}$ de classe C^n telle que $(u(t), u'(t), \dots, u^{(n-1)}(t)) \in E$, $\forall t \in I$. Une telle équation peut toujours être écrite sous forme d'une EDO vectorielle d'ordre 1 en introduisant les variables

$$u_1(t) = u(t), \quad u_2(t) = u'(t), \quad \dots, \quad u_n(t) = u^{(n-1)}(t),$$

ainsi que la fonction vectorielle

$$\mathbf{f} : I \times E \rightarrow \mathbb{R}^n, \quad \mathbf{f}(t, u_1, \dots, u_n) = \begin{pmatrix} u_2 \\ \vdots \\ u_n \\ f(t, u_1, \dots, u_n) \end{pmatrix},$$

ce qui permet d'écrire (9.3) comme $\mathbf{u}'(t) = \mathbf{f}(t, \mathbf{u}(t))$, $\forall t \in I$. Pour cette raison, une étude complète des systèmes d'EDO du premier ordre est suffisante pour pouvoir traiter aussi des équations d'ordres supérieurs.

9.1 Problème de Cauchy

Comme l'exemple 9.1 le montre, l'intégrale générale d'une EDO scalaire du premier ordre dépend en générale d'un paramètre arbitraire (n paramètres pour une EDO vectorielle en dimension n). Il est souvent pratique d'imposer des conditions supplémentaires pour obtenir une solution unique. Par exemple on peut demander que la solution passe par un point $(t_0, \mathbf{u}_0) \in I \times E$. Ceci porte au **Problème de Cauchy** :

Problème 9.3 (de Cauchy). *Étant donné $\mathbf{f} : I \times E \rightarrow \mathbb{R}^n$ continue, avec $I \subset \mathbb{R}$ un intervalle ouvert, $E \subset \mathbb{R}^n$ un ouvert et $(t_0, \mathbf{u}_0) \in I \times E$, trouver $\mathbf{u} : I \rightarrow E$ différentiable sur I t.q.*

$$\begin{cases} \mathbf{u}'(t) = \mathbf{f}(t, \mathbf{u}(t)), & t \in I, \\ \mathbf{u}(t_0) = \mathbf{u}_0. \end{cases} \quad (9.4)$$

Les équations différentielles ordinaires apparaissent souvent dans les applications pour décrire l'évolution d'un système en fonction du temps. Dans ce cas, la condition $\mathbf{u}(t_0) = \mathbf{u}_0$ peut être interprétée comme une *valeur initiale*, c.-à-d. elle décrit l'état du système au temps initial t_0 , et on cherche à prédire l'état $\mathbf{u}(t)$ du système pour des temps futures $t > t_0$. On parle alors d'un **Problème à valeur initiale** :

Problème 9.4 (à valeur initiale). *Étant donné $\mathbf{f} : I_+ \times E \rightarrow \mathbb{R}^n$ continue, avec $I_+ = [t_0, T[$ et $-\infty < t_0 < T \leq +\infty$, $E \subset \mathbb{R}^n$ un ouvert et $\mathbf{u}_0 \in E$, trouver $\mathbf{u} : I_+ \rightarrow E$ continue sur I_+ et différentiable sur $\overset{\circ}{I}_+$ t.q.*

$$\begin{cases} \mathbf{u}'(t) = \mathbf{f}(t, \mathbf{u}(t)), & t_0 < t < T, \\ \mathbf{u}(t_0) = \mathbf{u}_0. \end{cases} \quad (9.5)$$

Il arrive parfois qu'on souhaite modéliser l'évolution passée qui a porté à l'état actuel $\mathbf{u}(t_0) = \mathbf{u}_0$ du système. On parle alors d'un **Problème à valeur finale** :

Problème 9.5 (à valeur finale). *Étant donné $\mathbf{f} : I_- \times E \rightarrow \mathbb{R}^n$ continue, avec $I_- =]\tilde{T}, t_0]$ et $-\infty \leq \tilde{T} < t_0 < +\infty$, $E \subset \mathbb{R}^n$ un ouvert et $\mathbf{u}_0 \in E$, trouver $\mathbf{u} : I_- \rightarrow E$ continue sur I_- et différentiable sur $\overset{\circ}{I}_-$ t.q.*

$$\begin{cases} \mathbf{u}'(t) = \mathbf{f}(t, \mathbf{u}(t)), & \tilde{T} < t < t_0, \\ \mathbf{u}(t_0) = \mathbf{u}_0. \end{cases} \quad (9.6)$$

Soit $I =]\tilde{T}, T[= I_- \cup I_+$. Il est clair que si $\mathbf{u} : I \rightarrow E$ est une solution du problème de Cauchy (9.4), alors les restrictions $\mathbf{u}|_{I_+}$ et $\mathbf{u}|_{I_-}$ de \mathbf{u} à I_+ et I_- sont solutions des problèmes à valeur initiale (9.5) et finale (9.6), respectivement. Il est facile de montrer (exercice) l'implication inverse :

Lemme 9.6. *Si $\mathbf{u}_+ : I_+ \rightarrow E$ et $\mathbf{u}_- : I_- \rightarrow E$ sont solutions de (9.5), (9.6) et \mathbf{f} est continue sur $I \times E$ alors la fonction $\mathbf{u} : I \rightarrow E$ définie par $\mathbf{u}(t) = \mathbf{u}_-(t)$ si $t \in I_-$ et $\mathbf{u}(t) = \mathbf{u}_+(t)$ si $t \in I_+$ est de classe C^1 et est solution de (9.4).*

Bien que le problème de Cauchy 9.3 soit formulé sur l'intervalle ouvert I , sa solution (si elle existe) peut ne pas exister pour tout $t \in I$, comme l'exemple suivant le montre.

Exemple 9.7. *Soit $I, E = \mathbb{R}$, $f(t, u) = u^2$ et considérons le problème de Cauchy*

$$\begin{cases} u'(t) = u^2(t), & t \in \mathbb{R}, \\ u(0) = 1. \end{cases}$$

On vérifie facilement que $u(t) = \frac{1}{1-t}$ est une solution $\forall t \neq 1$. Toutefois, cette solution n'est pas définie pour $t = 1$. En particulier, elle présente une « explosion en temps fini » lorsque $t \rightarrow 1^-$. Du point de vue physique, prolonger cette solution pour $t > 1$ n'a pas vraiment de sens. On dit alors que la solution existe seulement pour $t < 1$ et elle n'est pas une solution globale du problème de Cauchy.

Les considérations précédentes nous portent à donner la définition suivante de solution locale et solution maximale du problème de Cauchy (une terminologie analogue s'applique aux problèmes à valeur initiale et à valeur finale).

Définition 9.8. On appelle solution locale du problème de Cauchy 9.3 un couple (J, \mathbf{u}) où $J \subset I$ est un intervalle ouvert contenant t_0 et où $\mathbf{u} \in C^1(J)$ satisfait (9.4) sur J .

- On dit qu’une solution locale (K, \mathbf{w}) du problème 9.3 prolonge strictement (J, \mathbf{u}) si $J \subset K$, $J \neq K$, et $\mathbf{u}(t) = \mathbf{w}(t)$, $\forall t \in J$.
- On dit qu’une solution locale (J, \mathbf{u}) est maximale s’il n’existe pas de solution locale qui la prolonge strictement.
- On dit qu’une solution maximale (J, \mathbf{u}) est une solution globale si $J = I$.
- Une solution maximale (J, \mathbf{u}) est dite unique si toute solution locale (K, \mathbf{w}) est telle que $K \subset J$ et $\mathbf{w}(t) = \mathbf{u}(t)$, $\forall t \in K$.

Si on s’intéresse à l’EDO $\mathbf{u}'(t) = \mathbf{f}(t, \mathbf{u}(t))$ pour t dans l’intervalle ouvert I , sans spécifier de condition initiale du type $\mathbf{u}(t_0) = \mathbf{u}_0$, on peut aussi introduire une terminologie analogue. Par exemple une solution locale est un couple (J, \mathbf{u}) où $J \subset I$ est un intervalle ouvert non-vide et où $\mathbf{u} \in C^1(J)$ est solution. Néanmoins, sans condition initiale, le concept d’unicité d’une solution maximale n’est plus pertinent et, à la place, on introduit la notion de *solution générale* : c’est l’ensemble de toutes les solutions maximales.

Remarque 9.9. Sans spécification de l’intervalle de définition d’une solution, il est souvent sous-entendu par “solution” le concept de solution maximale.

Il est important de savoir sous quelles conditions sur \mathbf{f} le problème de Cauchy admet des solutions locales ou globales et si elles sont uniques. On investiguera cette question dans la section 9.4.

9.2 Quelques méthodes de résolution d’EDO scalaires

On considère dans cette section le cas d’une EDO scalaire et le problème de Cauchy correspondant

$$u'(t) = f(t, u(t)), \quad t \in I, \quad u(t_0) = u_0,$$

avec $I \subset \mathbb{R}$ un intervalle ouvert contenant t_0 , $E \subset \mathbb{R}$ un ouvert contenant u_0 et $f : I \times E \rightarrow \mathbb{R}$ continue.

Avant d’illustrer quelques méthodes de résolution, on présente d’abord une interprétation géométrique de la solution du problème de Cauchy. Soit $u : J \rightarrow E$ une solution maximale du problème de Cauchy (J étant un intervalle ouvert contenant t_0 et inclus dans I), dont le graphe $\mathcal{G}(u) = \{(t, y) \in J \times E : y = u(t)\}$, appelé aussi une *courbe intégrale*, est contenu dans $I \times E$. Un vecteur tangent à $\mathcal{G}(u)$ en $(t, u(t))$ est donné par $\mathbf{v} = (1, u'(t)) = (1, f(t, u(t)))$. Définissons le champ vectoriel $\mathbf{v} : I \times E \rightarrow \mathbb{R}^2$, $(t, y) \mapsto \mathbf{v}(t, y) = (1, f(t, y))$. Alors, les courbes intégrales sont en tout point tangentes au champ vectoriel \mathbf{v} .

La figure 9.1 montre le champ vectoriel associé à la fonction $f(t, u) = u(1 - u)$, ainsi que les trois solutions des problèmes de Cauchy associés aux conditions initiales $u(0) = 2$, $u(0) = 0.25$, $u(0) = -0.25$. On voit bien que les graphes des solutions (courbes intégrales) sont en tout point tangentes au champ vectoriel $\mathbf{v}(t, y) = (1, y(1 - y))$.

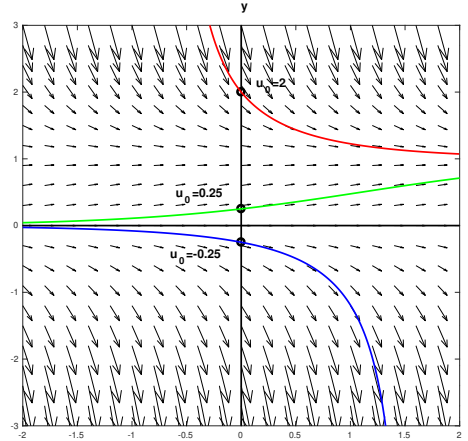


FIGURE 9.1 – Courbes intégrales associées aux conditions $u(0) = 2$, $u(0) = 0.25$, $u(0) = -0.25$ pour l'EDO $u' = u(1 - u)$.

9.2.1 Équations différentielles à variables séparées

Soit $f : I \times E \rightarrow \mathbb{R}$ de la forme $f(t, u) = g(t)k(u)$ avec $g : I \rightarrow \mathbb{R}$ et $k : E \rightarrow \mathbb{R}$ continues sur les intervalles ouverts I et E , et considérons le problème de Cauchy :

$$\begin{cases} u'(t) = g(t)k(u(t)), & t \in I, \\ u(t_0) = u_0, \end{cases} \quad (9.7)$$

avec $(t_0, u_0) \in I \times E$. Si $k(u_0) = 0$, le problème de Cauchy admet la solution globale $u(t) = u_0$, $\forall t \in I$. On verra dans la section 9.4 des conditions suffisantes sur k pour que cette solution soit unique. Si par contre, $k(u_0) \neq 0$, grâce à la continuité de k , il existe un intervalle ouvert $\tilde{E} \subset E$ contenant u_0 où la fonction k ne s'annule pas. Dans ce cas, le résultat suivant donne une procédure explicite pour calculer une solution locale du problème de Cauchy.

Théorème 9.10. Soient $I, \tilde{E} \subset \mathbb{R}$ deux intervalles ouverts, $g : I \rightarrow \mathbb{R}$ continue et $k : \tilde{E} \rightarrow \mathbb{R}$ continue telle que $k(u) \neq 0$, $\forall u \in \tilde{E}$. Pour tout $(t_0, u_0) \in I \times \tilde{E}$, notons $G(t) = \int_{t_0}^t g(s)ds : I \rightarrow \mathbb{R}$ et $F(u) = \int_{u_0}^u \frac{1}{k(v)}dv : \tilde{E} \rightarrow \mathbb{R}$. Alors il existe un intervalle ouvert $J \subset I$ contenant t_0 avec $G(J) \subset \text{Im}(F)$ et une fonction $u : J \rightarrow \mathbb{R}$ définie par

$$u(t) = F^{-1}(G(t)) \in \tilde{E}, \quad t \in J, \quad (9.8)$$

tels que (J, u) est une solution locale du problème de Cauchy (9.7). De plus, une telle solution locale (J, u) est unique au sens que toute autre solution locale (K, w) à valeurs dans E satisfait $w(t) = u(t) \in \tilde{E}$, $\forall t \in K \cap J$.

Démonstration. Notons tout d'abord que la fonction k , étant une fonction continue, ne change pas de signe sur \tilde{E} , ce qui implique que $F : \tilde{E} \rightarrow \mathbb{R}$ est continue strictement

monotone sur \tilde{E} , $\text{Im}(F)$ est ouvert et l'application inverse définie sur $\text{Im}(F)$ est continûment différentiable (F est un difféomorphisme entre \tilde{E} et $\text{Im}(F)$). Puisque $F(u_0) = 0$, il existe $\eta > 0$ tel que $]-\eta, \eta[\subset F(\tilde{E})$ et par la continuité de G en t_0 et le fait que $G(t_0) = 0$, il existe $\delta_\eta > 0$ tel que $J :=]t_0 - \delta_\eta, t_0 + \delta_\eta[\subset I$ et $G(t) \in]-\eta, \eta[, \forall t \in J$. On a donc montré l'existence d'un intervalle ouvert $J \subset I$ contenant t_0 et tel que $G(J) \subset \text{Im}(F)$.

Grâce au fait que F^{-1} et G sont de classe C^1 , on a que $u = F^{-1} \circ G \in C^1(J)$ et

$$u'(t) = \frac{G'(t)}{F'(F^{-1}(G(t)))} = g(t)k(u(t)), \quad \forall t \in J.$$

De plus, $u(t_0) = F^{-1}(0) = u_0$, donc (J, u) est une solution locale du problème de Cauchy.

