

**Fonctions de plusieurs variables,
analyse vectorielle,
intégrales multiples**

2M216

Jean-François Babadjian

Table des matières

1	Notions de topologie dans \mathbb{R}^n	5
1.1	Propriétés algébriques de \mathbb{R}^n	5
1.2	Normes	6
1.3	Suites	7
1.4	Topologie	8
1.4.1	Ensembles ouverts, fermés	8
1.4.2	Ensembles compacts	11
1.4.3	Ensembles connexes et convexes	12
2	Fonctions continues	13
2.1	Définitions, propriétés fondamentales	13
2.2	Exemples importants	15
2.2.1	Fonctions Lipschitziennes	15
2.2.2	Applications linéaires	16
2.2.3	Polynômes	17
2.2.4	Applications partielles	17
2.3	Continuité et compacité	18
3	Différentiabilité et dérivées partielles	21
3.1	Définitions	21
3.2	Propriétés	25
3.3	Fonctions de classe \mathcal{C}^1	29
3.4	Dérivées partielles d'ordre deux	31
3.5	Points critiques et extrema	34
4	Courbes et surfaces paramétrées	39
4.1	Théorème des fonctions implicites	39
4.2	Courbes paramétrées	40
4.3	Surfaces paramétrées	42
5	Equations aux dérivées partielles	47
5.1	Equations elliptiques	47
5.2	Equations paraboliques	47
5.3	Equations hyperboliques	48
5.3.1	Equation de transport	48
5.3.2	Equation des ondes	49

6	Intégrales multiples	51
6.1	Pavés, ensembles pavables et quarrables	51
6.2	Définition et propriétés de l'intégrale multiple	53
6.3	Théorème de Fubini	58
6.4	Formule de changement de variables	62
6.4.1	Coordonnées polaires	62
6.4.2	Coordonnées cylindriques	63
6.4.3	Coordonnées sphériques	63
7	Intégrales curvilignes et surfaciques	65
7.1	Intégrale curviligne et formules de Stokes-Ostrogradski	65
7.1.1	Longueur d'un arc orienté	65
7.1.2	Intégrale curviligne	67
7.1.3	Formules de Stokes-Ostrogradski	68
7.2	Intégrale surfacique et formule de la divergence	69
7.2.1	Aire d'une surface	69
7.2.2	Intégrale de surface	70
7.2.3	Formule de la divergence	71

Chapitre 1

Notions de topologie dans \mathbb{R}^n

1.1 Propriétés algébriques de \mathbb{R}^n

L'ensemble \mathbb{R}^n est un espace vectoriel sur \mathbb{R} de dimension finie égale à $n \in \mathbb{N}^*$. Ses éléments sont appelés des vecteurs. En tant qu'espace vectoriel de dimension n , il existe une base $\{e^1, \dots, e^n\}$, *i.e.* une famille libre et génératrice, de sorte que tout vecteur $x \in \mathbb{R}^n$ peut s'écrire

$$x = \sum_{i=1}^n x_i e^i,$$

où $x_1, \dots, x_n \in \mathbb{R}$ désignent les coordonnées de x dans la base $\{e^1, \dots, e^n\}$. Une fois fixée la base (la plupart du temps, on utilisera la base canonique définie par $e_j^i = \delta_{ij}$ pour tout $1 \leq i, j \leq n$), on identifiera le vecteur x avec la matrice colonne de taille $n \times 1$

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

et on écrira

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (x_1, \dots, x_n)^T.$$

AVERTISSEMENT : Le fait de représenter un vecteur de \mathbb{R}^n sous forme d'un vecteur colonne sera essentiel lorsque l'on devra effectuer des produits matrices/vecteurs. Dans les autres situations, la représentation d'un vecteur sous forme d'une colonne ou d'une ligne ne sera pas particulièrement importante et pour cette raison, nous ferons parfois des abus de notations.

En tant qu'espace vectoriel, \mathbb{R}^n possède

– une *loi interne* (l'addition) : si $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ et $y = (y_1, \dots, y_n)^T \in \mathbb{R}^n$, alors

$$x + y := (x_1 + y_1, \dots, x_n + y_n)^T \in \mathbb{R}^n;$$

– une *loi externe* (la multiplication par un réel) : si $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$, alors

$$\lambda x = (\lambda x_1, \dots, \lambda x_n)^T \in \mathbb{R}^n.$$

1.2 Normes

L'ensemble \mathbb{R}^n est un espace euclidien ce qui signifie qu'il possède un *produit scalaire* : si $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ et $y = (y_1, \dots, y_n)^T \in \mathbb{R}^n$, le produit scalaire euclidien entre x et y est défini par

$$x \cdot y := \sum_{i=1}^n x_i y_i,$$

et la *norme euclidienne* qui en est héritée est donnée par

$$\|x\| := \sqrt{x \cdot x} = \left(\sum_{i=1}^n x_i^2 \right)^{1/2}.$$

Notons que si $x \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$, alors

$$\|\lambda x\| = \left(\sum_{i=1}^n (\lambda x_i)^2 \right)^{1/2} = \sqrt{\lambda^2} \left(\sum_{i=1}^n x_i^2 \right)^{1/2} = |\lambda| \|x\|. \quad (1.2.1)$$

Par ailleurs, on a clairement que $\|0\| = 0$ et réciproquement, si $\|x\| = 0$ alors $x_i = 0$ pour tout $i = 1, \dots, n$ ce qui montre que

$$\|x\| = 0 \quad \text{si et seulement si} \quad x = 0. \quad (1.2.2)$$

Montrons deux inégalités fondamentales dont l'usage sera systématique dans la suite de ce cours.

Théorème 1.2.1 (Inégalité de Cauchy-Schwarz). *Pour tout x et $y \in \mathbb{R}^n$, on a*

$$|x \cdot y| \leq \|x\| \|y\|.$$

L'égalité $|x \cdot y| = \|x\| \|y\|$ a lieu si et seulement si $x = \lambda y$ pour un $\lambda \in \mathbb{R}$.

Démonstration. Pour tout $t \in \mathbb{R}$, on définit

$$P(t) := \|x + ty\|^2 = \sum_{i=1}^n (x_i + ty_i)^2 = \sum_{i=1}^n (x_i^2 + 2tx_i y_i + t^2 y_i^2) = \|x\|^2 + 2tx \cdot y + t^2 \|y\|^2.$$

La fonction P est donc un trinôme du second degré et son discriminant est donné par $\Delta = 4(x \cdot y)^2 - 4\|x\|^2 \|y\|^2$. Comme $P(t) \geq 0$ pour tout $t \in \mathbb{R}$, alors nécessairement $\Delta \leq 0$ car sinon le polynôme P posséderait deux racines réelles distinctes et il changerait de signe. La condition sur le discriminant implique donc $(x \cdot y)^2 \leq \|x\|^2 \|y\|^2$, soit $|x \cdot y| \leq \|x\| \|y\|$.

En ce qui concerne le cas d'égalité, on constate que $\Delta = 0$ si et seulement s'il existe un unique $t_0 \in \mathbb{R}$ tel que $P(t_0) = 0$, soit $\|x + t_0 y\|^2 = 0$, ce qui est encore équivalent au fait que $x = -t_0 y$. \square

Théorème 1.2.2 (Inégalité triangulaire). *Pour tout x et $y \in \mathbb{R}^n$, on a*

$$\|x + y\| \leq \|x\| + \|y\|.$$

Démonstration. On développe le carré et on utilise l'inégalité de Cauchy-Schwarz,

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

ce qui établit l'inégalité souhaitée par passage à la racine carrée. \square

De façon générale, on a la

Définition 1.2.3. On appelle norme sur \mathbb{R}^n une application $N : \mathbb{R}^n \rightarrow [0, +\infty[$ telle que

- i) (*Séparation*) $N(x) = 0$ si et seulement si $x = 0$;
- ii) (*Homogénéité*) $N(\lambda x) = |\lambda|N(x)$ pour tout $x \in \mathbb{R}^n$ et tout $\lambda \in \mathbb{R}$;
- iii) (*Inégalité triangulaire*) $N(x + y) \leq N(x) + N(y)$ pour tout x et $y \in \mathbb{R}^n$.

La conjonction de (1.2.1), (1.2.2) et de l'inégalité triangulaire montre bien que la norme euclidienne est effectivement une norme au sens de la définition 1.2.3. De façon générale, on considère les quantités

$$\|x\|_p := \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, \quad \text{pour tout } 1 \leq p < \infty, \quad \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

On peut alors montrer que les applications $\|\cdot\|_p$ ($1 \leq p \leq \infty$) définissent des normes sur \mathbb{R}^n . Notons au passage que la norme euclidienne correspond à $p = 2$.

La définition suivante fournit un critère de comparaison entre deux normes distinctes.

Définition 1.2.4. Soient N_1 et N_2 deux normes sur \mathbb{R}^n . On dit que N_1 et N_2 sont *équivalentes* s'il existe deux constantes $\alpha > 0$ et $\beta > 0$ telles que

$$\alpha N_1(x) \leq N_2(x) \leq \beta N_1(x), \quad \text{pour tout } x \in \mathbb{R}^n.$$

Cette propriété signifie intuitivement qu'un vecteur "petit" selon la norme N_1 le sera aussi selon l'autre norme N_2 . On peut vérifier à titre d'exercice (en utilisant l'inégalité de Cauchy-Schwarz) que pour tout $x \in \mathbb{R}^n$,

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2$$

et

$$\frac{1}{\sqrt{n}}\|x\|_2 \leq \|x\|_\infty \leq \|x\|_2.$$

En fait, la propriété suivante, que nous démontrerons à la fin du chapitre 2 nous permet de travailler indifféremment avec une norme ou une autre. Sauf mention du contraire, nous utiliserons systématiquement la norme euclidienne.

Théorème 1.2.5. *Toutes les normes sur \mathbb{R}^n sont équivalentes entre elles.*

1.3 Suites

Définition 1.3.1. Soit $(x^k)_{k \in \mathbb{N}}$ une suite de \mathbb{R}^n . On dit que $(x^k)_{k \in \mathbb{N}}$ converge vers $a \in \mathbb{R}^n$, et on note $x^k \rightarrow a$, si pour tout $\varepsilon > 0$ il existe un $N \in \mathbb{N}$ tel que pour tout $k \geq N$, on a $\|x^k - a\| < \varepsilon$.

AVERTISSEMENT : Si $(x^k)_{k \in \mathbb{N}}$ est une suite de \mathbb{R}^n , l'indice en haut désigne l'indice de la suite. En revanche, l'indice en bas désigne la composante du vecteur. Ainsi, x_i^k est la i -ème composante du vecteur x^k .

Proposition 1.3.2. *La limite, si elle existe, est unique.*

Démonstration. Soient a et $b \in \mathbb{R}^n$ deux limites de la suite $(x^k)_{k \in \mathbb{N}}$. Alors, pour tout $\varepsilon > 0$ il existe un $N_a \in \mathbb{N}$ tel que pour tout $k \geq N_a$, on a $\|x^k - a\| < \varepsilon$. De même, il existe un $N_b \in \mathbb{N}$ tel que pour tout $k \geq N_b$, on a $\|x^k - b\| < \varepsilon$. Notons $N := \max(N_a, N_b)$, en vertu de l'inégalité triangulaire, pour tout $k \geq N$, on a

$$\|a - b\| \leq \|a - x^k\| + \|x^k - b\| < 2\varepsilon,$$

ce qui montre que $a = b$, puisque ε est arbitraire. \square

Le résultat suivant montre que la convergence d'une suite de \mathbb{R}^n est équivalente à la convergence de chacune de ses composantes dans \mathbb{R} .

Proposition 1.3.3. *Soit $(x^k)_{k \in \mathbb{N}}$ une suite de \mathbb{R}^n et $a \in \mathbb{R}^n$. Alors*

$$x^k \rightarrow a \iff x_i^k \rightarrow a_i \quad \text{pour tout } 1 \leq i \leq n.$$

Démonstration. Si $x^k \rightarrow a$, alors pour tout $\varepsilon > 0$, il existe un $N \in \mathbb{N}$ tel que pour tout $k \geq N$, on a $\|x^k - a\| < \varepsilon$. Comme $|x_i^k - a_i| \leq \|x^k - a\|$ pour tout $1 \leq i \leq n$, on en déduit que $|x_i^k - a_i| < \varepsilon$ pour tout $k \geq N$ et tout $1 \leq i \leq n$, ce qui montre que $x_i^k \rightarrow a_i$ dans \mathbb{R} .

Réciproquement, supposons que pour tout $1 \leq i \leq n$, on a $x_i^k \rightarrow a_i$. Alors pour tout $\varepsilon > 0$, il existe un $N_i \in \mathbb{N}$ tel que pour tout $k \geq N_i$, on a $|x_i^k - a_i| < \varepsilon/\sqrt{n}$. Posons $N = \max(N_1, \dots, N_n)$ de sorte que pour tout $k \geq N$, on a $|x_i^k - a_i| < \varepsilon/\sqrt{n}$. En élevant cette inégalité au carré, puis en sommant par rapport à i , on obtient que pour tout $k \geq N$,

$$\|x^k - a\| = \left(\sum_{i=1}^n |x_i^k - a_i|^2 \right)^{1/2} < \varepsilon,$$

ce qui montre bien que $x^k \rightarrow a$. □

Tout comme dans \mathbb{R} , l'inconvénient majeur est que pour savoir qu'une suite converge, il est nécessaire de calculer et donc de connaître sa limite. Nous définissons maintenant un critère qui s'avère extrêmement pratique car il nous permet de nous affranchir de la connaissance *a priori* de la limite d'une suite pour savoir si celle-ci est convergente.

Définition 1.3.4. Une suite $(x^k)_{k \in \mathbb{N}}$ de \mathbb{R}^n est dite de *Cauchy* si pour tout $\varepsilon > 0$, il existe un $N \in \mathbb{N}$ tel que pour tout $k, l \geq N$, on a $\|x^k - x^l\| < \varepsilon$.

Proposition 1.3.5. *Une suite de \mathbb{R}^n converge si et seulement si elle est de Cauchy. On dit alors que \mathbb{R}^n est complet.*

Démonstration. Soit $(x^k)_{k \in \mathbb{N}}$ une suite de \mathbb{R}^n telle que $x^k \rightarrow a$. Alors, pour tout $\varepsilon > 0$, il existe un $N \in \mathbb{N}$ tel que pour tout $k \geq N$, $\|x^k - a\| < \varepsilon/2$. Si k et $l \geq N$, d'après l'inégalité triangulaire, on a

$$\|x^k - x^l\| \leq \|x^k - a\| + \|a - x^l\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

ce qui montre bien que $(x^k)_{k \in \mathbb{N}}$ est de Cauchy.

Réciproquement, supposons que $(x^k)_{k \in \mathbb{N}}$ est une suite de Cauchy dans \mathbb{R}^n . Pour tout $\varepsilon > 0$, il existe un $N \in \mathbb{N}$ tel que pour tout $k, l \geq N$, on a $\|x^k - x^l\| < \varepsilon$. Comme $|x_i^k - x_i^l| \leq \|x^k - x^l\|$ pour tout $1 \leq i \leq n$, on en déduit que $|x_i^k - x_i^l| < \varepsilon$ pour tout $k, l \geq N$ et tout $1 \leq i \leq n$. Par conséquent, la suite numérique $(x_i^k)_{k \in \mathbb{N}}$ est de Cauchy dans \mathbb{R} qui est lui même un espace complet. Pour tout $1 \leq i \leq n$, il existe donc un $a_i \in \mathbb{R}$ tel que $x_i^k \rightarrow a_i$ dans \mathbb{R} , et la proposition 1.3.3 permet de conclure que $x^k \rightarrow a = (a_1, \dots, a_n)^T$ dans \mathbb{R}^n . □

1.4 Topologie

1.4.1 Ensembles ouverts, fermés

La notion de boules dans \mathbb{R}^n remplace celle d'intervalles dans \mathbb{R} .

Définition 1.4.1. Soient N une norme sur \mathbb{R}^n , $x \in \mathbb{R}$ et $r > 0$. On définit la *boule ouverte* de centre x et rayon r par

$$B_N(x, r) = \{y \in \mathbb{R}^n : N(y - x) < r\}$$

et la *boule fermée* de centre x et rayon r par

$$\overline{B}_N(x, r) = \{y \in \mathbb{R}^n : N(y - x) \leq r\}.$$

Dans le cas où N est la norme euclidienne, on notera simplement $B(x, r)$ et $\overline{B}(x, r)$ les boules ouvertes et fermées.

Définition 1.4.2. Un ensemble $U \subset \mathbb{R}^n$ est dit *ouvert* si pour tout $x \in U$, il existe un $r > 0$ tel que $B(x, r) \subset U$. Par convention, l'ensemble vide \emptyset est ouvert.

Définition 1.4.3. Un ensemble $F \subset \mathbb{R}^n$ est dit *fermé* si son complémentaire ${}^c F$ est ouvert.

Remarque 1.4.4. Evidemment, l'ensemble tout entier \mathbb{R}^n est ouvert, ce qui montre que \mathbb{R}^n et \emptyset sont tout deux à la fois ouverts et fermés.

Proposition 1.4.5. 1. *Toute union d'ouverts est ouverte ;*

2. *Toute intersection finie d'ouverts est ouverte ;*

3. *Toute intersection de fermés est fermée ;*

4. *Toute union finie de fermés est fermée.*

Démonstration. Soit $(U_i)_{i \in I}$ une famille d'ensembles ouverts. Montrons que $V := \bigcup_{i \in I} U_i$ est ouvert. Pour ce faire, on considère un $x \in V$. Par définition de l'union, il existe un $i_0 \in I$ tel que $x \in U_{i_0}$. L'ensemble U_{i_0} étant lui même ouvert, il existe un $r > 0$ tel que $B(x, r) \subset U_{i_0}$ et comme $U_{i_0} \subset V$, on en déduit que $B(x, r) \subset V$, ce qui montre bien que V est ouvert.

Soient U_1, \dots, U_p des ouverts de \mathbb{R}^n montrons que $W := \bigcap_{i=1}^p U_i$ est ouvert. Soit $x \in W$, alors par définition de l'intersection $x \in U_i$ pour tout $1 \leq i \leq p$. Les ensembles U_1, \dots, U_p étant ouverts, pour tout $1 \leq i \leq p$, il existe un $r_i > 0$ tel que $B(x, r_i) \subset U_i$. Notons $r = \min(r_1, \dots, r_p)$ de sorte que $r \leq r_i$ pour tout $1 \leq i \leq p$ et donc $B(x, r) \subset B(x, r_i) \subset U_i$ pour tout $1 \leq i \leq p$. Par conséquent, $B(x, r) \subset U_i$ pour tout $1 \leq i \leq p$ et donc $B(x, r) \subset W$ ce qui montre que W est ouvert.

Pour les propriétés concernant les fermés, il suffit de remarquer que si $(F_i)_{i \in I}$ est une famille d'ensembles fermés, alors

$${}^c \left(\bigcap_{i \in I} F_i \right) = \bigcup_{i \in I} {}^c F_i$$

qui est ouvert d'après 1., et donc $\bigcap_{i \in I} F_i$ est fermé.

De même, si F_1, \dots, F_p sont des fermés de \mathbb{R}^n , alors

$${}^c \left(\bigcup_{i=1}^p F_i \right) = \bigcap_{i=1}^p {}^c F_i$$

qui est ouvert d'après 2., et donc $\bigcup_{i=1}^p F_i$ est fermé. □

Définition 1.4.6. Soit E un sous-ensemble quelconque de \mathbb{R}^n . On dit que x est un *point intérieur* à E s'il existe un $r > 0$ tel que $B(x, r) \subset E$. L'ensemble des points intérieurs à E est noté $\overset{\circ}{E}$ et est appelé *intérieur* de E .

Proposition 1.4.7. *Soit E un sous-ensemble quelconque de \mathbb{R}^n , alors $\overset{\circ}{E} \subset E$. De plus*

1. *$\overset{\circ}{E}$ est le plus grand ouvert inclus dans E ;*

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

Démonstration. Par définition, si x est un point intérieur à E , alors $x \in E$ ce qui montre que $\overset{\circ}{E} \subset E$

Étape 1 : $\overset{\circ}{E}$ est un ouvert. Si $x \in \overset{\circ}{E}$, il existe un $r > 0$ tel que $B(x, r) \subset E$. Soient $y \in B(x, r)$ et $z \in B(y, r - \|x - y\|)$. Alors par l'inégalité triangulaire, $\|z - x\| \leq \|z - y\| + \|y - x\| < (r - \|x - y\|) + \|y - x\| = r$, ce qui montre que $B(y, r - \|x - y\|) \subset B(x, r) \subset E$. Par conséquent, $y \in \overset{\circ}{E}$ pour tout $y \in B(x, r)$ et donc $B(x, r) \subset \overset{\circ}{E}$, ce qui montre $\overset{\circ}{E}$ est ouvert.

Étape 2 : $\overset{\circ}{E}$ est le plus grand ouvert inclus dans E . Soit U un ouvert contenu dans E . Si $x \in U$, il existe un $r > 0$ tel que $B(x, r) \subset U \subset E$ ce qui montre que x est un point intérieur de E et que $U \subset \overset{\circ}{E}$.

Étape 3 : E est ouvert si et seulement si $E = \overset{\circ}{E}$. On a déjà vu que $\overset{\circ}{E} \subset E$, il s'agit alors de montrer l'inclusion opposée. Si $x \in E$ et E est ouvert, alors il existe un $r > 0$ tel que $B(x, r) \subset E$ ce qui montre que $x \in \overset{\circ}{E}$ et donc $E \subset \overset{\circ}{E}$. \square

Définition 1.4.8. Soit E un sous-ensemble quelconque de \mathbb{R}^n . On dit que x est un *point adhérent* à E s'il existe une suite $(x^k)_{k \in \mathbb{N}}$ d'éléments de E telle que $x^k \rightarrow x$. L'ensemble des points adhérents de E est noté \overline{E} et est appelé *adhérence* ou *fermeture* de E .

Proposition 1.4.9. Soit E un sous-ensemble quelconque de \mathbb{R}^n , alors $E \subset \overline{E}$. De plus

1. \overline{E} est le plus petit fermé contenant E ;
2. E est fermé si et seulement si $E = \overline{E}$.

Démonstration. Si $x \in E$, on prend la suite stationnaire $x^k = x$ pour tout $k \in \mathbb{N}$ qui converge évidemment vers x , ce qui montre que $x \in \overline{E}$ et donc que $E \subset \overline{E}$.

Étape 1 : On montre que $x \in \overline{E} \iff B(x, r) \cap E \neq \emptyset$ pour tout $r > 0$.

\implies : Si $x \in \overline{E}$, il existe une suite $(x^k)_{k \in \mathbb{N}}$ dans E telle que $x^k \rightarrow x$. Donc pour tout $r > 0$, il existe un $N \in \mathbb{N}$ tel que pour tout $k \in \mathbb{N}$, on a $\|x^k - x\| < r$, autrement dit $x^k \in B(x, r)$. Ceci montre que $B(x, r) \cap E \neq \emptyset$ pour tout $r > 0$ puisque $x^k \in B(x, r) \cap E$ dès que $k \geq N$.

\impliedby : Choisissons $r = 1/k$ pour $k \in \mathbb{N}^*$. Comme $B(x, 1/k) \cap E \neq \emptyset$ pour tout $k \geq 1$, il existe un $x^k \in B(x, 1/k) \cap E$. Par conséquent, la suite $(x^k)_{k \in \mathbb{N}}$ est dans E et satisfait $\|x^k - x\| < 1/k \rightarrow 0$, d'où $x \in \overline{E}$.

Étape 2 : \overline{E} est le plus petit fermé contenant E . Notons $\tilde{E} := \bigcap \{F \text{ fermé}, F \supset E\}$, alors \tilde{E} est fermé car c'est une intersection de fermés et $E \subset \tilde{E}$.

${}^c\tilde{E} \subset {}^c\overline{E}$: Si $x \notin \tilde{E}$, comme ${}^c\tilde{E}$ est ouvert, il existe un $r > 0$ tel que $B(x, r) \subset {}^c\tilde{E} \subset {}^cE$ et donc $B(x, r) \cap E = \emptyset$. Ceci montre par l'étape 1 que $x \notin \overline{E}$.

${}^c\overline{E} \subset {}^c\tilde{E}$: Si $x \notin \overline{E}$, d'après l'étape 1, on a $B(x, r) \cap E = \emptyset$ pour un certain $r > 0$. Alors $E \subset {}^cB(x, r)$ et ${}^cB(x, r)$ est fermé, d'où $\tilde{E} \subset {}^cB(x, r)$ ce qui montre que $x \notin \tilde{E}$.

On en déduit que $\overline{E} = \tilde{E}$ et donc, si F est un fermé contenant E , on a $\overline{E} \subset F$.

Étape 3 : E est fermé si et seulement si $E = \overline{E}$. On a déjà vu que $E \subset \overline{E}$. Si E est fermé, d'après l'étape 2, on a $\overline{E} \subset E$. \square

La proposition 1.4.9 permet de caractériser séquentiellement le fait qu'un ensemble soit fermé.

Proposition 1.4.10. Un ensemble F de \mathbb{R}^n est fermé si et seulement si pour toute suite $(x^k)_{k \in \mathbb{N}}$ dans F telle que $x^k \rightarrow x$, alors $x \in F$.

Démonstration. \Leftarrow : Il s'agit de montrer que F est fermé, autrement dit que $F = \overline{F}$. On a toujours l'inclusion $F \subset \overline{F}$, montrons l'inclusion opposée. Soit $x \in \overline{F}$, alors il existe une suite $(x^k)_{k \in \mathbb{N}}$ de F telle que $x^k \rightarrow x$. Par hypothèse, on a donc que $x \in F$ soit $\overline{F} \subset F$.

\Rightarrow : Supposons F fermé de sorte que son complémentaire $U := {}^c F$ est ouvert. Considérons une suite $(x_k)_{k \in \mathbb{N}}$ de F telle que $x^k \rightarrow x$. Si $x \in U$, celui-ci étant ouvert, il existe un $r > 0$ tel que $B(x, r) \subset U$. Par définition de la limite, il existe un $N \in \mathbb{N}$ tel que pour tout $k \geq N$, on a $\|x^k - x\| < r$, soit $x^k \in B(x, r) \subset U$, ce qui est absurde puisque $x^k \in F$. Par conséquent, $x \in F$. \square

Définition 1.4.11. Soit E un sous-ensemble quelconque de \mathbb{R}^n . On définit la *frontière* de E par

$$\partial E := \overline{E} \setminus \overset{\circ}{E}.$$

1.4.2 Ensembles compacts

Définition 1.4.12. Un sous-ensemble K de \mathbb{R}^n est compact si toute suite d'éléments de K admet une sous-suite convergente dans K .

Théorème 1.4.13. Un ensemble K dans \mathbb{R}^n est compact si et seulement s'il est à la fois fermé et borné, i.e., s'il existe un $R > 0$ tel que $K \subset B(0, R)$.

Démonstration. \Rightarrow : Supposons que K est compact. Montrons d'abord que K est fermé. Soit $(x^k)_{k \in \mathbb{N}}$ une suite d'éléments de K telle que $x^k \rightarrow x$. De $(x^k)_{k \in \mathbb{N}}$ on peut extraire une sous-suite qui converge dans K . Par unicité de la limite, on en déduit que la limite de la sous-suite ne peut être que x ce qui implique que $x \in K$ et donc que K est fermé en vertu de la proposition 1.4.10.

Montrons à présent que K est borné. Supposons par l'absurde que tel n'est pas le cas. Alors pour tout $k \in \mathbb{N}$, il existerait un $x^k \in K$ tel que $\|x^k\| \geq k$. Or la suite $(x^k)_{k \in \mathbb{N}}$ ainsi construite ne peut admettre de sous-suite convergente ce qui est impossible. Par conséquent K est borné. Finalement on en déduit que K est compact en tant qu'ensemble fermé et borné.

\Leftarrow : Supposons K fermé et borné. Soit $(x^k)_{k \in \mathbb{N}}$ une suite d'éléments de K . L'ensemble K étant par définition borné, il existe un $R > 0$ tel que $K \subset B(0, R)$ et donc $\|x^k\| \leq R$. En particulier, comme $|x_i^k| \leq \|x^k\|$ pour tout $1 \leq i \leq n$, on en déduit que $|x_i^k| \leq R$ et les suites numériques $(x_i^k)_{k \in \mathbb{N}}$ sont bornées dans \mathbb{R} . Par conséquent elles admettent chacune une sous-suite convergente. Plus précisément :

- pour $i = 1$, il existe une extraction $\varphi_1 : \mathbb{N} \rightarrow \mathbb{N}$ strictement croissante et $a_1 \in \mathbb{R}$ tels que $x_1^{\varphi_1(k)} \rightarrow a_1$;
- pour $i = 2$, la suite $(x_2^{\varphi_1(k)})_{k \in \mathbb{N}}$ étant bornée dans \mathbb{R} , il existe une extraction $\varphi_2 : \mathbb{N} \rightarrow \mathbb{N}$ strictement croissante et $a_2 \in \mathbb{R}$ tels que $x_2^{\varphi_2 \circ \varphi_1(k)} \rightarrow a_2$;
- ...
- On suppose qu'il existe des extractions $\varphi_1, \dots, \varphi_{n-1} : \mathbb{N} \rightarrow \mathbb{N}$ strictement croissantes et des réels $a_1, \dots, a_{n-1} \in \mathbb{R}$ tels que pour tout $1 \leq i \leq n-1$ on a $x_i^{\varphi_i \circ \dots \circ \varphi_1(k)} \rightarrow a_i$. La suite $(x_n^{\varphi_{n-1} \circ \dots \circ \varphi_1(k)})_{k \in \mathbb{N}}$ étant bornée dans \mathbb{R} , il existe une extraction $\varphi_n : \mathbb{N} \rightarrow \mathbb{N}$ strictement croissante et $a_n \in \mathbb{R}$ tels que $x_n^{\varphi_n \circ \varphi_{n-1} \circ \dots \circ \varphi_1(k)} \rightarrow a_n$.