Montrons l'unicité de cette solution. En effet, soit (K, w) une autre solution locale à valeurs dans E . Puisque w est continue et \tilde{E} est ouvert, $w^{-1}(\tilde{E})$ est un ouvert dans K contenant t_0 . Soit \tilde{K} le plus grand intervalle ouvert inclus dans $w^{-1}(\tilde{E})$ et contenant t_0 . Alors, pour tout $t \in \tilde{K}$ on a $k(w(t)) \neq 0$ (car $w(t) \in \tilde{E}$) et $\int_{t_0}^t \frac{w'(s)}{k(w(s))} ds = \int_{t_0}^t g(s) ds = G(t)$. Mais, en faisant le changement de variable $v = w(t)$, la même intégrale devient $\int_{t_0}^t \frac{w'(s)}{k(w(s))} ds = \int_{u_0}^{w(t)} \frac{1}{k(v)} dv = F(w(t))$, ce qui implique, grâce à l'inversibilité de $F : \tilde{E} \rightarrow \text{Im}(F)$, que $w(t) = F^{-1}(G(t))$, $\forall t \in \tilde{K}$. Donc $w = u$ sur l'intervalle ouvert $J \cap \tilde{K} \subset J \cap K$. Par l'absurde, supposons qu'il existe $\tilde{t} \in \partial(J \cap \tilde{K}) \cap (J \cap K)$. Comme $u(\tilde{t}) \in \tilde{E}$, on obtiendrait par continuité que $w(\tilde{t}) = u(\tilde{t}) \in \tilde{E}$, ce qui conduirait à la contradiction $\tilde{t} \in J \cap \tilde{K}$. D'où $J \cap \tilde{K} = J \cap K$. \square

Le théorème précédent montre l'existence d'un intervalle ouvert $J \subset I$ contenant t_0 et tel que $G(J) \subset \text{Im}(F)$. Si maintenant on prend le plus grand intervalle $\tilde{J} \subset I$ contenant t_0 et tel que $G(\tilde{J}) \subset \text{Im}(F)$, qui sera non-vidé grâce au résultat du théorème précédent, on pourra calculer l'unique solution locale sur \tilde{J} par la formule

$$u(t) = F^{-1}(G(t)), \quad \forall t \in \tilde{J}.$$

Il est possible qu'il existe une solution locale sur un intervalle ouvert plus grand $\hat{J} \supset \tilde{J}$. Toutefois, le théorème 9.10 ne garantit plus l'unicité de la solution sur $\hat{J} \setminus \tilde{J}$. En effet, l'exemple 9.12 ci-dessous montre que dès que $k(u(t)) = 0$ pour quelque $t \in \hat{J}$, la solution peut ne plus être unique (observez que k n'est pas de classe C^1 dans cet exemple).

De façon informelle, on peut construire la solution d'un problème de Cauchy par séparation de variables en utilisant le procédé suivant :

$$\begin{aligned} \frac{du}{dt} = g(t)k(u) &\Rightarrow \frac{du}{k(u)} = g(t)dt \Rightarrow \int_{u_0}^u \frac{dv}{k(v)} = \int_{t_0}^t g(s)ds \\ \Rightarrow F(u) - F(u_0) = G(t) - G(t_0) &\Rightarrow u = F^{-1}(G(t) - G(t_0) + F(u_0)) \end{aligned}$$

où F est une primitive quelconque de $\frac{1}{k}$ et G une primitive quelconque de g .

Exemple 9.11. Soit $I, E = \mathbb{R}$, et $f(t, u) = u^2$, $(t, u) \in I \times E$. Considérons le problème de Cauchy

$$\begin{cases} u'(t) = u^2(t), & t \in I, \\ u(t_0) = u_0 > 0. \end{cases}$$

On note $g(t) = 1$, $t \in I$, et $k(u) = u^2$, $u \in E$. Pour pouvoir appliquer la méthode de séparation des variables, on se restreint à l'ensemble $\tilde{E} = \mathbb{R}_+^* \subset E$, puisque $u_0 > 0$. On a

$$F(u) = \int_{u_0}^u \frac{1}{y^2} dy = \frac{1}{u_0} - \frac{1}{u} : \tilde{E} \rightarrow \mathbb{R}$$

qui est bien inversible sur \tilde{E} , avec application inverse $F^{-1}(v) = \frac{u_0}{1-u_0v}$ continûment différentiable sur $] -\infty, \frac{1}{u_0}[$. De plus, $G(t) = \int_{t_0}^t 1 dt = t - t_0$ et donc

$$u(t) = \frac{u_0}{1 - (t - t_0)u_0}, \quad \forall t \in] -\infty, t_0 + \frac{1}{u_0}[$$

est l'unique solution du problème de Cauchy sur l'intervalle $] -\infty, t_0 + \frac{1}{u_0}[$. Cette solution est aussi la solution maximale car $\lim_{t \rightarrow (t_0 + \frac{1}{u_0})^-} = +\infty$ et elle ne peut pas être prolongée de façon continue sur un intervalle plus grand.

Exemple 9.12. Soit $I, E = \mathbb{R}$, et $f(t, u) = -\sqrt{|u|}$, $(t, u) \in I \times E$. Considérons le problème de Cauchy

$$\begin{cases} u'(t) = -\sqrt{|u(t)|}, & t \in I, \\ u(t_0) = u_0 > 0. \end{cases}$$

Observons déjà que toute solution locale u satisfait $u' \leq 0$. On note $g(t) = -1$, $t \in I$, et $k(u) = \sqrt{|u|}$, $u \in E$. De nouveau, puisque $u_0 > 0$, on se restreint à $\tilde{E} = \mathbb{R}_+^* \subset E$ et on cherche une primitive de $\frac{1}{\sqrt{|y|}}$ pour $y \in \tilde{E}$:

$$F(u) = \int_{u_0}^u \frac{1}{\sqrt{y}} dy = 2\sqrt{u} - 2\sqrt{u_0} : \tilde{E} \rightarrow \mathbb{R}$$

qui est bien inversible sur \tilde{E} avec application inverse différentiable $F^{-1}(v) = (\sqrt{u_0} + \frac{v}{2})^2$, $v \in] -2\sqrt{u_0}, +\infty[$. D'autre part, $G(t) = \int_{t_0}^t (-1) ds = t_0 - t$ et $G(t) \in] -2\sqrt{u_0}, +\infty[$ pour tout $t \in] -\infty, 2\sqrt{u_0} + t_0[$, et on a la solution locale

$$u(t) = \left(\sqrt{u_0} - \frac{t - t_0}{2} \right)^2, \quad t \in] -\infty, 2\sqrt{u_0} + t_0[.$$

La fonction $u(t) = (\sqrt{u_0} - \frac{t-t_0}{2})^2$, $t \in \mathbb{R}$, est bien définie pour $t \geq 2\sqrt{u_0} + t_0$, mais n'est pas solution car $u'(t) > 0$ pour tout $t > 2\sqrt{u_0} + t_0$. Toutefois il existe des solutions globales, mais pas uniques. En effet, les fonctions

$$\tilde{u}(t) = \begin{cases} (\sqrt{u_0} - \frac{t-t_0}{2})^2, & t < 2\sqrt{u_0} + t_0, \\ -(\sqrt{u_0} - \frac{t-t_0}{2})^2, & t \geq 2\sqrt{u_0} + t_0, \end{cases} \quad \hat{u}(t) = \begin{cases} (\sqrt{u_0} - \frac{t-t_0}{2})^2, & t < 2\sqrt{u_0} + t_0, \\ 0, & t \geq 2\sqrt{u_0} + t_0, \end{cases}$$

sont des solutions globales du même problème de Cauchy. On remarque que les deux fonctions coïncident pour $t < 2\sqrt{u_0} + t_0$, ce qui est consistant avec le résultat du théorème 9.10, et que la solution n'est plus unique pour $t > 2\sqrt{u_0} + t_0$, c.-à-d. une fois que la solution a touché la valeur critique $u_{cr} = 0$ pour laquelle $k(u_{cr}) = \sqrt{|u_{cr}|} = 0$. Ces deux solutions

globales peuvent être vues comme les membres extrêmes de la famille suivante de solutions globales, la famille étant paramétrée par $s \in]2\sqrt{u_0} + t_0, +\infty[$:

$$u_s(t) = \begin{cases} (\sqrt{u_0} - \frac{t-t_0}{2})^2, & t < 2\sqrt{u_0} + t_0, \\ 0, & 2\sqrt{u_0} + t_0 \leq t < s, \\ -(\frac{t-s}{2})^2, & t \geq s. \end{cases}$$

Exercice 9.13. Trouver, par la méthode de séparation des variables, la solution du problème de Cauchy

$$\begin{cases} u'(t) = \frac{3}{2}u(t)^{1/3}, & t \in \mathbb{R}, \\ u(t_0) = u_0 > 0. \end{cases}$$

Commenter sur l'unicité de la solution. Ici $v \mapsto v^{1/3}$ est la fonction impaire sur \mathbb{R} réciproque de la fonction $v \mapsto v^3$.

9.2.2 Équations avec fonction $f(t, u)$ homogène de degré zéro

Soit $f : \mathbb{R}^* \times \mathbb{R} \rightarrow \mathbb{R}$ continue telle que $f(\alpha t, \alpha y) = f(t, y)$, $\forall \alpha \in \mathbb{R}^*$. On appelle une telle fonction *homogène de degré zéro*. Si on prend, en particulier, $\alpha = \frac{1}{t}$, $t \neq 0$, on voit que f dépend uniquement du rapport y/t car $f(t, y) = f(1, \frac{y}{t})$. Introduisant la fonction $\phi(\frac{y}{t}) := f(1, \frac{y}{t})$, le problème de Cauchy pour $t_0 > 0$ prend la forme suivante :

$$\begin{cases} u'(t) = f(t, u(t)) = \phi\left(\frac{u(t)}{t}\right), & t \in]0, +\infty[, \\ u(t_0) = u_0. \end{cases}$$

On introduit, maintenant, le changement de variable $v(t) = \frac{u(t)}{t}$. Alors

$$\begin{aligned} v'(t) &= \frac{u'(t)}{t} - \frac{u(t)}{t^2} = \frac{1}{t} \phi\left(\frac{u(t)}{t}\right) - \frac{1}{t} \frac{u(t)}{t} \\ &= \frac{1}{t} [\phi(v(t)) - v(t)], \end{aligned}$$

ce qui nous donne une équation à variables séparées en (t, v) .

Le même raisonnement s'applique aussi si la fonction f est définie par exemple seulement sur $\mathbb{R}^* \times \mathbb{R}^*$ ou $\mathbb{R}_+^* \times \mathbb{R}_+^*$ et $u_0 \neq 0$.

Exemple 9.14. Considérons la fonction $f : \mathbb{R}_+^* \times \mathbb{R}_+^* \rightarrow \mathbb{R}$ donnée par $f(t, y) = \frac{t^3 + y^3}{ty^2}$ et le problème de Cauchy

$$u'(t) = \frac{t^3 + u^3(t)}{tu^2(t)}, \quad u(t) > 0, \quad t > 0, \quad u(1) = u_0 > 0.$$

On pose $v(t) = \frac{u(t)}{t}$ et on cherche la solution du problème transformé

$$v'(t) = \frac{1}{t} \left[\frac{1 + v^3(t)}{v^2(t)} - v(t) \right] = \frac{1}{t} \left[\frac{1}{v^2(t)} \right], \quad v(t) > 0, \quad t > 0, \quad v(1) = u_0.$$

En intégrant on trouve $F(v) = \int_{u_0}^v w^2 dw = \frac{1}{3}(v^3 - u_0^3)$, inversible entre \mathbb{R}_+^* et $]-\frac{1}{3}u_0^3, +\infty[$ (d'inverse C^1), et $G(t) = \int_1^t \frac{1}{s} ds = \ln t$, $t \in]e^{-u_0^3/3}, +\infty[$, et donc

$$v(t) = F^{-1}(G(t)) = (3 \ln t + u_0^3)^{1/3}, \quad t > e^{-u_0^3/3}.$$

Finalement on obtient la solution locale

$$u(t) = v(t)t = (3 \ln t + u_0^3)^{1/3}t, \quad t > e^{-u_0^3/3}.$$

9.3 EDO scalaires linéaires du premier ordre

On parle d'équation différentielle *linéaire* du premier ordre lorsque $f(t, y)$ est une fonction *affine* de la variable y . Soit $I \subset \mathbb{R}$ un intervalle ouvert, $g, p : I \rightarrow \mathbb{R}$ deux fonctions continues et $f(t, y) = g(t) - p(t)y$. Pour $t_0 \in I$, on pose le problème de Cauchy suivant :

$$\begin{cases} u'(t) + p(t)u(t) = g(t), & t \in I, \\ u(t_0) = u_0. \end{cases} \quad (9.9)$$

L'équation est dite *homogène* ou *sans second membre* si $g(t) = 0$, $\forall t \in I$, et *non-homogène*, *inhomogène* ou *avec second membre* autrement. (Faites attention que la terminologie "homogène" a déjà été utilisée dans la section 9.2.2, mais dans un sens différent !)

La méthode du facteur intégrant. Voici une première méthode courante de résolution de problème de Cauchy (9.9). Notons par $P : I \rightarrow \mathbb{R}$ n'importe quelle primitive fixée de p (la constante d'intégration est fixée librement) : $P(t) = \int^t p(s) ds$. On appelle la fonction $e^P : I \rightarrow \mathbb{R}$ un *facteur intégrant*. Cette terminologie provient de l'équivalence suivante :

$$(9.9) \Leftrightarrow \begin{cases} \left(e^{P(t)} u(t) \right)' = e^{P(t)} g(t), & t \in I, \\ u(t_0) = u_0. \end{cases}$$

Il est facile maintenant d'obtenir une formule pour la solution :

$$\begin{aligned} (9.9) &\Leftrightarrow e^{P(t)} u(t) - e^{P(t_0)} u_0 = \int_{t_0}^t e^{P(s)} g(s) ds, \quad t \in I, \\ &\Leftrightarrow u(t) = e^{-(P(t)-P(t_0))} u_0 + e^{-P(t)} \int_{t_0}^t e^{P(s)} g(s) ds, \quad t \in I. \end{aligned}$$

Avec le choix de la primitive $P(t) = \int_{t_0}^t p(s) ds$, ceci donne plus simplement

$$u(t) = e^{-P(t)} u_0 + e^{-P(t)} \int_{t_0}^t e^{P(s)} g(s) ds, \quad t \in I. \quad (9.10)$$

Ainsi (9.9) admet une unique solution *globale*.

Passons à une méthode standard, qui consiste à traiter d'abord le problème homogène et ensuite le problème non-homogène. Comme nous le verrons, cette seconde méthode s'appliquera aussi aux EDO du deuxième ordre linéaire.

Équation homogène. Il s'agit de trouver la solution générale du problème homogène associé (dans lequel on remplace g par 0) :

$$u'(t) + p(t)u(t) = 0, \quad t \in I,$$

sans condition initiale. La méthode du facteur intégrant s'applique et donne la solution générale

$$u(t) = Ce^{-P(t)}, \quad t \in I,$$

où $C \in \mathbb{R}$ est une constante arbitraire et $P : I \rightarrow \mathbb{R}$ est une primitive fixée de $p : I \rightarrow \mathbb{R}$. En fait additionner une constante à la primitive P est équivalent à multiplier C par une constante positive. Le choix $C = 0$ donne la solution "triviale" $u = 0$ sur I .

Une autre manière de résoudre le problème homogène associé est de l'écrire sous forme d'une EDO à variables séparées :

$$u'(t) = -p(t)u(t), \quad t \in I,$$

se souvenir qu'une primitive de $1/v$ est donnée par $\ln(|v|)$ et ne pas oublier la solution triviale $u = 0$ sur tout I .

Équation non-homogène. Considérons maintenant le cas non-homogène, pour lequel on a le résultat suivant dont la démonstration est immédiate :

Proposition 9.15 (Principe de superposition de solutions). *Soient $g_1, g_2 : I \rightarrow \mathbb{R}$ continues et deux constantes $\alpha_1, \alpha_2 \in \mathbb{R}$. Si $u_1 : I \rightarrow \mathbb{R}$ est une intégrale de l'équation $u_1'(t) + p(t)u_1(t) = g_1(t)$ et $u_2 : I \rightarrow \mathbb{R}$ une intégrale de l'équation $u_2'(t) + p(t)u_2(t) = g_2(t)$, alors $v = \alpha_1 u_1 + \alpha_2 u_2$ est une intégrale de l'équation $v'(t) + p(t)v(t) = \alpha_1 g_1(t) + \alpha_2 g_2(t)$.*

En particulier, en prenant $g_2 = 0$, $g_1 = g$ et $\alpha_1 = \alpha_2 = 1$, on obtient

Proposition 9.16. *Toute solution de l'équation différentielle*

$$u'(t) + p(t)u(t) = g(t), \quad t \in I \tag{9.11}$$

est de la forme

$$u(t) = w(t) + Ce^{-P(t)} \tag{9.12}$$

où $w(\cdot)$ est une solution particulière de l'équation non-homogène (i.e. une solution fixée du problème non-homogène qui ne satisfait pas forcément la condition $w(t_0) = u_0$) et $Ce^{-P(\cdot)}$ est la solution générale de l'équation homogène, avec $P(\cdot)$ une primitive quelconque de p et $C \in \mathbb{R}$ une constante arbitraire.

Démonstration. Il est clair que (9.12) est solution de (9.11) pour tout $C \in \mathbb{R}$. Montrons que (9.12) est bien l'intégrale générale. Par l'absurde, supposons qu'il existe une intégrale \tilde{u} de (9.11) qui n'est pas de la forme (9.12). Alors $\tilde{u} - w$ est solution de l'équation homogène et donc il existe $\tilde{C} \in \mathbb{R}$ telle que $\tilde{u}(t) - w(t) = \tilde{C}e^{-P(t)}$ ce qui contredit l'hypothèse que \tilde{u} n'est pas de la forme (9.12). \square

Si, après, on s'intéresse au problème de Cauchy, on peut toujours trouver sa solution (globale) en calculant la constante C dans (9.12) de telle sorte à avoir $u(t_0) = u_0$, ce qui donne $C = e^{P(t_0)}(u_0 - w(t_0))$.

Exemple 9.17. *Considérons l'équation différentielle linéaire $u'(t) = u(t) + 1$, ainsi que le problème de Cauchy avec condition initiale $u(t_0) = u_0$. On a*

- *solution générale de l'équation homogène : $z(t) = Ce^t$;*
- *solution particulière de l'équation non-homogène : $w(t) = -1$;*
- *solution générale : $u(t) = w(t) + z(t) = -1 + Ce^t$;*
- *solution globale (unique) du problème de Cauchy : $u(t) = -1 + (1 + u_0)e^{(t-t_0)}$.*

Expliquons une méthode importante pour trouver une solution particulière de l'équation non-homogène : la *méthode de variation des constantes*. Elle consiste à chercher une solution particulière de l'équation non-homogène sous la forme

$$w(t) = C(t)e^{-P(t)}, \quad \text{avec } C : I \rightarrow \mathbb{R} \text{ différentiable,}$$

c.-à-d. qu'on prend l'expression de l'intégrale générale de l'équation homogène $z(t) = Ce^{-P(t)}$ et on transforme la constante $C \in \mathbb{R}$ en une fonction $C : I \rightarrow \mathbb{R}$ différentiable. En remplaçant dans l'équation on a

$$w'(t) = -p(t)w(t) + g(t) = -p(t)C(t)e^{-P(t)} + g(t),$$

mais, d'autre part

$$w'(t) = \frac{d}{dt} \left(C(t)e^{-P(t)} \right) = C'(t)e^{-P(t)} - C(t)e^{-P(t)}p(t),$$

d'où $C'(t) = g(t)e^{P(t)}$ et donc $C(t) = \int^t g(s)e^{P(s)}ds$, où on peut fixer librement la primitive (c'est-à-dire, on peut fixer librement la constante "d'intégration"). On arrive finalement au résultat final donnant l'intégrale générale d'une EDO linéaire

$$u(t) = Ce^{-(P(t)-P(t_0))} + \int_{t_0}^t g(s)e^{-(P(t)-P(s))}ds, \quad t \in I,$$

et la solution du problème de Cauchy avec condition initiale $u(t_0) = u_0$ est donnée par la même formule (9.10) déjà obtenue. Comme la primitive a été fixée, qu'une seule constante C apparaît dans l'intégrale générale. Si par contre on avait écrit par exemple $C(t) = \int_{t_0}^t g(s)e^{P(s)}ds + D$ pour un certain $t_0 \in I$, on serait arrivé à

$$\begin{aligned} u(t) &= Ce^{-(P(t)-P(t_0))} + \int_{t_0}^t g(s)e^{-(P(t)-P(s))}ds + De^{-P(t)} \\ &= C'e^{-(P(t)-P(t_0))} + \int_{t_0}^t g(s)e^{-(P(t)-P(s))}ds \end{aligned}$$

avec de nouveau une seule constante C' , où $C' = C + e^{-P(t_0)}D$.

Exemple 9.18. Considérons le problème de Cauchy $u'(t) = u(t) + 1$, $t \in \mathbb{R}$ et $u(t_0) = u_0$. On a $P(t) = \int_{t_0}^t -1 ds = t_0 - t$ et

$$\begin{aligned} u(t) &= u_0 e^{t-t_0} + \int_{t_0}^t 1 \cdot e^{t-s} ds \\ &= u_0 e^{t-t_0} - e^t (e^{-t} - e^{-t_0}) \\ &= -1 + (1 + u_0) e^{t-t_0}. \end{aligned}$$

9.3.1 Équation différentielle linéaire à coefficient constant

Considérons l'équation différentielle linéaire à coefficient constant

$$u'(t) + \alpha u(t) = g(t), \quad t \in I,$$

avec $g : I \rightarrow \mathbb{R}$ continue et $\alpha \in \mathbb{R}$. Dans ce cas, la solution de l'équation homogène est tout simplement $z(t) = Ce^{-\alpha t}$ et l'intégrale générale prend donc la forme $u(t) = w(t) + Ce^{-\alpha t}$ où w est une solution particulière de l'équation non-homogène. Voyons quelques cas particuliers de fonctions g :

g est un polynôme de degré n et $\alpha \neq 0$: $g(t) = \sum_{j=0}^n a_j t^j$.

On cherche w sous la même forme : $w(t) = \sum_{j=0}^n \beta_j t^j$.

Exemple 9.19. Considérons l'EDO linéaire $u'(t) + \alpha u(t) = 1 + t$, $t \in \mathbb{R}$.

On cherche une solution particulière sous la forme $w(t) = \beta_0 + \beta_1 t$. On a donc

$$w'(t) + \alpha w(t) = \beta_1 + \alpha(\beta_0 + \beta_1 t) = 1 + t$$

ce qui implique

$$\begin{cases} \beta_1 + \alpha\beta_0 = 1 \\ \alpha\beta_1 = 1 \end{cases}$$

et donc $(\beta_0, \beta_1) = (\frac{1}{\alpha}(1 - \frac{1}{\alpha}), \frac{1}{\alpha})$. Finalement, l'intégrale générale est

$$u(t) = \frac{1}{\alpha} \left(1 - \frac{1}{\alpha}\right) + \frac{1}{\alpha} t + Ce^{-\alpha t}, \quad t \in \mathbb{R}, \quad C \in \mathbb{R}.$$

g est un polynôme multiplié par une fonction exponentielle : $g(t) = \left(\sum_{j=0}^n a_j t^j\right) e^{\delta t}$.

Si $\delta \neq -\alpha$ alors on cherche $w(t) = \left(\sum_{j=0}^n \beta_j t^j\right) e^{\delta t}$. Par exemple, si $g(t) = e^{\delta t}$, on recherche w sous la forme $w(t) = \beta e^{\delta t}$. On a alors

$$w'(t) = \beta \delta e^{\delta t} = -\alpha \beta e^{\delta t} + e^{\delta t}$$

ce qui implique $\beta = \frac{1}{\delta + \alpha}$ et donc l'intégrale générale est

$$u(t) = \frac{1}{\delta + \alpha} e^{\delta t} + Ce^{-\alpha t}, \quad t \in \mathbb{R}, \quad C \in \mathbb{R}.$$

Si $\delta = -\alpha$ alors on cherche $w(t) = t \left(\sum_{j=0}^n \beta_j t^j \right) e^{\delta t}$. Par exemple, si $g(t) = e^{\delta t} = e^{-\alpha t}$ (qui est solution du problème homogène associé), on recherche w sous la forme $w(t) = t\beta e^{\delta t} = t\beta e^{-\alpha t}$. On trouve que $w(t) = te^{-\alpha t}$ est une solution particulière et l'intégrale générale est

$$u(t) = (C + t)e^{-\alpha t}, \quad t \in \mathbb{R}, \quad C \in \mathbb{R}.$$

g est un polynôme multiplié par des fonctions trigonométriques-exponentielles :

$$g(t) = \left(\sum_{j=0}^n a_j t^j \right) e^{\delta t} \sin(\omega t) + \left(\sum_{j=0}^n \tilde{a}_j t^j \right) e^{\delta t} \cos(\omega t), \quad \omega \neq 0, \quad \delta \in \mathbb{R}.$$

On cherche $w(t)$ sous la même forme :

$$w(t) = \left(\sum_{j=0}^n \beta_j t^j \right) e^{\delta t} \sin(\omega t) + \left(\sum_{j=0}^n \tilde{\beta}_j t^j \right) e^{\delta t} \cos(\omega t).$$

En particulier, si $g(t) = b \sin(\omega t) + c \cos(\omega t)$ avec $b^2 + c^2 > 0$ et $\omega \neq 0$, on recherche w sous la forme $w(t) = \beta \sin(\omega t) + \gamma \cos(\omega t)$, ce qui donne

$$w'(t) = \beta \omega \cos(\omega t) - \gamma \omega \sin(\omega t) = -\alpha(\beta \sin(\omega t) + \gamma \cos(\omega t)) + b \sin(\omega t) + c \cos(\omega t)$$

et implique par identification

$$\begin{cases} \beta \omega = -\alpha \gamma + c \\ -\gamma \omega = -\alpha \beta + b \end{cases} \Rightarrow \begin{cases} \beta = \frac{1}{\alpha^2 + \omega^2} (\alpha b + \omega c) \\ \gamma = \frac{1}{\alpha^2 + \omega^2} (-\omega b + \alpha c) \end{cases}$$

Donc l'intégrale générale est

$$u(t) = \frac{1}{\alpha^2 + \omega^2} (\alpha b + \omega c) \sin(\omega t) + \frac{1}{\alpha^2 + \omega^2} (-\omega b + \alpha c) \cos(\omega t) + C e^{-\alpha t}, \quad t \in \mathbb{R}, \quad C \in \mathbb{R}.$$

Observons que $b = 0 \Rightarrow \beta \neq 0$, et $c = 0 \Rightarrow \gamma \neq 0$.

9.4 Existence et unicité de solutions

Dans cette section, on va étudier l'existence et unicité de solutions d'un problème de Cauchy pour un système d'EDO du premier ordre

$$\begin{cases} \mathbf{u}'(t) = \mathbf{f}(t, \mathbf{u}(t)), & t \in I, \\ \mathbf{u}(t_0) = \mathbf{u}_0, \end{cases} \quad (9.13)$$

où $I \subset \mathbb{R}$ est un intervalle ouvert contenant t_0 , $E \subset \mathbb{R}^n$ un ouvert contenant \mathbf{u}_0 et $\mathbf{f} : I \times E \rightarrow \mathbb{R}^n$ une fonction continue.

Pour cela, on va réécrire (9.13) sous forme intégrale : soit (J, \mathbf{u}) une solution locale (même globale si $J = I$). En particulier, $\mathbf{u} \in C^1(J, E)$ et

$$\mathbf{u}(t) = \mathbf{u}_0 + \int_{t_0}^t \mathbf{f}(s, \mathbf{u}(s)) ds, \quad \forall t \in J. \quad (9.14)$$

Le réciproque est aussi vrai. Si $\mathbf{u} \in C^0(J, E)$ satisfait l'équation intégrale (9.14), avec $J \subset I$ ouvert contenant t_0 , alors \mathbf{u} est différentiable sur J et (J, \mathbf{u}) est une solution locale de (9.13). On introduit alors l'application $\phi : C^0(J, E) \rightarrow C^0(J, \mathbb{R}^n)$ qui à toute fonction $\mathbf{v} \in C^0(J, E)$ fait correspondre la fonction $\phi(\mathbf{v}) \in C^0(J, \mathbb{R}^n)$ définie par

$$\phi(\mathbf{v})(t) = \mathbf{u}_0 + \int_{t_0}^t \mathbf{f}(s, \mathbf{v}(s)) ds, \quad \forall t \in J. \quad (9.15)$$

La solution de (9.14) est un *point fixe* de l'application ϕ , i.e. $\mathbf{u} = \phi(\mathbf{u})$. Avant de présenter les théorèmes d'existence et unicité de solutions de (9.13) (resp. (9.14)) on va introduire la notion abstraite d'espace de Banach et généraliser le théorème de point fixe dans un sous-ensemble fermé d'un espaces de Banach. Ceci nous permettra d'appliquer le théorème de point fixe à l'application ϕ dans l'espace $C^0(J, \mathbb{R}^n)$.