Posons $\varphi := \varphi_n \circ \dots \circ \varphi_1 : \mathbb{N} \rightarrow \mathbb{N}$ qui est strictement croissante. De plus, pour tout $1 \leq i \leq n$, la suite $(x_i^{\varphi(k)})_{k \in \mathbb{N}}$ est une sous-suite de $(x_i^{\varphi_i \circ \dots \circ \varphi_1(k)})_{k \in \mathbb{N}}$. Comme cette dernière converge vers a_i dans \mathbb{R} , on en déduit que $x_i^{\varphi(k)} \rightarrow a_i$ dans \mathbb{R} , et d'après la proposition 1.3.3, $x^{\varphi(k)} \rightarrow a = (a_1, \dots, a_n)^T$ dans \mathbb{R}^n . Enfin comme K est fermé, la proposition 1.4.10 assure que $a \in K$. \square

AVERTISSEMENT : La caractérisation des ensembles compacts établie dans le théorème 1.4.13 reste vraie dans un espace vectoriel de dimension finie mais est fautive en dimension infinie.

1.4.3 Ensembles connexes et convexes

Définition 1.4.14. Un sous-ensemble E de \mathbb{R}^n est dit *convexe* si pour tout x et $y \in E$, le segment

$$[x, y] := \{ty + (1 - t)x, t \in [0, 1]\} \subset E.$$

Définition 1.4.15. Un sous-ensemble E de \mathbb{R}^n est dit *connexe* s'il n'existe aucune paire d'ouverts non vides (U_1, U_2) tels que

$$E \subset U_1 \cup U_2, \quad (E \cap U_1) \cap (E \cap U_2) = \emptyset.$$

Autrement dit, un ensemble est connexe si on ne peut pas le séparer en deux parties disjointes en l'intersectant avec deux ouverts disjoints.

Remarque 1.4.16. En dimension $n = 1$, les notions de convexité et connexité coïncident.

Dans le cas d'un ensemble lui-même ouvert, cela se traduit par la propriété plus intuitive suivante.

Théorème 1.4.17. *Un ensemble ouvert U de \mathbb{R}^n est connexe si et seulement si, pour tout x et $y \in U$, il existe une ligne brisée contenue dans U qui les relie, i.e., il existe un nombre fini de points $x_1, \dots, x_p \in U$ tels que $x = x_1$, $y = x_p$ et les segments $[x_i, x_{i+1}] \subset U$ pour tout $1 \leq i \leq p - 1$.*

Démonstration. \implies : On suppose que $U \neq \emptyset$. Soit donc $x_0 \in U$ et V l'ensemble des points de U que l'on peut joindre à x_0 par une ligne brisée dans U .

- *V est ouvert :* Soit $a \in V$, alors il existe une ligne brisée $L \subset U$ qui relie x_0 et $a \in U$. L'ensemble U étant ouvert, il existe un $r > 0$ tel que $B(a, r) \subset U$. Donc pour tout $x \in B(a, r)$, le segment $[a, x] \subset B(a, r) \subset U$ et donc $L' = L \cup [a, x] \subset U$ est une ligne brisée joignant x_0 et x . Ceci montre que $x \in V$, soit $B(a, r) \subset V$, et donc V est ouvert.
- *V est fermé dans U :* Soit $b \in \overline{V} \cap U$, alors il existe un $r > 0$ tel que $B(b, r) \subset U$ car U est ouvert et il existe un $a \in V$ tel que $\|a - b\| < r$ car b est adhérent à V . Soit $L \subset U$ une ligne brisée joignant x_0 et a , alors $L' = L \cup [a, b]$ est une ligne brisée dans U joignant x_0 et b , ce qui montre que $b \in V$ et donc $V = \overline{V} \cap U$. La proposition 1.4.9 montre alors que V est fermé.

On écrit alors que $U = (U \setminus V) \cup V$ où V et $U \setminus V = U \setminus \overline{V}$ sont ouverts et disjoints. Comme U est connexe, alors soit $V = \emptyset$ soit $U \setminus V = \emptyset$. Comme $x_0 \in V$, on en déduit que $V \neq \emptyset$, et donc que $V = U$. Enfin, si x et $y \in U$, il existe une ligne brisée $L_x \subset U$ joignant x_0 et x , et une ligne brisée $L_y \subset U$ joignant x_0 et y . Finalement $L := L_x \cup L_y$ est une ligne brisée dans U qui relie x et y .

\impliedby : Soit U un ouvert de \mathbb{R}^n , supposons que U n'est pas connexe. Alors il existe deux ouverts non vides disjoints U_0 et U_1 tels que $U = U_0 \cup U_1$. Soient $x_0 \in U_0$, $x_1 \in U_1$ et $\gamma : [0, 1] \rightarrow U$ une fonction continue désignant la paramétrisation d'une ligne brisée telle que $\gamma(0) = x_0$ et $\gamma(1) = x_1$. Par continuité de γ , on a que $\gamma^{-1}(U_0)$ et $\gamma^{-1}(U_1)$ sont deux ouverts (voir la proposition 2.1.8 du chapitre 2) non vides disjoints de $[0, 1]$ et $[0, 1] = \gamma^{-1}(U_0) \cup \gamma^{-1}(U_1)$ ce qui implique que $[0, 1]$ n'est pas connexe. \square

Remarque 1.4.18. – En vertu du théorème 1.4.17, on en déduit que tout ensemble ouvert convexe est nécessairement connexe.

- En dimension $n = 1$, les ensembles convexes et connexes coïncident et ce sont précisément les intervalles. En effet, si $E \subset \mathbb{R}$ est connexe, et $x, y \in E$ (avec par exemple $x < y$), alors tout $a \in [x, y]$ appartient à E . Sinon, il existerait un $a \in [x, y]$ tel que $a \notin E$ et on pourrait écrire que $E \subset]-\infty, a[\cup]a, +\infty[$, où les ensembles $E \cap]-\infty, a[$ et $E \cap]a, +\infty[$ sont disjoints ce qui contredirait la connexité de E . Par conséquent, E est un intervalle et il est donc convexe.

Chapitre 2

Fonctions continues

2.1 Définitions, propriétés fondamentales

On s'intéresse à des fonctions de $n \geq 2$ variables à valeurs réelles (\mathbb{R}) ou vectorielles (\mathbb{R}^m , $m \geq 2$) :

$$\begin{aligned} f : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ x = (x_1, \dots, x_n)^T &\longmapsto (f_1(x), \dots, f_m(x)). \end{aligned}$$

Quand $m \geq 2$ on parle aussi de champ de vecteurs. Quand $n = 2$ (resp. 3), on notera souvent $f(x, y)$ (resp. $f(x, y, z)$) au lieu de $f(x_1, x_2)$ (resp. $f(x_1, x_2, x_3)$).

Une fonction f n'est pas forcément définie sur tout \mathbb{R}^n .

Définition 2.1.1. On appelle *ensemble de définition* d'une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ l'ensemble

$$\mathcal{D}_f := \{x \in \mathbb{R}^n : f(x) \text{ existe}\}.$$

Exemple 2.1.2. La fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $f(x, y) = \ln(3 - x^2 - y^2)$ est bien définie si et seulement si $3 - x^2 - y^2 > 0$, autrement dit si $(x, y) \in B(0, \sqrt{3})$. Dans ce cas, $\mathcal{D}_f = B(0, \sqrt{3})$.

Définition 2.1.3. Une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ admet une *limite* $\ell \in \mathbb{R}^m$ en $a \in \mathbb{R}^n$ si pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que pour tout $x \in \mathbb{R}^n$,

$$\|x - a\| < \delta \implies \|f(x) - \ell\| < \varepsilon.$$

On note alors

$$\ell = \lim_{x \rightarrow a} f(x).$$

Définition 2.1.4. Une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est *continue* en $a \in \mathcal{D}_f$ si

$$\lim_{x \rightarrow a} f(x) = f(a),$$

autrement dit, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que pour tout $x \in \mathbb{R}^n$,

$$\|x - a\| < \delta \implies \|f(x) - f(a)\| < \varepsilon. \tag{2.1.1}$$

Une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est continue sur un ensemble $E \subset \mathbb{R}^n$ si elle est continue en tout point de E . L'ensemble des fonctions continues de E dans \mathbb{R}^m sera noté $\mathcal{C}(E; \mathbb{R}^m)$ ou $\mathcal{C}^0(E; \mathbb{R}^m)$. Si $m = 1$, on notera simplement $\mathcal{C}(E)$ ou $\mathcal{C}^0(E)$.

AVERTISSEMENT : Dans la formule (2.1.1), la norme $\|x - a\|$ désigne la norme sur \mathbb{R}^n puisque $x - a$ est un élément de \mathbb{R}^n . En revanche la norme $\|f(x) - f(a)\|$ désigne la norme sur \mathbb{R}^m étant donné que $f(x) - f(a) \in \mathbb{R}^m$.

Remarque 2.1.5. Compte tenu de l'équivalence des normes sur \mathbb{R}^n , on peut remplacer la norme euclidienne $\|\cdot\|$ par n'importe quelle autre norme.

Donnons à présent deux caractérisations de la continuité, l'une séquentielle (*i.e.* à l'aide des suites) et l'autre topologique (*i.e.* en terme d'ouverts et de fermés).

Proposition 2.1.6. Soient $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ et $a \in \mathcal{D}_f$. Alors les propriétés suivantes sont équivalentes :

- i) f est continue en a ;
- ii) Pour toute suite $(x_k)_{k \in \mathbb{N}}$ de \mathbb{R}^n telle que $x^k \rightarrow a$ dans \mathbb{R}^n , alors $f(x^k) \rightarrow f(a)$ dans \mathbb{R}^m .

Démonstration. i) \Rightarrow ii) : Soient ε et δ comme dans la définition de la continuité en a et $(x^k)_{k \in \mathbb{N}}$ une suite de \mathbb{R}^n telle que $x^k \rightarrow a$. Alors, par définition de la limite d'une suite, il existe un $N \in \mathbb{N}$ tel que pour tout $k \geq N$, $\|x^k - a\| < \delta$, puis par définition de la continuité, $\|f(x^k) - f(a)\| < \varepsilon$. Ceci montre bien que $f(x^k) \rightarrow f(a)$ dans \mathbb{R}^m .

ii) \Rightarrow i) : Supposons que f n'est pas continue en a . Il existe alors un $\varepsilon_0 > 0$ tel que pour tout $\delta > 0$, on peut trouver un $x^\delta \in \mathbb{R}^n$ ayant les propriétés $\|x^\delta - a\| < \delta$ et $\|f(x^\delta) - f(a)\| \geq \varepsilon_0$. En prenant $\delta = 1/k$, avec $k \in \mathbb{N}^*$, on obtient une suite $(x^k)_{k \in \mathbb{N}}$ telle que $\|x^k - a\| < 1/k$ et $\|f(x^k) - f(a)\| \geq \varepsilon_0$ pour tout $k \in \mathbb{N}^*$. Par conséquent $x^k \rightarrow a$ dans \mathbb{R}^n mais $f(x^k) \not\rightarrow f(a)$ dans \mathbb{R}^m . \square

Cette caractérisation séquentielle de la continuité est particulièrement utile en pratique pour montrer la continuité d'une fonction en un point.

La caractérisation topologique, qui fait l'objet du résultat suivant, nécessite de définir la notion d'image réciproque d'un ensemble par une fonction.

Définition 2.1.7. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ et $E \subset \mathbb{R}^m$. On appelle *image réciproque de E par f* le sous-ensemble de \mathbb{R}^n défini par

$$f^{-1}(E) := \{x \in \mathbb{R}^n : f(x) \in E\}.$$

Proposition 2.1.8. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Alors les propriétés suivantes sont équivalentes :

- i) f est continue sur \mathbb{R}^n ;
- ii) Pour tout ouvert $U \subset \mathbb{R}^m$, $f^{-1}(U)$ est ouvert dans \mathbb{R}^n ;
- iii) Pour tout fermé $F \subset \mathbb{R}^m$, $f^{-1}(F)$ est fermé dans \mathbb{R}^n .

Démonstration. i) \Rightarrow ii) : Supposons f continue sur \mathbb{R}^n et soit U un ouvert de \mathbb{R}^m . Si $f^{-1}(U) = \emptyset$ c'est un ouvert de \mathbb{R}^n et il n'y a rien à montrer. Si en revanche $f^{-1}(U) \neq \emptyset$, il existe un $a \in f^{-1}(U)$, *i.e.*, $f(a) \in U$. L'ensemble U étant ouvert dans \mathbb{R}^m , il existe un $\varepsilon > 0$ tel que $B(f(a), \varepsilon) \subset U$. Soit $\delta > 0$ en correspondance avec cet ε donné par la définition de la continuité de f en a . Alors pour tout $x \in \mathbb{R}^n$ avec $\|x - a\| < \delta$, on a que $\|f(x) - f(a)\| < \varepsilon$, ce qui montre que pour tout $x \in B(a, \delta)$, $f(x) \in B(f(a), \varepsilon)$, *i.e.*, $x \in f^{-1}(B(f(a), \varepsilon))$. Par conséquent, $B(a, \delta) \subset f^{-1}(B(f(a), \varepsilon)) \subset f^{-1}(U)$ et donc $f^{-1}(U)$ est ouvert dans \mathbb{R}^n .

ii) \Rightarrow i) : Pour tout $\varepsilon > 0$ et pour tout $a \in \mathbb{R}^n$, l'ensemble $B(f(a), \varepsilon)$ est ouvert dans \mathbb{R}^m . Par conséquent $f^{-1}(B(f(a), \varepsilon))$ est ouvert dans \mathbb{R}^n et $a \in f^{-1}(B(f(a), \varepsilon))$ puisque $f(a) \in B(f(a), \varepsilon)$. Il existe donc un $\delta > 0$ tel que $B(a, \delta) \subset f^{-1}(B(f(a), \varepsilon))$, autrement dit, pour tout $x \in \mathbb{R}^n$ avec $\|x - a\| < \delta$, on a $x \in f^{-1}(B(f(a), \varepsilon))$ soit $f(x) \in B(f(a), \varepsilon)$ ou encore $\|f(x) - f(a)\| < \varepsilon$, ce qui établit la continuité de f en a .

ii) \Rightarrow iii) : Soit F un fermé de \mathbb{R}^m , alors $U := {}^c F$ est ouvert dans \mathbb{R}^m et $f^{-1}(U)$ est ouvert dans \mathbb{R}^n . Donc $f^{-1}(F) = f^{-1}({}^c U) = {}^c f^{-1}(U)$ est un fermé.

iii) \Rightarrow ii) : Idem. □

Cette caractérisation topologique de la continuité est fort utile en pratique pour montrer qu'un ensemble est ouvert ou fermé. Il suffit de montrer que c'est l'image réciproque d'un ouvert ou d'un fermé par une fonction continue.

Exemple 2.1.9. Soit $U := \{(x, y) \in \mathbb{R}^2 : x^2 - y^2 > 0\}$. La fonction $f : (x, y) \mapsto x^2 - y^2$ est continue sur \mathbb{R}^2 et $U = f^{-1}(]0, +\infty[)$ avec $]0, +\infty[$ ouvert dans \mathbb{R} . On en déduit que U est ouvert dans \mathbb{R}^2 .

Propriétés 2.1.10. i) Toute combinaison linéaire de fonctions continues est continue. Autrement dit, $\mathcal{C}(E; \mathbb{R}^m)$ est un espace vectoriel pour tout sous-ensemble E de \mathbb{R}^n ;
 ii) Si $m = 1$, le produit de deux fonctions continues est continu. Autrement dit $\mathcal{C}(E)$ est une algèbre pour tout sous-ensemble E de \mathbb{R}^n ;
 iii) Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est continue en $a \in \mathbb{R}^n$, et $g : \mathbb{R}^m \rightarrow \mathbb{R}^p$ est continue en $f(a)$, alors $g \circ f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ est continue en a ;

Démonstration. Nous allons utiliser la caractérisation séquentielle de la continuité établie dans la proposition 2.1.6.

i) Soient f et $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ continues en $a \in \mathbb{R}^n$ et $\lambda, \mu \in \mathbb{R}$. Si $(x^k)_{k \in \mathbb{N}}$ est une suite de \mathbb{R}^n telle que $x^k \rightarrow a$ dans \mathbb{R}^n , alors $f(x^k) \rightarrow f(a)$ et $g(x^k) \rightarrow g(a)$ dans \mathbb{R}^m et donc $\lambda f(x^k) + \mu g(x^k) \rightarrow \lambda f(a) + \mu g(a)$ dans \mathbb{R}^m ce qui montre que $\lambda f + \mu g$ est continue en a .

ii) Si f et $g : \mathbb{R}^n \rightarrow \mathbb{R}$ sont continues en $a \in \mathbb{R}^n$ et $(x^k)_{k \in \mathbb{N}}$ est une suite de \mathbb{R}^n telle que $x^k \rightarrow a$ dans \mathbb{R}^n , alors $f(x^k) \rightarrow f(a)$ et $g(x^k) \rightarrow g(a)$ dans \mathbb{R} et donc $f(x^k)g(x^k) \rightarrow f(a)g(a)$ dans \mathbb{R} ce qui montre que fg est continue en a .

iii) Si $(x^k)_{k \in \mathbb{N}}$ est une suite de \mathbb{R}^n telle que $x^k \rightarrow a$ dans \mathbb{R}^n , par continuité de f en a , on a que $f(x^k) \rightarrow f(a)$ dans \mathbb{R}^m , puis par continuité de g en $f(a)$, il vient $g(f(x^k)) \rightarrow g(f(a))$ dans \mathbb{R}^p , ce qui montre que $g \circ f$ est continue en a . □

Remarque 2.1.11. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est continue en $a \in \mathcal{D}_f$ et $f(a) \neq 0$, alors la fonction $1/f$ est bien définie et continue en a .

Exemple 2.1.12. On a déjà vu que la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto \ln(3 - x^2 - y^2)$ est bien définie sur $B(0, \sqrt{3})$. Elle y est également continue comme composée de fonctions continues.

2.2 Exemples importants

2.2.1 Fonctions Lipschitziennes

Définition 2.2.1. Une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est dite *Lipschitzienne* s'il existe une constante $L > 0$ telle que

$$\|f(x) - f(y)\| \leq L\|x - y\|, \quad \text{pour tout } x, y \in \mathbb{R}^n. \quad (2.2.1)$$

Si $L < 1$, on dit que f est *contractante*.

Si L est la plus petite constante vérifiant (2.2.1), on dit que f est L -Lipschitzienne et L est appelée la constante de Lipschitz). On constate que toute fonction L -Lipschitzienne est continue. En effet, si $\varepsilon > 0$, on pose $\delta = \varepsilon/L$ de sorte que si x et $y \in \mathbb{R}^n$ et $\|x - y\| < \delta$, alors $\|f(x) - f(y)\| < \varepsilon$. Ceci montre de plus que f est uniformément continue (voir la définition 2.3.1 plus loin).

Un exemple typique est la fonction norme qui définit une fonction 1-Lipschitzienne. En effet, pour tout x et $y \in \mathbb{R}^n$, l'inégalité triangulaire assure que $\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\|$ ce

qui montre que $\|x\| - \|y\| \leq \|x - y\|$. De même en échangeant les rôles de x et y , on obtient que $\|y\| - \|x\| \leq \|x - y\|$ ce qui établit

$$\| \|x\| - \|y\| \| \leq \|x - y\|.$$

Le résultat suivant d'existence et d'unicité d'un point fixe a des applications importantes en analyse. C'est dessus qu'est basée la démonstration du théorème d'inversion locale et du théorème de Cauchy-Lipschitz qui assure l'existence et l'unicité d'une solution à certaines équations différentielles.

Théorème 2.2.2 (Picard). *Soit $F \subset \mathbb{R}^n$ un ensemble fermé et $f : F \rightarrow F$ une fonction contractante. Alors f admet un unique point fixe, i.e., il existe un unique $x \in F$ tel que $f(x) = x$.*

Démonstration. Existence : Soit $x^0 \in F$, on définit par récurrence la suite $(x^k)_{k \in \mathbb{N}}$ par $x^{k+1} = f(x^k)$ pour tout $k \in \mathbb{N}$. Par construction, pour tout $k \in \mathbb{N}$, on a que $x^k \in F$ et

$$\|x^{k+1} - x^k\| = \|f(x^k) - f(x^{k-1})\| \leq L\|x^k - x^{k-1}\| \leq L^k\|x^1 - x^0\|.$$

Ainsi, pour tout $l \geq k$, l'inégalité triangulaire montre que

$$\|x^l - x^k\| \leq \sum_{i=k}^{l-1} \|x^{i+1} - x^i\| \leq \|x^1 - x^0\| \sum_{i=k}^{l-1} L^i = \|x^1 - x^0\| \frac{L^k - L^l}{1 - L}.$$

Comme $L < 1$, on en déduit que $L^k \rightarrow 0$ quand $k \rightarrow +\infty$, et donc il existe un $N \in \mathbb{N}$ tel que pour tout $l \geq k \geq N$, on a $\|x^l - x^k\| < \varepsilon$. Ceci montre que la suite $(x^k)_{k \in \mathbb{N}}$ est de Cauchy dans \mathbb{R}^n et donc, par la proposition 1.3.5, qu'elle converge vers un x . Comme F est fermé, la proposition 1.4.10 assure que $x \in F$. Enfin, par passage à la limite dans la relation de récurrence $x^{k+1} = f(x^k)$, on en déduit par continuité de f que $x = f(x)$, ce qui montre que x est un point fixe de f .

Unicité : Soient x et $y \in F$ deux points fixes de f , alors $f(x) = x$ et $f(y) = y$, d'où

$$\|x - y\| = \|f(x) - f(y)\| \leq L\|x - y\|,$$

ce qui montre que $x = y$ puisque $L < 1$. □

2.2.2 Applications linéaires

Définition 2.2.3. On note $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ l'ensemble des *applications linéaires* de \mathbb{R}^n dans \mathbb{R}^m , i.e., l'ensemble des fonctions $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ telles que $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$ pour tout $x, y \in \mathbb{R}^n$ et $\lambda, \mu \in \mathbb{R}$.

Si $\{e^1, \dots, e^n\}$ désigne la base canonique de \mathbb{R}^n , i.e., $e_j^i = \delta_{ij}$ pour tout $1 \leq i, j \leq n$, alors

$$x = \sum_{i=1}^n x_i e^i \implies f(x) = \sum_{i=1}^n x_i f(e^i),$$

et d'après les inégalités triangulaire et de Cauchy-Schwarz, on a que

$$\|f(x)\| \leq \sum_{i=1}^n |x_i| \|f(e^i)\| \leq \|x\| \left(\sum_{i=1}^n \|f(e^i)\|^2 \right)^{1/2} = M\|x\|,$$

où l'on a posé $M := \left(\sum_{i=1}^n \|f(e^i)\|^2 \right)^{1/2}$. En appliquant cette formule à $x - y$, on obtient

$$\|f(x) - f(y)\| = \|f(x - y)\| \leq M\|x - y\|,$$

et donc f est Lipschitzienne donc continue. On note, pour tout $f \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$,

$$\|f\| := \sup_{x \neq 0} \frac{\|f(x)\|}{\|x\|},$$

et on peut montrer qu'il s'agit de la plus petite constante $C > 0$ telle que $\|f(x)\| \leq C\|x\|$ pour tout $x \in \mathbb{R}^n$. Autrement dit, d'après la linéarité de f , $\|f\|$ n'est autre que la constante de Lipschitz de f .

2.2.3 Polynômes

Définition 2.2.4. Un *polynôme* est une fonction $p : \mathbb{R}^n \rightarrow \mathbb{R}$ de la forme

$$p(x) = \sum_{\substack{k \in \mathbb{N}^n, \\ k_1 + \dots + k_n \leq m}} a_k x_1^{k_1} \dots x_n^{k_n}, \quad \text{pour tout } x \in \mathbb{R}^n,$$

où les $a_k \in \mathbb{R}$. On note $\deg(p) = m$ le degré du polynôme p .

Un polynôme est continu sur \mathbb{R}^n comme produit et combinaison linéaire de fonctions continues.

Exemple 2.2.5. En dimension $n = 2$, un polynôme de degré 2 s'écrit sous la forme

$$p(x, y) = a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5y^2, \quad \text{pour tout } (x, y) \in \mathbb{R}^2,$$

où $a_0, \dots, a_5 \in \mathbb{R}$.

2.2.4 Applications partielles

Une façon naturelle d'étudier une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ consiste à fixer $n - 1$ variables et à étudier ensuite la fonction par rapport à l'unique variable restante. Ainsi, on définit la i -ème application partielle au point a par

$$f_{i,a}(t) = f(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_n) = f(a + te^i), \quad \text{pour tout } t \in \mathbb{R}.$$

Proposition 2.2.6. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est continue sur \mathbb{R}^n , alors l'application partielle $f_{i,a}$ est continue sur \mathbb{R} .

Démonstration. Soient $a \in \mathbb{R}$ et $t \in \mathbb{R}$. Si $t^k \rightarrow t$, alors le vecteur $a + t^k e_i \rightarrow a + te_i$ quand $k \rightarrow +\infty$, et donc, par continuité de f ,

$$f_{i,a}(t^k) = f(a_1, \dots, a_{i-1}, a_i + t^k, a_{i+1}, \dots, a_n) \rightarrow f(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_n) = f_{i,a}(t),$$

ce qui montre que $f_{i,a}$ est continue en t . □

En général, la réciproque est fautive comme le montre l'exemple suivant :

Exemple 2.2.7. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, définie par

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

Si $y = 0$, l'application $x \mapsto f(x, 0) = 0$ est constante donc continue sur \mathbb{R} . Si $y \neq 0$, comme dans ce cas $x^2 + y^2 > 0$, on en déduit que $x \mapsto f(x, y)$ est également continue sur \mathbb{R} . Comme x et y

jouent des rôles symétriques, on en déduit que $y \mapsto f(x, y)$ est continue sur \mathbb{R} pour tout $x \in \mathbb{R}$. En revanche, nous allons montrer que f n'est pas continue $(0, 0)$. En effet, pour tout $x \neq 0$,

$$f(x, x) = \frac{x^2}{2x^2} = \frac{1}{2}$$

et donc

$$\lim_{x \rightarrow 0} f(x, x) = \frac{1}{2} \neq 0 = f(0, 0),$$

ce qui montre que f n'est pas continue en $(0, 0)$.

2.3 Continuité et compacité

Définition 2.3.1. Une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est dite *uniformément continue* sur un sous-ensemble E de \mathbb{R}^n si, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que pour tout $x, y \in E$,

$$\|x - y\| < \delta \implies \|f(x) - f(y)\| < \varepsilon.$$

La différence entre continuité et uniforme continuité est subtile et réside dans l'ordre des quantificateurs. Dans la définition de la continuité en un point a , le δ dépend à la fois de a et de ε . En revanche, dans la définition de l'uniforme continuité, le δ est indépendant du point ; il ne dépend que de ε .

Théorème 2.3.2 (Heine). Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ une fonction continue sur un compact K de \mathbb{R}^n . Alors f est uniformément continue sur K .

Démonstration. On raisonne par l'absurde. Supposons que f n'est pas uniformément continue sur K . Alors il existe un $\varepsilon_0 > 0$ tel que pour tout $\delta > 0$, on peut trouver x^δ et $y^\delta \in K$ avec les propriétés $\|x^\delta - y^\delta\| < \delta$ et $\|f(x^\delta) - f(y^\delta)\| \geq \varepsilon_0$. En choisissant $\delta = 1/k$, avec $k \in \mathbb{N}^*$, on obtient alors deux suites $(x^k)_{k \in \mathbb{N}}$ et $(y^k)_{k \in \mathbb{N}}$ dans K telles que $\|x^k - y^k\| < 1/k$ et $\|f(x^k) - f(y^k)\| \geq \varepsilon_0$. L'ensemble K étant compact, le théorème de Bolzano-Weierstrass assure que $(x^k)_{k \in \mathbb{N}}$ et $(y^k)_{k \in \mathbb{N}}$ admettent des sous-suites convergentes. Plus précisément, il existe une extraction $\varphi_1 : \mathbb{N} \rightarrow \mathbb{N}$ strictement croissante et $x \in K$ tels que $x^{\varphi_1(k)} \rightarrow x$ dans K . Ensuite $(y^{\varphi_1(k)})_{k \in \mathbb{N}}$ étant dans K , il existe une nouvelle extraction $\varphi_2 : \mathbb{N} \rightarrow \mathbb{N}$ strictement croissante et $y \in K$ tels que $y^{\varphi_2 \circ \varphi_1(k)} \rightarrow y$ dans K . On définit alors $\varphi := \varphi_2 \circ \varphi_1 : \mathbb{N} \rightarrow \mathbb{N}$ qui est strictement croissante. Alors $y^{\varphi(k)} \rightarrow y$ dans K et comme $(x^{\varphi(k)})_{k \in \mathbb{N}}$ est une sous-suite de $(x^{\varphi_1(k)})_{k \in \mathbb{N}}$, on en déduit que $x^{\varphi(k)} \rightarrow x$ dans K . Comme f est continue, on en déduit que $f(x^{\varphi(k)}) - f(y^{\varphi(k)}) \rightarrow f(x) - f(y)$ et donc que $\|f(x) - f(y)\| \geq \varepsilon_0 > 0$. Par ailleurs, comme $\|x^{\varphi(k)} - y^{\varphi(k)}\| \leq 1/\varphi(k) \rightarrow 0$, on en déduit que $\|x - y\| = 0$, soit $x = y$ et donc $f(x) = f(y)$ ce qui est impossible. \square

Les résultats qui suivent représentent une première incursion vers la recherche d'extrema d'une fonction à valeurs réelles. Nous avons besoin, tout d'abord d'introduire un peu de vocabulaire.