9.4.1 Espace de Banach et complétude de $C^0(J)$

Définition 9.20. Soit $(V, \|\cdot\|)$ un espace vectoriel réel normé et $\{v^{(k)}\}_{k \in \mathbb{N}} \subset V$ une suite d'éléments de V . On dit que

— $\{v^{(k)}\}_{k \in \mathbb{N}}$ converge vers $v \in V$ si $\lim_{k \rightarrow \infty} \|v - v^{(k)}\| = 0$, i.e. si

$$\forall \epsilon > 0, \quad \exists N > 0 : \quad \forall k \geq N, \quad \|v - v^{(k)}\| \leq \epsilon;$$

il est facile de vérifier que, si elle existe, la limite $v \in V$ est unique.

— $\{v^{(k)}\}_{k \in \mathbb{N}}$ est une suite de Cauchy si

$$\forall \epsilon > 0, \quad \exists N > 0 : \quad \forall k, \ell \geq N, \quad \|v^{(k)} - v^{(\ell)}\| \leq \epsilon.$$

On a vu au chapitre 1 que \mathbb{R}^n a une structure d'espace vectoriel (réel) normé et que toute suite de Cauchy est convergente. Cette propriété n'est pas toujours vraie dans un espace vectoriel normé quelconque.

Définition 9.21 (Espace complet). On dit qu'un espace vectoriel normé $(V, \|\cdot\|)$ est complet si toute suite de Cauchy de V converge dans V . Autrement dit, pour toute suite $\{v^{(k)}\}_{k \in \mathbb{N}} \subset V$ de Cauchy, $\exists v^* \in V$ tel que $\lim_{k \rightarrow \infty} \|v^* - v^{(k)}\| = 0$. Un espace vectoriel normé complet est appelé **espace de Banach**.

On dit qu'un sous-ensemble $K \subset V$ est complet si toute suite de Cauchy de K converge dans K .

Exercice 9.22. Considérons l'espace vectoriel réel $C^0([-1, 1])$ de fonctions réelles continues sur le compact $[-1, 1]$ et l'application $\|\cdot\|_1 : C^0([-1, 1]) \rightarrow \mathbb{R}_+$ donnée par $\|f\|_1 = \int_{-1}^1 |f(x)| dx$. Vérifier que $\|\cdot\|_1$ est une norme. L'espace $(C^0([-1, 1]), \|\cdot\|_1)$ est-il complet ?

Dans un espace de Banach $(V, \|\cdot\|)$, les sous-ensembles ouverts sont définis à l'aide de la norme de la même manière que dans \mathbb{R}^n et les sous-ensembles fermés sont définis comme les complémentaires des ouverts. Comme dans \mathbb{R}^n , un sous-ensemble $E \subset V$ est fermé si, et seulement si, toute suite d'éléments de E qui converge dans V a sa limite dans E . Il en résulte immédiatement le lemme suivant qui caractérise les sous-ensembles complets d'un espace de Banach.

Lemme 9.23. *Soit $(V, \|\cdot\|)$ un espace de Banach et $K \subset V$ un sous-ensemble de V . Alors, K est complet si et seulement si K est fermé.*

Par cette caractérisation, on a que si K est un sous-ensemble fermé d'un espace de Banach $(V, \|\cdot\|)$, alors (K, d) est un espace métrique complet, avec distance $d(u, v) = \|u - v\|$ induite par la norme de V .

Le théorème de point fixe s'applique en fait à tout espace de Banach et même plus généralement, à tout espace métrique complet. La preuve est essentiellement la même que celle vue dans \mathbb{R}^n .

Théorème 9.24 (Point fixe de Banach). *Soit $(V, \|\cdot\|)$ un espace de Banach, $K \subset V$ un sous-ensemble fermé non-vidé et $\phi : K \rightarrow K$ une application contractante, i.e.*

$$\exists \rho \in]0, 1[: \forall v, w \in K \quad \|\phi(v) - \phi(w)\| \leq \rho \|v - w\|.$$

Alors il existe un unique $v_ \in K$ tel que $\phi(v_*) = v_*$ (c'est-à-dire, l'application admet un point fixe, qui est unique).*

Pour démontrer les théorèmes de la section suivante, on va travailler avec l'espace des fonctions continues sur un compact, qui a une structure d'espace de Banach s'il est muni d'une norme appropriée, comme le théorème suivant le montre.

Théorème 9.25. *Soit $\Omega \subset \mathbb{R}^n$ compact. L'espace vectoriel $C^0(\Omega, \mathbb{R}^m)$, muni de la norme $\|\mathbf{v}\|_{C^0(\Omega)} = \max_{\mathbf{x} \in \Omega} \|\mathbf{v}(\mathbf{x})\|$, est un espace de Banach. (Ici, $\|\cdot\|$ est n'importe quelle norme sur \mathbb{R}^m).*

Démonstration. Voir cours d'Analyse I. □

En fait, on travaillera plutôt sur des boules fermées de $C^0(\Omega, \mathbb{R}^m)$:

$$K_{b, \mathbf{u}}(\Omega) = \{\mathbf{v} \in C^0(\Omega, \mathbb{R}^m) : \|\mathbf{v} - \mathbf{u}\|_{C^0(\Omega)} \leq b\}, \quad b > 0, \mathbf{u} \in C^0(\Omega, \mathbb{R}^m)$$

qui sont donc des sous-ensembles complets de $C^0(\Omega, \mathbb{R}^m)$.

9.4.2 Théorèmes d'existence et unicité locale

On considère d'abord la question de l'existence de solutions du problème de Cauchy (9.13). Le théorème suivant, qu'on ne va pas démontrer dans ce cours, montre que la seule hypothèse de continuité de $\mathbf{f} : I \times E \rightarrow \mathbb{R}^n$ est suffisante pour garantir l'existence d'une solution locale (mais non pas l'unicité, en générale).

Théorème 9.26 (Cauchy–Peano). *Soit $I \subset \mathbb{R}$ un intervalle ouvert, $E \subset \mathbb{R}^n$ un ouvert et $\mathbf{f} : I \times E \rightarrow \mathbb{R}^n$ continue. Alors, pour tout $(t_0, \mathbf{u}_0) \in I \times E$, le problème de Cauchy (9.13) admet au moins une solution locale (J, \mathbf{u}) , avec $J \subset I$.*

Pour avoir unicité de solutions locales, il faut demander un peu plus de régularité pour la fonction \mathbf{f} . Il s'avère que la bonne condition est une propriété de Lipschitz locale par rapport au deuxième argument :

Définition 9.27. *Soit $I \subset \mathbb{R}$ un intervalle ouvert, $E \subset \mathbb{R}^n$ un ouvert et $\mathbf{f} : I \times E \rightarrow \mathbb{R}^n$ une fonction continue. On dit que \mathbf{f} est localement lipschitzienne par rapport au deuxième argument si, $\forall (t_0, \mathbf{u}_0) \in I \times E$, il existe $a, b > 0$ avec $[t_0 - a, t_0 + a] \times \bar{B}(\mathbf{u}_0, b) \subset I \times E$ et une constante $L > 0$ tels que*

$$\forall t \in [t_0 - a, t_0 + a] \quad \forall \mathbf{x}, \mathbf{y} \in \bar{B}(\mathbf{u}_0, b) \quad \|\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \mathbf{y})\| \leq L \|\mathbf{x} - \mathbf{y}\|,$$

où a, b et L dépendent éventuellement de (t_0, \mathbf{u}_0) .

On remarque, en particulier, que si $\mathbf{f} : I \times E \rightarrow \mathbb{R}^n$, $(t, \mathbf{y}) \mapsto \mathbf{f}(t, \mathbf{y})$, est continue avec dérivées partielles $\frac{\partial f_i}{\partial y_j} : I \times E \rightarrow \mathbb{R}$ continues $\forall i, j = 1, \dots, n$, alors \mathbf{f} est localement lipschitzienne par rapport au deuxième argument.

Théorème 9.28 (Cauchy–Lipschitz — version locale). *Soit $I \subset \mathbb{R}$ un intervalle ouvert, $E \subset \mathbb{R}^n$ ouvert et $\mathbf{f} : I \times E \rightarrow \mathbb{R}^n$ une fonction continue et localement lipschitzienne par rapport au deuxième argument.*

Alors, pour tout $(t_0, \mathbf{u}_0) \in I \times E$, il existe $\delta > 0$ tel que $J_\delta = [t_0 - \delta, t_0 + \delta] \subset I$ et une fonction $\mathbf{u} : J_\delta \rightarrow E$ de classe C^1 solution du problème de Cauchy (9.13).

Cette solution est aussi unique au sens suivant (“unicité locale”) : toute autre solution locale (K, \mathbf{w}) du problème de Cauchy, avec K un intervalle ouvert contenant t_0 et inclus dans I , vérifie $\mathbf{w} = \mathbf{u}$ sur $K \cap J_\delta$.

Remarques. L'énoncé est formulé pour l'intervalle fermé J_δ car ceci est plus naturel pour la preuve. Néanmoins il en découle les mêmes conséquences pour son intérieur \mathring{J}_δ , par exemple l'unicité locale : toute solution locale (K, \mathbf{w}) du problème de Cauchy vérifie $\mathbf{w} = \mathbf{u}$ sur $K \cap \mathring{J}_\delta$ (K étant un intervalle ouvert contenant t_0 et inclus dans I). Les mêmes conséquences seront aussi valables pour toutes les valeurs de $\delta > 0$ plus petites.

Il y a deux manières, équivalentes ici, de définir le concept de solution \mathbf{u} de classe C^1 sur l'intervalle compact $J_\delta \subset I$. On peut d'une part demander que (i) \mathbf{u} soit continue sur J_δ et solution de classe C^1 sur \mathring{J}_δ , ou d'autre part que (ii) \mathbf{u} soit continue sur J_δ , solution de classe C^1 sur \mathring{J}_δ et \mathbf{u}' admette un prolongement par continuité sur J_δ . On voit qu'ici (i) implique (ii) car, en supposant (i), $\mathbf{u}'(t) = \mathbf{f}(t, \mathbf{u}(t))$ sur \mathring{J}_δ , \mathbf{u} est continue sur J_δ et donc $\mathbf{f}(t, \mathbf{u}(t))$ est continue sur J_δ . Clairement (ii) implique (i).

Démonstration. Puisque \mathbf{f} est localement lipschitzienne par rapport au deuxième argument, il existe a, b, L tels que $[t_0 - a, t_0 + a] \times \bar{B}(\mathbf{u}_0, b) \subset I \times E$ et

$$\|\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \mathbf{y})\| \leq L \|\mathbf{x} - \mathbf{y}\|, \quad \forall t \in [t_0 - a, t_0 + a], \quad \forall \mathbf{x}, \mathbf{y} \in \bar{B}(\mathbf{u}_0, b).$$

Choisissons

$$M > \max_{(t, \mathbf{y}) \in [t_0 - a, t_0 + a] \times \bar{B}(\mathbf{u}_0, b)} \|\mathbf{f}(t, \mathbf{y})\|.$$

Soit maintenant $\delta \in]0, a]$, $J_\delta = [t_0 - \delta, t_0 + \delta] \subset I$ et considérons

$$K_{b, \mathbf{u}_0}(J_\delta) = \{\mathbf{v} : J_\delta \rightarrow \mathbb{R}^n \text{ continue} : \max_{t \in J_\delta} \|\mathbf{v}(t) - \mathbf{u}_0\| \leq b\},$$

qui est un sous-ensemble fermé (et donc complet) de $C^0(J_\delta, \mathbb{R}^n)$. On va étudier l'application $\phi : K_{b, \mathbf{u}_0}(J_\delta) \rightarrow C^0(J_\delta, \mathbb{R}^n)$ définie en (9.15) :

$$\phi(\mathbf{v})(t) = \mathbf{u}_0 + \int_{t_0}^t \mathbf{f}(s, \mathbf{v}(s)) ds, \quad \forall t \in J_\delta.$$

Pour tout $\mathbf{v} \in K_{b, \mathbf{u}_0}(J_\delta)$ et $t \in J_\delta$ on a

$$\|\phi(\mathbf{v})(t) - \mathbf{u}_0\| = \left\| \int_{t_0}^t \mathbf{f}(s, \mathbf{v}(s)) ds \right\| \leq \left| \int_{t_0}^t \|\mathbf{f}(s, \mathbf{v}(s))\| ds \right| \leq M\delta,$$

donc, si on prend $\delta \leq \frac{b}{M}$ on a $\|\phi(\mathbf{v}) - \mathbf{u}_0\|_{C^0(J_\delta)} \leq b$, i.e. $\phi(\mathbf{v}) \in K_{b, \mathbf{u}_0}(J_\delta)$, $\forall \mathbf{v} \in K_{b, \mathbf{u}_0}(J_\delta)$ et ϕ envoie $K_{b, \mathbf{u}_0}(J_\delta)$ dans lui-même. De plus, $\forall \mathbf{v}_1, \mathbf{v}_2 \in K_{b, \mathbf{u}_0}(J_\delta)$ et $\forall t \in J_\delta$

$$\begin{aligned} \|\phi(\mathbf{v}_1)(t) - \phi(\mathbf{v}_2)(t)\| &= \left\| \int_{t_0}^t \mathbf{f}(s, \mathbf{v}_1(s)) - \mathbf{f}(s, \mathbf{v}_2(s)) ds \right\| \leq \left| \int_{t_0}^t \|\mathbf{f}(s, \mathbf{v}_1(s)) - \mathbf{f}(s, \mathbf{v}_2(s))\| ds \right| \\ &\leq \left| \int_{t_0}^t L \|\mathbf{v}_1(s) - \mathbf{v}_2(s)\| ds \right| \leq L\delta \max_{t \in J_\delta} \|\mathbf{v}_1(t) - \mathbf{v}_2(t)\| \end{aligned}$$

ce qui implique $\|\phi(\mathbf{v}_1) - \phi(\mathbf{v}_2)\|_{C^0(J_\delta)} \leq L\delta \|\mathbf{v}_1 - \mathbf{v}_2\|_{C^0(J_\delta)}$. Donc, pour $\delta < \min\{a, \frac{b}{M}, \frac{1}{L}\}$, l'application $\phi : K_{b, \mathbf{u}_0}(J_\delta) \rightarrow K_{b, \mathbf{u}_0}(J_\delta)$ est contractante et il existe un unique $\mathbf{u} \in K_{b, \mathbf{u}_0}(J_\delta)$ point fixe de ϕ . Alors (J_δ, \mathbf{u}) est l'unique solution du problème de Cauchy (9.13) dans $K_{b, \mathbf{u}_0}(J_\delta)$.

On pourrait encore se demander s'il existe d'autres solutions $\tilde{\mathbf{u}} \notin K_{b, \mathbf{u}_0}(J_\delta)$ définies sur J_δ . Montrons que ceci n'est pas le cas en considérant une solution $\tilde{\mathbf{u}} : J_\delta \rightarrow E$ du problème de Cauchy. Puisque $\tilde{\mathbf{u}}(t_0) = \mathbf{u}_0$ et $\tilde{\mathbf{u}}_0$ est continue en t_0 , il existe $0 < \tilde{\delta} \leq \delta$ tel que $\|\tilde{\mathbf{u}}(t) - \mathbf{u}_0\| \leq b$, $\forall t \in J_{\tilde{\delta}} = [t_0 - \tilde{\delta}, t_0 + \tilde{\delta}]$. Soit

$$\beta = \max\{\tilde{\delta} \in]0, \delta] : \|\tilde{\mathbf{u}}(t) - \mathbf{u}_0\| \leq b, \forall t \in [t_0 - \tilde{\delta}, t_0 + \tilde{\delta}]\}$$

(se convaincre que le maximum existe) et supposons par l'absurde que $\beta < \delta$ et donc $\beta < \frac{b}{M}$. Puisque $\tilde{\mathbf{u}}$ satisfait l'équation $\tilde{\mathbf{u}}(t) - \mathbf{u}_0 = \int_{t_0}^t \mathbf{f}(s, \tilde{\mathbf{u}}(s)) ds$, on a

$$\|\tilde{\mathbf{u}}(t) - \mathbf{u}_0\| \leq \left| \int_{t_0}^t \|\mathbf{f}(s, \tilde{\mathbf{u}}(s))\| ds \right| \leq M\beta < b, \quad \forall t \in [t_0 - \beta, t_0 + \beta].$$

Mais alors, il existe un $\hat{\delta} \in]0, \delta - \beta[$ tel que $\|\tilde{\mathbf{u}}(t) - \mathbf{u}_0\| \leq b$, $\forall t \in [t_0 - \beta - \hat{\delta}, t_0 + \beta + \hat{\delta}]$, ce qui contredit la définition de β . Ainsi $\tilde{\mathbf{u}} = \mathbf{u}$ sur J_δ .

Soit, maintenant, (K, \mathbf{w}) une solution locale du problème de Cauchy et $\tilde{J} \subset K \cap J_\delta$ un intervalle fermé contenant t_0 dans son intérieur. Le même raisonnement que celui ci-dessus, mais appliqué à l'intervalle \tilde{J} , donne que $\mathbf{w}(t) = \mathbf{u}(t)$, $\forall t \in \tilde{J}$. Comme on obtient ainsi l'égalité sur tout intervalle fermé $\tilde{J} \subset K \cap J_\delta$ contenant t_0 dans son intérieur, on a prouvé l'égalité sur $K \cap J_\delta$.

□

Exemple 9.29. *Considérons l'edo*

$$u'(t) = u^2(t), \quad t \in \mathbb{R}.$$

Puisque $f(t, y) = y^2 : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ est localement lipschitzienne par rapport au deuxième argument, le problème de Cauchy avec condition initiale $u(t_0) = u_0$ admet, pour tout $(t_0, u_0) \in \mathbb{R}^2$, une solution locale, localement unique.

Exemple 9.30. *Considérons l'edo*

$$u'(t) = -\sqrt{|u(t)|}, \quad t \in \mathbb{R}.$$

La fonction $f(t, y) = \sqrt{|y|} : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ est continue sur \mathbb{R}^2 mais pas localement lipschitzienne par rapport au deuxième argument. Par contre, si f est restreinte à $\mathbb{R} \times (\mathbb{R} \setminus \{0\})$, f est localement lipschitzienne par rapport au deuxième argument.

On conclut du théorème 9.28 appliqué à $\mathbb{R} \times (\mathbb{R} \setminus \{0\})$ que le problème de Cauchy avec condition initiale $u(t_0) = u_0$ admet une solution locale, localement unique, pour tout $(t_0, u_0) \in \mathbb{R} \times (\mathbb{R} \setminus \{0\})$. Si la condition initiale est $u(t_0) = 0$, on aura l'existence de solutions locales grâce au théorème de Cauchy-Peano 9.26 appliqué à $\mathbb{R} \times \mathbb{R}$ mais on ne peut pas garantir l'unicité locale. En fait, en raisonnant comme dans l'exemple 9.12, on voit qu'aucune solution locale n'est localement unique.

Le théorème 9.28 donne existence et unicité d'une solution $\mathbf{u} : J \rightarrow E$ du problème de Cauchy sur un intervalle fermé $J = [t_0 - \delta, t_0 + \delta] \subset I$, pourvu que la fonction \mathbf{f} soit localement lipschitzienne par rapport au deuxième argument autour de la condition initiale (t_0, \mathbf{u}_0) . Cette solution est toujours prolongeable. En effet, prenons $t_1 = t_0 + \delta$ et $\mathbf{u}_1 = \mathbf{u}(t_0 + \delta)$. Puisque $(t_1, \mathbf{u}_1) \in I \times E$ et \mathbf{f} est localement lipschitzienne par rapport au deuxième argument sur $I \times E$, on peut appliquer encore le théorème 9.28, cette fois-ci au problème de Cauchy avec condition initiale $\mathbf{u}(t_1) = \mathbf{u}_1$. On aura alors l'existence d'une solution $\tilde{\mathbf{u}} : \tilde{J} \rightarrow E$ définie sur un intervalle fermé $\tilde{J} = [t_1 - \tilde{\delta}, t_1 + \tilde{\delta}] \subset I$ avec $\tilde{\delta} > 0$. Soit la fonction $\hat{\mathbf{u}} : J \cup [t_1, t_1 + \tilde{\delta}] = [t_0 - \delta, t_0 + \delta + \tilde{\delta}] \rightarrow E$ définie par $\hat{\mathbf{u}}(t) = \mathbf{u}(t)$ si $t \in J$ et $\hat{\mathbf{u}}(t) = \tilde{\mathbf{u}}(t)$ si $t \in]t_1, t_1 + \tilde{\delta}]$. Alors $\hat{\mathbf{u}}$ est solution (vérification directe dans l'esprit du Lemme 9.6). La solution $\hat{\mathbf{u}}$ est un « prolongement à droite » de la solution \mathbf{u} . De même, on peut prolonger \mathbf{u} à gauche, et à la fois à droite et à gauche.

Lemme 9.31. *Sous les mêmes hypothèses du théorème 9.28, soient des intervalles $J_1, J_2 \subset I$ contenant t_0 dans leurs intérieurs et soient des fonctions continues $\mathbf{u}_1 : J_1 \rightarrow E$ et $\mathbf{u}_2 : J_2 \rightarrow E$, solutions du problème de Cauchy (9.13) sur respectivement $\overset{\circ}{J}_1$ et $\overset{\circ}{J}_2$. Alors $\mathbf{u}_1(t) = \mathbf{u}_2(t)$, $\forall t \in J_1 \cap J_2$.*

Démonstration. Notons $J_1 \cap J_2 = J$ et $\overset{\circ}{J} =]A, B[$ avec $-\infty \leq A < t_0 < B \leq +\infty$. Le théorème 9.28 garantit l'existence d'un intervalle fermé $J_\delta = [t_0 - \delta, t_0 + \delta] \subset \overset{\circ}{J}$ (avec $\delta > 0$) où la solution existe et est unique. Sur un tel intervalle on doit avoir nécessairement $\mathbf{u}_1 = \mathbf{u}_2$. Soit alors

$$\alpha = \inf \left\{ \tilde{\alpha} \in]A, t_0 - \delta] : \mathbf{u}_1 = \mathbf{u}_2 \text{ sur } [\tilde{\alpha}, t_0 + \delta] \right\} \geq A$$

et

$$\beta = \sup \left\{ \tilde{\beta} \in [t_0 + \delta, B[: \mathbf{u}_1 = \mathbf{u}_2 \text{ sur } [t_0 - \delta, \tilde{\beta}] \right\} \leq B.$$

Donc $\tilde{J} =]\alpha, \beta[$ est le plus grand intervalle ouvert inclus dans \mathring{J} , contenant t_0 et où $\mathbf{u}_1 = \mathbf{u}_2$. On veut montrer que $\tilde{J} = \mathring{J}$. Supposons par l'absurde que $\beta \neq B$, et donc $\beta \in \mathring{J} \setminus \tilde{J}$. Puisque \mathbf{u}_1 et \mathbf{u}_2 sont continues sur $J \supset \tilde{J}$ et coïncident sur \tilde{J} on a

$$\mathbf{u}_1(\beta) = \lim_{t \rightarrow \beta^-} \mathbf{u}_1(t) = \lim_{t \rightarrow \beta^-} \mathbf{u}_2(t) = \mathbf{u}_2(\beta) := \mathbf{z}_0$$

et on pourra encore trouver un intervalle $J_{\beta, \hat{\delta}} = [\beta - \hat{\delta}, \beta + \hat{\delta}] \subset \mathring{J}$ où la solution du problème de Cauchy avec condition initiale (β, \mathbf{z}_0) existe et est unique. Mais cette solution doit coïncider avec \mathbf{u}_1 et \mathbf{u}_2 , et donc $\mathbf{u}_1 = \mathbf{u}_2$ sur $\tilde{J} \cup J_{\beta, \hat{\delta}}$, ce qui contredit l'hypothèse que \tilde{J} était le plus grand intervalle ouvert contenant t_0 où $\mathbf{u}_1 = \mathbf{u}_2$. Le même raisonnement s'applique à gauche si $\alpha \neq A$, donc on conclut que $\tilde{J} = \mathring{J}$, $\mathbf{u}_1 = \mathbf{u}_2$ sur \mathring{J} , et donc sur J (par continuité de \mathbf{u}_1 et \mathbf{u}_2 sur J). \square

Grâce au Lemme précédent, on peut toujours construire une unique solution maximale.

Théorème 9.32. *Sous les hypothèses du théorème 9.28, pour tout $(t_0, \mathbf{u}_0) \in I \times E$ il existe une unique solution maximale (J_{\max}, \mathbf{u}) du problème de Cauchy (9.13) avec $J_{\max} \subset I$ ouvert contenant t_0 .*

Démonstration. Soit $\{(J_\eta, \mathbf{u}_\eta)\}_\eta$ l'ensemble des solutions locales du problème de Cauchy (9.13), avec $J_\eta \subset I$ intervalle ouvert contenant t_0 . Ici chaque solution locale est paramétrée par η appartenant à un certain ensemble Γ en bijection avec l'ensemble des solutions locales.

On note $J_{\max} = \bigcup_\eta J_\eta$ et on définit la fonction $\mathbf{u} : J_{\max} \rightarrow \mathbb{R}^n$ de la façon suivante : pour $t \in J_{\max}$ il existe au moins un η t.q. $t \in J_\eta$; on définit alors $\mathbf{u}(t) = \mathbf{u}_\eta(t)$. Cette définition de $\mathbf{u}(t)$ ne dépend pas du choix de η car si η' est aussi tel que $t \in J_{\eta'}$, par le lemme 9.31 on a $\mathbf{u}_\eta(t) = \mathbf{u}_{\eta'}(t)$. De plus, la fonction \mathbf{u} ainsi définie est solution du problème de Cauchy et elle est la solution maximale car toute solution locale $(J_\eta, \mathbf{u}_\eta)$ est telle que $J_\eta \subset J_{\max}$ et $\mathbf{u}_\eta(t) = \mathbf{u}(t)$, $\forall t \in J_\eta$. \square

Remarque 9.33. *Si $\mathbf{u} : J \rightarrow E$ est solution du problème de Cauchy sur un intervalle compact contenant t_0 dans son intérieur, on peut toujours prolonger la solution \mathbf{u} sur un ouvert $J_\eta \supset J$ comme expliqué juste avant l'énoncé du Lemme 9.31. Comme $J_\eta \subset J_{\max}$, on a aussi $J \subset J_{\max}$.*

Par le théorème précédent, sous les hypothèses du théorème 9.28 le problème de Cauchy (9.13) a une solution maximale unique. Alors, ou bien $J_{\max} = I$ et la solution est globale, ou bien J_{\max} est strictement inclus dans I et on aura seulement une solution maximale mais non globale. On peut se poser encore la question de savoir que se passe-t-il aux extrémités de l'intervalle maximale si celui ci est strictement inclus dans I . Le théorème suivant répond à cette question.

Théorème 9.34. *Sous les mêmes hypothèses du théorème 9.28, soit (J_{\max}, \mathbf{u}) la solution maximale du problème de Cauchy (9.13), définie sur un intervalle ouvert $J_{\max} =]\alpha, \beta[\subset I$ contenant t_0 , où $-\infty \leq \alpha < \beta \leq +\infty$. Si $\beta \in I$ (et donc β est fini), alors*

- pour toute suite $\{t_n\}_{n \in \mathbb{N}}$ dans J_{\max} t.q. $\lim_{n \rightarrow \infty} t_n = \beta$ et $\lim_{n \rightarrow \infty} \mathbf{u}(t_n) = \boldsymbol{\xi} \in \overline{E} \subset \mathbb{R}^n$, on a $\boldsymbol{\xi} \in \partial E$;
- en particulier, si $E = \mathbb{R}^n$, alors $\lim_{t \rightarrow \beta^-} \|\mathbf{u}(t)\| = +\infty$.

On a des conclusions analogues si $\alpha \in I$.

La démonstration de ce théorème va au delà des objectifs de ce cours.

Une des conséquences du Lemme 9.31 est que deux solutions locales (maximales ou non) de l'EDO sans spécification d'une condition initiale, soit coïncident sur la partie commune de leurs domaines de définition, soit leurs graphes ne peuvent jamais « se rencontrer ». Plus précisément : considérons l'équation différentielle $\mathbf{u}'(t) = \mathbf{f}(t, \mathbf{u}(t))$, $t \in I$, où \mathbf{f} satisfait les hypothèses du théorème 9.28. Soit (J, \mathbf{v}) une solution locale du problème de Cauchy avec condition initiale $\mathbf{v}(t_0) = \mathbf{v}_0$ et $(\tilde{J}, \tilde{\mathbf{v}})$ une solution locale du problème de Cauchy avec condition initiale $\tilde{\mathbf{v}}(\tilde{t}_0) = \tilde{\mathbf{v}}_0$. Si \mathbf{v} et $\tilde{\mathbf{v}}$ prennent la même valeur en un certain $\hat{t} \in J \cap \tilde{J}$, sous les hypothèses du théorème 9.28 de Cauchy–Lipschitz local, on doit avoir nécessairement $\mathbf{v}(t) = \tilde{\mathbf{v}}(t)$, $\forall t \in J \cap \tilde{J}$. Voir le Lemme 9.31.