Soient $f : \mathbb{R}^n \rightarrow \mathbb{R}$ et E un sous-ensemble de \mathbb{R}^n . On dit que f est *minorée* sur E s'il existe un réel $m \in \mathbb{R}$ tel que $f(x) \geq m$ pour tout $x \in E$. Le plus grand des minorants de f , appelé *borne inférieure* ou *infimum*, est noté

$$\alpha := \inf_{x \in E} f(x).$$

Il est caractérisé par les deux propriétés suivantes :

- pour tout $y \in E$, $f(y) \geq \alpha$;
- il existe une suite $(x^k)_{k \in \mathbb{N}}$ de E telle que $f(x^k) \rightarrow \alpha$.

Si l'infimum est atteint, *i.e.*, s'il existe un $a \in E$ tel que $f(a) = \alpha$, alors on dit que f admet un *minimum* sur E en a et on note

$$\alpha = f(a) = \min_{x \in E} f(x).$$

De façon similaire, on dit que f est *majorée* sur E s'il existe un réel $M \in \mathbb{R}$ tel que $f(x) \leq M$ pour tout $x \in E$. Le plus petit des majorants de f , appelé *borne supérieure* ou *supremum*, est noté

$$\beta := \sup_{x \in E} f(x).$$

Il est caractérisé par les deux propriétés suivantes :

- pour tout $y \in E$, $f(y) \leq \beta$;
- il existe une suite $(x^k)_{k \in \mathbb{N}}$ de E telle que $f(x^k) \rightarrow \beta$.

Si le supremum est atteint, *i.e.*, s'il existe un $b \in E$ tel que $f(b) = \beta$, alors on dit que f admet un *maximum* sur E en b et on note

$$\beta = f(b) = \max_{x \in E} f(x).$$

Théorème 2.3.3. *Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continue sur un compact K de \mathbb{R}^n . Alors f admet un maximum et un minimum sur K . Autrement dit, il existe un $a \in K$ et un $b \in K$ tels que*

$$f(a) = \min_{x \in K} f(x), \quad f(b) = \max_{x \in K} f(x).$$

Démonstration. Nous montrons ici seulement l'existence du minimum. Des arguments similaires permettent de traiter le cas du maximum. Par définition de la borne inférieure, il existe une suite $(x^k)_{k \in \mathbb{N}}$ de K telle que $f(x^k) \rightarrow \inf_{x \in K} f(x)$. Comme K est compact, d'après le théorème de Bolzano-Weierstrass, on peut extraire de $(x^k)_{k \in \mathbb{N}}$ une sous-suite convergente : il existe une extraction $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ strictement croissante et $a \in K$ tels que $x^{\varphi(k)} \rightarrow a$ dans K . Comme f est continue, on a que $f(x^{\varphi(k)}) \rightarrow f(a)$, et par unicité de la limite, on a nécessairement que

$$f(a) = \inf_{x \in K} f(x).$$

Ceci montre que l'infimum de f sur K est atteint, et donc que f admet un minimum sur K en a . □

Pour finir ce chapitre, nous présentons à présent une démonstration du théorème 1.2.5 sur l'équivalence des normes dans \mathbb{R}^n .

Démonstration du théorème 1.2.5. Soit N une norme sur \mathbb{R}^n . Nous allons montrer qu'elle est équivalente à la norme euclidienne $\|\cdot\|$.

Étape 1 : Montrons que l'application $N : (\mathbb{R}^n, \|\cdot\|) \rightarrow (\mathbb{R}^n, N)$ est Lipschitzienne. En effet, pour tout x et $y \in \mathbb{R}^n$, d'après l'inégalité triangulaire et l'homogénéité de N et l'inégalité de Cauchy-Schwarz, on a

$$N(x - y) = N\left(\sum_{i=1}^n (x_i - y_i)e^i\right) \leq \sum_{i=1}^n N((x_i - y_i)e^i) = \sum_{i=1}^n |x_i - y_i|N(e^i) \leq M\|x - y\|,$$

où l'on a noté $M := (\sum_{i=1}^n N(e^i)^2)^{1/2}$. Une nouvelle application de l'inégalité triangulaire montre que $|N(x) - N(y)| \leq N(x - y)$, et donc

$$|N(x) - N(y)| \leq M\|x - y\|, \quad \text{pour tout } x, y \in \mathbb{R}^n.$$

Étape 2 : Montrons que l'ensemble $S := \{x \in \mathbb{R}^n : \|x\| = 1\}$ est compact. Tout d'abord, comme $S \subset B(0, 2)$, on en déduit que S est borné. Ensuite, on introduit la fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ définie par

$f(x) = \|x\|$ pour tout $x \in \mathbb{R}^n$ qui est une fonction continue. Comme $S = f^{-1}(\{1\})$ et $\{1\}$ est un ensemble fermé dans \mathbb{R} , la proposition 2.1.8 montre que S est fermé et donc finalement compact.

Étape 3 : Comme N est continue sur \mathbb{R}^n , elle l'est en particulier sur le compact S . Le théorème 2.3.3 montre alors l'existence d'un minimum $a \in S$ et d'un maximum $b \in S$ tels que

$$\|a\| = \|b\| = 1, \quad N(a) \leq N(x) \leq N(b), \quad \text{pour tout } x \in S.$$

Notons $m = N(a)$ et $M = N(b)$. Comme $\|a\| = 1$, on en déduit que $a \neq 0$ (car sinon $\|a\| = 0$) et donc $N(a) > 0$ (car sinon $a = 0$). Par conséquent, $m > 0$, $M > 0$ et pour tout $y \in \mathbb{R}^n$ ($y \neq 0$), on a $y/\|y\| \in S$, ce qui implique

$$m \leq N\left(\frac{y}{\|y\|}\right) \leq M.$$

Par la propriété d'homogénéité de la norme N , on en déduit que

$$m\|y\| \leq N(y) \leq M\|y\|, \quad \text{pour tout } y \neq 0. \quad (2.3.1)$$

Cette inégalité reste bien évidemment vraie si $y = 0$, ce qui montre que les normes $\|\cdot\|$ et N sont équivalentes.

Enfin si N_1 et N_2 sont deux normes quelconques sur \mathbb{R}^n , en combinant les inégalités (2.3.1) appliquées à N_1 et N_2 , on en déduit que N_1 et N_2 sont effectivement équivalentes. \square

Chapitre 3

Différentiabilité et dérivées partielles

3.1 Définitions

Toujours dans l'esprit d'étudier les variations d'une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ en fixant $n - 1$ variables et en étudiant les applications partielles, nous introduisons la notion de dérivée partielle. Dans la suite de ce chapitre, on notera $\{e^1, \dots, e^n\}$ la base canonique de \mathbb{R}^n et $\{\eta^1, \dots, \eta^m\}$ la base canonique de \mathbb{R}^m .

Définition 3.1.1. Soit U un ouvert de \mathbb{R}^n et $f : U \rightarrow \mathbb{R}$. On dit que f admet une *dérivée partielle* en $a \in U$ par rapport à sa i -ème variable si la i -ème application partielle $f_{i,a}$ est dérivable en $t = 0$. On note alors

$$\frac{\partial f}{\partial x_i}(a) := f'_{i,a}(0) = \lim_{h \rightarrow 0} \frac{f(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_n) - f(a)}{h} = \lim_{h \rightarrow 0} \frac{f(a + he^i) - f(a)}{h}.$$

Le calcul de $\frac{\partial f}{\partial x_i}(a)$ consiste donc à ne dériver l'expression de f que par rapport à la variable x_i .

Exemple 3.1.2. Si $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ est définie par $f(x, y, z) = -2x \cos y$, on a que

$$\frac{\partial f}{\partial x}(x, y, z) = -2 \cos y, \quad \frac{\partial f}{\partial y}(x, y, z) = 2x \sin y, \quad \frac{\partial f}{\partial z}(x, y, z) = 0.$$

Le calcul de la i -ème dérivée partielle consiste à regarder les variations de la fonction f suivant la direction e^i . L'inconvénient majeur de la notion de dérivée partielle est qu'elle nécessite *a priori* le choix d'une base dans \mathbb{R}^n . La différentielle introduite ci-dessous permet au contraire d'introduire une notion intrinsèque permettant de quantifier les variations d'une fonction au voisinage d'un point. Nous donnons ici la définition directement pour les fonctions à valeurs vectorielles.

Définition 3.1.3. Soit U un ouvert de \mathbb{R}^n et $f : U \rightarrow \mathbb{R}^m$. On dit que f est *différentiable* en $a \in U$ s'il existe une application linéaire $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ telle que

$$\lim_{\|h\| \rightarrow 0} \frac{\|f(a+h) - f(a) - L(h)\|}{\|h\|} = 0.$$

Autrement dit, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que si $h \in \mathbb{R}^n$ est tel que $\|h\| < \delta$, alors

$$\|f(a+h) - f(a) - L(h)\| < \varepsilon \|h\|.$$

Dans ce cas, l'application linéaire L est noté $df(a)$ et est appelée *différentielle* ou *application linéaire tangente* de f en a . Nous renvoyons au chapitre 4 pour une interprétation géométrique de cet objet.

Remarque 3.1.4. Une autre façon de vérifier la différentiabilité d'une fonction f en un point a est de montrer l'existence d'une fonction $\varepsilon : \mathbb{R} \rightarrow \mathbb{R}^m$ satisfaisant $\varepsilon(t) \rightarrow 0$ quand $t \rightarrow 0$ et telle que

$$f(a+h) = f(a) + df(a)(h) + \|h\|\varepsilon(\|h\|), \quad \text{pour tout } h \text{ de norme assez petite,}$$

ou encore

$$f(a+h) = f(a) + df(a)(h) + o(\|h\|), \quad \text{pour tout } h \text{ de norme assez petite.}$$

Notons que cette dernière expression n'est autre que le développement limité de f à l'ordre 1 au voisinage de a .

AVERTISSEMENT : L'application $a \mapsto df(a)$ est une fonction de \mathbb{R}^n dans $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Il faudra bien se garder de croire que cette application est linéaire. Dans la définition de la différentielle, le point a est fixé et la différentielle au point a est une application linéaire qui dépend de a dont la dépendance peut être totalement arbitraire. Dans cette définition, l'application qui est linéaire est $h \mapsto df(a)(h)$; c'est l'application linéaire $df(a)$ appliquée au vecteur h .

Remarque 3.1.5. Dans le cas d'une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$, la différentielle, qui est une application linéaire de \mathbb{R} dans \mathbb{R} , peut être identifiée à un réel noté $f'(a)$ et on a

$$df(a)(h) = f'(a)h, \quad \text{pour tout } h \in \mathbb{R}.$$

Autrement dit, l'application linéaire $h \mapsto df(a)(h)$ n'est autre que la multiplication du réel h par le réel $f'(a)$.

Proposition 3.1.6. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est différentiable en $a \in \mathbb{R}^n$, alors f est continue en a .

Démonstration. On a pour tout $h \in \mathbb{R}^n$, $f(a+h) - f(a) = df(a)(h) + o(\|h\|) \rightarrow 0$. □

Remarque 3.1.7. Notons qu'en général, l'existence des seules dérivées partielles n'implique pas la continuité. Pour s'en convaincre on peut considérer la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par

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

En effet,

$$\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0,$$

mais f n'est pas continue en $(0, 0)$ car, par exemple, $f(x, x^2) = 1/2$ pour tout $x \neq 0$.

Le résultat suivant nous fournit un premier lien entre les notions de différentielle et dérivées partielles.

Proposition 3.1.8. Soit U un ouvert de \mathbb{R}^n et $f : U \rightarrow \mathbb{R}^m$. Si f est différentiable en $a \in U$, alors pour tout $1 \leq i \leq m$, les fonctions f_i admettent des dérivées partielles et on a

$$df(a)(h) = \sum_{i=1}^m \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(a) h_j \eta^i, \quad \text{pour tout } h \in \mathbb{R}^n.$$

Démonstration. Comme pour tout $1 \leq i \leq m$ et tout $h \in \mathbb{R}^n$, on a $|f_i(a+h) - f_i(a) - [df(a)(h)]_i| \leq \|f(a+h) - f(a) - df(a)(h)\|$, on en déduit que

$$\lim_{\|h\| \rightarrow 0} \frac{|f_i(a+h) - f_i(a) - [df(a)(h)]_i|}{\|h\|} = 0.$$

En particulier, en prenant $h = te^j$ pour $1 \leq j \leq n$, il vient que

$$\lim_{t \rightarrow 0} \frac{|f_i(a + te^j) - f_i(a) - t[df(a)(e^j)]_i|}{t} = 0,$$

soit

$$[df(a)(e^j)]_i = \lim_{t \rightarrow 0} \frac{f_i(a + te^j) - f_i(a)}{t}.$$

Ceci montre que, pour tout $1 \leq i \leq m$, les fonctions f_i admettent des dérivées partielles et on a $\frac{\partial f_i}{\partial x_j}(a) = [df(a)(e^j)]_i$. Comme $h \in \mathbb{R}^n$ et $df(a)(h) \in \mathbb{R}^m$, on écrit que

$$h = \sum_{j=1}^n h_j e^j, \quad df(a)(h) = \sum_{i=1}^m [df(a)(h)]_i \eta^i,$$

et par linéarité de $h \mapsto df(a)(h)$, il vient que

$$df(a)(h) = \sum_{i=1}^m \sum_{j=1}^n h_j [df(a)(e^j)]_i \eta^i = \sum_{i=1}^m \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(a) h_j \eta^i,$$

ce qui conclut la preuve du résultat. \square

Ainsi, une fonction différentiable en un point a admet des dérivées partielles. La réciproque est fautive en général comme le montre l'exemple suivant.

Exemple 3.1.9. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie, pour tout $(x, y) \in \mathbb{R}^2$, par

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

Alors

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

ce qui montre que f admet des dérivées partielles en $(0, 0)$. Si f était différentiable en $(0, 0)$, la proposition 3.1.8 montrerait que sa différentielle $df(0, 0)$ est l'application linéaire nulle. Or

$$\frac{f(h, h) - f(0, 0) - df(0, 0)(h, h)}{h} = 1$$

qui ne tend pas vers 0 lorsque $h \rightarrow 0$. On en déduit que f n'est pas différentiable en $(0, 0)$.

Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est une fonction différentiable en $a \in \mathbb{R}^n$, sa différentielle $df(a)$ est une application linéaire de \mathbb{R}^n dans \mathbb{R}^m . Les bases $\{e^1, \dots, e^n\}$ de \mathbb{R}^n et $\{\eta^1, \dots, \eta^m\}$ de \mathbb{R}^m étant fixées, l'application linéaire $df(a)$ peut être représentée par une matrice de taille $m \times n$ notée $Df(a)$. D'après la proposition 3.1.8, la composante (i, j) de cette matrice est donnée par la j -ème dérivée partielle de la i -ème composante de f .

Définition 3.1.10. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ une fonction qui admet des dérivées partielles au point $a \in \mathbb{R}^n$. On appelle la *matrice jacobienne* de f en a la matrice

$$Df(a) := \left(\frac{\partial f_i}{\partial x_j} \right)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}.$$

Lorsque $m = n$, $Df(a)$ est une matrice carrée et on note

$$J_f(a) := \det Df(a),$$

le *jacobien* de f en a .

Remarque 3.1.11. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est différentiable en $a \in \mathbb{R}^n$, la proposition 3.1.8 montre que pour tout $h \in \mathbb{R}^n$, on a

$$[Df(a)h]_i = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j} h_j = [df(a)(h)]_i, \quad \text{pour tout } 1 \leq i \leq m.$$

Autrement dit,

$$Df(a)h = df(a)(h),$$

où dans le membre de gauche, on note $Df(a)h$ le produit de la matrice $Df(a)$ par le vecteur colonne h , et dans le membre de droite on note $df(a)(h)$ l'application linéaire $df(a)$ appliquée à h . Par conséquent, on a la formule de Taylor suivante

$$f(a+h) = f(a) + Df(a)h + o(\|h\|), \quad \text{pour tout } h \text{ de norme assez petite.}$$

Les fonctions $f : \mathbb{R}^n \rightarrow \mathbb{R}$ à valeurs scalaires ($m = 1$) est un cas particulier qu'il convient de distinguer. En effet, dans ce cas, la différentielle $df(a)$ est une application linéaire de \mathbb{R}^n dans \mathbb{R} (i.e. une forme linéaire sur \mathbb{R}^n) et la matrice jacobienne est une matrice ligne à n composantes qui sont les dérivées partielles de f que l'on peut représenter sous forme d'un vecteur colonne.

Définition 3.1.12. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction qui admet des dérivées partielles au point $a \in \mathbb{R}^n$. Le vecteur colonne

$$\nabla f(a) := \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right)^T$$

est appelé *gradient* de f en a .

Remarque 3.1.13. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est différentiable en $a \in \mathbb{R}^n$, la proposition 3.1.8 montre que pour tout $h \in \mathbb{R}^n$, on a

$$df(a)(h) = Df(a)h = \nabla f(a) \cdot h,$$

où $\nabla f(a) \cdot h$ désigne le produit scalaire entre les vecteurs $\nabla f(a)$ et h . De plus, on a la formule de Taylor suivante

$$f(a+h) = f(a) + \nabla f(a) \cdot h + o(\|h\|), \quad \text{pour tout } h \text{ de norme assez petite.}$$

Il est fréquent d'identifier le symbole ∇ à un "vecteur" de coordonnées $\left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right)^T$ que l'on "applique" à une fonction f pour obtenir ∇f . Notons bien que cette identification est purement symbolique et n'a rien de rigoureux, mais elle peut être très utile d'un point de vue mnémotechnique.

Définition 3.1.14. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ une fonction qui admet des dérivées partielles en $a \in \mathbb{R}^n$. On définit

– la *divergence* : Si $m = n$, on note

$$\operatorname{div} f(a) := \nabla \cdot f(a) = \sum_{j=1}^n \frac{\partial f_j}{\partial x_j}(a) = \operatorname{tr} Df(a).$$

– le *rotationnel* : Si $m = n = 3$, on note

$$\operatorname{rot} f(a) := \nabla \wedge f(a) = \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}, \frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}, \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right)(a),$$

et si $m = n = 2$,

$$\operatorname{rot} f(a) := \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}.$$

3.2 Propriétés

Comme le calcul des dérivées partielles se ramène au calcul de dérivées classiques, celles-ci jouissent des mêmes règles que celles connues pour les fonctions d'une seule variable. Les propriétés suivantes sont présentées sans démonstration.

Propriétés 3.2.1. Soient $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ deux fonctions admettant une dérivée partielle par rapport à la i -ème variable en $a \in \mathbb{R}^n$ et soient $\lambda, \mu \in \mathbb{R}$. Alors

- i) si f est indépendante de x_i , $\frac{\partial f}{\partial x_i}(a) = 0$;
- ii) $\lambda f + \mu g$ admet une dérivée partielle en a par rapport à la i -ème variable et

$$\frac{\partial(\lambda f + \mu g)}{\partial x_i}(a) = \lambda \frac{\partial f}{\partial x_i}(a) + \mu \frac{\partial g}{\partial x_i}(a);$$

- iii) fg admet une dérivée partielle en a par rapport à la i -ème variable et

$$\frac{\partial(fg)}{\partial x_i}(a) = f(a) \frac{\partial g}{\partial x_i}(a) + g(a) \frac{\partial f}{\partial x_i}(a).$$

Montrons à présent des propriétés analogues pour la différentielle.

Propriétés 3.2.2. Soient $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ deux fonctions différentiables en $a \in \mathbb{R}^n$ et soient $\lambda, \mu \in \mathbb{R}$. Alors

- i) si f est constante, alors $df(a) = 0$ pour tout $a \in \mathbb{R}^n$;
- ii) si $f \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ est une application linéaire, alors $df(a) = f$ pour tout $a \in \mathbb{R}^n$;
- iii) $\lambda f + \mu g$ est différentiable en a et $d(\lambda f + \mu g)(a) = \lambda df(a) + \mu dg(a)$;
- iv) Si $m = 1$, alors fg est différentiable en a et $d(fg)(a) = f(a)dg(a) + g(a)df(a)$.

Démonstration. i) Si f est constante, alors il existe $c \in \mathbb{R}^m$ tel que $f(x) = c$ pour tout $x \in \mathbb{R}^n$ et donc $f(a+h) - f(a) = 0$ pour tout $a, h \in \mathbb{R}^n$, ce qui montre que $df(a)(h) = 0$.

ii) Si $f \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ est une application linéaire, alors $f(a+h) - f(a) = f(h)$ pour tout $a, h \in \mathbb{R}^n$. Comme f est linéaire, ceci montre que $df(a)(h) = f(h)$.

iii) Si $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ sont deux fonctions différentiables en $a \in \mathbb{R}^n$, alors il existe ε_1 et $\varepsilon_2 : \mathbb{R} \rightarrow \mathbb{R}^m$ telles que $\varepsilon_1(t) \rightarrow 0$ et $\varepsilon_2(t) \rightarrow 0$ quand $t \rightarrow 0$ telles que pour tout $h \in \mathbb{R}^n$ assez petit,

$$\begin{aligned} f(a+h) - f(a) &= df(a)(h) + \|h\| \varepsilon_1(\|h\|), \\ g(a+h) - g(a) &= dg(a)(h) + \|h\| \varepsilon_2(\|h\|), \end{aligned}$$

et par conséquent,

$$\lambda f(a+h) + \mu g(a+h) = \lambda f(a) + \mu g(a) + \lambda df(a)(h) + \mu dg(a)(h) + \|h\|[\lambda \varepsilon_1(\|h\|) + \mu \varepsilon_2(\|h\|)].$$

Comme $h \mapsto \lambda df(a)(h) + \mu dg(a)(h)$ est linéaire et $\varepsilon(t) := \lambda \varepsilon_1(t) + \mu \varepsilon_2(t) \rightarrow 0$ quand $t \rightarrow 0$, on en déduit que $\lambda f + \mu g$ est différentiable en a avec $d(\lambda f + \mu g)(a) = \lambda df(a) + \mu dg(a)$.

iv) Si de plus $m = 1$, alors

$$\begin{aligned} f(a+h)g(a+h) - f(a)g(a) &= f(a+h)[g(a+h) - g(a)] + g(a)[f(a+h) - f(a)] \\ &= f(a+h)[dg(a)(h) + \|h\|\varepsilon_2(\|h\|)] + g(a)[df(a)(h) + \|h\|\varepsilon_1(\|h\|)]. \end{aligned}$$

Par conséquent,

$$\begin{aligned} &\frac{\|f(a+h)g(a+h) - f(a)g(a) - f(a)dg(a)(h) - g(a)df(a)(h)\|}{\|h\|} \\ &= \frac{\|[f(a+h) - f(a)]dg(a)(h) + f(a+h)\|h\|\varepsilon_2(\|h\|) + g(a)\|h\|\varepsilon_1(\|h\|)\|}{\|h\|} \\ &\leq \|f(a+h) - f(a)\| \|dg(a)\| + |f(a+h)|\varepsilon_2(\|h\|) + |g(a)|\varepsilon_1(\|h\|) \rightarrow 0, \end{aligned}$$

car, d'après la proposition 3.1.6, f est continue en a . Ceci montre que fg est différentiable en a et $d(fg)(a) = f(a)dg(a) + g(a)df(a)$. □

Nous établissons à présent une formule de différentiation des fonctions composées.

Proposition 3.2.3. *Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ une fonction différentiable en $a \in \mathbb{R}^n$ et $g : \mathbb{R}^m \rightarrow \mathbb{R}^p$ une fonction différentiable en $f(a) \in \mathbb{R}^m$. Alors $g \circ f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ est différentiable en $a \in \mathbb{R}^n$ et*

$$d(g \circ f)(a) = dg(f(a)) \circ df(a). \quad (3.2.1)$$

Démonstration. Notons $b = f(a)$. Il existe alors des fonctions $\varepsilon_1 : \mathbb{R} \rightarrow \mathbb{R}^m$ et $\varepsilon_2 : \mathbb{R} \rightarrow \mathbb{R}^p$ satisfaisant $\varepsilon_1(t) \rightarrow 0$ et $\varepsilon_2(t) \rightarrow 0$ quand $t \rightarrow 0$ et telles que pour tout $h \in \mathbb{R}^n$ et $k \in \mathbb{R}^m$ assez petits,

$$f(a+h) - f(a) = df(a)(h) + \|h\|\varepsilon_1(\|h\|), \quad (3.2.2)$$

$$g(b+k) - g(b) = dg(b)(k) + \|k\|\varepsilon_2(\|k\|). \quad (3.2.3)$$

Utilisons (3.2.3) avec $k_h = f(a+h) - f(a)$. Par (3.2.2), on voit qu'il existe $c > 0$ tel que $\|k_h\| \leq c\|h\|$ pour h assez petit. Il vient alors, en reportant dans (3.2.3), que

$$\begin{aligned} g(f(a+h)) &= g(f(a) + k_h) = g(f(a)) + dg(f(a))(k_h) + \|k_h\|\varepsilon_2(\|k_h\|) \\ &= g(f(a)) + dg(f(a))[df(a)(h) + \|h\|\varepsilon_1(\|h\|)] + \|k_h\|\varepsilon_2(\|k_h\|). \end{aligned}$$

Par conséquent,

$$\begin{aligned} &\frac{\|g(f(a+h)) - g(f(a)) - [dg(f(a)) \circ df(a)](h)\|}{\|h\|} \\ &= \frac{\|[h\|dg(f(a))\varepsilon_1(\|h\|) + \|k_h\|\varepsilon_2(\|k_h\|)\|}{\|h\|} \\ &\leq \|dg(f(a))\| \varepsilon_1(\|h\|) + c\varepsilon_2(\|h\|) \rightarrow 0, \end{aligned}$$

ce qui conclut la preuve de la proposition. □

AVERTISSEMENT : Contrairement aux fonctions d'une seule variable l'ordre d'apparition des différentielles dans la formule (3.2.1) est extrêmement important. En effet, comme $g \circ f : \mathbb{R}^n \rightarrow \mathbb{R}^p$, sa différentielle en $a \in \mathbb{R}^n$ (quand elle existe) doit être une application linéaire de \mathbb{R}^n dans \mathbb{R}^p . Comme $df(a) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, on en déduit que pour tout $h \in \mathbb{R}^n$, $df(a)(h) \in \mathbb{R}^m$, et comme $dg(f(a)) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^p)$, alors $dg(f(a))(df(a)(h)) \in \mathbb{R}^p$ ce qui montre bien que $dg(f(a)) \circ df(a) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$. Ceci peut se voir également en terme de matrices jacobiniennes. En effet, on a

$$D(g \circ f)(a) = Dg(f(a))Df(a).$$

Comme $Df(a)$ est une matrice $m \times n$ et $Dg(f(a))$ est une matrice $p \times m$, la multiplication de matrices $Dg(f(a))Df(a)$ a un sens car $Dg(f(a))$ a m colonnes et $Df(a)$ a m lignes. En revanche, le produit de matrices dans l'ordre inverse n'est pas bien défini.

Remarque 3.2.4. Il est intéressant d'écrire la formule (3.2.1) en terme de dérivées partielles. Si l'on note $u \in \mathbb{R}^m$ la variable de g , alors pour tout $1 \leq i \leq p$ et $1 \leq j \leq n$, on a

$$\frac{\partial (g \circ f)_i}{\partial x_j}(a) = \sum_{k=1}^m \frac{\partial g_i}{\partial u_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a).$$

Le calcul consiste donc à dériver g_i par rapport à la variable u_k et d'évaluer le résultat en $f(a)$, de multiplier par la dérivée de f_k par rapport à la variable x_j évaluée en a , puis de sommer par rapport à k .

Exemple 3.2.5. Soient $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction différentiable et $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ la fonction définie par

$$\Phi(x, y) := \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}, \quad \text{pour tout } (x, y) \in \mathbb{R}^2.$$

La fonction Φ est différentiable sur \mathbb{R}^2 car c'est une application linéaire. On définit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ par $f = g \circ \Phi$, i.e.,

$$f(x, y) = g(ax + by, cx + dy), \quad \text{pour tout } (x, y) \in \mathbb{R}^2.$$

Alors f est différentiable sur \mathbb{R}^2 et

$$\begin{cases} \frac{\partial f}{\partial x}(x, y) = a \frac{\partial g}{\partial u}(ax + by, cx + dy) + c \frac{\partial g}{\partial v}(ax + by, cx + dy), \\ \frac{\partial f}{\partial y}(x, y) = b \frac{\partial g}{\partial u}(ax + by, cx + dy) + d \frac{\partial g}{\partial v}(ax + by, cx + dy). \end{cases}$$

Le théorème des accroissements finis, bien connu pour les fonctions d'une seule variable, se généralise aux fonctions de plusieurs variables. Dans les résultats qui suivent, les notions de convexité et connexité introduites au chapitre 1 vont se montrer particulièrement utiles.

Théorème 3.2.6 (Accroissements finis). Soit $U \subset \mathbb{R}^n$ un ouvert convexe et $f : U \rightarrow \mathbb{R}^m$ une fonction différentiable. Alors, pour tout $x, y \in U$,

$$\|f(x) - f(y)\| \leq \sup_{z \in [x, y]} \|df(z)\| \|x - y\|,$$

où $[x, y] := \{ty + (1-t)x, t \in [0, 1]\}$ désigne le segment dans \mathbb{R}^n d'extrémités x et y .