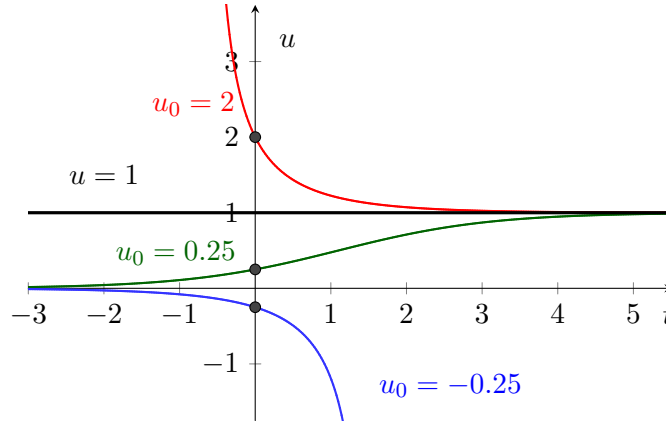
Exemple 9.35. *Considérons le problème de Cauchy suivant*

$$\begin{cases} u'(t) = u(t)(1 - u(t)), & t \in \mathbb{R} \\ u(t_0) = u_0. \end{cases} \quad (9.16)$$

La fonction $f(t, y) = y(1 - y)$ est localement lipschitzienne par rapport au deuxième argument pour tout $(t, y) \in \mathbb{R} \times \mathbb{R}$. Donc le problème de Cauchy admet une solution maximale unique. Clairement, $u(t) = 0$ et $u(t) = 1 \ \forall t \in \mathbb{R}$ sont deux solutions (globales). Le graphe de toute autre solution maximale ne peut pas rencontrer les droites $u = 0, u = 1$. De plus,

- Si $u_0 \in]0, 1[$ alors la solution maximale correspondante satisfait $\text{Im}(u) \in]0, 1[$ et est croissante car $f(t, y) > 0$ pour $0 < y < 1$; par le Théorème 9.34, u est définie sur tout \mathbb{R} ;
- Si $u_0 \in]1, +\infty[$ alors $\text{Im}(u) \subset]1, +\infty[$ et est décroissante car $f(t, y) < 0$ pour $y > 1$; par le Théorème 9.34, u est définie sur un intervalle ouvert contenant $[t_0, +\infty[$;
- Si $u_0 \in]-\infty, 0[$ alors $\text{Im}(u) \subset]-\infty, 0[$ et est décroissante car $f(t, y) < 0$ pour $y < 0$; par le Théorème 9.34, u est définie sur un intervalle ouvert contenant $] -\infty, t_0]$.

La figure ci dessous montre trois solutions maximales correspondantes aux trois problèmes de Cauchy avec $u_0 \in]0, 1[$, $u_0 \in]1, +\infty[$, $u_0 \in]-\infty, 0[$ et $t_0 = 0$.



Le théorème 9.34 nous garantit que pour $u_0 \in]0, 1[$ la solution est globale (autrement elle devrait aller à l'infini en un temps fini mais ceci impliquerait de franchir une des deux barrières $u = 0$ ou $u = 1$.) On ne peut par arriver à la même conclusion si $u_0 \in]1, +\infty[$ ou $u_0 \in]-\infty, 0[$. En effet, en faisant les calculs, on voit que la solution est maximale mais non globale.

9.4.3 Théorèmes d'existence globale

Dans cette section on présente des conditions suffisantes sur la fonction \mathbf{f} qui garantissent l'existence et unicité d'une solution globale. La première condition qu'on considère est une condition de lipschitzianité globale par rapport au deuxième argument qui permet de donner une version globale du théorème de Cauchy-Lipschitz local 9.28.

Définition 9.36. Soit $I \subset \mathbb{R}$ un intervalle ouvert et $\mathbf{f} : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ une fonction continue (attention : \mathbf{f} doit être définie sur tout $I \times \mathbb{R}^n$). On dit que \mathbf{f} est globalement lipschitzienne par rapport au deuxième argument s'il existe une fonction continue $\ell : I \rightarrow \mathbb{R}_+$ non négative telle que

$$\forall t \in I \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \quad \|\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \mathbf{y})\| \leq \ell(t) \|\mathbf{x} - \mathbf{y}\|.$$

Théorème 9.37 (Cauchy-Lipschitz — version globale). Soit $I \subset \mathbb{R}$ un intervalle ouvert et $\mathbf{f} : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ continue et globalement lipschitzienne par rapport au deuxième argument. Alors, pour tout $(t_0, \mathbf{u}_0) \in I \times \mathbb{R}^n$, le problème de Cauchy (9.13) a une solution globale unique $\mathbf{u} \in C^1(I, \mathbb{R}^n)$.

Démonstration. Puisque \mathbf{f} est globalement lipschitzienne par rapport au deuxième argument, elle est aussi localement lipschitzienne par rapport au deuxième argument. En effet, pour tout intervalle compact $K \subset I$ non vide, on a

$$\forall t \in K \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \quad \|\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \mathbf{y})\| \leq L \|\mathbf{x} - \mathbf{y}\|,$$

où $L = \max_{s \in K} \ell(s)$. Par le théorème de Cauchy-Lipschitz local 9.28 et le Théorème 9.32, il existe une unique solution maximale (J_{\max}, \mathbf{u}) du problème de Cauchy. Pour montrer le théorème, il suffit donc de vérifier que $J_{\max} = I$.

Raisonnons par l'absurde en supposant que J_{max} est strictement inclus dans I , autrement dit, $\sup J_{max} \in I$ ou $\inf J_{max} \in I$ (ou les deux à la fois). Traitons le cas $\sup J_{max} \in I$, l'autre cas étant analogue. Pour $T := \sup J_{max}$ (qui est fini puisque dans I), choisissons $\mu > 0$ tel que $J := [T - \mu, T + \mu] \subset I$, et soit $\delta \in]0, \mu]$ tel que $\delta \max_{t \in J} \ell(t) \leq \frac{1}{3}$.

Posons $J_0 = [T - \delta, T + \delta]$. Alors, l'application $\phi : C^0(J_0, \mathbb{R}^n) \rightarrow C^0(J_0, \mathbb{R}^n)$ définie par $\phi(\mathbf{v})(t) = \mathbf{u}(T - \frac{\delta}{2}) + \int_{T-\frac{\delta}{2}}^t \mathbf{f}(s, \mathbf{v}(s))ds$, $t \in J_0$, est contractante. En effet, pour tout $\mathbf{v}_1, \mathbf{v}_2 \in C^0(J_0, \mathbb{R}^n)$ on a

$$\begin{aligned} \max_{t \in J_0} \|\phi(\mathbf{v}_1)(t) - \phi(\mathbf{v}_2)(t)\| &\leq \max_{t \in J_0} \left| \int_{T-\frac{\delta}{2}}^t \|\mathbf{f}(s, \mathbf{v}_1(s)) - \mathbf{f}(s, \mathbf{v}_2(s))\| ds \right| \\ &\leq \max_{t \in J_0} \left| \int_{T-\frac{\delta}{2}}^t \ell(s) \|\mathbf{v}_1(s) - \mathbf{v}_2(s)\| ds \right| \\ &\leq \frac{3\delta}{2} \left(\max_{t \in J_0} \ell(t) \right) \left(\max_{t \in J_0} \|\mathbf{v}_1(t) - \mathbf{v}_2(t)\| \right) \\ &\leq \frac{1}{2} \max_{t \in J_0} \|\mathbf{v}_1(t) - \mathbf{v}_2(t)\|. \end{aligned}$$

Par le théorème de point fixe de Banach on a alors une fonction $\mathbf{u}^{(0)} \in C^0(J_0, \mathbb{R}^n)$ point fixe de ϕ qui est, en particulier, de classe C^1 et solution du problème de Cauchy sur J_0 .

Grâce au Lemme 9.31 appliqué aux intervalles $J_1 = J_2 = [T - \delta, T[$ et à la relation $\mathbf{u}^{(0)}(T - \frac{\delta}{2}) = \mathbf{u}(T - \frac{\delta}{2})$, on a $\mathbf{u}^{(0)} = \mathbf{u}$ sur $[T - \delta, T[$. On obtient une solution locale $\tilde{\mathbf{u}}$ du problème de Cauchy de départ définie sur $J_{max} \cup [T, T + \delta[$ en posant $\tilde{\mathbf{u}} = \mathbf{u}$ sur J_{max} et $\tilde{\mathbf{u}} = \mathbf{u}^{(0)}$ sur $[T, T + \delta[$. Ceci contredit la maximalité de J_{max} . \square

Exemple 9.38. *Considérons le problème de Cauchy*

$$\begin{cases} u'(t) = \sin(u(t)), & t \in \mathbb{R}. \\ u(t_0) = u_0. \end{cases}$$

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction $f(t, y) = \sin(y)$. Puisque $|\frac{\partial f}{\partial y}(t, y)| \leq 1$, $\forall (t, y) \in \mathbb{R}^2$ on a que f est globalement lipschitzienne par rapport à y avec $\ell(t) = 1$ et le problème de Cauchy admet une unique solution globale, i.e. définie sur tout \mathbb{R} .

Exemple 9.39. *Soit $I \subset \mathbb{R}$ un intervalle ouvert et $g, p \in C^0(I)$. Considérons le problème de Cauchy*

$$\begin{cases} u'(t) = g(t) - p(t)u(t), & t \in I \\ u(t_0) = u_0. \end{cases}$$

La méthode du facteur intégrant nous a montré que ce problème admet une unique solution globale (définie sur tout I). On peut aussi le voir comme suit. La fonction $f(t, y) = g(t) - p(t)y$ est globalement lipschitzienne par rapport au deuxième argument avec $\ell(t) = |p(t)|$, qui est une fonction continue, et donc le problème de Cauchy a une solution unique globale définie sur tout l'intervalle I . Ceci remontre que une edo linéaire scalaire avec fonctions p, g continues a toujours une solution unique globale satisfaisant une condition initiale $u(t_0) = u_0$.

La condition de Lipschitz globale demandée par le théorème 9.37 de Cauchy–Lipschitz global est assez restrictive. L'exemple 9.35 montre qu'un problème de Cauchy peut avoir des solutions globales même si la fonction f n'est pas globalement lipschitzienne par rapport au deuxième argument. On présente ci-après d'autres théorèmes d'existence globale qui demandent des conditions moins restrictives sur \mathbf{f} .

Théorème 9.40. *Soit $I \subset \mathbb{R}$ un intervalle ouvert contenant t_0 et $\mathbf{f} : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ une fonction continue et localement lipschitzienne par rapport au deuxième argument. Supposons de plus qu'il existe $\ell : I \rightarrow \mathbb{R}$ continue telle que*

$$\mathbf{y} \cdot \mathbf{f}(t, \mathbf{y}) \leq \ell(t)(1 + \|\mathbf{y}\|^2) \quad \forall t \in I, \forall \mathbf{y} \in \mathbb{R}^n. \quad (9.17)$$

Alors, le problème à valeur initiale

$$\mathbf{u}'(t) = \mathbf{f}(t, \mathbf{u}(t)), \quad t \in I_+ = I \cap [t_0, +\infty[, \quad \mathbf{u}(t_0) = \mathbf{u}_0$$

admet une solution globale unique $\mathbf{u} \in C^1(I_+, \mathbb{R}^n)$. Si, de plus, ℓ est non négative et

$$|\mathbf{y} \cdot \mathbf{f}(t, \mathbf{y})| \leq \ell(t)(1 + \|\mathbf{y}\|^2) \quad \forall t \in I, \forall \mathbf{y} \in \mathbb{R}^n, \quad (9.18)$$

alors le problème de Cauchy (9.13) admet une solution globale unique sur I .

Démonstration. Par le théorème de Cauchy–Lipschitz local 9.28 et le Théorème 9.32, il existe une solution maximale (J_{max}, \mathbf{u}) du problème de Cauchy. Si $J_{max,+} = J_{max} \cap [t_0, +\infty[$ est strictement contenu dans I_+ alors $J_{max,+} = [t_0, \beta[$, avec $\beta \in \dot{I}_+$ et $\lim_{t \rightarrow \beta^-} \|\mathbf{u}(t)\| = +\infty$ selon le théorème 9.34. Montrons que ceci ne peut pas arriver. Pour tout $t \in J_{max}$ on a

$$\mathbf{u}(t) \cdot \mathbf{u}'(t) = \frac{1}{2} \frac{d}{dt} (\|\mathbf{u}(t)\|^2) = \mathbf{u}(t) \cdot \mathbf{f}(t, \mathbf{u}(t)) \leq \ell(t)(1 + \|\mathbf{u}(t)\|^2).$$

Soit $h(t) = \|\mathbf{u}(t)\|^2$. Alors $h'(t) \leq 2\ell(t)(1 + h(t))$ ce qui implique, pour tout $t \in [t_0, \beta[$

$$\int_{t_0}^t \frac{h'(s)}{1 + h(s)} ds = \ln \left(\frac{1 + h(t)}{1 + h(t_0)} \right) \leq 2 \int_{t_0}^t \ell(s) ds$$

et donc

$$h(t) \leq -1 + (1 + h(t_0)) \exp \left(2 \int_{t_0}^t \ell(s) ds \right).$$

Mais $\ell \in C^0([t_0, \beta])$ est bornée donc $\left| \int_{t_0}^t \ell(s) ds \right| \leq \int_{t_0}^\beta |\ell(s)| ds < +\infty$, $\forall t \in [t_0, \beta[$, et $h(t)$ est bornée uniformément sur $[t_0, \beta[$. Par conséquent, si on note $M = \sqrt{1 + \|\mathbf{u}_0\|^2} e^{\int_{t_0}^\beta |\ell(s)| ds}$ on a $M < +\infty$ et

$$\|\mathbf{u}(t)\| \leq M, \quad \forall t \in [t_0, \beta[,$$

ce qui contredit l'hypothèse que $\lim_{t \rightarrow \beta^-} \|\mathbf{u}(t)\| = +\infty$.

Dans le cas de la condition bilatérale (9.18) on a aussi

$$h'(t) = 2\mathbf{u}(t) \cdot \mathbf{u}'(t) = 2\mathbf{u}(t) \cdot \mathbf{f}(t, \mathbf{u}(t)) \geq -2\ell(t)(1 + \|\mathbf{u}(t)\|^2) = -2\ell(t)(1 + h(t))$$

et donc, pour tout $t \in J_{\max} \cap]-\infty, t_0]$,

$$\int_t^{t_0} \frac{h'(s)}{1+h(s)} ds = \ln \left(\frac{1+h(t_0)}{1+h(t)} \right) \geq -2 \int_t^{t_0} \ell(s) ds$$

ce qui implique de nouveau $h(t) \leq -1 + (1+h(t_0)) \exp \left(2 \int_t^{t_0} \ell(s) ds \right)$. Si $J_{\max} \cap]-\infty, t_0] =]\alpha, t_0]$ et $\alpha \in I$, alors on doit avoir $\lim_{t \rightarrow \alpha^+} \|\mathbf{u}(t)\| = +\infty$ mais, d'un autre côté, on a $\|\mathbf{u}(t)\| \leq \sqrt{1 + \|\mathbf{u}_0\|^2} e^{\int_\alpha^{t_0} \ell(s) ds} < +\infty, \forall t \in]\alpha, t_0]$, ce qui est contradictoire. On conclut alors que $J_{\max} = I$.