Démonstration. Soient x et $y \in U$, l'ensemble U étant convexe, le segment $[x, y]$ est bien inclus dans U . Pour tout $t \in [0, 1]$, on pose

$$g(t) := f(ty + (1-t)x) \cdot [f(y) - f(x)] = \sum_{i=1}^m f_i(ty + (1-t)x)[f_i(y) - f_i(x)].$$

Notons que g est bien définie sur $[0, 1]$ puisque $ty + (1-t)x \in U$ et elle y est par ailleurs continue. De plus, f étant différentiable sur U et $t \mapsto ty + (1-t)x$ étant dérivable sur $]0, 1[$, on en déduit par la formule de différentiation des fonctions composées que g est dérivable sur $]0, 1[$ et

$$g'(t) = [df(ty + (1-t)x)(y-x)] \cdot [f(y) - f(x)], \quad \text{pour tout } t \in]0, 1[.$$

D'après le théorème des accroissements finis en dimension 1, pour tout $1 \leq i \leq m$, il existe un $t_0 \in]0, 1[$ tel que $g(1) - g(0) = g'(t_0)$, soit

$$\|f(y) - f(x)\|^2 = [df(t_0y + (1-t_0)x)(y-x)] \cdot [f(y) - f(x)].$$

Par conséquent, d'après l'inégalité de Cauchy-Schwarz,

$$\|f(y) - f(x)\| \leq \|df(t_0y + (1-t_0)x)\| \|y - x\| \leq \sup_{z \in [x, y]} \|df(z)\| \|x - y\|,$$

ce qui établit le résultat. □

Théorème 3.2.7. Soit $U \subset \mathbb{R}^n$ un ouvert connexe et $f : U \rightarrow \mathbb{R}^m$ telle que $df(x) = 0$ pour tout $x \in U$. Alors f est constante sur U , i.e., il existe $c \in \mathbb{R}^m$ tel que $f(x) = c$ pour tout $x \in U$.

Démonstration. Soient x et $y \in U$. L'ensemble U étant connexe, le théorème 1.4.17 assure l'existence d'une ligne brisée contenue dans U qui relie x et y : il existe $x_1, \dots, x_p \in U$ tels que $x = x_1$, $y = x_p$ et les segments $[x_i, x_{i+1}] \subset U$ pour tout $1 \leq i \leq p-1$. D'après le théorème des accroissements finis,

$$\|f(x_{i+1}) - f(x_i)\| \leq \sup_{z \in [x_i, x_{i+1}]} \|df(z)\| \|x_{i+1} - x_i\| = 0,$$

puisque $df(z) = 0$ pour tout $z \in [x_{i+1}, x_i] \subset U$. D'après l'inégalité triangulaire,

$$\|f(x) - f(y)\| \leq \sum_{i=1}^{p-1} \|f(x_{i+1}) - f(x_i)\| = 0,$$

ce qui montre que f est constante dans U . □

L'hypothèse de connexité est réellement nécessaire comme le montre l'exemple suivant.

Exemple 3.2.8. On note $x_0 = (-1, 0)$ et $y_0 = (1, 0)$. On définit l'ensemble $U = B(x_0, 1/2) \cup B(y_0, 1/2)$. Il s'agit d'un ensemble ouvert car c'est la réunion de deux boules ouvertes mais non connexe puisque U est lui-même l'union de deux ouverts disjoints. Si l'on définit la fonction $f : U \rightarrow \mathbb{R}$ par $f(x, y) = -1$ si $(x, y) \in B(x_0, 1/2)$ et $f(x, y) = 1$ si $(x, y) \in B(y_0, 1/2)$ alors $df(x, y) = 0$ pour tout $(x, y) \in U$, mais pourtant f n'est pas constante sur U .

3.3 Fonctions de classe \mathcal{C}^1

Nous avons vu précédemment qu'une fonction dont toutes les dérivées partielles sont partout définies n'est pas forcément différentiable. La situation est différente lorsque les dérivées partielles sont continues.

Définition 3.3.1. Soient $U \subset \mathbb{R}^n$ un ouvert et $f : U \rightarrow \mathbb{R}^m$. On dit que f est de classe \mathcal{C}^1 sur U si f admet des dérivées partielles et $\frac{\partial f_i}{\partial x_j}$ sont continues sur U pour tout $1 \leq i \leq m$ et tout $1 \leq j \leq n$. On note $\mathcal{C}^1(U; \mathbb{R}^m)$ l'ensemble des fonctions de classe \mathcal{C}^1 sur U à valeurs dans \mathbb{R}^m et, plus simplement $\mathcal{C}^1(U)$ quand $m = 1$.

Théorème 3.3.2. Soient $U \subset \mathbb{R}^n$ un ouvert et $f : U \rightarrow \mathbb{R}^m$. Si $f \in \mathcal{C}^1(U; \mathbb{R}^m)$, alors f est différentiable sur U .

Démonstration. Soit $a \in U$, montrons que f est différentiable en a . Par continuité des dérivées partielles, pour tout $\varepsilon > 0$, il existe un $r > 0$ tel que $\overline{B}(a, r) \subset U$ et

$$\left| \frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a) \right| < \frac{\varepsilon}{n}, \quad \text{pour tout } x \in \overline{B}(a, r), \quad (3.3.1)$$

pour tout $1 \leq i \leq m$ et tout $1 \leq j \leq n$. Soit $h \in \mathbb{R}^n$ avec $\|h\| \leq r$, on pose

$$v^{(0)} = 0, \quad v^{(j)} = h_1 e^1 + \dots + h_j e^j, \quad \text{pour tout } 1 \leq j \leq n$$

de sorte que

$$f(a+h) - f(a) = \sum_{j=1}^n [f(a+v^{(j)}) - f(a+v^{(j-1)})]. \quad (3.3.2)$$

Comme $\|v^{(j)}\| \leq r$ pour tout $0 \leq j \leq n$, on en déduit que $a+v^{(j)} \in \overline{B}(a, r)$ pour tout $0 \leq j \leq n$. Par ailleurs, la boule $\overline{B}(a, r)$ étant convexe, les segments $[a+v^{(j-1)}, a+v^{(j)}]$ sont contenus dans $\overline{B}(a, r)$ pour tout $1 \leq j \leq n$. Posons

$$g_j(t) = f(a+v^{(j-1)} + th_j e^j) \cdot [f(a+v^{(j-1)}) - f(a+v^{(j)})], \quad \text{pour tout } t \in [0, 1].$$

Notons que g_j est bien définie car $a+v^{(j-1)} + th_j e^j = a+tv^{(j)} + (1-t)v^{(j-1)} = t(a+v^{(j)}) + (1-t)(a+v^{(j-1)}) \in \overline{B}(a, r)$. De plus g_j est continue sur le fermé $[0, 1]$ et, par définition des dérivées partielles, g_j est dérivable sur l'ouvert $]0, 1[$ avec

$$g'_j(t) = \sum_{i=1}^m \frac{\partial f_i}{\partial x_j}(a+v^{(j-1)} + th_j e^j) h_j [f_i(a+v^{(j-1)}) - f_i(a+v^{(j)})], \quad \text{pour tout } t \in]0, 1[.$$

En vertu du théorème des accroissements finis en dimension 1, on en déduit l'existence d'un $\theta_j \in [0, 1]$ tel que $g_j(1) - g_j(0) = g'_j(\theta_j)$, soit

$$\|f(a+v^{(j-1)}) - f(a+v^{(j)})\|^2 = \sum_{i=1}^m \frac{\partial f_i}{\partial x_j}(a+v^{(j-1)} + \theta_j h_j e^j) h_j [f_i(a+v^{(j-1)}) - f_i(a+v^{(j)})].$$

Comme $\|a+v^{(j-1)} + \theta_j h_j e^j - a\| = \|v^{(j-1)} + \theta_j h_j e^j\| \leq \|h\| \leq r$, (3.3.1) montre que

$$\left| \frac{\partial f_i}{\partial x_j}(a+v^{(j-1)} + \theta_j h_j e^j) - \frac{\partial f_i}{\partial x_j}(a) \right| < \frac{\varepsilon}{n}.$$

Par conséquent, d'après (3.3.2), il vient

$$\begin{aligned} \left\| f(a+h) - f(a) - \sum_{j=1}^n h_j \frac{\partial f}{\partial x_j}(a) \right\| &\leq \sum_{j=1}^n |h_j| \left\| \frac{\partial f}{\partial x_j}(a) - \frac{\partial f}{\partial x_j}(a + v^{(j-1)} + \theta_j h_j e^j) \right\| \\ &\leq \frac{\varepsilon}{\sqrt{n}} \sum_{j=1}^n |h_j| \leq \varepsilon \|h\|, \quad \text{pour tout } h \in \overline{B}(0, r). \end{aligned}$$

Ceci montre que f est différentiable en a et $df(a)(h) = \sum_{j=1}^n h_j \frac{\partial f}{\partial x_j}(a)$ pour tout $h \in \mathbb{R}^n$. \square

Le théorème 3.3.2 est fort utile en pratique. En effet, il est généralement difficile de montrer directement qu'une fonction f est différentiable sur un ouvert U . Une possibilité consiste donc à montrer que f admet des dérivées partielles sur U et que celles-ci sont continues sur U . Le théorème précédent assure alors la différentiabilité de f sur U .

Exemple 3.3.3. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction définie par

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

La fonction f est bien définie, continue et admet des dérivées partielles sur $\mathbb{R}^2 \setminus \{(0, 0)\}$. De plus, pour tout $(x, y) \neq (0, 0)$, on a

$$\frac{\partial f}{\partial x}(x, y) = y \frac{x^4 + 4x^2y^2 - y^4}{(x^2 + y^2)^2}, \quad \frac{\partial f}{\partial y}(x, y) = x \frac{x^4 - 4x^2y^2 - y^4}{(x^2 + y^2)^2}.$$

Par ailleurs, le calcul des dérivées partielles en $(0, 0)$ donne

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

Les fonctions $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont continues $\mathbb{R}^2 \setminus \{(0, 0)\}$. Pour étudier la continuité des dérivées partielles en $(0, 0)$ on pose $x = r \cos \theta$ et $y = r \sin \theta$ où $r = \sqrt{x^2 + y^2} > 0$ et $\theta \in [0, 2\pi[$. Notons que $(x, y) \rightarrow 0$ si et seulement si $r \rightarrow 0$. Il vient

$$\left| \frac{\partial f}{\partial x}(x, y) \right| \leq r |\sin \theta| |\cos^4 \theta + 4 \cos^2 \theta \sin^2 \theta - \sin^4 \theta| \leq 6r \rightarrow 0 = \frac{\partial f}{\partial x}(0, 0)$$

quand $(x, y) \rightarrow (0, 0)$ ce qui prouve que $\frac{\partial f}{\partial x}$ est continue en $(0, 0)$. On montre de même que $\frac{\partial f}{\partial y}$ est continue en $(0, 0)$. Finalement les dérivées partielles de f sont continues sur \mathbb{R}^2 ce qui assure que f est différentiable sur \mathbb{R}^2 .

Pour les fonctions de \mathbb{R}^n dans lui-même, on introduit la notion suivante.

Définition 3.3.4. Soit $U \subset \mathbb{R}^n$ un ouvert, $f : U \rightarrow \mathbb{R}^n$ de classe \mathcal{C}^1 sur U . On dit que f réalise un \mathcal{C}^1 -difféomorphisme s'il existe un ouvert $V \subset \mathbb{R}^n$ tel que $f : U \rightarrow V$ soit bijective et si sa fonction réciproque $f^{-1} : V \rightarrow U$ est également de classe \mathcal{C}^1 sur V .

Comme

$$\begin{cases} f \circ f^{-1}(y) = y, & \text{pour tout } y \in V, \\ f^{-1} \circ f(x) = x, & \text{pour tout } x \in U, \end{cases}$$

d'après la formule de différentiation des fonctions composées, on en déduit que

$$\begin{cases} df(f^{-1}(y)) \circ d(f^{-1})(y) = \text{Id}, & \text{pour tout } y \in V, \\ d(f^{-1})(f(x)) \circ df(x) = \text{Id}, & \text{pour tout } x \in U, \end{cases}$$

ce qui montre que $df(x)$ est une application linéaire inversible, et son inverse est donnée par

$$[df(x)]^{-1} = d(f^{-1})(f(x)), \quad \text{pour tout } x \in U.$$

En particulier le déterminant jacobien $J_f(x) \neq 0$ pour tout $x \in U$.

AVERTISSEMENT : Si la notion de fonctions de classe \mathcal{C}^1 est définie pour les fonctions de $\mathbb{R}^n \rightarrow \mathbb{R}^m$, celle de \mathcal{C}^1 -difféomorphisme ne l'est que pour des fonctions de \mathbb{R}^n dans lui-même. La dimension de l'espace de départ et d'arrivée doit être la même.

En dimension $n = 1$, si $f : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction de classe \mathcal{C}^1 et $f'(a) \neq 0$, alors f est strictement monotone dans un voisinage de a . Autrement dit, il existe un intervalle ouvert I contenant a , un intervalle ouvert J contenant $f(a)$ tels que f réalise un \mathcal{C}^1 -difféomorphisme de I sur J . Le théorème d'inversion locale étend ce résultat en n'importe quelle dimension. Il s'agit d'un résultat difficile dont la démonstration est admise.

Théorème 3.3.5 (Inversion locale). *Soit U un ouvert de \mathbb{R}^n et $f : U \rightarrow \mathbb{R}^n$ une fonction de classe \mathcal{C}^1 . On suppose qu'il existe un $a \in U$ tel que $J_f(a) \neq 0$. Alors il existe deux ouverts V et W de \mathbb{R}^n tels que $a \in V \subset U$, $f(a) \in W$ et f réalise un \mathcal{C}^1 -difféomorphisme de V sur W .*

Il existe une version globale de ce résultat permettant d'établir qu'une fonction est un \mathcal{C}^1 -difféomorphisme global.

Théorème 3.3.6 (Inversion globale). *Soit U un ouvert de \mathbb{R}^n et $f : U \rightarrow \mathbb{R}^n$ une fonction injective et de classe \mathcal{C}^1 . On suppose que pour tout $x \in U$, $J_f(x) \neq 0$. Alors $f(U)$ est un ouvert de \mathbb{R}^n et f réalise un \mathcal{C}^1 -difféomorphisme de U sur $f(U)$.*

Démonstration. Montrons tout d'abord que $f(U)$ est ouvert dans \mathbb{R}^n . Soit $y \in f(U)$, alors il existe un $x \in U$ tel que $y = f(x)$. Comme par hypothèse $J_f(x) \neq 0$, le théorème d'inversion locale assure l'existence de deux ouverts V_x et W_y tels que $x \in V_x \subset U$, $y \in W_y$ et f réalise un \mathcal{C}^1 -difféomorphisme de V_x sur W_y . En particulier, pour tout $y' \in W_y$, il existe un $x' \in V_x$ tel que $y' = f(x')$, ce qui montre que $y' \in f(U)$ et donc $W_y \subset f(U)$. Comme W_y est un ouvert qui contient y , il existe un $r > 0$ tel que $B(y, r) \subset W_y \subset f(U)$, ce qui établit que $f(U)$ est ouvert.

Montrons maintenant que f réalise un \mathcal{C}^1 -difféomorphisme de U sur $f(U)$. On sait tout d'abord que f est de classe \mathcal{C}^1 sur U . Comme f est injective, on en déduit que $f : U \rightarrow f(U)$ est bijective. Il reste donc à voir que son application réciproque f^{-1} est de classe \mathcal{C}^1 sur $f(U)$. Or, par le théorème d'inversion locale, on sait que pour tout $y \in f(U)$, il existe deux ouverts V et W de \mathbb{R}^n tels que $V \subset U$, $y \in W$ et f réalise un \mathcal{C}^1 -difféomorphisme de V sur W , ce qui montre que f^{-1} est de classe \mathcal{C}^1 sur W . Ceci étant vrai pour tout $y \in f(U)$, on en déduit que f^{-1} est de classe \mathcal{C}^1 sur $f(U)$. \square

3.4 Dérivées partielles d'ordre deux

Tout comme pour les fonctions d'une seule variable, nous avons aussi des notions de dérivées d'ordre supérieur pour les fonctions de plusieurs variables. Nous nous limitons ici à la notion de dérivée partielle d'ordre deux même s'il est possible de définir la différentielle d'ordre $k \geq 2$.

Définition 3.4.1. Soit $U \subset \mathbb{R}^n$ un ouvert et $f : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 . On dit que f admet des dérivées partielles d'ordre 2 en $a \in U$ si les fonctions $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$ admettent des dérivées partielles en a . On note alors pour tout $1 \leq i, j \leq n$,

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) := \frac{\partial}{\partial x_i} \left[\frac{\partial f}{\partial x_j} \right](a)$$

la i -ème dérivée partielle de la j -ème dérivée partielle de f en a . La matrice formée des dérivées partielles d'ordre 2 en a est notée

$$D^2 f(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)_{1 \leq i, j \leq n}$$

est appelée *matrice hessienne* de f au point a .

Un résultat fondamental est la relation de Schwarz qui affirme que lorsque les dérivées partielles secondes sont continues, l'ordre des variables n'importe pas.

Théorème 3.4.2 (Schwarz). Soit $U \subset \mathbb{R}^n$ un ouvert et $f : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 qui admet des dérivées partielles secondes sur U qui sont continues en $a \in U$. Alors pour tout $1 \leq i, j \leq n$,

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

En particulier, la matrice $D^2 f(a)$ est symétrique.

Démonstration. L'ensemble U étant ouvert, il existe un $r > 0$ tel que $\overline{B}(a, r) \subset U$.

Étape 1 : Le cas $n = 2$. Pour tout $h \in \mathbb{R}^2$ avec $\|h\| \leq r$, on pose

$$\Delta_h := [f(a+h) - f(a_1+h_1, a_2)] - [f(a_1, a_2+h_2) - f(a)].$$

Soit $\varphi : [a_1 - r, a_1 + r] \rightarrow \mathbb{R}$ la fonction définie par $\varphi(x) = f(x, a_2 + h_2) - f(x, a_2)$. La fonction φ est continue sur le fermé $[a_1 - r, a_1 + r]$ et dérivable sur l'ouvert $]a_1 - r, a_1 + r[$ avec

$$\varphi'(x) = \frac{\partial f}{\partial x}(x, a_2 + h_2) - \frac{\partial f}{\partial x}(x, a_2), \quad \text{pour tout } x \in]a_1 - r, a_1 + r[.$$

Le théorème des accroissements finis montre alors l'existence d'un $\theta_1 \in [0, 1]$ tel que $\varphi(a_1 + h_1) - \varphi(a_1) = h_1 \varphi'(a_1 + \theta_1 h_1)$, soit

$$\begin{aligned} \Delta_h &= [f(a+h) - f(a_1+h_1, a_2)] - [f(a_1, a_2+h_2) - f(a)] \\ &= h_1 \left[\frac{\partial f}{\partial x}(a_1 + \theta_1 h_1, a_2 + h_2) - \frac{\partial f}{\partial x}(a_1 + \theta_1 h_1, a_2) \right]. \end{aligned}$$

Soit $\psi : [a_2 - r, a_2 + r] \rightarrow \mathbb{R}$ la fonction définie par $\psi(y) = \frac{\partial f}{\partial x}(a_1 + \theta_1 h_1, y)$. La fonction ψ est continue sur le fermé $[a_2 - r, a_2 + r]$ et dérivable sur l'ouvert $]a_2 - r, a_2 + r[$ avec

$$\psi'(y) = \frac{\partial^2 f}{\partial y \partial x}(a_1 + \theta_1 h_1, y), \quad \text{pour tout } y \in]a_2 - r, a_2 + r[.$$

Le théorème des accroissements finis montre alors l'existence d'un $\theta_2 \in [0, 1]$ tel que $\psi(a_2 + h_2) - \psi(a_2) = h_2 \psi'(a_2 + \theta_2 h_2)$, soit

$$\Delta_h = h_1 h_2 \frac{\partial^2 f}{\partial y \partial x}(a_1 + \theta_1 h_1, a_2 + \theta_2 h_2).$$

En échangeant les rôles de x et y , le même argument montre que

$$\Delta_h = h_1 h_2 \frac{\partial^2 f}{\partial x \partial y}(a_1 + \eta_1 h_1, a_2 + \eta_2 h_2),$$

où η_1 et $\eta_2 \in [0, 1]$. Ainsi, pour tout $h \neq (0, 0)$ avec $\|h\| \leq r$, il existe $\theta_1, \theta_2, \eta_1$ et $\eta_2 \in [0, 1]$ tels que

$$\frac{\partial^2 f}{\partial y \partial x}(a_1 + \theta_1 h_1, a_2 + \theta_2 h_2) = \frac{\partial^2 f}{\partial x \partial y}(a_1 + \eta_1 h_1, a_2 + \eta_2 h_2).$$

Comme $\|(\theta_1 h_1, \theta_2 h_2)\| \leq \|h\|$, $\|(\eta_1 h_1, \eta_2 h_2)\| \leq \|h\|$ et les dérivées partielles secondes $\frac{\partial^2 f}{\partial y \partial x}$ et $\frac{\partial^2 f}{\partial x \partial y}$ sont continues en a , on en déduit par passage à la limite quand $h \rightarrow (0, 0)$ que

$$\frac{\partial^2 f}{\partial y \partial x}(a) = \frac{\partial^2 f}{\partial x \partial y}(a).$$

Étape 2 : Le cas $n \geq 3$. Soient $1 \leq i, j \leq n$. Pour tout $(x, y) \in \mathbb{R}^2$ avec $\|(x, y)\| \leq r$, on pose $g(x, y) := f(a + xe^i + ye^j)$. La fonction g est de classe \mathcal{C}^1 sur $B(0, r)$ et elle admet des dérivées partielles première et seconde données, pour tout $(x, y) \in B(0, r)$, par

$$\frac{\partial g}{\partial x}(x, y) = \frac{\partial f}{\partial x_i}(a + xe^i + ye^j), \quad \frac{\partial g}{\partial y}(x, y) = \frac{\partial f}{\partial x_j}(a + xe^i + ye^j)$$

et

$$\frac{\partial^2 g}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a + xe^i + ye^j), \quad \frac{\partial^2 g}{\partial y \partial x}(x, y) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a + xe^i + ye^j).$$

Comme, $\frac{\partial^2 f}{\partial x_i \partial x_j}$ et $\frac{\partial^2 f}{\partial x_j \partial x_i}$ sont continues en a , on en déduit que $\frac{\partial^2 g}{\partial x \partial y}$ et $\frac{\partial^2 g}{\partial y \partial x}$ sont continues en $(0, 0)$. L'étape 1 montre alors que

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 g}{\partial x \partial y}(0, 0) = \frac{\partial^2 g}{\partial y \partial x}(0, 0) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a),$$

ce qui conclut la preuve du théorème. \square

Définition 3.4.3. Soit $U \subset \mathbb{R}^n$ un ouvert et $f : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 . On dit que f est de classe \mathcal{C}^2 sur U si f admet des dérivées partielles secondes en tout point de U qui sont continues sur U . On note $\mathcal{C}^2(U; \mathbb{R}^m)$ l'ensemble des fonctions de classe \mathcal{C}^2 sur U à valeurs dans \mathbb{R}^m et, plus simplement $\mathcal{C}^2(U)$ quand $m = 1$.

De façon générale, on a la définition suivante :

Définition 3.4.4. Soit $U \subset \mathbb{R}^n$ un ouvert et $f : U \rightarrow \mathbb{R}$ une fonction. On dit que f est de classe \mathcal{C}^k ($k \in \mathbb{N}$) si f admet des dérivées partielles sur U d'ordre inférieur ou égal à k qui sont continues sur U . On dit que f est de classe \mathcal{C}^∞ sur U si f est de classe \mathcal{C}^k sur U pour tout $k \in \mathbb{N}$. On note $\mathcal{C}^k(U; \mathbb{R}^m)$ (resp. $\mathcal{C}^\infty(U; \mathbb{R}^m)$) l'ensemble des fonctions de classe \mathcal{C}^k (resp. \mathcal{C}^∞) sur U à valeurs dans \mathbb{R}^m et, plus simplement $\mathcal{C}^k(U)$ (resp. $\mathcal{C}^\infty(U)$) quand $m = 1$.

Définition 3.4.5. Soit $U \subset \mathbb{R}^n$ un ouvert et $f : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^2 sur U . On définit le *Laplacien* de f en a par

$$\Delta f(a) = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}(a) = \text{tr} D^2 f(a).$$

La connaissance de dérivées partielles d'ordre 2 permet de donner une formule de Taylor à l'ordre 2 pour les fonctions de plusieurs variables.

Théorème 3.4.6 (Formule de Taylor). Soit $U \subset \mathbb{R}^n$ un ouvert et $f : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^2 sur U . Pour tout $a \in U$, on a pour h assez petit,

$$f(a+h) = f(a) + \nabla f(a) \cdot h + \frac{1}{2}[D^2 f(a)h] \cdot h + o(\|h\|^2),$$

AVERTISSEMENT : Le terme $[D^2 f(a)h] \cdot h$ est le produit scalaire du vecteur $D^2 f(a)h$ avec h . En effet, comme la matrice hessienne $D^2 f(a)$ est une matrice $n \times n$, le produit matrice/vecteur $D^2 f(a)h$ est un vecteur de \mathbb{R}^n . En terme de dérivées partielles, on peut écrire

$$[D^2 f(a)h] \cdot h = \sum_{i=1}^n [D^2 f(a)h]_i h_i = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_j h_i.$$

Démonstration. L'ensemble U étant ouvert, il existe un $r > 0$ tel que $\overline{B}(a, r) \subset U$. Pour $h \in \mathbb{R}^n$ avec $h \neq 0$ et $\|h\| < r$, on pose $g(t) = f(a + th/\|h\|)$ pour tout $t \in [0, r]$. La fonction g est de classe \mathcal{C}^2 sur $]0, r[$ et d'après la formule de Taylor, on a pour tout $t \in]0, r[$

$$g(t) = g(0) + tg'(0) + \frac{t^2}{2}g''(0) + o(t^2).$$

Calculons les dérivées première et seconde de g en fonction de f . D'après la formule de différentiation des fonctions composées, on a pour tout $t \in [0, r]$,

$$g'(t) = \sum_{i=1}^n \frac{h_i}{\|h\|} \frac{\partial f}{\partial x_i} \left(a + t \frac{h}{\|h\|} \right) = \nabla f \left(a + t \frac{h}{\|h\|} \right) \cdot \frac{h}{\|h\|}$$

et

$$\begin{aligned} g''(t) &= \sum_{i=1}^n \sum_{j=1}^n \frac{h_i h_j}{\|h\|^2} \frac{\partial^2 f}{\partial x_i \partial x_j} \left(a + t \frac{h}{\|h\|} \right) \\ &= \sum_{i=1}^n \frac{h_i}{\|h\|} \left[D^2 f \left(a + t \frac{h}{\|h\|} \right) \frac{h}{\|h\|} \right]_i = \left[D^2 f \left(a + t \frac{h}{\|h\|} \right) \frac{h}{\|h\|} \right] \cdot \frac{h}{\|h\|}. \end{aligned}$$

Par conséquent, en prenant $t = \|h\| < r$, il vient

$$f(a+h) = f(a) + \nabla f(a) \cdot h + \frac{1}{2}[D^2 f(a)h] \cdot h + o(\|h\|^2),$$

ce qui conclut la preuve de la formule de Taylor. \square

3.5 Points critiques et extrema

Définition 3.5.1. Soit $U \subset \mathbb{R}^n$ et $f : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 sur U . On dit que $a \in U$ est un *point critique* de f si $\nabla f(a) = 0$.

Définition 3.5.2. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction et E un sous-ensemble de \mathbb{R}^n . On dit que f admet un *minimum global* sur E au point $a \in E$ si $f(a) \leq f(y)$ pour tout $y \in E$. On dit que f admet un *minimum local* sur E en $a \in E$ s'il existe un ouvert $V \subset \mathbb{R}^n$ contenant a tel que $f(a) \leq f(y)$ pour tout $y \in E \cap V$.

Définition 3.5.3. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction et E un sous-ensemble de \mathbb{R}^n . On dit que f admet un *maximum global* sur E au point $a \in E$ si $f(a) \geq f(y)$ pour tout $y \in E$. On dit que f admet un *maximum local* sur E en $a \in E$ s'il existe un ouvert $V \subset \mathbb{R}^n$ contenant a tel que $f(a) \geq f(y)$ pour tout $y \in E \cap V$.

Nous utilisons la dénomination d'*extremum* pour désigner sans distinction un maximum ou un minimum. Nous avons déjà établi l'existence d'un maximum et d'un minimum global pour toute fonction continue sur un compact (voir le théorème 2.3.3). Nous montrons à présent une condition nécessaire qui assure que tout extremum local est forcément un point critique

Théorème 3.5.4. Soit $U \subset \mathbb{R}^n$ et $f : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 sur U . Si f admet un extremum local sur U en $a \in U$, alors $\nabla f(a) = 0$.

Démonstration. Supposons que f admet un minimum local sur U en $a \in U$. Alors il existe un ouvert V de \mathbb{R}^n contenant a tel que $f(a) \leq f(y)$ pour tout $y \in U \cap V$. L'ensemble $U \cap V$ étant ouvert, il existe un $r > 0$ tel que $\overline{B}(a, r) \subset U \cap V$. Pour tout $t \in]-1, 1[$ et $h \in \mathbb{R}^n$ avec $\|h\| \leq r$, on a que $a + th \in \overline{B}(a, r)$ de sorte que la fonction $g :]-1, 1[\rightarrow \mathbb{R}$ définie par $g(t) = f(a + th)$ pour tout $t \in]-1, 1[$ est bien définie et de classe \mathcal{C}^1 sur $] - 1, 1[$. De plus g admet un minimum local sur $] - 1, 1[$ en 0 car

$$g(0) = f(a) \leq f(a + th) = g(t), \quad \text{pour tout } t \in] - 1, 1[.$$

Si $t > 0$, on a alors

$$\frac{g(t) - g(0)}{t} \geq 0,$$

et par passage à la limite quand $t \rightarrow 0^+$, on obtient $g'(0) \geq 0$. De même si $t < 0$,

$$\frac{g(t) - g(0)}{t} \leq 0,$$

et par passage à la limite quand $t \rightarrow 0^-$, on obtient $g'(0) \leq 0$. Finalement, il vient que $g'(0) = 0$ et la formule de différentiation des fonctions composées montre que $g'(t) = df(a + th)(h)$ pour tout $t \in] - 1, 1[$, et donc

$$df(a)(h) = 0, \quad \text{pour tout } h \in \overline{B}(0, r).$$

Par conséquent, pour tout $v \in \mathbb{R}^n$, en posant $h = rv/\|v\| \in \overline{B}(0, r)$, on obtient que $df(a)(v) = 0$ et donc $df(a) = 0$, ou encore $\nabla f(a) = 0$. \square

Etre un point critique est donc une condition nécessaire pour être un extremum. Elle n'est cependant pas suffisante au regard de l'exemple suivant.

Exemple 3.5.5. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction définie par $f(x, y) = x^2 - y^2$ pour tout $(x, y) \in \mathbb{R}^2$. Il s'agit d'une fonction de classe \mathcal{C}^∞ puisque c'est un polynôme. Cherchons d'abord les points critiques de f . Pour ce faire, on résoud l'équation $\nabla f(x, y) = 0$, i.e.,

$$\frac{\partial f}{\partial x}(x, y) = 2x = 0, \quad \frac{\partial f}{\partial y}(x, y) = -2y = 0,$$

et on trouve que $(x, y) = (0, 0)$. La fonction f admet donc un unique point critique qui est $(0, 0)$. Par ailleurs, on constate que f n'admet ni un minimum local ni un maximum local sur \mathbb{R}^2 en $(0, 0)$. En effet, si U est un ouvert de \mathbb{R}^2 contenant $(0, 0)$, alors pour $\varepsilon > 0$ assez petit $\overline{B}((0, 0), \varepsilon) \subset U$ et

$$(\varepsilon, 0) \in U, \quad f(\varepsilon, 0) = \varepsilon^2 > 0 = f(0, 0) \implies (0, 0) \text{ n'est pas un point de maximum local,}$$

et

$$(0, \varepsilon) \in U, \quad f(0, \varepsilon) = -\varepsilon^2 < 0 = f(0, 0) \implies (0, 0) \text{ n'est pas un point de minimum local.}$$

Afin d'identifier plus précisément les extrema parmi les points critiques, il convient d'utiliser le développement de Taylor à l'ordre 2 pour les fonctions de classe \mathcal{C}^2 .

Théorème 3.5.6. Soit $U \subset \mathbb{R}^n$ et $f : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^2 sur U . Alors

- Si f admet un minimum local sur U au point $a \in U$, alors les valeurs propres de la matrice hessienne $D^2f(a)$ sont toutes positives ou nulles ;
- Si f admet un maximum local sur U au point $a \in U$, alors les valeurs propres de la matrice hessienne $D^2f(a)$ sont toutes négative ou nulles.

Démonstration. D'après le théorème de Schwarz, la matrice hessienne $D^2f(a)$ est symétrique. Un résultat d'algèbre linéaire montre qu'elle est diagonalisable dans une base orthonormée de vecteurs propres : il existe des valeurs propres $\lambda_1 \leq \dots \leq \lambda_n$ et des vecteurs propres $v^1, \dots, v^n \in \mathbb{R}^n$ tels que la famille $\{v^1, \dots, v^n\}$ est une base orthonormée de \mathbb{R}^n et $D^2f(a)v^i = \lambda_i v^i$ pour tout $1 \leq i \leq n$. D'après la formule de Taylor, on a

$$f(a + tv^i) = f(a) + t\nabla f(a) \cdot v^i + \frac{t^2}{2}[D^2f(a)v^i] \cdot v^i + o(t^2), \quad \text{pour tout } 1 \leq i \leq n.$$

Comme f admet un extremum local sur U en $a \in U$, le théorème 3.5.4 montre que $\nabla f(a) = 0$. De plus en utilisant le fait que les vecteurs v^i sont unitaires, on a que $[D^2f(a)v^i] \cdot v^i = \lambda_i v^i \cdot v^i = \lambda_i \|v^i\|^2 = \lambda_i$ et il vient

$$f(a + tv^i) - f(a) = \frac{t^2}{2}\lambda_i + o(t^2), \quad \text{pour tout } 1 \leq i \leq n.$$

Si f admet un minimum local sur U en $a \in U$, il existe un ouvert V de \mathbb{R}^n tel que $f(a) \leq f(x)$ pour tout $x \in U \cap V$. L'ensemble $U \cap V$ étant ouvert, on peut trouver un $r > 0$ tel que $B(a, r) \subset U \cap V$. Pour tout $t \in]-r, r[$ et tout $1 \leq i \leq n$, on a que $a + tv^i \in B(a, r)$ et $f(a) \leq f(a + tv^i)$ ce qui montre, après division par t^2 que $\lambda_i + o(t^2)/t^2 \geq 0$ pour tout $1 \leq i \leq n$. Enfin, par passage à la limite quand $t \rightarrow 0$, il vient que $\lambda_i \geq 0$ pour tout $1 \leq i \leq n$.

Si f admet un maximum local sur U en $a \in U$, un argument similaire montre que les valeurs propres $\lambda_i \leq 0$ pour tout $1 \leq i \leq n$. \square

Dans le cas où la matrice hessienne est non dégénérée, on peut aussi obtenir une condition suffisante assurant qu'un point critique est un point d'extremum local.

Théorème 3.5.7. Soit $U \subset \mathbb{R}^n$, $f : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^2 sur U et $a \in U$ un point critique de f .

- Si les valeurs propres de la matrice hessienne $D^2f(a)$ sont toutes strictement positives, alors f admet un minimum local sur U en $a \in U$;
- Si les valeurs propres de la matrice hessienne $D^2f(a)$ sont toutes strictement négatives, alors f admet un maximum local sur U en $a \in U$.

Démonstration. Comme l'ensemble U est ouvert et $a \in U$, il existe un $r > 0$ tel que $B(a, r) \subset U$. Par ailleurs, soient $\lambda_1 \leq \dots \leq \lambda_n$ les valeurs propres de la matrice hessienne $D^2f(a)$ et $v^1, \dots, v^n \in \mathbb{R}^n$ les vecteurs propres associés tels que la famille $\{v^1, \dots, v^n\}$ est une base orthonormée de \mathbb{R}^n et $D^2f(a)v^i = \lambda_i v^i$ pour tout $1 \leq i \leq n$.

Pour tout $h \in B(0, r)$, il existe donc des réels $h_1, \dots, h_n \in \mathbb{R}$ tels que $h = \sum_{i=1}^n h_i v^i$ et donc $[D^2f(a)h] \cdot h = \sum_{i=1}^n h_i [D^2f(a)v^i] \cdot h = \sum_{i=1}^n \lambda_i h_i v^i \cdot h = \sum_{i=1}^n \lambda_i |h_i|^2$. Le point a étant un point critique de f , il vient d'après la formule de Taylor

$$\begin{aligned} f(a + h) - f(a) &= \nabla f(a) \cdot h + \frac{1}{2}[D^2f(a)h] \cdot h + o(\|h\|^2) \\ &= \frac{1}{2} \sum_{i=1}^n \lambda_i |h_i|^2 + o(\|h\|^2). \end{aligned}$$

Si les valeurs propres sont toutes strictement positives, alors

$$f(a+h) - f(a) \geq \|h\|^2 \left(\frac{\lambda_1}{2} + \varepsilon(\|h\|) \right),$$

où $\varepsilon(t) \rightarrow 0$ quand $t \rightarrow 0$. Il existe donc un $\delta > 0$ tel que pour tout $0 \leq t < \delta$, on a $|\varepsilon(t)| \leq \lambda_1/2$, ce qui implique que

$$f(a+h) - f(a) \geq 0, \quad \text{pour tout } h \in B(0, \delta).$$

En posant $V := B(a, \delta)$, on a donc montré que $f(a) \leq f(y)$ pour tout $y \in V$, ce qui assure que f admet un minimum local sur U en $a \in U$.

Si les valeurs propres sont toutes strictement négatives, un argument similaire montre que f admet un maximum local sur U en $a \in U$. \square

Remarque 3.5.8. En dimension $n = 2$, la matrice hessienne est une matrice symétrique 2×2 et il est facile de connaître le signe des valeurs propres λ_1 et λ_2 . Il suffit de remarquer que $\det D^2 f(a) = \lambda_1 \lambda_2$ et que $\text{tr } D^2 f(a) = \lambda_1 + \lambda_2$. On constate alors que si a est un point critique et $\det D^2 f(a) > 0$ alors, les deux valeurs propres sont non nulles et de même signe, ce qui implique que a est un extremum local. Si $\text{tr } D^2 f(a) > 0$ alors les deux valeurs propres sont strictement positives et a est un point de minimum local, et si $\text{tr } D^2 f(a) < 0$ alors les deux valeurs propres sont strictement négatives et a est un point de maximum local.

Chapitre 4

Courbes et surfaces paramétrées

4.1 Théorème des fonctions implicites

Si $g : \mathbb{R} \rightarrow \mathbb{R}$, alors le graphe de g

$$\mathcal{G} := \{(x, y) \in \mathbb{R}^2 : y = g(x)\}$$

définit une courbe dans le plan donné par les solutions $(x, y) \in \mathbb{R}^2$ de l'équation $f(x, y) = y - g(x) = 0$. Réciproquement, si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ est une fonction de classe \mathcal{C}^1 , on considère l'équation cartésienne

$$f(x, y) = 0$$

et on cherche à comprendre dans quelle mesure cette équation est équivalente à $y = g(x)$, où $g : \mathbb{R} \rightarrow \mathbb{R}$. L'exemple suivant montre qu'en général, ceci n'est possible que localement.

Exemple 4.1.1. Pour tout $(x, y) \in \mathbb{R}^2$, on pose $f(x, y) = x^2 + y^2 - 1$. L'ensemble des solutions à l'équation $f(x, y) = 0$ n'est autre que le cercle unité. Si $y > 0$, alors on peut écrire que $y = \sqrt{1 - x^2}$, si $y < 0$, alors $y = -\sqrt{1 - x^2}$. Par contre, si $y = 0$, alors nécessairement $x = \pm 1$ et les portions du cercle unité au voisinage de ces points ne peuvent pas s'écrire comme un graphe.

Le théorème des fonctions implicites donne certaines conditions suffisantes assurant que tout l'ensemble de niveau d'une fonction de n variables peut être localement donné par le graphe d'une fonction de $n - 1$ variables.

Théorème 4.1.2 (Fonctions implicites). Soit U un ouvert de \mathbb{R}^n et $f : U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 . On suppose qu'il existe $x_0 \in \mathbb{R}^{n-1}$ et $y_0 \in \mathbb{R}$ tels que

$$(x_0, y_0) \in U, \quad f(x_0, y_0) = 0 \quad \text{et} \quad \frac{\partial f}{\partial x_n}(x_0, y_0) \neq 0.$$

Alors, il existe un ouvert $V \subset \mathbb{R}^{n-1}$, un ouvert $W \subset \mathbb{R}$ tels que

$$V \times W \subset U, \quad x_0 \in V, \quad y_0 \in W$$

et il existe une fonction $g : V \rightarrow W$ de classe \mathcal{C}^1 telle que pour tout $(x, y) \in V \times W$,

$$f(x, y) = 0 \quad \iff \quad y = g(x).$$

De plus, on a que pour tout $x \in V$ et pour tout $1 \leq i \leq n - 1$,

$$\frac{\partial g}{\partial x_i}(x) = -\frac{\frac{\partial f}{\partial x_i}(x, g(x))}{\frac{\partial f}{\partial x_n}(x, g(x))}.$$

Démonstration. Pour simplifier, nous présentons la preuve uniquement en dimension $n = 2$. On définit la fonction $F : U \rightarrow \mathbb{R}^2$ par $F(x, y) = (x, f(x, y))$ pour tout $(x, y) \in U$. Il est clair que F est de classe \mathcal{C}^1 sur U et pour tout $(x, y) \in U$,

$$DF(x, y) = \begin{pmatrix} 1 & 0 \\ \frac{\partial f}{\partial x}(x, y) & \frac{\partial f}{\partial y}(x, y) \end{pmatrix}, \quad J_F(x, y) = \frac{\partial f}{\partial y}(x, y).$$

En particulier, $J_F(x_0, y_0) \neq 0$ et le théorème d'inversion locale montre l'existence d'un ouvert $U' \subset U$ tel que $(x_0, y_0) \in U'$ et F réalise un \mathcal{C}^1 -difféomorphisme de U' sur son image. Par ailleurs, comme $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$ et $\frac{\partial f}{\partial y}$ est continue, on peut supposer que $\frac{\partial f}{\partial y}(x, y) \neq 0$ pour tout $(x, y) \in U'$.

Soient I' et J deux intervalles ouverts de \mathbb{R} tels que $x_0 \in I'$, $y_0 \in J$ et $I' \times J \subset U'$. Comme $F(I' \times J)$ est un ouvert de \mathbb{R}^2 qui contient le point $F(x_0, y_0) = (x_0, f(x_0, y_0)) = (x_0, 0)$, on peut encore trouver un intervalle ouvert $I \subset I'$ tel que $x_0 \in I$ et $I \times \{0\} \subset F(I' \times J)$. Par conséquent, pour tout $x \in I$, il existe un unique $(x', y) \in I' \times J$ tel que $(x, 0) = F(x', y) = (x', f(x', y))$, ce qui implique que $x = x'$ et $f(x, y) = 0$. On pose alors $g(x) := y$ de sorte que pour tout $(x, y) \in I \times J$,

$$y = g(x) \iff f(x, y) = 0.$$

Montrons à présent que g est de classe \mathcal{C}^1 sur I . Comme pour tout $x \in I$, $(x, 0) = F(x, g(x))$ et F réalise un \mathcal{C}^1 -difféomorphisme de $I \times J$ sur son image, on en déduit que $(x, g(x)) = (F|_{I \times J})^{-1}(x, 0)$. Ceci montre bien que g est de classe \mathcal{C}^1 sur I . De plus comme

$$f(x, g(x)) = 0, \quad \text{pour tout } x \in I,$$

la formule de dérivation des fonctions composées montre que

$$\frac{\partial f}{\partial x}(x, g(x)) + \frac{\partial f}{\partial y}(x, g(x))g'(x) = 0,$$

soit

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

ce qui conclut la preuve du théorème. □

4.2 Courbes paramétrées

Définition 4.2.1. Un *arc orienté* de classe \mathcal{C}^1 est défini par la donnée d'un couple (I, γ) , où I est un intervalle fermé de \mathbb{R} et $\gamma : t \in \mathbb{R} \mapsto \gamma(t) = (x(t), y(t)) \in \mathbb{R}^2$ est une fonction de classe \mathcal{C}^1 telle que $\gamma|_I$ est injective et $\dot{\gamma}(t) \neq (0, 0)$ pour tout $t \in I$. On note alors

$$\Gamma := \{\gamma(t) \in \mathbb{R}^2 : t \in I\} = \gamma(I).$$

On dit que γ est un paramétrage de Γ . Si $I = [a, b]$, $\gamma(a) = \gamma(b)$ et $\gamma|_{[a, b]}$ est injective, on parle alors de *courbe fermée orientée de classe \mathcal{C}^1* .

Le vecteur $\dot{\gamma}(t)$ est alors tangent à Γ en $\gamma(t)$. On appelle *tangente* en $(x_0, y_0) := \gamma(t_0)$ la droite passant par ce point et parallèle au vecteur $\dot{\gamma}(t_0)$.

Proposition 4.2.2. Soit (I, γ) un arc orienté de classe \mathcal{C}^1 , $t_0 \in I$ et $(x_0, y_0) = \gamma(t_0) \in \Gamma$. Alors la tangente à Γ en (x_0, y_0) admet comme équation paramétrique

$$t \in \mathbb{R} \mapsto \gamma(t_0) + t\dot{\gamma}(t_0).$$

Remarque 4.2.3. – Si $\dot{\gamma}(t) = (0, 0)$, la courbe n'admet pas forcément de tangente. Par exemple la courbe dont un paramétrage est donné par

$$\gamma(t) := \begin{pmatrix} t^3 \\ t^2 \end{pmatrix}, \quad \text{pour tout } t \in [-1, 1]$$

n'admet pas de tangente en $(0, 0)$ qui est un point singulier.

- L'hypothèse d'injectivité signifie que la courbe Γ ne se croise pas elle-même : $t \neq t' \implies \gamma(t) \neq \gamma(t')$.
- L'orientation de l'arc Γ précise le sens dans lequel on le parcourt : de $A = \gamma(a)$ vers $B = \gamma(b)$ si $I = [a, b]$. Si l'on définit $\mu(t) = \gamma(a+b-t)$ pour tout $t \in [a, b]$, alors $\mu([a, b]) = \gamma([a, b]) = \Gamma$, $\mu(a) = \gamma(b) = B$ et $\mu(b) = \gamma(a) = A$. On obtient ainsi un autre paramétrage du même arc, orienté dans le sens opposé (de B vers A). Il n'existe que deux sens de parcours possible : le sens inverse "des aiguilles d'une montre" dit aussi aussi sens *direct*, et le sens inverse dit *indirect*.

Comme le montre le résultat suivant, une paramétrisation n'est pas forcément unique.

Proposition 4.2.4. Soient $([a, b], \gamma)$ un arc orienté de classe \mathcal{C}^1 et $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 telle que $\varphi|_{[c, d]}$ est strictement croissante, $\varphi([c, d]) = [a, b]$ et $\varphi(c) = a$, $\varphi(d) = b$. On définit la fonction $\mu : \mathbb{R} \rightarrow \mathbb{R}^2$ par $\mu = \gamma \circ \varphi$. Alors $([c, d], \mu)$ définit le même arc orienté que $([a, b], \gamma)$ avec la même orientation.

Démonstration. Si $\mu = \gamma \circ \varphi : \mathbb{R} \rightarrow \mathbb{R}^2$, alors μ est de classe \mathcal{C}^1 et $\mu|_{[c, d]}$ est injective. Par ailleurs, φ étant strictement croissante sur $[c, d]$, on a $\varphi'(t) > 0$ pour tout $t \in [c, d]$ et on en déduit que $\dot{\mu}(t) = \dot{\gamma}(\varphi(t))\varphi'(t) \neq (0, 0)$ pour tout $t \in [c, d]$. Par conséquent, $([c, d], \mu)$ définit le même arc orienté que $([a, b], \gamma)$. De plus, comme $\mu(c) = \gamma(\varphi(c)) = \gamma(a)$ et $\mu(d) = \gamma(\varphi(d)) = \gamma(b)$, on en déduit que l'orientation est préservée. \square

Exemple 4.2.5. 1. L'ellipse d'équation cartésienne $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ est un arc orienté de classe \mathcal{C}^1 dont une paramétrisation est

$$t \in [0, 2\pi] \mapsto \begin{pmatrix} a \cos t \\ b \sin t \end{pmatrix}.$$

2. L'hyperbole d'équation cartésienne $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ est la réunion de deux arcs orientés de classe \mathcal{C}^1 dont les paramétrisations sont

$$t \in \mathbb{R} \mapsto \begin{pmatrix} a \cosh t \\ b \sinh t \end{pmatrix}, \quad t \in \mathbb{R} \mapsto \begin{pmatrix} -a \cosh t \\ b \sinh t \end{pmatrix}.$$

3. La parabole d'équation cartésienne $y^2 - x = 0$ est un arc orienté de classe \mathcal{C}^1 dont la paramétrisation est

$$t \in \mathbb{R} \mapsto \begin{pmatrix} t^2 \\ t \end{pmatrix}.$$

Remarque 4.2.6. 1. Le cas des graphes : si $f : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction de classe \mathcal{C}^1 , on définit

$$\gamma(t) := \begin{pmatrix} t \\ f(t) \end{pmatrix}, \quad \text{pour tout } t \in \mathbb{R}.$$

Alors

$$\dot{\gamma}(t) := \begin{pmatrix} 1 \\ f'(t) \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ce qui montre que le graphe de f est un arc orienté de classe \mathcal{C}^1 .

2. *Le cas des ensembles de niveau* : Soit $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 . On définit

$$\mathcal{C} := \{(x, y) \in \mathbb{R}^2 : g(x, y) = 0\}$$

l'ensemble de niveau 0 de la fonction g . On suppose que $(x_0, y_0) \in \mathcal{C}$ et que $\frac{\partial g}{\partial y}(x_0, y_0) \neq 0$. Le théorème des fonctions implicites assure alors l'existence d'intervalles ouverts I et J tels que $x_0 \in I$, $y_0 \in J$ et d'une fonction $f : I \rightarrow J$ de classe \mathcal{C}^1 telle que pour tout $(x, y) \in I \times J$,

$$(x, y) \in \mathcal{C} \iff y = f(x).$$

Autrement dit, localement, l'ensemble de niveau 0 de g est le graphe d'une fonction f et donc un arc paramétré. De plus comme $g(x, f(x)) = 0$ pour tout $x \in I$, on en déduit par la formule de dérivation des fonctions composées que

$$\frac{\partial g}{\partial x}(x, f(x)) + f'(x) \frac{\partial g}{\partial y}(x, f(x)) = \nabla g(x, f(x)) \cdot \begin{pmatrix} 1 \\ f'(x) \end{pmatrix} = 0, \quad \text{pour tout } x \in I.$$

Comme le vecteur $(1, f'(x))^T$ est tangent à \mathcal{C} en $(x, f(x))$, on en déduit que $\nabla g(x, f(x))$ est un vecteur orthogonal à \mathcal{C} au point $(x, f(x))$.

4.3 Surfaces paramétrées

La définition des surfaces paramétrées dans l'espace \mathbb{R}^3 s'inspire de celle des courbes paramétrées, à la différence notable qu'elles ne sont pas paramétrées par des intervalles, mais par des sous-ensembles de \mathbb{R}^2 . On rappelle que le produit vectoriel de deux vecteurs x et y dans \mathbb{R}^3 est donné par

$$x \wedge y = \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}.$$

Quand x et y sont linéairement indépendants, le vecteur $x \wedge y$ est non nul et orthogonal à x et y , i.e. $x \cdot (x \wedge y) = y \cdot (x \wedge y) = 0$.

Définition 4.3.1. Une *surface orientée* de classe \mathcal{C}^1 est définie par la donnée d'un couple (D, Φ) , où D est un fermé de \mathbb{R}^2 et $(s, t) \in \mathbb{R}^2 \mapsto \Phi(s, t) = (x(s, t), y(s, t), z(s, t))$ est une fonction de classe \mathcal{C}^1 telle que $\Phi|_D$ est injective, et pour tout $(s, t) \in D$, les vecteurs

$$\frac{\partial \Phi}{\partial s}(s, t), \quad \frac{\partial \Phi}{\partial t}(s, t)$$

sont linéairement indépendants dans \mathbb{R}^3 . On note

$$\Sigma := \{\Phi(s, t) \in \mathbb{R}^3 : (s, t) \in D\} = \Phi(D).$$

On dit que Φ est une paramétrisation de Σ .

Pour tout $(s, t) \in D$, les vecteurs $\frac{\partial \Phi}{\partial s}(s, t)$ et $\frac{\partial \Phi}{\partial t}(s, t)$ sont tangents à la surface Σ au point $\Phi(s, t)$. Par ailleurs, comme ils sont linéairement indépendants, ils sont non nuls et donc le vecteur

$$\frac{\partial \Phi}{\partial s}(s, t) \wedge \frac{\partial \Phi}{\partial t}(s, t)$$

est également non nul et est orthogonal à la surface Σ en $\Phi(s, t)$. On note

$$N(s, t) := \frac{\frac{\partial \Phi}{\partial s}(s, t) \wedge \frac{\partial \Phi}{\partial t}(s, t)}{\left\| \frac{\partial \Phi}{\partial s}(s, t) \wedge \frac{\partial \Phi}{\partial t}(s, t) \right\|}$$

le vecteur normal unitaire. Comme Φ est de classe \mathcal{C}^1 , ce vecteur dépend continûment de (s, t) . Nous dirons que la surface Σ est orientée suivant le vecteur N .

On appelle *plan tangent* en $(x_0, y_0, z_0) := F(s_0, t_0)$ le plan affine passant par ce point et parallèle au sous-espace vectoriel engendré par les vecteurs

$$\frac{\partial \Phi}{\partial s}(s_0, t_0), \quad \frac{\partial \Phi}{\partial t}(s_0, t_0).$$

Il est donc orthogonal au vecteur N au point (x_0, y_0, z_0) .

Proposition 4.3.2. *Soit (D, Φ) une surface orientée de classe \mathcal{C}^1 . Soit $(s_0, t_0) \in D$ et $(x_0, y_0, z_0) = \Phi(s_0, t_0) \in \Sigma$. On note $N_0 := N(s_0, t_0)$ le vecteur unitaire orthogonal Σ en (x_0, y_0, z_0) . Alors le plan tangent à Σ en (x_0, y_0, z_0) admet comme équation*

$$(x - x_0, y - y_0, z - z_0) \cdot N_0 = 0.$$

Remarque 4.3.3. – Tout comme dans le cas des courbes, l'hypothèse d'injectivité empêche la surface de se croiser.

- Quand les vecteurs $\frac{\partial \Phi}{\partial s}(s, t)$ et $\frac{\partial \Phi}{\partial t}(s, t)$ sont liés, alors leur produit vectoriel est nul et la normale n'est pas définie.
- C'est le vecteur normal N qui précise l'orientation de Σ . Pour changer l'orientation, il suffit de changer N en $-N$ et donc de définir un nouveau paramétrage. Par exemple, on peut permuter les variables s et t , de sorte que la surface est paramétrée par $(t, s) \mapsto \Phi(s, t)$ en sens inverse.

Tout comme dans le cas des courbes, la paramétrisation d'une surface orientée n'est pas unique.

Proposition 4.3.4. *Soient (D_1, Φ_1) une surface orientée de classe \mathcal{C}^1 et $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ un \mathcal{C}^1 -difféomorphisme tel que $\varphi(D_2) = D_1$ et $J_\varphi(s, t) > 0$ pour tout $(s, t) \in D_2$. On définit la fonction $\Phi_2 : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ par $\Phi_2 := \Phi_1 \circ \varphi$. Alors (D_2, Φ_2) définit la même surface orientée que (D_1, Φ_1) avec la même orientation.*

Démonstration. Si $\Phi_2 = \Phi_1 \circ \varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, alors Φ_2 est de classe \mathcal{C}^1 et $\Phi_2|_{D_2}$ est injective. De plus la formule de dérivation des fonctions composées montre que pour tout $(s, t) \in D_2$,

$$\begin{cases} \frac{\partial \Phi_2}{\partial s}(s, t) = \frac{\partial \Phi_1}{\partial u}(\varphi(s, t)) \frac{\partial \varphi_1}{\partial s}(s, t) + \frac{\partial \Phi_1}{\partial v}(\varphi(s, t)) \frac{\partial \varphi_2}{\partial s}(s, t), \\ \frac{\partial \Phi_2}{\partial t}(s, t) = \frac{\partial \Phi_1}{\partial u}(\varphi(s, t)) \frac{\partial \varphi_1}{\partial t}(s, t) + \frac{\partial \Phi_1}{\partial v}(\varphi(s, t)) \frac{\partial \varphi_2}{\partial t}(s, t). \end{cases} \quad (4.3.1)$$

Soient λ et $\mu \in \mathbb{R}$ tels que

$$\lambda \frac{\partial \Phi_2}{\partial s}(s, t) + \mu \frac{\partial \Phi_2}{\partial t}(s, t) = 0,$$

alors, on en déduit que

$$\left(\lambda \frac{\partial \varphi_1}{\partial s}(s, t) + \mu \frac{\partial \varphi_1}{\partial t}(s, t) \right) \frac{\partial \Phi_1}{\partial u}(\varphi(s, t)) + \left(\lambda \frac{\partial \varphi_2}{\partial s}(s, t) + \mu \frac{\partial \varphi_2}{\partial t}(s, t) \right) \frac{\partial \Phi_1}{\partial v}(\varphi(s, t)) = 0,$$

ce qui implique, comme $\frac{\partial \Phi_1}{\partial u}(\varphi(s, t))$ et $\frac{\partial \Phi_1}{\partial v}(\varphi(s, t))$ sont linéairement indépendants, que

$$\lambda \frac{\partial \varphi_1}{\partial s}(s, t) + \mu \frac{\partial \varphi_1}{\partial t}(s, t) = \lambda \frac{\partial \varphi_2}{\partial s}(s, t) + \mu \frac{\partial \varphi_2}{\partial t}(s, t) = 0.$$

Par suite, comme $J_\varphi(s, t) > 0$, alors les vecteurs colonnes de la matrice jacobienne de φ sont linéairement indépendants, ce qui montre que $\lambda = \mu = 0$, autrement dit que $\frac{\partial \Phi_2}{\partial s}(s, t)$ et $\frac{\partial \Phi_2}{\partial t}(s, t)$

sont linéairement indépendants pour tout $(s, t) \in D_2$. On a donc montré que (D_2, Φ_2) définit la même surface orientée que (D_1, Φ_1) . Pour montrer que l'orientation est préservée, pour tout $(s, t) \in D_2$, on calcule à l'aide de (4.3.1)

$$\frac{\partial \Phi_2}{\partial s}(s, t) \wedge \frac{\partial \Phi_2}{\partial t}(s, t) = J_\varphi(s, t) \left[\frac{\partial \Phi_1}{\partial u}(\varphi(s, t)) \wedge \frac{\partial \Phi_1}{\partial v}(\varphi(s, t)) \right],$$

et donc, comme $J_\varphi(s, t) > 0$, on en déduit que $N_2(s, t) = N_1(\varphi(s, t))$. \square

Exemple 4.3.5. 1. La *sphère* centrée à l'origine et de rayon R représentée par l'équation cartésienne $x^2 + y^2 + z^2 = R^2$ est une surface orientée de classe \mathcal{C}^1 dont la paramétrisation est donnée par

$$(\theta, \phi) \in [0, 2\pi] \times [-\pi/2, \pi/2] \mapsto \begin{pmatrix} R \cos \phi \cos \theta \\ R \cos \phi \sin \theta \\ R \sin \phi \end{pmatrix}.$$

2. Le *cylindre* centré à l'origine, d'axe vertical et de rayon R , représenté par l'équation cartésienne $x^2 + y^2 = R^2$ est une surface orientée de classe \mathcal{C}^1 dont la paramétrisation est donnée par

$$(\theta, z) \in [0, 2\pi] \times \mathbb{R} \mapsto \begin{pmatrix} R \cos \theta \\ R \sin \theta \\ z \end{pmatrix}.$$

Remarque 4.3.6. 1. *Le cas des graphes* : si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ est une fonction de classe \mathcal{C}^1 , on définit

$$\Phi(s, t) := \begin{pmatrix} s \\ t \\ f(s, t) \end{pmatrix}, \quad \text{pour tout } (s, t) \in \mathbb{R}^2.$$

Alors les vecteurs

$$\frac{\partial \Phi}{\partial s}(s, t) = \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial s}(s, t) \end{pmatrix}, \quad \frac{\partial \Phi}{\partial t}(s, t) = \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial t}(s, t) \end{pmatrix}$$

sont linéairement indépendants et le vecteur normal est donné par

$$N(s, t) = \frac{1}{\sqrt{1 + \left(\frac{\partial f}{\partial s}(s, t)\right)^2 + \left(\frac{\partial f}{\partial t}(s, t)\right)^2}} \begin{pmatrix} -\frac{\partial f}{\partial s}(s, t) \\ -\frac{\partial f}{\partial t}(s, t) \\ 1 \end{pmatrix},$$

ce qui montre que le graphe de f est une surface orientée de classe \mathcal{C}^1 .