Il reste à prouver l'unicité. Soit l'unique solution maximale (J_{\max}, \mathbf{u}) pour le problème de Cauchy. Considérons une solution $\tilde{\mathbf{u}} : I_+ \rightarrow \mathbb{R}^n$ pour le problème à valeur initiale. Elle peut se prolonger sur la gauche en une solution $\hat{\mathbf{u}} :]t_0 - \epsilon, t_0[\cup I_+ \rightarrow \mathbb{R}^n$ pour un certain $\epsilon > 0$. Par unicité de la solution maximale, $]t_0 - \epsilon, t_0[\cup I_+ \subset I_{\max}$ et $\hat{\mathbf{u}} = \mathbf{u}$ sur $]t_0 - \epsilon, t_0[\cup I_+$, et donc forcément que $\tilde{\mathbf{u}} = \mathbf{u}$ sur I_+ .

Considérons une solution globale $\bar{\mathbf{u}} : I \rightarrow \mathbb{R}^n$ pour le problème de Cauchy. Par unicité de (J_{\max}, \mathbf{u}) , on a $I = J_{\max}$ et $\bar{\mathbf{u}} = \mathbf{u}$ sur I ; ainsi $\bar{\mathbf{u}}$ est uniquement déterminée. \square

Remarque 9.41. La condition bilatérale (9.18) du théorème 9.40 est garantie si par exemple $\exists k_1, k_2 : I \rightarrow \mathbb{R}_+$ continues et non négatives telles que

$$\|\mathbf{f}(t, \mathbf{y})\| \leq k_1(t) + k_2(t)\|\mathbf{y}\| \quad \forall t \in I, \forall \mathbf{y} \in \mathbb{R}^n.$$

En effet,

$$\begin{aligned} |\mathbf{y} \cdot \mathbf{f}(t, \mathbf{y})| &\leq \|\mathbf{y}\| \|\mathbf{f}(t, \mathbf{y})\| \leq k_1(t)\|\mathbf{y}\| + k_2(t)\|\mathbf{y}\|^2 \\ &\leq \frac{k_1(t)}{2}(1 + \|\mathbf{y}\|^2) + k_2(t)(1 + \|\mathbf{y}\|^2) \leq \left(\frac{k_1(t)}{2} + k_2(t) \right) (1 + \|\mathbf{y}\|^2). \end{aligned}$$

Ceci montre que l'on a des solutions globales du problème de Cauchy si la norme de $\mathbf{f}(t, \mathbf{y})$ croît au plus linéairement par rapport à la norme de \mathbf{y} .

Exemple 9.42. Considérons le problème à valeur initiale

$$\begin{cases} u'(t) = -u(t)e^{u(t)}, & t \geq t_0, \\ u(t_0) = u_0. \end{cases}$$

La fonction $f(t, y) = -ye^y$ est localement lipschitzienne par rapport au deuxième argument, donc ce problème admet une solution maximale unique. Toutefois, f n'est pas globalement lipschitzienne par rapport au deuxième argument, donc on n'est pas garanti à priori que la solution maximale soit définie sur tout $[t_0, +\infty[$. Toutefois, $yf(t, y) = -y^2e^y \leq 0$, $\forall y \in \mathbb{R}, t \in [t_0, \infty[$, donc la solution maximale est globale.

(Attention : on ne peut pas déduire la même conclusion si on définit le problème sur tout \mathbb{R} au lieu de $[t_0, +\infty[$. On a donc l'existence globale garantie seulement « à droite ».)

Le théorème prochain donne une autre condition suffisante pour l'existence et unicité de solutions globales unilatérales.

Théorème 9.43. Soit $I \subset \mathbb{R}$ un intervalle ouvert contenant t_0 et $\mathbf{f} : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ une fonction continue et localement lipschitzienne par rapport au deuxième argument. Supposons qu'il existe $\ell : I \rightarrow \mathbb{R}$ continue telle que

$$(\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \mathbf{y})) \cdot (\mathbf{x} - \mathbf{y}) \leq \ell(t) \|\mathbf{x} - \mathbf{y}\|^2 \quad \forall t \in I, \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n. \quad (9.19)$$

Alors, le problème à valeur initiale

$$\mathbf{u}'(t) = \mathbf{f}(t, \mathbf{u}(t)), \quad t \in I_+ = I \cap [t_0, +\infty[, \quad \mathbf{u}(t_0) = \mathbf{u}_0$$

a une solution globale unique $\mathbf{u} \in C^1(I_+, \mathbb{R}^n)$.

Démonstration. La condition (9.19) est plus forte que la condition (9.17) du théorème 9.40. En effet, si on prend $\mathbf{y} = \mathbf{0}$ en (9.19) on obtient pour tout $\mathbf{x} \in \mathbb{R}^n$

$$\begin{aligned} \mathbf{f}(t, \mathbf{x}) \cdot \mathbf{x} &\leq \mathbf{f}(t, \mathbf{0}) \cdot \mathbf{x} + \ell(t) \|\mathbf{x}\|^2 \\ &\leq \|\mathbf{f}(t, \mathbf{0})\| \|\mathbf{x}\| + |\ell(t)| \|\mathbf{x}\|^2 \\ &\leq \left(\frac{\|\mathbf{f}(t, \mathbf{0})\|}{2} + |\ell(t)| \right) (1 + \|\mathbf{x}\|^2). \end{aligned}$$

où dans la dernière inégalité on a utilisé que $\|\mathbf{x}\| \leq \frac{1}{2}(1 + \|\mathbf{x}\|^2)$. Donc le résultat suit par le théorème 9.40. □

Remarque 9.44. Si, de plus, ℓ est non négative et

$$\left| (\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \mathbf{y})) \cdot (\mathbf{x} - \mathbf{y}) \right| \leq \ell(t) \|\mathbf{x} - \mathbf{y}\|^2 \quad \forall t \in I, \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \quad (9.20)$$

alors le problème de Cauchy (9.13) admet une solution globale unique sur I . En effet, on vérifie de la même manière que (9.20) est une condition plus forte que (9.18). Comme être globalement Lipschitz par rapport au deuxième argument implique (9.20), ceci donne une nouvelle preuve du Théorème 9.37.

9.5 EDO scalaires linéaires du second ordre

Soit $I \subset \mathbb{R}$ un intervalle ouvert, $E \subset \mathbb{R}^2$ un ouvert et $f : I \times E \rightarrow \mathbb{R}$ une fonction continue. Une équation différentielle scalaire du second ordre est une équation du type

$$u''(t) = f(t, u(t), u'(t)), \quad t \in I,$$

dont l'inconnue est une fonction $u : I \rightarrow \mathbb{R}$ de classe C^2 telle que $(u(t), u'(t)) \in E, \forall t \in I$. Comme on l'a vu dans l'introduction du chapitre, une telle équation peut toujours être réécrite sous forme d'un système de deux équations différentielles ordinaires du premier ordre en introduisant les nouvelles inconnues $u_1(t) = u(t)$, $u_2(t) = u'(t)$. On a alors

$$\begin{cases} u_1'(t) = u_2(t), \\ u_2'(t) = f(t, u_1(t), u_2(t)), \end{cases} \quad t \in I.$$

Si on note $\mathbf{u} = (u_1, u_2)$ et $\mathbf{f}(t, \mathbf{y}) = (y_2, f(t, y_1, y_2)) : I \times E \rightarrow \mathbb{R}^2$, alors le système précédent peut s'écrire sous forme compacte

$$\mathbf{u}'(t) = \mathbf{f}(t, \mathbf{u}(t)), \quad t \in I$$

et peut être analysé du point de vue théorique utilisant les résultats de la section 9.4. Le problème de Cauchy correspondant s'écrit pour un $t_0 \in I$ et $\mathbf{u}_0 = (u_0, v_0) \in E$:

$$\begin{cases} \mathbf{u}'(t) = \mathbf{f}(t, \mathbf{u}(t)), & t \in I, \\ \mathbf{u}(t_0) = \mathbf{u}_0, \end{cases} \quad \Longleftrightarrow \quad \begin{cases} u''(t) = f(t, u(t), u'(t)), & t \in I, \\ u(t_0) = u_0, \quad u'(t_0) = v_0. \end{cases}$$

On voit donc que, pour une edo du second ordre, on a de façon naturelle *deux* conditions initiales, une sur la solution $u(t_0) = u_0$ et une sur la dérivée première $u'(t_0) = v_0$.

Dans le reste de la section, on restreint notre étude aux équations différentielle scalaires linéaires du second ordre, pour lesquelles la fonction $\mathbf{f}(t, \mathbf{y})$ est une fonction *affine* de \mathbf{y} . Soient $a, b, g : I \rightarrow \mathbb{R}$ trois fonctions continues. Une *équation différentielle linéaire du second ordre* est une équation de la forme

$$u''(t) + a(t)u'(t) + b(t)u(t) = g(t), \quad t \in I,$$

ou, sous forme de système de deux équations du premier ordre,

$$\mathbf{u}'(t) = \mathbf{f}(t, \mathbf{u}(t)), \quad t \in I, \quad \text{avec} \quad \mathbf{f}(t, \mathbf{y}) = \begin{pmatrix} y_2 \\ g(t) - b(t)y_1 - a(t)y_2 \end{pmatrix}.$$

Il est facile de montrer (exercice) que cette fonction $\mathbf{f} : I \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$ est toujours continue et globalement lipschitzienne par rapport au deuxième argument (voir définition 9.36). Donc, pour tout $(t_0, u_0, v_0) \in I \times E$, le problème de Cauchy avec condition initiale $u(t_0) = u_0$, $u'(t_0) = v_0$ a toujours une solution unique globale.

Pour la construction de l'intégrale générale et de la solution du problème de Cauchy, on procède comme dans la section 9.3 en analysant séparément le cas homogène ($g = 0$) et le cas non-homogène ($g \neq 0$).

9.5.1 Solution générale de l'équation homogène

Par les considérations précédentes, le problème de Cauchy associé à l'équation homogène

$$\begin{cases} u''(t) + a(t)u'(t) + b(t)u(t) = 0, & t \in I, \\ u(t_0) = u_0, \quad u'(t_0) = v_0, \end{cases} \quad (9.21)$$

admet une solution unique globale pour toute donnée initiale $\mathbf{u}_0 = (u_0, v_0)$. En particulier, si $\mathbf{u}_0 = (0, 0)$, la solution (unique) est identiquement nulle.

Définition 9.45. On dit que deux solutions $z_1, z_2 : I \rightarrow \mathbb{R}$ de l'équation homogène

$$u''(t) + a(t)u'(t) + b(t)u(t) = 0, \quad t \in I. \quad (9.22)$$

sont linéairement indépendantes si, pour tout choix des constantes $\alpha, \beta \in \mathbb{R}$, l'implication suivante est vraie :

$$\alpha z_1(t) + \beta z_2(t) = 0, \forall t \in I \quad \implies \quad \alpha = \beta = 0.$$

Inversement, on dit que z_1, z_2 sont linéairement dépendantes s'il existe deux constantes réelles α, β non simultanément nulles telles que $\alpha z_1(t) + \beta z_2(t) = 0, \forall t \in I$. Dans ce cas, on peut exprimer une de deux solutions en fonction de l'autre, i.e. $z_1(t) = -\frac{\beta}{\alpha} z_2(t)$ ou bien $z_2(t) = -\frac{\alpha}{\beta} z_1(t)$.

Définition 9.46. Soit z_1, z_2 deux solutions de l'équation homogène (9.22). On appelle wronskien de z_1, z_2 , noté $W[z_1, z_2]$ la fonction $W[z_1, z_2] : I \rightarrow \mathbb{R}$ définie par

$$W[z_1, z_2](t) = \det \begin{pmatrix} z_1(t) & z_2(t) \\ z_1'(t) & z_2'(t) \end{pmatrix} = z_1(t)z_2'(t) - z_2(t)z_1'(t), \quad t \in I.$$

Le théorème suivant montre que deux solutions z_1, z_2 de l'équation homogène sont linéairement indépendantes si et seulement si leur wronskien est non nul pour tout $t \in I$.

Théorème 9.47. Deux solutions z_1, z_2 de l'équation homogène sont linéairement indépendantes si et seulement si pour tout $t \in I$ les deux vecteurs $\mathbf{z}_1(t) = (z_1(t), z_1'(t))$ et $\mathbf{z}_2(t) = (z_2(t), z_2'(t))$ de \mathbb{R}^2 sont linéairement indépendants, ce qui équivaut à dire que

$$W[z_1, z_2](t) \neq 0, \quad \forall t \in I.$$

Démonstration.

« \Rightarrow » : Soient $z_1, z_2 : I \rightarrow \mathbb{R}$ linéairement indépendants. On veut montrer que pour tout $t \in I$ les deux vecteurs $\mathbf{z}_1(t), \mathbf{z}_2(t) \in \mathbb{R}^2$ sont linéairement indépendants. Par l'absurde supposons qu'il existe $t_0 \in I$ tel que $\mathbf{z}_1(t_0), \mathbf{z}_2(t_0)$ sont linéairement dépendants, c'est-à-dire $\exists \alpha, \beta$ non simultanément nuls tels que $\alpha \mathbf{z}_1(t_0) + \beta \mathbf{z}_2(t_0) = \mathbf{0}$, i.e.

$$\alpha z_1(t_0) + \beta z_2(t_0) = 0, \quad \alpha z_1'(t_0) + \beta z_2'(t_0) = 0.$$

Mais alors, la fonction $v(t) = \alpha z_1(t) + \beta z_2(t)$ satisfait le problème de Cauchy

$$\begin{cases} v''(t) + a(t)v'(t) + b(t)v(t) = 0, & t \in I, \\ v(t_0) = 0, v'(t_0) = 0, \end{cases}$$

dont on connaît l'unique solution globale : $v(t) = 0, \forall t \in I$. On obtient la contradiction que $\alpha z_1(t) + \beta z_2(t) = 0$ pour tout $t \in I$.