2. *Le cas des ensembles de niveau* : Soit $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 . On définit

$$\mathcal{S} := \{(x, y, z) \in \mathbb{R}^3 : g(x, y, z) = 0\}$$

l'ensemble de niveau 0 de la fonction g . On suppose que $(x_0, y_0, z_0) \in \mathcal{S}$ et que $\frac{\partial g}{\partial z}(x_0, y_0, z_0) \neq 0$. Le théorème des fonctions implicites assure alors l'existence d'ouverts $U \subset \mathbb{R}^2$ et $V \subset \mathbb{R}$ tels que $(x, y) \in U$, $z_0 \in V$ et d'une fonction $f : U \rightarrow V$ de classe \mathcal{C}^1 telle que pour tout $(x, y, z) \in U \times V$,

$$(x, y, z) \in \mathcal{S} \iff z = f(x, y).$$

Autrement dit, localement, l'ensemble de niveau 0 de g est le graphe d'une fonction f et donc une surface orientée de classe \mathcal{C}^1 . De plus comme $g(x, y, f(x, y)) = 0$ pour tout $(x, y) \in U$, on en déduit par la formule de dérivation des fonctions composées que pour tout $(x, y) \in U$,

$$\frac{\partial g}{\partial x}(x, y, f(x, y)) + \frac{\partial g}{\partial z}(x, y, f(x, y)) \frac{\partial f}{\partial x}(x, y) = \nabla g(x, y, f(x, y)) \cdot \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x}(x, y) \end{pmatrix} = 0$$

et

$$\frac{\partial g}{\partial y}(x, y, f(x, y)) + \frac{\partial g}{\partial z}(x, y, f(x, y)) \frac{\partial f}{\partial y}(x, y) = \nabla g(x, y, f(x, y)) \cdot \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y}(x, y) \end{pmatrix} = 0.$$

On en déduit que $\nabla g(x, y, f(x, y))$ est un vecteur orthogonal à \mathcal{S} au point $(x, y, f(x, y))$.

Chapitre 5

Equations aux dérivées partielles

De nombreux phénomènes mécaniques, physiques, biologiques ou économiques se modélisent à l'aide d'équations dont l'inconnue est une fonction de plusieurs variables (on peut penser à la variable temporelle t et à la variable d'espace x) qui fait intervenir les dérivées partielles de la fonction. De telles équations s'appellent *équations aux dérivées partielles*. Dans ce chapitre, nous présentons quelques exemples d'équations aux dérivées partielles ainsi que des méthodes de résolution dans des cas très particuliers.

5.1 Equations elliptiques

Le prototype des équations elliptiques est l'*équation de Laplace*. Etant donnée une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$, on cherche une fonction $u : \mathbb{R}^n \rightarrow \mathbb{R}$ qui satisfait

$$-\Delta u(x) = f(x), \quad \text{pour tout } x \in \mathbb{R}^n. \quad (5.1.1)$$

En électrostatique, f désigne la densité de charge électrique et u est le potentiel électrique. Quand $f = 0$, les solutions de l'équation $\Delta u(x) = 0$ pour tout $x \in \mathbb{R}^n$ sont appelées *fonctions harmoniques*. En dimension $n = 1$, les fonctions harmoniques sont tout simplement les solutions de l'équation différentielle $u''(x) = 0$ pour tout $x \in \mathbb{R}$, et il est immédiat de voir que toutes les solutions sont des fonctions affines de la forme $u(x) = ax + b$ pour tout $x \in \mathbb{R}$, où a et $b \in \mathbb{R}$. L'ensemble des solutions forme donc un espace vectoriel de dimension 2. En revanche, la structure des fonctions harmoniques se complexifie considérablement quand $n \geq 2$. Par exemple pour $n = 2$, si l'on identifie \mathbb{R}^2 avec l'ensemble des nombres complexes, on peut montrer que les parties réelles et imaginaires de toute fonction holomorphe (*i.e.* dérivable au sens des nombres complexes) sont harmoniques.

5.2 Equations paraboliques

L'*équation de la chaleur* est un exemple typique d'équation parabolique. Etant donnée une fonction $u_0 : \mathbb{R}^n \rightarrow \mathbb{R}$, on cherche $u : \mathbb{R}^n \times [0, +\infty[\rightarrow \mathbb{R}$ telle que

$$\begin{cases} \frac{\partial u}{\partial t}(x, t) - \Delta u(x, t) = 0, & \text{pour tout } (x, t) \in \mathbb{R}^n \times]0, +\infty[, \\ u(x, 0) = u_0(x), & \text{pour tout } x \in \mathbb{R}^n. \end{cases}$$

L'inconnue $u(x, t)$ représente la température mesurée à la position $x \in \mathbb{R}^n$ et au temps $t > 0$. En dimension $n = 1$ d'espace, on peut montrer que si u_0 est une fonction de classe \mathcal{C}^2 sur \mathbb{R} telle que

$$\sup_{x \in \mathbb{R}} \{|u_0(x)| + |u_0'(x)|\} < +\infty,$$

alors la solution est donnée pour tout $(x, t) \in \mathbb{R} \times [0, +\infty[$ par la formule

$$u(x, t) = \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{+\infty} u_0(x - y) e^{-\frac{y^2}{4t}} dy.$$

Remarquons que l'intégrale impropre converge puisque u_0 est bornée sur \mathbb{R} et $y \mapsto e^{-\frac{y^2}{4t}}$ est intégrable sur \mathbb{R} . En effectuant le changement de variable $z = y/\sqrt{4t}$ dans l'intégrale précédente, il vient que

$$u(x, t) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} u_0(x - 2z\sqrt{t}) e^{-z^2} dz,$$

ce qui montre que $u(x, 0) = \left(\frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-z^2} dz\right) u_0(x) = u_0(x)$. Par ailleurs, en dérivant (formellement) sous le signe somme, on montre que

$$\frac{\partial^2 u}{\partial x^2}(x, t) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} u_0''(x - 2z\sqrt{t}) e^{-z^2} dz,$$

et

$$\frac{\partial u}{\partial t}(x, t) = -\frac{1}{\sqrt{\pi t}} \int_{-\infty}^{+\infty} u_0'(x - 2z\sqrt{t}) z e^{-z^2} dz.$$

En intégrant par parties (les termes de bord s'annulent car u_0' est bornée sur \mathbb{R} et l'exponentielle s'annule à l'infini), on obtient que

$$\frac{\partial u}{\partial t}(x, t) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} u_0''(x - 2z\sqrt{t}) e^{-z^2} dz = \frac{\partial^2 u}{\partial x^2}(x, t),$$

ce qui montre que u est une solution de l'équation de la chaleur.

5.3 Equations hyperboliques

5.3.1 Equation de transport

Soit $a \in \mathbb{R}^n \setminus \{0\}$ et $u_0 : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 donnée. L'équation de transport consiste à chercher une fonction $u : \mathbb{R}^n \times [0, +\infty[\rightarrow \mathbb{R}$ telle que

$$\begin{cases} \frac{\partial u}{\partial t}(x, t) - a \cdot \nabla u(x, t) = 0, & \text{pour tout } (x, t) \in \mathbb{R}^n \times]0, +\infty[, \\ u(x, 0) = u_0(x), & \text{pour tout } x \in \mathbb{R}^n. \end{cases}$$

Pour résoudre cette équation, on utilise la *méthode des caractéristiques* que nous présentons pour $n = 1$. Dans ce cas, $a \in \mathbb{R} \setminus \{0\}$ et l'équation de transport devient

$$\frac{\partial u}{\partial t}(x, t) - a \frac{\partial u}{\partial x}(x, t) = 0, \quad \text{pour tout } (x, t) \in \mathbb{R} \times]0, +\infty[.$$

Pour tout $(x, t) \in \mathbb{R}^2$, on pose

$$\Phi(x, t) = \begin{pmatrix} x - at \\ x + at \end{pmatrix}.$$

On vérifie aisément que Φ est une application linéaire dont le déterminant de la matrice est donné par $2a \neq 0$. Par conséquent, Φ est inversible et

$$\begin{pmatrix} y \\ z \end{pmatrix} = \Phi(x, t) \iff \Phi^{-1}(y, z) = \begin{pmatrix} \frac{z+y}{2} \\ \frac{z-y}{2a} \end{pmatrix}.$$

Par conséquent, Φ réalise un \mathcal{C}^1 -difféomorphisme de \mathbb{R}^2 dans \mathbb{R}^2 . Posons $v(y, z) = u \circ \Phi^{-1}(y, z)$ pour tout $(y, z) \in \mathbb{R}^2$. La formule de dérivation des fonctions composées montre que

$$\frac{\partial v}{\partial z}(y, z) = \nabla u(\Phi^{-1}(y, z)) \cdot \frac{\partial \Phi^{-1}}{\partial z}(y, z) = \nabla u(\Phi^{-1}(y, z)) \cdot \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2a} \end{pmatrix} = \frac{1}{2a} \left(a \frac{\partial u}{\partial x} + \frac{\partial u}{\partial t} \right) (\Phi^{-1}(y, z)).$$

Ainsi, u est solution de l'équation de transport si et seulement $\frac{\partial v}{\partial z}(y, z) = 0$. Autrement dit, si et seulement s'il existe une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ telle que $v(y, z) = f(y)$ pour tout $(y, z) \in \mathbb{R}^2$. Par conséquent, les solutions de l'équation de transport sont les fonctions u de la forme $u(x, t) = v(\Phi(x, t)) = f(x - at)$ pour tout $(x, t) \in \mathbb{R} \times [0, +\infty[$. Pour déterminer la fonction f , on utilise la condition initiale qui assure que pour tout $x \in \mathbb{R}$, on a que $u(x, 0) = u_0(x)$ et donc $f(x) = u_0(x)$. Ceci montre que la solution de l'équation de transport est donnée par

$$u(x, t) = u_0(x - at), \quad \text{pour tout } (x, t) \in \mathbb{R} \times [0, +\infty[.$$

5.3.2 Equation des ondes

L'équation des ondes régit la propagation d'une onde électromagnétique. Etant donné un réel $a \neq 0$ et des fonctions $u_0 : \mathbb{R}^n \rightarrow \mathbb{R}$ continue et $u_1 : \mathbb{R}^n \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 , elle s'écrit

$$\begin{cases} \frac{\partial^2 u}{\partial t^2}(x, t) - a^2 \Delta u(x, t) = 0, & \text{pour tout } (x, t) \in \mathbb{R}^n \times]0, +\infty[, \\ u(x, 0) = u_0(x), & \text{pour tout } x \in \mathbb{R}^n, \\ \frac{\partial u}{\partial t}(x, 0) = u_1(x), & \text{pour tout } x \in \mathbb{R}^n. \end{cases}$$

Tout comme l'équation de transport, elle peut se résoudre grâce à la méthode des caractéristiques. De nouveau, en dimension $n = 1$, l'équation des ondes devient

$$\frac{\partial^2 u}{\partial t^2}(x, t) - a^2 \frac{\partial^2 u}{\partial x^2}(x, t) = 0, \quad \text{pour tout } (x, t) \in \mathbb{R} \times]0, +\infty[.$$

Avec les mêmes notations qu'à la section précédente, on calcule

$$\frac{\partial^2 v}{\partial y \partial z}(y, z) = \nabla \left(\frac{1}{2} \frac{\partial u}{\partial x} + \frac{1}{2a} \frac{\partial u}{\partial t} \right) (\Phi^{-1}(y, z)) \cdot \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2a} \end{pmatrix} = \frac{1}{4a^2} \left(\frac{\partial^2 u}{\partial t^2}(x, t) - a^2 \frac{\partial^2 u}{\partial x^2}(x, t) \right) (\Phi^{-1}(y, z)).$$

Ainsi, u est solution de l'équation des ondes si et seulement $\frac{\partial^2 v}{\partial y \partial z}(y, z) = 0$. Il existe donc une fonction $g : \mathbb{R} \rightarrow \mathbb{R}$ telle que $\frac{\partial^2 v}{\partial y \partial z}(y, z) = g(z)$. En intégrant cette nouvelle équation différentielle, on trouve que $v(y, z) = F(y) + G(z)$, où F et $G : \mathbb{R} \rightarrow \mathbb{R}$ et donc

$$u(x, t) = F(x - at) + G(x + at), \quad \text{pour tout } (x, t) \in \mathbb{R} \times [0, +\infty[.$$

De nouveau, on utilise les conditions initiales pour déterminer F et G . En effet, pour tout $x \in \mathbb{R}$,

$$u_0(x) = u(x, 0) = F(x) + G(x), \quad u_1(x) = \frac{\partial u}{\partial t}(x, 0) = -aF'(x) + aG'(x).$$

On peut résoudre facilement ce système différentiel en dérivant la première équation, en la multipliant par a , puis en l'additionnant et la soustrayant à la seconde équation. Il vient alors que

$$G'(x) = \frac{au_0'(x) + u_1(x)}{2a}, \quad F'(x) = \frac{au_0'(x) - u_1(x)}{2a},$$

soit

$$G(x) = \frac{1}{2}u_0(x) + \frac{1}{2a} \int_0^x u_1(y) dy + \beta, \quad F(x) = \frac{1}{2}u_0(x) - \frac{1}{2a} \int_0^x u_1(y) dy + \alpha,$$

où α et $\beta \in \mathbb{R}$ sont des constantes. Par conséquent, pour tout $(x, t) \in \mathbb{R} \times [0, +\infty[$

$$u(x, t) = \frac{1}{2}u_0(x + at) + \frac{1}{2a} \int_0^{x+at} u_1(y) dy + \frac{1}{2}u_0(x - at) - \frac{1}{2a} \int_0^{x-at} u_1(y) dy + c,$$

où $c = \alpha + \beta$. En utilisant de nouveau la condition initiale $u(x, 0) = u_0(x)$, on trouve que $c = 0$ et donc que la solution de l'équation des ondes est donnée par la fonction

$$u(x, t) = \frac{1}{2}u_0(x + at) + \frac{1}{2a} \int_0^{x+at} u_1(y) dy + \frac{1}{2}u_0(x - at) - \frac{1}{2a} \int_0^{x-at} u_1(y) dy$$

pour tout $(x, t) \in \mathbb{R} \times [0, +\infty[$.

Chapitre 6

Intégrales multiples

6.1 Pavés, ensembles pavables et quarrables

Nous introduisons tout d'abord une notion d'ensemble élémentaire dans \mathbb{R}^n .

Définition 6.1.1. On dit qu'une partie P de \mathbb{R}^n est un *pavé* s'il existe $a = (a_1, \dots, a_n)$, $b = (b_1, \dots, b_n) \in \mathbb{R}^n$ tel que

$$P := \prod_{i=1}^n [a_i, b_i] = \{x \in \mathbb{R}^n : a_i \leq x_i \leq b_i \text{ pour tout } 1 \leq i \leq n\}.$$

Remarque 6.1.2. 1. Un pavé est un ensemble fermé et borné, donc compact.

2. En général, si P_1 et P_2 sont deux pavés, $P_1 \cup P_2$ et $P_1 \setminus P_2$ ne sont pas de pavés (faire un dessin pour s'en convaincre). Par contre l'intersection de deux pavés en est toujours un. En effet, si $P_1 := \prod_{i=1}^n [a_i^1, b_i^1]$ et $P_2 := \prod_{i=1}^n [a_i^2, b_i^2]$ sont deux pavés, alors

$$P_1 \cap P_2 = \prod_{i=1}^n [\max(a_i^1, a_i^2), \min(b_i^1, b_i^2)].$$

Nous définissons à présent la mesure d'un pavé qui correspond intuitivement à la longueur d'un intervalle quand $n = 1$, à l'aire d'un rectangle quand $n = 2$ et au volume d'un parallélépipède rectangle quand $n = 3$.

Définition 6.1.3. Soit $P = \prod_{i=1}^n [a_i, b_i]$ un pavé de \mathbb{R}^n . On appelle *mesure* de P le réel

$$m(P) := \prod_{i=1}^n (b_i - a_i).$$

On pose, par convention, $m(\emptyset) = 0$.

La seule notion de pavé est insuffisante pour décrire une classe raisonnable de sous-ensembles de \mathbb{R}^n dont on souhaite définir la notion de mesure car celle-ci n'est pas stable par les opérations ensemblistes élémentaires (voir la Remarque 6.1.2–2). Une classe plus satisfaisante est celle formée de toutes les unions finies de pavés.

Définition 6.1.4. Une partie A de \mathbb{R}^n est dite *pavable* s'il existe des pavés P_1, \dots, P_N d'intérieurs deux à deux disjoints, i.e., $\overset{\circ}{P}_i \cap \overset{\circ}{P}_j = \emptyset$ pour tout $i \neq j$, et tels que

$$A = \bigcup_{i=1}^N P_i.$$

On définit alors la mesure de A par

$$m(A) := \sum_{i=1}^N m(P_i).$$

Remarque 6.1.5. La représentation d'un ensemble pavable en une union finie de pavés d'intérieurs deux à deux disjoints n'est pas unique. On peut alors démontrer que la définition de la mesure d'un ensemble pavable est indépendante de la représentation.

Comme le montre le résultat suivant, les ensembles pavables sont stables par les opérations ensemblistes usuelles, ce qui les rend beaucoup plus flexibles à manipuler que les pavés.

Proposition 6.1.6. Soient A et B deux ensembles pavables de \mathbb{R}^n . Alors $A \cup B$, $A \cap B$ et $A \setminus \overset{\circ}{B}$ sont pavables.

Démonstration. Soient A et B deux ensembles pavables de \mathbb{R}^n . Alors il existe des pavés P_1, \dots, P_N et Q_1, \dots, Q_M tels que $\overset{\circ}{P}_i \cap \overset{\circ}{P}_{i'} = \emptyset$ si $i \neq i'$, $\overset{\circ}{Q}_j \cap \overset{\circ}{Q}_{j'} = \emptyset$ si $j \neq j'$ et

$$A = \bigcup_{i=1}^N P_i, \quad B = \bigcup_{j=1}^M Q_j.$$

Par conséquent,

$$A \cup B = \left(\bigcup_{i=1}^N P_i \right) \cup \left(\bigcup_{j=1}^M Q_j \right), \quad A \cap B = \left(\bigcup_{i=1}^N P_i \right) \cap \left(\bigcup_{j=1}^M Q_j \right) = \bigcup_{i=1}^N \bigcup_{j=1}^M P_i \cap Q_j$$

où $P_i \cap Q_j$ est un pavé de \mathbb{R}^n , ce qui montre que $A \cup B$ et $A \cap B$ sont pavables. Par ailleurs,

$$A \setminus \overset{\circ}{B} = \bigcup_{i=1}^N \bigcap_{j=1}^M P_i \setminus \overset{\circ}{Q}_j$$

et comme P_i et Q_j sont des pavés, alors $P_i \setminus \overset{\circ}{Q}_j$ est pavable, ce qui montre que $A \setminus \overset{\circ}{B}$ est pavable. \square

Intuitivement, une partie de \mathbb{R}^n a une mesure que l'on peut "calculer" si l'on peut l'approcher par la mesure d'une union de pavés. La définition suivante clarifie cette approximation.

Définition 6.1.7. Soit D un ensemble compact de \mathbb{R}^n tel que $\overset{\circ}{D} \neq \emptyset$. On note

$$\begin{aligned} m^+(D) &:= \inf\{m(B) : B \supset A \text{ avec } B \text{ pavable}\}, \\ m^-(D) &:= \sup\{m(A) : A \subset D \text{ avec } A \text{ pavable}\}. \end{aligned}$$

On dit que D est *quarrable* si $m^+(D) = m^-(D) \in \mathbb{R}$. La valeur commune est notée $|D|$ et est appelée *mesure* de D .

Remarque 6.1.8. Un ensemble pavable est quarrable et on a $|D| = m(D)$. Autrement dit, les deux notions de mesures introduites précédemment coïncident pour les ensembles pavables.

Exemple 6.1.9. Soient f et $g : [a, b] \rightarrow [0, +\infty[$ deux fonctions continues telles que $f \leq g$ sur $[a, b]$ et

$$D := \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, f(x) \leq y \leq g(x)\}.$$

Alors D est quarrable et par définition de l'intégrale de Riemann $|D| = \int_a^b [g(x) - f(x)] dx$.

Nous montrons ci-dessous que tout ensemble quarrable a une frontière de mesure arbitrairement petite. Il s'agit en fait d'une caractérisation, néanmoins comme dans la suite nous n'aurons besoin que de la condition nécessaire, nous n'énonçons et ne démontrons que cette implication.

Proposition 6.1.10. Soit $D \subset \mathbb{R}^n$ un ensemble quarrable. Alors, pour tout $\varepsilon > 0$ il existe des pavés P_1, \dots, P_N tels que $\mathring{P}_i \cap \mathring{P}_j = \emptyset$ pour tout $i \neq j$, $\partial D \subset \bigcup_{i=1}^N P_i$ et

$$\sum_{i=1}^N m(P_i) \leq \varepsilon.$$

Démonstration. Comme D est quarrable, alors $m^+(D) = m^-(D) \in \mathbb{R}$. Par conséquent, pour tout $\varepsilon > 0$ il existe deux ensembles pavables A et B tels que $A \subset D \subset B$ et $m(B) - m(A) \leq \varepsilon$. D'après la Proposition 6.1.6, $B \setminus \mathring{A}$ est un ensemble pavable et $\partial D \subset B \setminus \mathring{A}$. Soient P_1, \dots, P_N des pavés tels que $\mathring{P}_i \cap \mathring{P}_j = \emptyset$ pour tout $i \neq j$ et $B \setminus \mathring{A} = \bigcup_{i=1}^N P_i$. Comme $B = A \cup (B \setminus \mathring{A}) = A \cup \bigcup_{i=1}^N P_i$, on a que

$$m(B) = m(A) + \sum_{i=1}^N m(P_i)$$

ce qui montre effectivement que

$$\sum_{i=1}^N m(P_i) \leq \varepsilon,$$

comme annoncé. □

6.2 Définition et propriétés de l'intégrale multiple

Définition 6.2.1. Soit $U \subset \mathbb{R}^n$ un ouvert, $f : U \rightarrow \mathbb{R}$ une fonction et $D \subset U$ un ensemble quarrable. On définit

$$I^-(f, D) = \sup \left\{ \sum_{i=1}^N \left(\inf_{P_i} f \right) m(P_i) : P_1, \dots, P_N \text{ pavés avec } \mathring{P}_i \cap \mathring{P}_j = \emptyset \text{ si } i \neq j \text{ et } \bigcup_{i=1}^N P_i \subset D \right\}$$

$$I^+(f, D) = \inf \left\{ \sum_{i=1}^N \left(\sup_{P_i} f \right) m(P_i) : P_1, \dots, P_N \text{ pavés avec } \mathring{P}_i \cap \mathring{P}_j = \emptyset \text{ si } i \neq j \text{ et } \bigcup_{i=1}^N P_i \supset D \right\}.$$

On dit que f est *Riemann-intégrable* sur D si

$$I^+(f, D) = I^-(f, D) \in \mathbb{R}.$$

Dans ce cas, la valeur commune est appelée *intégrale (de Riemann)* de f sur D et est notée

$$\int \cdots \int_D f(x_1, \dots, x_n) dx_1 \cdots dx_n$$

ou plus simplement

$$\int_D f(x) dx.$$

Si $n = 2$, on note aussi

$$\iint_D f(x, y) dx dy$$

et si $n = 3$,

$$\iiint_D f(x, y, z) dx dy dz.$$

Nous énonçons tout d'abord des premières propriétés élémentaires qui découlent directement de la définition.

Propriétés 6.2.2. i) Si $f(x) = c$ pour tout $x \in \mathbb{R}^n$ est une fonction constante, alors les Définitions 6.1.7 et 6.2.1 de la mesure et de l'intégrale montrent que

$$\int_D c dx = c|D|.$$

ii) D'après la définition, il est clair que

$$I^+(-f, D) = -I^-(f, D), \quad I^-(-f, D) = -I^+(f, D)$$

de sorte que si f est Riemann-intégrable sur D , alors

$$\int_D (-f(x)) dx = - \int_D f(x) dx.$$

iii) *Monotonie.* Si f et $g : U \rightarrow \mathbb{R}$ sont des fonctions Riemann-intégrables sur D telles que $f(x) \leq g(x)$ pour tout $x \in U$, on obtient immédiatement que

$$\int_D f(x) dx \leq \int_D g(x) dx.$$

Vérifier l'intégrabilité au sens de Riemann d'une fonction à partir de la définition seule est une chose difficile en général. Le résultat suivant montre que les fonctions continues sont toujours Riemann-intégrables.

Théorème 6.2.3. Soit $U \subset \mathbb{R}^n$ un ouvert, $f : U \rightarrow \mathbb{R}$ une fonction continue et $D \subset U$ un ensemble quarrable. Alors f est Riemann-intégrable sur D . De plus, si, pour tout $\delta > 0$, $\{P_i^\delta\}_{i \in \mathbb{N}}$ désigne une famille de pavés de \mathbb{R}^n tels que

$$\mathbb{R}^n = \bigcup_{i \in \mathbb{N}} P_i^\delta, \quad \text{diam}(P_i^\delta) \leq \delta \text{ pour tout } i \in \mathbb{N}, \quad \overset{\circ}{P}_i^\delta \cap \overset{\circ}{P}_j^\delta = \emptyset \text{ pour tout } i \neq j,$$

et

$$I_\delta := \{i \in \mathbb{N} \text{ tels que } P_i^\delta \subset D\} \quad \text{et} \quad J_\delta := \{i \in \mathbb{N} \text{ tels que } P_i^\delta \cap D \neq \emptyset\},$$

alors

$$\int_D f(x) dx = \lim_{\delta \rightarrow 0} \sum_{i \in I_\delta} \left(\inf_{P_i^\delta} f \right) m(P_i^\delta) = \lim_{\delta \rightarrow 0} \sum_{i \in J_\delta} \left(\sup_{P_i^\delta} f \right) m(P_i^\delta).$$

Démonstration. **Étape 1 :** La fonction f étant continue sur D compact, f est bornée sur D et donc il existe une constante $C > 0$ telle que $|f(x)| \leq C$ pour tout $x \in D$. On en déduit alors de la Remarque 6.2.2 que

$$-C|D| \leq I^\pm(f, D) \leq C|D|$$

d'où

$$I^\pm(f, D) \in \mathbb{R}.$$

Étape 2 : Comme f est continue et D est compact, f est uniformément continue sur D . Par conséquent, pour tout $\varepsilon > 0$ il existe un $\delta_0 > 0$ tel que pour tout x et $y \in D$, si $\|x - y\| < \delta_0$ alors $|f(x) - f(y)| < \varepsilon$.

Pour tout $\delta < \delta_0$, soit $\{P_i^\delta\}_{i \in \mathbb{N}}$ une décomposition de \mathbb{R}^n comme dans l'énoncé du théorème. Alors les ensembles pavables

$$A_\delta := \bigcup_{i \in I_\delta} P_i^\delta \quad \text{et} \quad B_\delta := \bigcup_{i \in J_\delta} P_i^\delta$$

satisfont les propriétés $A_\delta \subset D \subset B_\delta \subset U$ pour δ assez petit et $\partial D \subset B_\delta \setminus \overset{\circ}{A}_\delta = \bigcup_{i \in J_\delta \setminus I_\delta} P_i^\delta$.

La fonction f étant continue sur P_i^δ qui est compact, elle atteint ses bornes : il existe donc x_i^δ et $y_i^\delta \in P_i^\delta$ tels que

$$f(x_i^\delta) = \inf_{x \in P_i^\delta} f = m_i^\delta, \quad f(y_i^\delta) = \sup_{x \in P_i^\delta} f = M_i^\delta.$$

Par conséquent, comme $M_i^\delta - m_i^\delta = f(y_i^\delta) - f(x_i^\delta) < \varepsilon$ pour tout $i \in I_\delta$ et $M_i^\delta \leq C$ pour tout $i \in J_\delta \setminus I_\delta$, il vient alors que

$$0 \leq I^+(f, D) - I^-(f, D) \leq \sum_{i \in J_\delta} M_i^\delta |P_i^\delta| - \sum_{i \in I_\delta} m_i^\delta |P_i^\delta| \leq C \sum_{i \in J_\delta \setminus I_\delta} |P_i^\delta| + \varepsilon \sum_{i \in I_\delta} |P_i^\delta|. \quad (6.2.1)$$

D'après la Proposition 6.1.10, comme D est quarrable et $\partial D \subset B_\delta \setminus \overset{\circ}{A}_\delta = \bigcup_{i \in J_\delta \setminus I_\delta} P_i^\delta$, on en déduit que

$$\lim_{\delta \rightarrow 0} \sum_{i \in J_\delta \setminus I_\delta} |P_i^\delta| = 0.$$

Par ailleurs, la Définition 6.1.7 de la mesure d'un ensemble pavable montre que

$$\sum_{i \in I_\delta} |P_i^\delta| \leq |D|.$$

En regroupant les deux informations précédent, on obtient par passage à la limite quand $\delta \rightarrow 0$ que

$$I^+(f, D) - I^-(f, D) \leq \varepsilon |D|,$$

soit $I^+(f, D) = I^-(f, D) \in \mathbb{R}$ puisque $\varepsilon > 0$ est arbitraire, ce qui montre que f est Riemann-intégrable sur D . De plus, d'après (6.2.1), on a également

$$I^+(f, D) - \sum_{i \in I_\delta} m_i^\delta |P_i^\delta| \leq C \sum_{i \in J_\delta \setminus I_\delta} |P_i^\delta| + \varepsilon |D|, \quad \sum_{i \in J_\delta} M_i^\delta |P_i^\delta| - I^-(f, D) \leq C \sum_{i \in J_\delta \setminus I_\delta} |P_i^\delta| + \varepsilon |D|$$

soit par passage à la limite d'abord quand $\delta \rightarrow 0$ puis $\varepsilon \rightarrow 0$,

$$\int_D f(x) dx = \lim_{\delta \rightarrow 0} \sum_{i \in I_\delta} m_i^\delta |P_i^\delta| = \lim_{\delta \rightarrow 0} \sum_{i \in J_\delta} M_i^\delta |P_i^\delta|,$$

ce qui termine la preuve du théorème. \square

Nous établissons maintenant un certain nombre de propriétés utiles en pratique.

Propriétés 6.2.4. 1. *Additivité.* Soient $U \subset \mathbb{R}^n$ un ouvert, D_1 et $D_2 \subset U$ deux ensembles quarrables tels que $\overset{\circ}{D}_1 \cap \overset{\circ}{D}_2 = \emptyset$, et $f : U \rightarrow \mathbb{R}$ une fonction Riemann-intégrable sur D_1 et D_2 . Alors f est Riemann-intégrable sur $D_1 \cup D_2$ et

$$\int_{D_1 \cup D_2} f(x) dx = \int_{D_1} f(x) dx + \int_{D_2} f(x) dx.$$

2. *Linéarité.* Soient $U \subset \mathbb{R}^n$ un ouvert, $D \subset U$ un ensemble quarrable, f et $g : U \rightarrow \mathbb{R}$ des fonctions Riemann-intégrables sur D et $\lambda \in \mathbb{R}$. Alors $f + \lambda g$ est Riemann-intégrable sur D et

$$\int_D [f(x) + \lambda g(x)] dx = \int_D f(x) dx + \lambda \int_D g(x) dx.$$

3. *Inégalité triangulaire.* Soient $U \subset \mathbb{R}^n$ un ouvert, $D \subset U$ un ensemble quarrable et $f : U \rightarrow \mathbb{R}$ une fonction Riemann-intégrable sur D telle que $|f|$ est aussi Riemann-intégrable sur D . Alors,

$$\left| \int_D f(x) dx \right| \leq \int_D |f(x)| dx.$$

Démonstration. 1. Pour $k = 1, 2$, soient $A_k = \bigcup_{i=1}^{M_k} P_i^k$ et $B_k = \bigcup_{i=1}^{N_k} Q_i^k$ des ensembles pavables tels que $A_k \subset D_k \subset B_k \subset U$ et

$$\overset{\circ}{P}_i^k \cap \overset{\circ}{P}_j^k = \overset{\circ}{Q}_i^k \cap \overset{\circ}{Q}_j^k = \emptyset \text{ si } i \neq j.$$

Alors $A_1 \cup A_2 = \bigcup_{i=1}^{M_1} P_i^1 \cup \bigcup_{j=1}^{M_2} P_j^2$ est pavable et $A_1 \cup A_2 \subset D_1 \cup D_2$. Comme $A_1 \cap A_2 = \emptyset$, on a que $\overset{\circ}{P}_i^1 \cap \overset{\circ}{P}_j^2 = \emptyset$ pour tout $1 \leq i \leq M_1$ et $1 \leq j \leq M_2$ de sorte que

$$I^-(f, D_1 \cup D_2) \geq \sum_{i=1}^{M_1} \left(\inf_{P_i^1} f \right) m(P_i^1) + \sum_{j=1}^{M_2} \left(\inf_{P_j^2} f \right) m(P_j^2),$$

puis par passage au sup en $\{P_i^1\}_{1 \leq i \leq M_1}$ et $\{P_j^2\}_{1 \leq j \leq M_2}$, il vient

$$I^-(f, D_1 \cup D_2) \geq \int_{D_1} f(x) dx + \int_{D_2} f(x) dx.$$

L'ensemble $B_1 \cup B_2 = \bigcup_{i=1}^{N_1} Q_i^1 \cup \bigcup_{j=1}^{N_2} Q_j^2$ est pavable et $B_1 \cup B_2 \supset D_1 \cup D_2$. Cependant comme $B_1 \cap B_2 \neq \emptyset$, les cubes Q_i^1 et Q_j^2 ne sont pas forcément d'intérieurs deux à deux disjoints. On écrit alors que

$$B_1 \cup B_2 = (B_1 \setminus \overset{\circ}{B}_2) \cup B_2$$

où $B_1 \setminus \overset{\circ}{B}_2 = \bigcup_{i=1}^{N_1} Q_i^1$ est un ensemble quarrable avec $\overset{\circ}{Q}_i^1 \cap \overset{\circ}{Q}_j^1 = \emptyset$ si $i \neq j$ de sorte que, maintenant $\overset{\circ}{Q}_i^1 \cap \overset{\circ}{Q}_j^2 = \emptyset$ pour tout $1 \leq i \leq N_1$ et $1 \leq j \leq N_2$. Par conséquent,

$$I^+(f, D_1 \cup D_2) \leq \sum_{i=1}^{N_1} \left(\sup_{Q_i^1} f \right) m(Q_i^1) + \sum_{j=1}^{N_2} \left(\sup_{Q_j^2} f \right) m(Q_j^2)$$

et comme, pour tout $1 \leq i \leq N_1$ il existe un $1 \leq i' \leq N_1$ tel que $Q_i^1 \subset Q_{i'}^1$, on en déduit que

$$I^+(f, D_1 \cup D_2) \leq \sum_{i=1}^{N_1} \left(\sup_{Q_i^1} f \right) m(Q_i^1) + \sum_{j=1}^{N_2} \left(\sup_{Q_j^2} f \right) m(Q_j^2).$$

Par passage à l'inf en $\{Q_i^1\}_{1 \leq i \leq N_1}$ et $\{Q_j^2\}_{1 \leq j \leq N_2}$, il vient

$$I^+(f, D_1 \cup D_2) \leq \int_{D_1} f(x) dx + \int_{D_2} f(x) dx.$$

On en déduit alors que f est Riemann-intégrable sur $D_1 \cup D_2$ car

$$I^+(f, D_1 \cup D_2) = I^-(f, D_1 \cup D_2) = \int_{D_1} f(x) dx + \int_{D_2} f(x) dx$$

et aussi que

$$\int_{D_1 \cup D_2} f(x) dx = \int_{D_1} f(x) dx + \int_{D_2} f(x) dx.$$

2. D'après la Remarque 6.2.2, il suffit donc de montrer la linéarité pour des $\lambda > 0$. Soient $B_1 = \bigcup_{i=1}^{N_1} Q_i^1$ et $B_2 = \bigcup_{i=1}^{N_2} Q_i^2$ des ensembles pavables tels que $D \subset B_1 \subset U$ et $D \subset B_2 \subset U$ et

$$\dot{Q}_i^1 \cap \dot{Q}_j^1 = \dot{Q}_i^2 \cap \dot{Q}_j^2 = \emptyset \text{ si } i \neq j.$$

Alors $B_1 \cap B_2$ est toujours un ensemble pavable de la forme $\bigcup_{i=1}^N Q_j$ avec $\dot{Q}_i \cap \dot{Q}_j = \emptyset$ si $i \neq j$. En particulier, pour tout $1 \leq i \leq N$, il existe $1 \leq i_1 \leq N_1$ et $1 \leq i_2 \leq N_2$ tels que $Q_i \subset Q_{i_1}^1$ et $Q_i \subset Q_{i_2}^2$. Par conséquent,

$$I^+(f + \lambda g, D) \leq \sum_{j=1}^N \left(\sup_{Q_j} (f + \lambda g) \right) m(Q_j) \leq \sum_{j=1}^{N_1} \left(\sup_{Q_j^1} f \right) m(Q_j^1) + \lambda \sum_{j=1}^{N_2} \left(\sup_{Q_j^2} g \right) m(Q_j^2)$$

et par passage à l'inf en $\{Q_i^1\}_{1 \leq i \leq N_1}$ et $\{Q_i^2\}_{1 \leq i \leq N_2}$, il vient

$$I^+(f + \lambda g, D) \leq I^+(f, D) + \lambda I^+(g, D).$$

On montre de façon identique que

$$I^-(f + \lambda g, D) \geq I^-(f, D) + \lambda I^-(g, D).$$

La combinaison des deux inégalités précédentes et le fait que f et g sont toutes deux Riemann-intégrables sur D montre que

$$I^+(f + \lambda g, D) = I^-(f + \lambda g, D) = \int_D f(x) dx + \lambda \int_D g(x) dx$$

et donc que $f + \lambda g$ est Riemann-intégrable sur D avec

$$\int_D [f(x) + \lambda g(x)] dx = \int_D f(x) dx + \lambda \int_D g(x) dx.$$

3. Comme $-|f(x)| \leq f(x) \leq |f(x)|$ pour tout $x \in U$, d'après la monotonie et la linéarité de l'intégrale de Riemann, on obtient que

$$-\int_D |f(x)| dx \leq \int_D f(x) dx \leq \int_D |f(x)| dx,$$

ce qui montre que

$$\left| \int_D f(x) dx \right| \leq \int_D |f(x)| dx.$$

□

6.3 Théorème de Fubini

En dimension $n = 1$, le calcul d'intégrales est lié à la recherche d'une primitive de la fonction que l'on cherche à intégrer. Le calcul d'intégrales multiples se heurte au fait qu'il n'y a pas de notion de primitive en dimension $n \geq 2$. Le(s) théorème(s) de Fubini permet(tent) de ramener le calcul d'une intégrale multiple à celui de plusieurs intégrales simples, et donc d'utiliser les règles classiques de calcul intégral.

Théorème 6.3.1 (Fubini sur un pavé). Soit $P = \prod_{i=1}^n [a_i, b_i]$ un pavé de \mathbb{R}^n et $f : P \rightarrow \mathbb{R}$ une fonction continue sur P . Alors

$$\int_P f(x) dx = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} \cdots \left(\int_{a_{n-1}}^{b_{n-1}} \left(\int_{a_n}^{b_n} f(x_1, \dots, x_n) dx_n \right) dx_{n-1} \right) \cdots dx_2 \right) dx_1$$

et l'ordre d'intégration peut être arbitrairement permuté sans altérer le résultat.

Démonstration. Nous donnons une preuve de ce résultat dans le cas simplifié de la dimension $n = 2$. On note $P = [a, b] \times [c, d]$. Pour tout $k \in \mathbb{N}^*$ et tout $0 \leq i, j \leq k$, on pose

$$a_i := a + i \frac{b-a}{k}, \quad c_j := c + j \frac{d-c}{k},$$

et pour tout $1 \leq i, j \leq k$

$$P_{ij} = [a_{i-1}, a_i] \times [c_{j-1}, c_j],$$

$$m_{ij} = \inf_{P_{ij}} f, \quad M_{ij} = \sup_{P_{ij}} f, \quad \mu_{ij} = \frac{1}{c_j - c_{j-1}} \int_{c_{j-1}}^{c_j} f(a_i, y) dy = \frac{k}{d-c} \int_{c_{j-1}}^{c_j} f(a_i, y) dy.$$

Alors

$$\sum_{i,j=1}^k \frac{(b-a)(d-c)}{k^2} m_{ij} \leq S_k := \sum_{i,j=1}^k \frac{(b-a)(d-c)}{k^2} \mu_{ij} \leq \sum_{i,j=1}^k \frac{(b-a)(d-c)}{k^2} M_{ij}.$$

Comme d'après le Théorème 6.2.3 f est Riemann-intégrable sur P , on en déduit que

$$\lim_{k \rightarrow +\infty} \sum_{i,j=1}^k \frac{(b-a)(d-c)}{k^2} m_{ij} = \lim_{k \rightarrow +\infty} \sum_{i,j=1}^k \frac{(b-a)(d-c)}{k^2} M_{ij} = \iint_P f(x, y) dx dy,$$

ce qui implique que

$$S_k \rightarrow \iint_P f(x, y) dx dy.$$

Par ailleurs,

$$S_k = \sum_{i=1}^k \frac{b-a}{k} \sum_{j=1}^k \int_{c_{j-1}}^{c_j} f(a_i, y) dy = \sum_{i=1}^k \frac{b-a}{k} \int_c^d f(a_i, y) dy.$$

Pour tout $x \in [a, b]$, posons

$$\varphi(x) := \int_c^d f(x, y) dy.$$

Notons que φ est bien définie car f étant continue sur P , alors pour tout $x \in [a, b]$, la fonction partielle $y \mapsto f(x, y)$ est continue sur $[c, d]$ et donc intégrable sur $[c, d]$. Montrons maintenant que

φ est continue sur $[a, b]$. Comme f est continue sur le compact P , elle est uniformément continue. Par conséquent, pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que pour tout $(x, y), (x', y') \in P$ avec $\|(x, y) - (x', y')\| < \delta$, alors $|f(x, y) - f(x', y')| < \varepsilon$. En particulier, pour tout $y \in [c, d]$ et pour tout $x, x' \in [a, b]$ avec $|x - x'| < \delta$, on a $|f(x, y) - f(x', y)| < \varepsilon$. En intégrant cette dernière inégalité par rapport à $y \in [c, d]$ et en appliquant l'inégalité triangulaire, on obtient que

$$|\varphi(x) - \varphi(x')| \leq \int_c^d |f(x, y) - f(x', y)| dy \leq \varepsilon(d - c),$$

ce qui montre que φ est continue sur $[a, b]$. Par définition de l'intégrale de Riemann en dimension $n = 1$, il vient alors que

$$S_k = \sum_{i=1}^k \frac{b-a}{k} \varphi(a_i) \rightarrow \int_a^b \varphi(x) dx = \int_a^b \left(\int_c^d f(x, y) dy \right) dx,$$

puis, par unicité de la limite, que

$$\iint_P f(x, y) dx dy = \int_a^b \left(\int_c^d f(x, y) dy \right) dx.$$

On montre de même que

$$\iint_P f(x, y) dx dy = \int_c^d \left(\int_a^b f(x, y) dx \right) dy,$$

ce qui conclut la preuve du théorème. \square

AVERTISSEMENT : Dans le calcul d'intégrales multiples, il sera essentiel d'écrire explicitement les dx, dy, dz, dt, \dots . En effet, contrairement aux intégrales simples en dimension $n = 1$, une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est susceptible d'être intégrée par rapport à plusieurs variables. L'omission de la mention explicite de la variable d'intégration pourrait donc donner lieu à ambiguïté.

Une version plus générale du théorème de Fubini est donnée par le résultat suivant.

Théorème 6.3.2 (Fubini). Soit $D \subset \mathbb{R}^n$ un ensemble quarrable de la forme

$$D = \{x = (\underbrace{x_1, \dots, x_{n-1}}_{x'}, x_n)^T \in \mathbb{R}^n : x' \in D', \varphi(x') \leq x_n \leq \psi(x')\},$$

où $D' \subset \mathbb{R}^{n-1}$ est un ensemble quarrable et $\varphi, \psi : D' \rightarrow \mathbb{R}$ sont des fonctions continues telles que $\varphi(x') \leq \psi(x')$ pour tout $x' \in D'$. Soit également $f : D \rightarrow \mathbb{R}$ une fonction continue. Alors

$$\int_D f(x) dx = \int_{D'} \left(\int_{\varphi(x_1, \dots, x_{n-1})}^{\psi(x_1, \dots, x_{n-1})} f(x_1, \dots, x_{n-1}, x_n) dx_n \right) dx_1 \cdots dx_{n-1}.$$

Démonstration. De nouveau, nous montrons ce résultat en dimension $n = 2$ où $D = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, \varphi(x) \leq y \leq \psi(x)\}$. L'idée consiste à approcher l'ensemble D par une union finie de pavés. Pour ce faire, pour tout $k \in \mathbb{N}^*$, on subdivise $[a, b]$ en k sous-intervalles de longueur $(b - a)/k$. Pour tout $0 \leq i \leq k$, on pose

$$a_i = a + i \frac{b - a}{k},$$

et on définit les fonctions constantes par morceaux $\varphi_k, \psi_k : [a, b] \rightarrow \mathbb{R}$ par

$$\varphi_k(x) = \sup_{[a_i, a_{i+1}]} \varphi, \quad \psi_k(x) = \inf_{[a_i, a_{i+1}]} \psi, \quad \text{pour tout } x \in [a_i, a_{i+1}[\text{ si } 0 \leq i \leq k-1$$

et $\varphi_k(b) = \varphi(b)$ et $\psi_k(b) = \psi(b)$.

Étape 1. Montrons que φ_k converge uniformément vers φ sur $[a, b]$ et que ψ_k converge uniformément vers ψ sur $[a, b]$. La fonction φ étant continue sur le compact $[a, b]$ elle est uniformément continue. Pour tout $\varepsilon > 0$ il existe donc un $\delta > 0$ tel que pour tout x et $x' \in [a, b]$, si $|x - x'| < \delta$ alors $|\varphi(x) - \varphi(x')| < \varepsilon$. Par ailleurs φ atteint ses bornes sur le compact $[a_i, a_{i+1}]$ et donc en particulier, il existe un $c_i \in [a_i, a_{i+1}]$ tel que

$$\varphi(c_i) = \sup_{[a_i, a_{i+1}]} \varphi = \varphi_k(x) \quad \text{pour tout } x \in [a_i, a_{i+1}[.$$

Soit $k_0 \in \mathbb{N}$ (qui ne dépend que de δ et donc ε) tel que $(b-a)/k_0 < \delta$ et $k \geq k_0$. Pour tout $x \in [a, b]$, il existe un $0 \leq i \leq k-1$ tel que $x \in [a_i, a_{i+1}[$ et donc $|c_i - x| \leq a_{i+1} - a_i = (b-a)/k < \delta$. Par conséquent, $|\varphi(x) - \varphi_k(x)| = |\varphi(x) - \varphi(c_i)| < \varepsilon$ et par passage au sup en $x \in [a, b]$, il vient que

$$\sup_{x \in [a, b]} |\varphi(x) - \varphi_k(x)| \leq \varepsilon, \quad \text{pour tout } k \geq k_0,$$

ce qui montre bien que φ_k converge uniformément vers φ sur $[a, b]$. On démontre de même que ψ_k converge uniformément vers ψ sur $[a, b]$.

Étape 2. Notons

$$D_k = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, \varphi_k(x) \leq y \leq \psi_k(x)\}.$$

Comme les fonctions φ_k et ψ_k sont constantes sur chaque sous-intervalle $[a_i, a_{i+1}[$ de $[a, b]$, on en déduit que D_k est un ensemble pavable de la forme

$$D_k = \bigcup_{i=0}^{k-1} [a_i, a_{i+1}] \times [b_i, b_{i+1}],$$

où $b_i = \sup_{[a_i, a_{i+1}]} \varphi$ et $b_{i+1} = \inf_{[a_i, a_{i+1}]} \psi$. De plus, comme $\varphi(x) \leq \varphi_k(x)$ et $\psi_k(x) \leq \psi(x)$ pour tout $x \in [a, b]$, on en déduit que $D_k \subset D$. Par conséquent,

$$\begin{aligned} |D \setminus D_k| &= \int_a^b (\varphi(x) - \varphi_k(x)) dx + \int_a^b (\psi_k(x) - \psi(x)) dx \\ &\leq (b-a) \sup_{[a, b]} |\varphi - \varphi_k| + (b-a) \sup_{[a, b]} |\psi - \psi_k| \rightarrow 0, \end{aligned}$$

d'après l'étape 1.

Étape 3. Par additivité de l'intégrale double, on a que

$$\int_{D_k} f(x, y) dx dy = \sum_{i=0}^{k-1} \int_{[a_i, a_{i+1}] \times [b_i, b_{i+1}]} f(x, y) dx dy.$$

Par suite, le théorème de Fubini sur un pavé montre que

$$\int_{[a_i, a_{i+1}] \times [b_i, b_{i+1}]} f(x, y) dx dy = \int_{a_i}^{a_{i+1}} \left(\int_{b_i}^{b_{i+1}} f(x, y) dy \right) dx = \int_{a_i}^{a_{i+1}} \left(\int_{\varphi_k(x)}^{\psi_k(x)} f(x, y) dy \right) dx$$

et donc

$$\int_{D_k} f(x, y) dx dy = \sum_{\bar{i}=0}^{k-1} \int_{a_i}^{a_{i+1}} \left(\int_{\varphi_k(x)}^{\psi_k(x)} f(x, y) dy \right) dx = \int_a^b \left(\int_{\varphi_k(x)}^{\psi_k(x)} f(x, y) dy \right) dx. \quad (6.3.1)$$

Étape 4. Comme f est continue sur le compact D , il existe une constante $M > 0$ telle que $|f(x, y)| \leq M$ pour tout $(x, y) \in D$. Par conséquent, d'après l'inégalité triangulaire, le fait que $D_k \subset D$ et l'étape 2, il vient

$$\begin{aligned} \left| \int_D f(x, y) dx dy - \int_{D_k} f(x, y) dx dy \right| &= \left| \int_{D \setminus D_k} f(x, y) dx dy \right| \\ &\leq \int_{D \setminus D_k} |f(x, y)| dx dy \leq M |D \setminus D_k| \rightarrow 0. \end{aligned} \quad (6.3.2)$$

De même, pour tout $x \in [a, b]$, l'inégalité triangulaire montre que

$$\begin{aligned} \left| \int_{\varphi(x)}^{\psi(x)} f(x, y) dy - \int_{\varphi_k(x)}^{\psi_k(x)} f(x, y) dy \right| &= \left| \int_{\varphi(x)}^{\varphi_k(x)} f(x, y) dy + \int_{\psi_k(x)}^{\psi(x)} f(x, y) dy \right| \\ &\leq M \sup_{[a, b]} |\varphi - \varphi_k| + M \sup_{[a, b]} |\psi - \psi_k|, \end{aligned}$$

puis en intégrant par rapport à $x \in [a, b]$, on obtient que

$$\begin{aligned} \left| \int_a^b \left(\int_{\varphi(x)}^{\psi(x)} f(x, y) dy \right) dx - \int_a^b \left(\int_{\varphi_k(x)}^{\psi_k(x)} f(x, y) dy \right) dx \right| \\ \leq \int_a^b \left| \int_{\varphi(x)}^{\psi(x)} f(x, y) dy - \int_{\varphi_k(x)}^{\psi_k(x)} f(x, y) dy \right| dx \\ \leq M(b-a) \sup_{[a, b]} |\varphi - \varphi_k| + M(b-a) \sup_{[a, b]} |\psi - \psi_k| \rightarrow 0, \end{aligned} \quad (6.3.3)$$

d'après l'étape 1. Finalement, par passage à la limite dans (6.3.1) et en utilisant (6.3.2) et (6.3.3), on obtient que

$$\int_D f(x, y) dx dy = \int_a^b \left(\int_{\varphi(x)}^{\psi(x)} f(x, y) dy \right) dx,$$

ce qui conclut la preuve du théorème. \square

AVERTISSEMENT : Dans l'énoncé précédent, l'ordre d'intégration est important. Comme la variable x_n est donnée en fonction des variables x_1, \dots, x_{n-1} , il faut d'abord intégrer par rapport à x_n sur un ensemble qui dépend de x_1, \dots, x_{n-1} . Le résultat du calcul est alors une quantité qui dépend de x_1, \dots, x_{n-1}

$$\int_{\varphi(x_1, \dots, x_{n-1})}^{\psi(x_1, \dots, x_{n-1})} f(x_1, \dots, x_{n-1}, x_n) dx_n$$

que l'on intègre ensuite par rapport aux variables x_1, \dots, x_{n-1} .

Exemple 6.3.3. Calculons $I := \iint_D xy dx dy$ où D est la partie du plan limitée par les paraboles d'équation $y = x^2$ et $x = y^2$. On a donc que $D = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, x^2 \leq y \leq \sqrt{x}\}$ et d'après le théorème de Fubini,

$$I = \int_0^1 \left(\int_{x^2}^{\sqrt{x}} xy dy \right) dx = \int_0^1 x \left(\left[\frac{y^2}{2} \right]_{x^2}^{\sqrt{x}} \right) dx = \int_0^1 x \left(\frac{x}{2} - \frac{x^4}{2} \right) dx = \left[\frac{x^3}{6} - \frac{x^6}{12} \right]_0^1 = \frac{1}{12}.$$

6.4 Formule de changement de variables

Outre le théorème de Fubini, le second outil majeur dans le calcul d'intégrales multiples est la formule de changement de variables que nous admettrons.

Théorème 6.4.1 (Changement de variables). *Soient U et V des ouverts de \mathbb{R}^n et $\Phi : U \rightarrow V$ un \mathcal{C}^1 -difféomorphisme. Soit également $D \subset V$ un ensemble quarrable. Alors $\Phi^{-1}(D) \subset U$ est quarrable et pour toute fonction $f : V \rightarrow \mathbb{R}$ Riemann-intégrable sur D , la fonction $f \circ \Phi$ est Riemann-intégrable sur $\Phi^{-1}(D)$ et on a*

$$\int_D f(y) dy = \int_{\Phi^{-1}(D)} f \circ \Phi(x) |J_\Phi(x)| dx.$$

Remarque 6.4.2. – On écrit symboliquement que $y = \Phi(x)$ et l'élément de volume $dy = |J_\Phi(x)| dx$. Notons que c'est la valeur absolue du déterminant jacobien qui intervient dans l'expression de la formule de changement de variables.