« \Leftarrow » : Soit $\mathbf{z}_1(t), \mathbf{z}_2(t)$ linéairement indépendants pour tout $t \in I$. On veut montrer que z_1, z_2 sont linéairement indépendants. Par l'absurde, si z_1, z_2 sont linéairement dépendants alors $\exists \alpha, \beta$ non simultanément nuls tels que $\alpha z_1(t) + \beta z_2(t) = 0, \forall t \in I$. Mais ceci implique $\alpha z_1'(t) + \beta z_2'(t) = 0, \forall t \in I$ et donc $\alpha \mathbf{z}_1(t) + \beta \mathbf{z}_2(t) = \mathbf{0}$, ce qui est une contradiction. \square

L'ensemble $S = \{z \in C^2(I) : z'' + az' + bz = 0 \text{ sur } I\}$ est un espace vectoriel : si $z_1, z_2 \in S$ et $\alpha, \beta \in \mathbb{R}$, alors $\alpha z_1 + \beta z_2 \in S$. Dans ce cadre, $z_1, z_2 \in S$ sont linéairement

indépendants s'il n'existe pas deux constantes réelles α, β non simultanément nulles telles que $\alpha z_1 + \beta z_2 = 0$ sur I , comme nous l'avons déjà mentionné. De plus, pour $t_0 \in I$ fixé, l'application $z \mapsto (z(t_0), z'(t_0)) \in \mathbb{R}^2$ pour $z \in S$ est linéaire. Cette application est aussi bijective car, pour chaque $(u_0, v_0) \in \mathbb{R}^2$, il existe une unique fonction $z \in S$ telle que $z(t_0) = u_0$ et $z'(t_0) = v_0$, comme nous l'avons vu. Il en résulte que S et \mathbb{R}^2 sont deux espaces vectoriels isomorphes, ce qui donne une nouvelle preuve du théorème précédent. En particulier ils ont la même dimension. Ainsi S est de dimension 2 et toute paire $z_1, z_2 \in S$ linéairement indépendante en est une base : $S = \{C_1 z_1 + C_2 z_2 : C_1, C_2 \in \mathbb{R}\}$.

Ces considérations permettent de conclure que la **solution générale de l'équation homogène** a la forme

$$u(t) = C_1 z_1(t) + C_2 z_2(t), \quad t \in I, \quad (9.23)$$

où z_1, z_2 sont deux solutions linéairement indépendantes de l'équation homogène et $C_1, C_2 \in \mathbb{R}$ deux constantes arbitraires. Si on souhaite résoudre le problème de Cauchy (9.21) il faudra encore trouver les bonnes valeurs des constantes pour satisfaire la condition initiale $u(t_0) = u_0, u'(t_0) = v_0$.

La méthode de variation des constantes, déjà présentée dans la section 9.3, permet, à partir d'une solution de l'équation homogène, d'en construire une deuxième linéairement indépendante. Étant donnée une solution non identiquement nulle $z_1 : I \rightarrow \mathbb{R}$ de l'équation homogène, l'idée est d'en chercher une deuxième sous la forme $z_2(t) = C(t)z_1(t)$, où $C : I \rightarrow \mathbb{R}$ est une fonction non constante de classe C^2 . Les calculs sont laissés comme exercice.

9.5.2 Solution générale de l'équation non-homogène

Considérons maintenant l'équation non-homogène

$$u''(t) + a(t)u'(t) + b(t)u(t) = g(t), \quad t \in I.$$

En utilisant le principe de superposition de solutions (voir section 9.3) la **solution générale de l'équation non-homogène** peut s'écrire comme

$$u(t) = w(t) + C_1 z_1(t) + C_2 z_2(t), \quad t \in I,$$

où w est une *solution particulière* de l'équation non-homogène et $C_1 z_1(t) + C_2 z_2(t)$ est la solution générale de l'équation homogène, avec z_1, z_2 linéairement indépendantes.

De nouveau, on peut utiliser la *méthode de variation des constantes* pour construire une solution particulière de l'équation non-homogène à partir des solutions de l'équation homogène. Étant donné deux solutions z_1, z_2 linéairement indépendantes de l'équation homogène, on peut toujours trouver une solution particulière de l'équation non-homogène de la forme

$$w(t) = C_1(t)z_1(t) + C_2(t)z_2(t) = \sum_{i=1}^2 C_i(t)z_i(t)$$

où $C_1, C_2 : I \rightarrow \mathbb{R}$ sont deux fonctions de classe C^2 à déterminer. On a

$$w' = \sum_{i=1}^2 (C_i' z_i + C_i z_i'), \quad w'' = \sum_{i=1}^2 (C_i'' z_i + 2C_i' z_i' + C_i z_i'').$$

Donc

$$\begin{aligned}
 w'' + aw' + bw &= \sum_{i=1}^2 (C_i'' z_i + 2C_i' z_i' + \cancel{C_i z_i''} + aC_i' z_i + \cancel{aC_i z_i'} + b\cancel{C_i z_i}) \\
 &= \sum_{i=1}^2 (C_i' z_i' + aC_i' z_i + \frac{d}{dt}(C_i' z_i)) \\
 &= C_1' z_1' + C_2' z_2' + a \left(\sum_{i=1}^2 C_i' z_i \right) + \frac{d}{dt} \left(\sum_{i=1}^2 C_i' z_i \right) = g.
 \end{aligned}$$

C'est une seule équation pour les deux fonctions inconnues C_1 et C_2 . On peut donc en principe imposer une relation supplémentaire entre C_1 et C_2 . L'équation précédente se simplifie beaucoup si on impose la condition $\sum_{i=1}^2 C_i' z_i = 0$. On obtient alors le système de deux équations différentielles

$$\begin{cases} C_1' z_1 + C_2' z_2 = 0 \\ C_1' z_1' + C_2' z_2' = g \end{cases}$$

qui peut être écrit sous forme matricielle comme $\begin{bmatrix} z_1 & z_2 \\ z_1' & z_2' \end{bmatrix} \begin{bmatrix} C_1' \\ C_2' \end{bmatrix} = \begin{bmatrix} 0 \\ g \end{bmatrix}$ ce qui implique

$$\begin{bmatrix} C_1' \\ C_2' \end{bmatrix} = \begin{bmatrix} z_1 & z_2 \\ z_1' & z_2' \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ g \end{bmatrix} = \frac{1}{W[z_1, z_2]} \begin{bmatrix} z_2' & -z_2 \\ -z_1' & z_1 \end{bmatrix} \begin{bmatrix} 0 \\ g \end{bmatrix}$$

et donc $C_1' = -\frac{z_2 g}{W[z_1, z_2]}$ et $C_2' = \frac{z_1 g}{W[z_1, z_2]}$. Par intégration on obtient finalement

$$\begin{cases} C_1(t) = \int_{t_0}^t \frac{-z_2(s)g(s)}{W[z_1, z_2](s)} ds + \kappa_1, \\ C_2(t) = \int_{t_0}^t \frac{z_1(s)g(s)}{W[z_1, z_2](s)} ds + \kappa_2, \end{cases}$$

avec $\kappa_1, \kappa_2 \in \mathbb{R}$ et $t_0 \in I$ fixés librement. Par conséquent, on peut choisir comme solution particulière

$$w(t) = \int_{t_0}^t K(t, s) g(s) ds, \quad K(t, s) = \frac{z_1(s)z_2(t) - z_1(t)z_2(s)}{W[z_1, z_2](s)}. \quad (9.24)$$

La fonction $K : I \times I \rightarrow \mathbb{R}$ a l'interprétation suivante : pour chaque $s \in I$ fixé, la fonction $t \mapsto K(t, s) := K_s(t)$ est solution du problème homogène associé :

$$K_s''(t) + a(t)K_s'(t) + b(t)K_s(t) = 0, \quad t \in I,$$

sous les conditions

$$K_s(s) = K(s, s) = 0, \quad K_s'(s) = \frac{\partial K}{\partial t}(t, s)|_{t=s} = 1.$$

9.5.3 EDO linéaires du second ordre à coefficients constants

On s'intéresse ici à des équations de la forme :

$$u''(t) + au'(t) + bu(t) = g(t), \quad t \in I,$$

où $a, b \in \mathbb{R}$ sont des coefficients constants (ne dépendent pas de t).

Équation homogène : $u''(t) + au'(t) + bu(t) = 0$, $t \in I$.

On cherche des solutions exponentielles $z(t) = e^{\lambda t}$. Remplaçant dans l'équation on obtient $\lambda^2 e^{\lambda t} + a\lambda e^{\lambda t} + be^{\lambda t} = 0$ qui donne l'**équation caractéristique**

$$\lambda^2 + a\lambda + b = 0.$$

On distingue les trois cas correspondants au signe du discriminant $\Delta := a^2 - 4b$ de l'équation caractéristique.

Cas $\Delta \geq 0$. On a deux solutions réelles distinctes :

$$\lambda_1 = \frac{-a + \sqrt{a^2 - 4b}}{2}, \quad \lambda_2 = \frac{-a - \sqrt{a^2 - 4b}}{2}.$$

On vérifie facilement que $z_1(t) = e^{\lambda_1 t}$ et $z_2(t) = e^{\lambda_2 t}$ sont deux solutions linéairement indépendantes. La solution générale de l'équation homogène est donc

$$u(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}$$

avec $C_1, C_2 \in \mathbb{R}$ constantes arbitraires.

Cas $\Delta < 0$. On a deux solutions complexes conjuguées :

$$\lambda_1 = \frac{-a + i\sqrt{4b - a^2}}{2}, \quad \lambda_2 = \frac{-a - i\sqrt{4b - a^2}}{2},$$

ce qui conduit à considérer les fonctions à valeurs complexes $e^{\lambda_1 t}$ et $e^{\lambda_2 t}$, c'est-à-dire, $\tilde{z}_1(t) = e^{-\frac{a}{2}t} e^{i\mu t}$ et $\tilde{z}_2(t) = e^{-\frac{a}{2}t} e^{-i\mu t}$ avec $\mu = \frac{\sqrt{4b - a^2}}{2}$. Puisque l'on cherche des solutions réelles, on pose

$$\begin{aligned} z_1(t) &= \frac{\tilde{z}_1(t) + \tilde{z}_2(t)}{2} = e^{-\frac{a}{2}t} \cos(\mu t) \\ z_2(t) &= \frac{\tilde{z}_1(t) - \tilde{z}_2(t)}{2i} = e^{-\frac{a}{2}t} \sin(\mu t) \end{aligned}$$

qui sont effectivement des solutions linéairement indépendantes. La solution générale de l'équation homogène est donc

$$u(t) = e^{-\frac{a}{2}t} (C_1 \cos(\mu t) + C_2 \sin(\mu t))$$

avec $C_1, C_2 \in \mathbb{R}$ constantes arbitraires.

Cas $\Delta = 0$. On a ici une seule solution réelle caractérisée par $\lambda = -\frac{a}{2}$ ce qui donne la solution $z_1(t) = e^{-\frac{a}{2}t}$ de l'équation homogène. On cherche une deuxième solution sous la forme $z_2(t) = C(t)z_1(t)$. On trouve $z_2(t) = te^{-\frac{a}{2}t}$ (vérifiez-le comme exercice) qui est linéairement indépendante de $z_1(t)$. La solution générale de l'équation homogène est donc

$$u(t) = (C_1 + C_2t)e^{-\frac{a}{2}t}$$

avec $C_1, C_2 \in \mathbb{R}$ constantes arbitraires.

Équation non-homogène : $u''(t) + au'(t) + bu(t) = g(t), \forall t \in I$.

Pour trouver une solution particulière de l'équation non-homogène, on peut toujours utiliser la formule générale (9.24). Toutefois, si la fonction g a une forme particulière (polynomiale, exponentielle, etc.) on peut souvent chercher la solution particulière dans une famille appropriée de fonctions paramétrées. Voici quelques exemples.

g est un polynôme de degré n : $g(t) = \sum_{j=0}^n a_j t^j$.

Si $b \neq 0$, on cherche w sous la même forme : $w(t) = \sum_{j=0}^n \beta_j t^j$.

Si $b = 0$ et $a \neq 0$, on cherche w sous la forme : $w(t) = t \sum_{j=0}^n \beta_j t^j$ (remarquez que dans ce cas la fonction $t \mapsto 1$ est solution de l'équation homogène associée, mais pas la fonction $t \mapsto t$).

Si $b = 0$ et $a = 0$, on cherche w sous la forme : $w(t) = t^2 \sum_{j=0}^n \beta_j t^j$ (remarquez que dans ce cas les fonctions $t \mapsto 1$ et $t \mapsto t$ sont solutions de l'équation homogène associée). En fait, dans ce cas, on peut résoudre l'équation $u''(t) = g(t)$ directement par une double intégration.

g est un polynôme multiplié par des fonctions trigonométriques-exponentielles :

$$g(t) = \left(\sum_{j=0}^n a_j t^j \right) e^{\delta t} \sin(\omega t) + \left(\sum_{j=0}^n \tilde{a}_j t^j \right) e^{\delta t} \cos(\omega t), \quad \omega \neq 0, \delta \in \mathbb{R}.$$

Si la fonction $t \mapsto e^{\delta t} \cos(\omega t)$ n'est pas solution du problème homogène associé (et donc $t \mapsto e^{\delta t} \sin(\omega t)$ non plus), on cherche $w(t)$ sous la même forme :

$$w(t) = \left(\sum_{j=0}^n \beta_j t^j \right) e^{\delta t} \sin(\omega t) + \left(\sum_{j=0}^n \tilde{\beta}_j t^j \right) e^{\delta t} \cos(\omega t).$$

Si la fonction $t \mapsto e^{\delta t} \cos(\omega t)$ est solution du problème homogène associé, on cherche $w(t)$ sous la forme

$$w(t) = t \left(\sum_{j=0}^n \beta_j t^j \right) e^{\delta t} \sin(\omega t) + t \left(\sum_{j=0}^n \tilde{\beta}_j t^j \right) e^{\delta t} \cos(\omega t).$$

Plus explicitement, la fonction $t \mapsto e^{\delta t} \cos(\omega t)$ est solution du problème homogène associé si et seulement si à la fois $\delta = -a/2$, $4b - a^2 > 0$ et $\omega = \pm \sqrt{4b - a^2}/2$.

g est un polynôme multiplié par une fonction exponentielle : $g(t) = \left(\sum_{j=0}^n a_j t^j\right) e^{\delta t}$.

Si $\delta = 0$, la discussion qui suit redonne celle que l'on a déjà vu pour g un polynôme.

Si la fonction $t \mapsto e^{\delta t}$ n'est pas solution du problème homogène associé, on cherche $w(t)$ sous la même forme :

$$w(t) = \left(\sum_{j=0}^n \beta_j t^j\right) e^{\delta t}.$$

On est dans ce cas exactement lorsque, à la fois, $\delta \neq \lambda_1$ et $\delta \neq \lambda_2$.

Si la fonction $t \mapsto e^{\delta t}$ est solution du problème homogène associé, mais pas la fonction $t \mapsto te^{\delta t}$, on cherche $w(t)$ sous la forme

$$w(t) = t \left(\sum_{j=0}^n \beta_j t^j\right) e^{\delta t}.$$

On est dans ce cas exactement lorsque à la fois $\lambda_1 \neq \lambda_2$ et $\delta \in \{\lambda_1, \lambda_2\}$.

Si les fonction $t \mapsto e^{\delta t}$ et $t \mapsto te^{\delta t}$ sont solutions du problème homogène associé, on cherche $w(t)$ sous la forme

$$w(t) = t^2 \left(\sum_{j=0}^n \beta_j t^j\right) e^{\delta t}.$$

On est dans ce cas exactement lorsque $\lambda_1 = \lambda_2 = \delta$.