- Soit $\Phi(x) = Ax$ une application linéaire où A est la matrice diagonale $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ avec $\lambda_i \neq 0$ pour tout $1 \leq i \leq n$. Alors A est inversible ce qui assure que $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ est un \mathcal{C}^1 -difféomorphisme. Soit $P = [0, 1]^n$ le pavé unité, alors $y \in \Phi(P)$ si et seulement si $y_i = \lambda_i x_i$ pour tout $1 \leq i \leq n$, où $x \in P$, ce qui est encore équivalent à écrire que $0 \leq y_i \leq \lambda_i$ si $\lambda_i > 0$ et $0 \leq y_i \leq -\lambda_i$ si $\lambda_i < 0$. Autrement dit,

$$\Phi(P) = \prod_{i=1}^n [0, |\lambda_i|]$$

est un pavé et dont la mesure est donnée par

$$m(\Phi(P)) = \prod_{i=1}^n |\lambda_i| = |\det A| = |J_\Phi(x)| m(P).$$

6.4.1 Coordonnées polaires

On définit l'application $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ par $\Phi(r, \theta) = (r \cos \theta, r \sin \theta)$ pour tout $(r, \theta) \in \mathbb{R}^2$. La fonction Φ est de classe \mathcal{C}^1 sur \mathbb{R}^2 , de plus, pour tout $(r, \theta) \in \mathbb{R}^2$, sa matrice jacobienne est donnée par

$$D\Phi(r, \theta) = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix},$$

et donc $J_\Phi(r, \theta) = r$. Par conséquent, pour tout $(r, \theta) \in]0, +\infty[\times \mathbb{R}$, le déterminant jacobien $J_\Phi(r, \theta) > 0$. Enfin si (r, θ) et $(r', \theta') \in]0, +\infty[\times \mathbb{R}$ sont tels que $\Phi(r, \theta) = \Phi(r', \theta')$, alors $r^2 = r'^2$, ce qui implique que $r = r'$. Par suite, on a que $\cos \theta = \cos \theta'$ et $\sin \theta = \sin \theta'$ ce qui implique $\theta = \theta'$ modulo 2π . On en déduit que Φ est injective sur $]0, +\infty[\times]0, 2\pi[$. Le théorème d'inversion globale assure donc que Φ est un \mathcal{C}^1 -difféomorphisme de $]0, +\infty[\times]0, 2\pi[$ sur son image $\mathbb{R}^2 \setminus \{(0, 0)\}$. La formule de changement de variable montre alors que

$$\int_D f(x, y) dx dy = \int_{\Phi^{-1}(D)} f(r \cos \theta, r \sin \theta) r dr d\theta.$$

Exemple 6.4.3. Calculons $I = \iint_D \frac{1}{1+x^2+y^2} dx dy$ où D est le disque fermé de centre $(0, 0)$ et de rayon 1. On a donc que $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$. En effectuant le changement de variables

en coordonnées polaires, on remarque que $(x, y) \in D$ si et seulement si $(r, \theta) \in [0, 1] \times [0, 2\pi]$. Il vient alors que

$$I = \int_{[0,1] \times [0,2\pi]} \frac{r}{1+r^2} dr d\theta.$$

D'après le théorème de Fubini sur un pavé, on peut intégrer d'abord par rapport à θ . Comme l'intégrande est indépendante de θ , on obtient

$$I = 2\pi \int_0^1 \frac{r}{1+r^2} dr = \pi [\ln(1+r^2)]_0^1 = \pi \ln 2.$$

6.4.2 Coordonnées cylindriques

On définit l'application $\Phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ par $\Phi(r, \theta, z) = (r \cos \theta, r \sin \theta, z)$ pour tout $(r, \theta, z) \in \mathbb{R}^3$. La fonction Φ est de classe \mathcal{C}^1 sur \mathbb{R}^3 , de plus, pour tout $(r, \theta, z) \in \mathbb{R}^3$, sa matrice jacobienne est donnée par

$$D\Phi(r, \theta) = \begin{pmatrix} \cos \theta & -r \sin \theta & 0 \\ \sin \theta & r \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

et donc $J_\Phi(r, \theta) = r$. Par conséquent, pour tout $(r, \theta, z) \in]0, +\infty[\times \mathbb{R} \times \mathbb{R}$, le déterminant jacobien $J_\Phi(r, \theta, z) > 0$. Enfin si (r, θ, z) et $(r', \theta', z') \in]0, +\infty[\times \mathbb{R} \times \mathbb{R}$ sont tels que $\Phi(r, \theta, z) = \Phi(r', \theta', z')$, alors évidemment $z = z'$ et on montre comme dans le cas des coordonnées polaires que $r = r'$ et $\theta = \theta'$ modulo 2π . On en déduit que Φ est injective sur $]0, +\infty[\times]0, 2\pi[\times \mathbb{R}$. Le théorème d'inversion globale assure donc que Φ est un \mathcal{C}^1 -difféomorphisme de $]0, +\infty[\times]0, 2\pi[\times \mathbb{R}$ sur son image $\mathbb{R}^3 \setminus ([0, +\infty[\times \{0\} \times \mathbb{R})$. La formule de changement de variable montre alors que

$$\int_D f(x, y, z) dx dy dz = \int_{\Phi^{-1}(D)} f(r \cos \theta, r \sin \theta, z) r dr d\theta dz.$$

Exemple 6.4.4. Soit $D = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 \leq 1, z^2 \leq 4(x^2 + y^2)\}$, on veut calculer l'intégrale triple $I = \iiint_D \sqrt{x^2 + y^2} dx dy dz$. On passe en coordonnées cylindriques et on note que $(x, y, z) \in D$ si et seulement si $r \in [0, 1]$, $\theta \in [0, 2\pi]$ et $-2r \leq z \leq 2r$. Par conséquent, la formule de changement de variable couplée au théorème de Fubini donne

$$I = \int_0^1 \left(\int_0^{2\pi} \left(\int_{-2r}^{2r} r \cdot r dz \right) d\theta \right) dr = 2\pi \int_0^1 4r^3 dr = 2\pi.$$

6.4.3 Coordonnées sphériques

On définit l'application $\Phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ par $\Phi(r, \theta, \varphi) = (r \cos \theta \sin \varphi, r \sin \theta \sin \varphi, r \cos \varphi)$ pour tout $(r, \theta, \varphi) \in \mathbb{R}^3$. La fonction Φ est de classe \mathcal{C}^1 sur \mathbb{R}^3 , de plus, pour tout $(r, \theta, \varphi) \in \mathbb{R}^3$, sa matrice jacobienne est donnée par

$$D\Phi(r, \theta) = \begin{pmatrix} \cos \theta \sin \varphi & -r \sin \theta \sin \varphi & r \cos \theta \cos \varphi \\ \sin \theta \sin \varphi & r \cos \theta \sin \varphi & r \sin \theta \cos \varphi \\ \cos \varphi & 0 & -r \sin \varphi \end{pmatrix},$$

et donc $J_\Phi(r, \theta) = r^2 \sin \varphi$. Par conséquent, pour tout $(r, \theta, \varphi) \in]0, +\infty[\times \mathbb{R} \times]0, \pi[$, le déterminant jacobien $J_\Phi(r, \theta, \varphi) > 0$. Enfin si (r, θ, φ) et $(r', \theta', \varphi') \in]0, +\infty[\times \mathbb{R} \times]0, \pi[$ sont tels que $\Phi(r, \theta, \varphi) = \Phi(r', \theta', \varphi')$, alors $r^2 = r'^2$, soit $r = r'$. De plus, on a $r^2 \sin^2 \varphi = r'^2 \sin^2 \varphi'$ et $r \cos \varphi = r' \cos \varphi'$, ce qui implique, comme φ et $\varphi' \in]0, \pi[$ que $\cos \varphi = \cos \varphi'$ et $\sin \varphi = \sin \varphi'$ soit $\varphi = \varphi'$. Enfin,

comme $\sin \varphi \neq 0$ et $\sin \varphi' \neq 0$, il vient que $\cos \theta = \cos \theta'$ et $\sin \theta = \sin \theta'$ ce qui implique que $\theta = \theta'$ modulo 2π . Par conséquent, le théorème d'inversion globale assure que Φ est un \mathcal{C}^1 -difféomorphisme de $]0, +\infty[\times]0, 2\pi[\times]0, \pi[$ sur son image $\mathbb{R}^3 \setminus ([0, +\infty[\times \{0\} \times \mathbb{R})$. La formule de changement de variable montre alors que

$$\int_D f(x, y, z) dx dy dz = \int_{\Phi^{-1}(D)} f(r \cos \theta \sin \varphi, r \sin \theta \sin \varphi, r \cos \varphi) r^2 \sin \varphi dr d\theta d\varphi.$$

Exemple 6.4.5. Soit $D = \{(x, y, z) \in \mathbb{R}^3 : 1 \leq x^2 + y^2 + z^2 \leq 4, z \geq 0\}$. Calculons l'intégrale triple $I = \iiint_D z dx dy dz$ à l'aide d'un changement de variables en coordonnées sphériques. Remarquons que $(x, y, z) \in D$ si et seulement si $(r, \theta, \varphi) \in [1, 2] \times [0, 2\pi] \times [0, \pi/2]$. Il vient donc par la formule de changement de variables et le théorème de Fubini

$$I = \iiint_{[1,2] \times [0,2\pi] \times [0,\pi/2]} r \cos \varphi \cdot r^2 \sin \varphi dr d\theta d\varphi = 2\pi \left(\int_1^2 r^3 dr \right) \left(\int_0^{\pi/2} \sin \varphi \cos \varphi d\varphi \right) = \frac{15\pi}{4}.$$

Chapitre 7

Intégrales curvilignes et surfaciques

L'objet de ce chapitre est de définir la notion d'intégrale sur des courbes dans le plan ou des surfaces dans l'espace et de relier ces notions aux intégrales multiples introduites au chapitre précédent.

7.1 Intégrale curviligne et formules de Stokes-Ostrogradski

A la section 4.2, nous avons défini ce qu'est un arc orienté de classe \mathcal{C}^1 . Nous allons présenter une notion d'intégrale sur de tels objets géométriques appelée intégrale curviligne.

7.1.1 Longueur d'un arc orienté

Soit $([a, b], \gamma)$ un arc orienté. Pour calculer la longueur L de $\Gamma = \gamma([a, b])$, nous allons adopter le point de vue du numéricien en subdivisant la courbe en un nombre fini de points et approcher la longueur de la courbe entre deux de ces points par leur distance.

Soit donc $a = t_0 < t_1 < \dots < t_N = b$ une subdivision de l'intervalle $[a, b]$. Les points $\gamma(t_0), \dots, \gamma(t_N)$ sont donc des points de Γ . Une propriété classique de géométrie euclidienne affirme que la distance minimale entre deux points est la ligne droite. Par conséquent, la longueur de la courbe Γ qui joint les points $\gamma(t_i)$ et $\gamma(t_{i+1})$ est plus grande que $\|\gamma(t_{i+1}) - \gamma(t_i)\|$. En sommant ces quantités, on obtient que la longueur de Γ est plus grande que

$$\sum_{i=0}^{N-1} \|\gamma(t_{i+1}) - \gamma(t_i)\|.$$

Pour approcher L encore faut-il "faire tendre" la distances entre chacun des points de la subdivision vers zéro, à ceci près que rien n'assure l'existence de la limite. Par contre, l'inégalité précédente ayant lieu pour n'importe quelle subdivision, on peut passer au sup parmi toutes les subdivisions possibles. Ceci motive la définition suivante.

Définition 7.1.1. Soit $([a, b], \gamma)$ un arc orienté. La longueur de $\Gamma = \gamma([a, b])$ est définie par

$$\ell(\Gamma) := \sup \left\{ \sum_{i=0}^{N-1} \|\gamma(t_{i+1}) - \gamma(t_i)\| : N \in \mathbb{N}, a = t_0 < t_1 < \dots < t_N = b \right\}.$$

Proposition 7.1.2. Soit $([a, b], \gamma)$ un arc orienté de classe \mathcal{C}^1 . Alors

$$\ell(\Gamma) = \int_a^b \|\dot{\gamma}(s)\| ds.$$

Démonstration. Soit $a = t_0 < t_1 < \dots < t_N = b$ une subdivision de l'intervalle $[a, b]$. Comme γ est de classe \mathcal{C}^1 , on a pour tout $0 \leq i \leq N-1$,

$$\gamma(t_{i+1}) - \gamma(t_i) = \int_{t_i}^{t_{i+1}} \dot{\gamma}(s) ds$$

et donc d'après l'inégalité triangulaire pour les intégrales vectorielles,

$$\|\gamma(t_{i+1}) - \gamma(t_i)\| \leq \int_{t_i}^{t_{i+1}} \|\dot{\gamma}(s)\| ds,$$

ce qui implique que

$$\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \|\dot{\gamma}(s)\| ds = \int_a^b \|\dot{\gamma}(s)\| ds.$$

Par passage au sup par rapport à toutes les subdivisions possibles de $[a, b]$ dans le membre de gauche de l'inégalité précédente, on obtient que

$$\ell(\Gamma) \leq \int_a^b \|\dot{\gamma}(s)\| ds.$$

Montrons à présent l'inégalité opposée. Comme γ est de classe \mathcal{C}^1 , $\dot{\gamma}$ est uniformément continu sur le compact $[a, b]$. Pour tout $\varepsilon > 0$ il existe donc un $\delta > 0$ tel que pour tout $s, t \in [a, b]$, si $|s - t| < \delta$, alors $\|\dot{\gamma}(s) - \dot{\gamma}(t)\| < \varepsilon$. Soit $a = t_0 < t_1 < \dots < t_N = b$ une subdivision de l'intervalle $[a, b]$ telle que

$$\max_{0 \leq i \leq N-1} (t_{i+1} - t_i) < \delta.$$

D'après le théorème des accroissements finis, pour tout $0 \leq i \leq N-1$, il existe un $s_i \in]t_i, t_{i+1}[$ tel que $\gamma(t_{i+1}) - \gamma(t_i) = \dot{\gamma}(s_i)(t_{i+1} - t_i)$. De plus, pour tout $s \in]t_i, t_{i+1}[$, on a que $|s - s_i| \leq |t_{i+1} - t_i| < \delta$ ce qui montre que $\|\dot{\gamma}(s) - \dot{\gamma}(s_i)\| < \varepsilon$, ou encore

$$\begin{aligned} \left\| \int_{t_i}^{t_{i+1}} \dot{\gamma}(s) ds - \dot{\gamma}(s_i)(t_{i+1} - t_i) \right\| &= \left\| \int_{t_i}^{t_{i+1}} (\dot{\gamma}(s) - \dot{\gamma}(s_i)) ds \right\| \\ &\leq \int_{t_i}^{t_{i+1}} \|\dot{\gamma}(s) - \dot{\gamma}(s_i)\| ds \leq \varepsilon(t_{i+1} - t_i). \end{aligned}$$

Par conséquent,

$$\begin{aligned} \ell(\Gamma) &\geq \sum_{i=0}^{N-1} \|\gamma(t_{i+1}) - \gamma(t_i)\| = \sum_{i=0}^{N-1} \|\dot{\gamma}(s_i)\|(t_{i+1} - t_i) \\ &\geq \sum_{i=0}^{N-1} \left(\int_{t_i}^{t_{i+1}} \|\dot{\gamma}(s)\| ds - \varepsilon(t_{i+1} - t_i) \right) = \int_a^b \|\dot{\gamma}(s)\| ds - \varepsilon(b - a). \end{aligned}$$

Comme $\varepsilon > 0$ est arbitraire, on en déduit que

$$\ell(\Gamma) \geq \int_a^b \|\dot{\gamma}(s)\| ds,$$

ce qui conclut la preuve de la proposition. \square

Remarque 7.1.3. 1. La longueur est indépendante de la paramétrisation. En effet, si $\mu = \gamma \circ \varphi$, où $\varphi : [c, d] \rightarrow [a, b]$ est une fonction strictement croissante et de classe \mathcal{C}^1 , alors d'après la formule de changement de variables,

$$\int_c^d \|\dot{\mu}(s)\| ds = \int_c^d \|\dot{\mu}(\varphi(s))\| \dot{\varphi}(s) ds = \int_a^b \|\dot{\gamma}(t)\| dt.$$

2. La longueur est également indépendante de l'orientation. En effet, si $\mu(t) = \gamma(a + b - t)$, μ est une paramétrisation de Γ en sens inverse car $\mu(a) = \gamma(b)$ et $\mu(b) = \gamma(a)$ et la formule de changement de variable montre de nouveau que

$$\int_a^b \|\dot{\mu}(t)\| dt = \int_a^b \|\dot{\gamma}(a + b - t)\| dt = - \int_b^a \|\dot{\gamma}(s)\| ds = \int_a^b \|\dot{\gamma}(s)\| ds.$$

7.1.2 Intégrale curviligne

Définition 7.1.4. Soient $([a, b], \gamma)$ un arc orienté de classe \mathcal{C}^1 et $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction continue. On définit l'intégrale curviligne de f sur Γ par

$$\int_{\Gamma} f(x, y) dl := \int_a^b g(\gamma(t)) \|\dot{\gamma}(t)\| dt.$$

Remarque 7.1.5. Tout comme la longueur d'une courbe, l'intégrale curviligne ne dépend ni de la paramétrisation, ni de l'orientation de la courbe.

Grâce aux propriétés classiques de l'intégrale simple on peut montrer facilement les propriétés suivantes.

Propriétés 7.1.6. Soient (I, γ) un arc orienté de classe \mathcal{C}^1 , $\Gamma = \gamma(I)$, $f, g : \mathbb{R}^2 \rightarrow \mathbb{R}$ des fonctions continues et $\lambda \in \mathbb{R}$.

1. *Linéarité.*

$$\int_{\Gamma} [f(x, y) + \lambda g(x, y)] dl = \int_{\Gamma} f(x, y) dl + \lambda \int_{\Gamma} g(x, y) dl.$$

2. *Monotonie.* Si $f(x, y) \leq g(x, y)$ pour tout $(x, y) \in \Gamma$, alors

$$\int_{\Gamma} f(x, y) dl \leq \int_{\Gamma} g(x, y) dl.$$

3. *Additivité.* Si $\Gamma = \Gamma_1 \cup \Gamma_2$ avec $\Gamma_1 \cap \Gamma_2 = \emptyset$, alors

$$\int_{\Gamma} f(x, y) dl = \int_{\Gamma_1} f(x, y) dl + \int_{\Gamma_2} f(x, y) dl.$$

Définition 7.1.7. Soit $([a, b], \gamma)$ un arc orienté de classe \mathcal{C}^1 , $\Gamma = \gamma([a, b])$. On note $\tau(x, y) = \frac{\dot{\gamma}(t)}{\|\dot{\gamma}(t)\|}$ le vecteur tangent unitaire à Γ au point $(x, y) = \gamma(t)$. Si $V : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ est un champ de vecteurs continu, la *circulation* de V le long de Γ est définie par l'intégrale curviligne de la composante tangentielle de V sur Γ :

$$\int_{\Gamma} V(x, y) \cdot \tau(x, y) dl := \int_a^b V(\gamma(t)) \cdot \dot{\gamma}(t) dt.$$

AVERTISSEMENT : Bien qu'étant définie à l'aide de l'intégrale curviligne, la circulation d'un champ de vecteur dépend de l'orientation de l'arc car il dépend de la tangente. En effet si l'on définit $\mu(t) = \gamma(a + b - t)$ pour tout $t \in [a, b]$, alors μ est une paramétrisation de Γ en sens inverse et un simple changement de variable montre que

$$\int_a^b V(\mu(t)) \cdot \dot{\mu}(t) dt = - \int_a^b V(\gamma(t)) \cdot \dot{\gamma}(t) dt.$$

7.1.3 Formules de Stokes-Ostrogradski

Les formules de Stokes-Ostrogradski sont d'autres outils permettant de calculer des intégrales doubles à l'aide d'intégrales curvilignes.

Théorème 7.1.8 (Formule de Stokes). *Soit D un ensemble quarrable dont la frontière $\Gamma = \partial D$ est une courbe fermée orientée de classe \mathcal{C}^1 dans le sens direct. Soit également $V = (v_1, v_2) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ un champ de vecteurs de classe \mathcal{C}^1 . Alors*

$$\iint_D \text{rot } V(x, y) dx dy = \int_{\Gamma} V(x, y) \cdot \tau(x, y) dl.$$

Démonstration. Nous démontrons le théorème dans le cas simplifié où

$$\begin{aligned} D &= \{(x, y) \in \mathbb{R}^2 : a_1 \leq x \leq b_1, \alpha_1(x) \leq y \leq \beta_1(x)\} \\ &= \{(x, y) \in \mathbb{R}^2 : a_2 \leq y \leq b_2, \alpha_2(y) \leq x \leq \beta_2(y)\}, \end{aligned}$$

où $a_1, a_2, b_1, b_2 \in \mathbb{R}$ et $\alpha_1, \alpha_2, \beta_1$ et $\beta_2 : \mathbb{R} \rightarrow \mathbb{R}$ sont des fonctions de classe \mathcal{C}^1 . On rappelle que $\text{rot } V(x, y) = \frac{\partial v_2}{\partial x}(x, y) - \frac{\partial v_1}{\partial y}(x, y)$ pour tout $(x, y) \in \mathbb{R}^2$. On calcul l'intégrale double à l'aide du théorème de Fubini

$$\iint_D \frac{\partial v_1}{\partial y}(x, y) dx dy = \int_{a_1}^{b_1} \left(\int_{\alpha_1(x)}^{\beta_1(x)} \frac{\partial v_1}{\partial y}(x, y) dy \right) dx = \int_{a_1}^{b_1} (v_1(x, \beta_1(x)) - v_1(x, \alpha_1(x))) dx.$$

Par ailleurs, notons Γ_{α_1} et Γ_{β_1} les deux portions de Γ définis par les paramétrisation $\gamma_{\alpha_1}(t) = (t, \alpha_1(t))$ et $\gamma_{\beta_1}(t) = (t, \beta_1(t))$ pour tout $t \in [a_1, b_1]$. L'arc Γ_{α_1} est orienté dans le sens direct, alors que Γ_{β_1} est orienté dans le sens indirect. Par conséquent, comme Γ est orienté dans le sens direct, il vient

$$\int_{\Gamma} v_1(x, y) \tau_1(x, y) dl = \int_{a_1}^{b_1} v_1(t, \alpha_1(t)) dt - \int_{a_1}^{b_1} v_1(t, \beta_1(t)) dt = - \iint_D \frac{\partial v_1}{\partial y}(x, y) dx dy. \quad (7.1.1)$$

On montre de même que

$$\int_{\Gamma} v_2(x, y) \tau_2(x, y) dl = \iint_D \frac{\partial v_2}{\partial x}(x, y) dx dy, \quad (7.1.2)$$

et le résultat suit en additionnant (7.1.1) et (7.1.2). \square

AVERTISSEMENT : Il est important de noter la dissymétrie des rôles joués par les variables x et y par la présence du signe moins dans l'intégrale double. Dans le cas d'une courbe orientée dans le sens indirect il faut changer le signe devant l'intégrale curviligne.

Théorème 7.1.9 (Formule d'Ostrogradski). *Sous les mêmes hypothèses que le théorème de Stokes, on a*

$$\iint_D \text{div } V(x, y) dx dy = \int_{\Gamma} V(x, y) \cdot \nu(x, y) dl,$$

où $\nu(x, y) = \frac{1}{\|\dot{\gamma}(t)\|} (\dot{y}(t), -\dot{x}(t))^T$ est la normale unitaire extérieure à Γ au point $(x, y) = \gamma(t)$.

Démonstration. Il suffit d'appliquer la formule de Stokes au champ de vecteur $\tilde{V} : (x, y) \mapsto (-v_2(x, y), v_1(x, y))$ et de remarquer que $\tilde{V}(x, y) \cdot \tau(x, y) = V(x, y) \cdot \nu(x, y)$ pour tout $(x, y) \in \Gamma$. \square

Exemple 7.1.10. Soit Γ la courbe fermée orientée dans le sens direct, constituée des deux portions de courbes comprises entre les points d'intersection des paraboles d'équations $x = y^2$ et $y = x^2$. Soit également le champ de vecteurs $V : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ défini par $V(x, y) = (xy, 0)$ pour tout $(x, y) \in \mathbb{R}^2$. On se propose de calculer de deux manières différentes la circulation de V le long de Γ .

- Tout d'abord, l'application $\gamma : [0, 2] \rightarrow \mathbb{R}^2$ définie par $\gamma(t) = (t, t^2)$ si $0 \leq t \leq 1$ et $\gamma(t) = ((2-t)^2, 2-t)$ si $1 \leq t \leq 2$ est une paramétrisation de Γ . Par conséquent, la définition de la circulation de V donne

$$\int_{\Gamma} V(x, y) \cdot \tau(x, y) dl = \int_0^2 V(\gamma(t)) \cdot \dot{\gamma}(t) dt = \int_0^1 t^3 dt - 2 \int_1^2 (2-t)^4 dt = -\frac{3}{20}.$$

- On cherche par ailleurs à appliquer la formule de Stokes. Pour ce faire, on note $D := \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, x^2 \leq y \leq \sqrt{x}\}$ le domaine du plan délimité par la courbe Γ . La formule de Stokes et le théorème de Fubini montrent alors que

$$\begin{aligned} \int_{\Gamma} V(x, y) \cdot \tau(x, y) dl &= \iint_D \operatorname{rot} V(x, y) dx dy \\ &= \int_0^1 \left(\int_{x^2}^{\sqrt{x}} -x dy \right) dx = - \int_0^1 (x^{3/2} - x^3) dx = -\frac{3}{20}. \end{aligned}$$

7.2 Intégrale surfacique et formule de la divergence

Dans cette section, nous introduisons les notions d'aire et d'intégrale de surface de façon analogue au point de vue utilisé pour les courbes. Toutefois, la difficulté technique nécessaire pour rendre rigoureuses les notions introduites ci-après nous conduit à ne présenter que formellement les outils en question.

7.2.1 Aire d'une surface

Soit (D, Φ) et $\Sigma = \Phi(D)$ une surface orientée de classe \mathcal{C}^1 et $(s, t) \in D$. Nous désignons par R un rectangle élémentaire contenu dans D de côtés (s, t) , $(s + \delta s, t)$, $(s, t + \delta t)$ et $(s + \delta s, t + \delta t)$. D'après la formule de Taylor, on a pour tout $(u, v) \in [0, 1]^2$,

$$\Phi(s + u\delta s, t + v\delta t) = \Phi(s, t) + u\delta s \frac{\partial \Phi}{\partial s}(s, t) + v\delta t \frac{\partial \Phi}{\partial t}(s, t) + o(\|(\delta s, \delta t)\|).$$

Par conséquent, au premier ordre, l'image de R par Φ est "proche" du parallélogramme d'origine $\Phi(s, t)$ engendré par les vecteurs

$$\delta s \frac{\partial \Phi}{\partial s}(s, t) \quad \text{et} \quad \delta t \frac{\partial \Phi}{\partial t}(s, t)$$

et dont l'aire est donnée par

$$\delta s \delta t \left\| \frac{\partial \Phi}{\partial s}(s, t) \wedge \frac{\partial \Phi}{\partial t}(s, t) \right\|.$$

Il en résulte que si l'on subdivise D en petits rectangles de côté δs et δt , que l'on somme sur tous les rectangles et enfin que l'on fait tendre δs et δt vers zéro, on obtient l'aire de la surface Σ .

Définition 7.2.1. Soit (D, Φ) et $\Sigma = \Phi(D)$ une surface orientée de classe \mathcal{C}^1 où $D \subset \mathbb{R}^2$ est un ensemble quarrable. L'aire de $\Sigma = \Phi(D)$ est définie par

$$\sigma(\Sigma) := \iint_D \left\| \frac{\partial \Phi}{\partial s}(s, t) \wedge \frac{\partial \Phi}{\partial t}(s, t) \right\| ds dt.$$

Remarque 7.2.2. L'aire d'une surface est indépendante de la paramétrisation et de l'orientation. En effet, si (D_1, Φ_1) est une surface orientée de classe \mathcal{C}^1 et $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ un \mathcal{C}^1 -difféomorphisme tel que $\varphi(D_2) = D_1$ et $J_\varphi(s, t) \neq 0$ pour tout $(s, t) \in D_2$, on définit la fonction $\Phi_2 : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ par $\Phi_2 := \Phi_1 \circ \varphi$ de sorte que (D_2, Φ_2) définit la même surface orientée que (D_1, Φ_1) . La formule de différentiation des fonctions composées implique que pour tout $(s, t) \in D_2$,

$$\frac{\partial \Phi_2}{\partial s}(s, t) \wedge \frac{\partial \Phi_2}{\partial t}(s, t) = J_\varphi(s, t) \left[\frac{\partial \Phi_1}{\partial u}(\varphi(s, t)) \wedge \frac{\partial \Phi_1}{\partial v}(\varphi(s, t)) \right],$$

il vient

$$\iint_{D_2} \left\| \frac{\partial \Phi_2}{\partial s}(s, t) \wedge \frac{\partial \Phi_2}{\partial t}(s, t) \right\| ds dt = \iint_{D_2} \left\| \frac{\partial \Phi_1}{\partial u}(\varphi(s, t)) \wedge \frac{\partial \Phi_1}{\partial v}(\varphi(s, t)) \right\| |J_\varphi(s, t)| ds dt$$

et en effectuant le changement de variables $(u, v) = \varphi(s, t)$, on en déduit que

$$\iint_{D_2} \left\| \frac{\partial \Phi_2}{\partial s}(s, t) \wedge \frac{\partial \Phi_2}{\partial t}(s, t) \right\| ds dt = \iint_{D_1} \left\| \frac{\partial \Phi_1}{\partial u}(u, v) \wedge \frac{\partial \Phi_1}{\partial v}(u, v) \right\| du dv.$$

7.2.2 Intégrale de surface

Définition 7.2.3. Soient (D, Φ) et $\Sigma = \Phi(D)$ une surface orientée de classe \mathcal{C}^1 où $D \subset \mathbb{R}^2$ est un ensemble quarrable, et $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ une fonction continue. On définit l'intégrale de surface de f sur Σ par

$$\iint_\Sigma f(x, y, z) d\sigma := \iint_D f(\Phi(s, t)) \left\| \frac{\partial \Phi}{\partial s}(s, t) \wedge \frac{\partial \Phi}{\partial t}(s, t) \right\| ds dt.$$

Remarque 7.2.4. De nouveau, la formule de changement de variable montre que l'intégrale de surface est invariante par changement de paramétrisation qui préserve l'orientation.

Grâce aux propriétés classiques de l'intégrale double on peut montrer les propriétés suivantes.

Propriétés 7.2.5. Soient (D, Φ) une surface orientée de classe \mathcal{C}^1 où $D \subset \mathbb{R}^2$ est un ensemble quarrable, $\Sigma = \Phi(D)$, $f, g : \mathbb{R}^3 \rightarrow \mathbb{R}$ des fonctions continues et $\lambda \in \mathbb{R}$.

1. *Linéarité.*

$$\iint_\Sigma [f(x, y, z) + \lambda g(x, y, z)] d\sigma = \iint_\Sigma f(x, y, z) d\sigma + \lambda \iint_\Sigma g(x, y, z) d\sigma$$

2. *Monotonie.* Si $f(x, y, z) \leq g(x, y, z)$ pour tout $(x, y, z) \in \Sigma$, alors

$$\iint_\Sigma f(x, y, z) d\sigma \leq \iint_\Sigma g(x, y, z) d\sigma.$$

3. *Additivité.* Si $\Sigma = \Sigma_1 \cup \Sigma_2$ avec $\Sigma_1 \cap \Sigma_2 = \emptyset$, alors

$$\iint_\Sigma f(x, y, z) d\sigma = \iint_{\Sigma_1} f(x, y, z) d\sigma + \iint_{\Sigma_2} f(x, y, z) d\sigma.$$

7.2.3 Formule de la divergence

Nous avons vu dans le théorème 7.1.9 la formule d'Ostrogradski qui relie le flux d'un champ de vecteurs sur une courbe de \mathbb{R}^2 à l'intégrale (double) de la divergence de ce même champ de vecteur sur la partie du plan délimitée par cette courbe. Nous allons à présent généraliser cette formule en dimension $n = 3$.

Théorème 7.2.6 (Formule de la divergence). *Soient $\Omega \subset \mathbb{R}^3$ un ensemble quarrable tel que son bord $\partial\Omega = \Sigma$ est une surface orientée de classe \mathcal{C}^1 . Soit également $V : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ un champ de vecteurs de classe \mathcal{C}^1 . Alors*

$$\iiint_{\Omega} \operatorname{div} V(x, y, z) \, dx \, dy \, dz = \iint_{\Sigma} V(x, y, z) \cdot N(x, y, z) \, d\sigma,$$

où $N(x, y, z)$ est la normale unitaire sortante à Σ en (x, y, z) définie (si Φ est une paramétrisation de Σ) par

$$N(x, y, z) := \frac{\frac{\partial\Phi}{\partial s}(s, t) \wedge \frac{\partial\Phi}{\partial t}(s, t)}{\left\| \frac{\partial\Phi}{\partial s}(s, t) \wedge \frac{\partial\Phi}{\partial t}(s, t) \right\|} \quad \text{pour } (x, y, z) = \Phi(s, t).$$